

## Quality of papers, gender balance, accessibility, environmental sustainability

### Editorial board

---

JLIS.it, in full respect of the originality, rigor, and scientific quality of each of the papers published, places among its constitutive values the dissemination of science within the broad context of open science. This mission intends to guarantee conditions of accessibility and equitable results to researchers, collaborators, reviewers, board members, and readers, regardless of their affiliation, gender, age, geographical origin, or economic condition.

JLIS.it aims to publish essays of high quality and accessibility for all, reducing the criticalities of an international publishing system in which some researchers are marginalized, discriminated, and isolated as belonging to personal, social, and geographical areas that are not included in the typical western canons. The journal's philosophy aims not only to the full accessibility to papers (i.e. the possibility of reading, downloading or sharing the research papers), but it wants to participate in constructing a democratic process aimed at generating inclusive, accessible, and open knowledge. Our belief is that diversity, equity, and inclusion guarantee impartiality in editorial processes and represent an element of strengthening and consolidating knowledge.

Furthermore, JLIS.it aims to become more sustainable by creating a system of projects, actions, and processes that reduce inequalities and discrimination, promoting environmental sustainability actions to face the global ecological crisis; therefore, it fully adheres to the objectives of the UN 2030 Agenda.

These purposes will be pursued by adopting guidelines for gender balance and inclusion in editorial processes and products and the signing and publishing declarations and policies for authors, editors, and reviewers that the magazine undertakes to adopt in 2022.

JLIS.it also aims to increase and improve the accessibility of its website and editorial products so that the contents (articles, tables, metadata, supplementary materials, media, etc.) can be entirely usable by those who use adaptive technologies, goals that will require an additional effort by the board and editorial staff.

Looking at the near future, JLIS.it assumes knowledge as a common good, peer review, open access, gender diversity, and environmental sustainability as its founding and development principles and tools to guarantee quality. JLIS.it will undertake a series of actions and projects built and shared with the scientific communities, the paradigm of a renewed editorial commitment.

## Qualità dei contributi, equilibrio di genere, accessibilità, sostenibilità ambientale

JLIS.it, nel pieno rispetto dell'originalità del rigore e della qualità scientifica dei singoli contributi pubblicati, pone tra i suoi valori costitutivi la diffusione della scienza nell'ampia accezione dell'*open science*. Tale missione intende garantire condizioni di accessibilità e risultati equi a ricercatrici e ricercatori, collaboratrici e collaboratori, *reviewer*, membri dei comitati, lettrici e lettori indipendentemente dalla loro affiliazione, genere, età, provenienza geografica o condizione economica.

JLIS.it si pone l'obiettivo di pubblicare saggi che siano di elevata qualità e accessibilità per tutti, riducendo le criticità di un sistema editoriale internazionale in cui alcuni ricercatrici e ricercatori sono emarginati, discriminati e isolati in quanto appartenenti a dimensioni personali, sociali e geografiche non inserite nei canoni comuni occidentali. Infatti, la filosofia della rivista non riguarda solo l'accessibilità ai contributi (cioè la possibilità di leggere, scaricare o condividere gli articoli), ma partecipa alla costruzione di un processo democratico volto a generare una conoscenza inclusiva, accessibile e aperta, nella convinzione che *diversità, equità e inclusione* siano garanzia di imparzialità nei processi editoriali ed elemento di rafforzamento e consolidamento della conoscenza. JLIS.it, inoltre, mira a diventare più sostenibile tramite la messa a sistema di progetti, azioni e processi che riducano le disuguaglianze e le discriminazioni, promuovendo azioni di sostenibilità ambientale per fronteggiare la crisi ecologica globale; pertanto, essa aderisce pienamente agli obiettivi dell'*Agenda ONU 2030*.

Tali finalità saranno perseguite con l'adozione di linee guida per l'equilibrio di genere e l'inclusione nei processi e prodotti editoriali, nonché con la sottoscrizione e pubblicazione di dichiarazioni e policy per autori, redattori e revisori che la rivista si impegna ad adottare nel corso del 2022.

JLIS.it si prefigge, inoltre, di aumentare e migliorare l'accessibilità del sito web e dei prodotti editoriali, affinché i contenuti (articoli, tabelle, metadati, materiali supplementari, media ecc.) possano essere interamente fruibili da coloro che utilizzano tecnologie adattive e assistenziali, obiettivi che richiederanno un ulteriore sforzo da parte del board e della redazione.

Guardando, quindi, a un futuro prossimo, JLIS.it assume come principi fondativi e di sviluppo la *conoscenza come bene comune*, la *peer review*, l'*open access*, la diversità di genere e la sostenibilità ambientale, intesi come strumenti e garanzia di qualità; e intraprenderà una serie di azioni e progetti, costruiti e condivisi con le comunità scientifiche, che segneranno il paradigma di un rinnovato impegno editoriale.

## The updating of ISBD and its transformation

Elena Escolano Rodríguez<sup>(a)</sup>

a) Ministerio de Economía, Industria y Competitividad. España, <https://orcid.org/0000-0001-8296-9859>

---

**Contact:** Elena Escolano Rodríguez, [elena.escolano@yahoo.es](mailto:elena.escolano@yahoo.es)

**Received:** 20 January 2022; **Accepted:** 28 February 2022; **First Published:** 15 May 2022

---

### ABSTRACT

This article describes the difficult work of revising the ISBD standard. It reviews the previous work and the conceptual and structural problems found in the process of updating the ISBD standard. After the recognition of the difficulties in debates, two points of view were distinguished among the members of the group that ended with the split of the ISBD Editorial Group into two subgroups. The article also gives an account of the changes carried out in the updating of the content of ISBD, to end with the discussion topics of the second working group that will change the structure of ISBD as we know it today.

### KEYWORDS

ISBD; International Standard Bibliographic Description; IFLA LRM; IFLA Library Reference Model; manuscripts; celestial cartographic resources.

## Introduction

The International Standard Bibliographic Description (ISBD) is a standard that was created and is maintained by the International Federation of Library Associations and Institutions (IFLA), intended to serve as the principal standard to promote universal bibliographic control; that is, to make universally and promptly available, in a form that is internationally acceptable, basic bibliographic data.

A long time has passed since 2011, when the ISBD Consolidated Edition was published, more time than what is stipulated for the revision of IFLA standards every five years. At that time, when it was close to being published, some last minute requests were received from user communities that could not be taken into account at that time and were left for the next review, believing it would be in five years.

The work to prepare the next revision did not start immediately, as the ISBD Review Group was busy with other parallel work that would adapt ISBD to the more developed technological environment and that would allow ISBD elements to be published as linked data. The representation of the ISBD in the W3C Resource Description Framework (RDF) standard enables its use as Linked Open Data (LOD) and enhances developments of Semantic Web bibliographic tools and services (IFLA 2021). In addition, the necessary guidelines were created to help the application in the libraries in the way that best suits their possibilities and resources (IFLA ISBD Linked Data Study Group 2016). All these works contribute to pursuing an IFLA's fundamental objective, which is to help in the evolution and adaptation of all types of libraries to technological developments.

The ISBD Consolidated edition, published in 2011, was compliant and aligned with the bibliographic model FRBR (IFLA Study Group on the Functional Requirements for Bibliographic Records 1998). However, the IFLA FRBR Review Group was working on consolidating the family of models into the Library reference model LRM (IFLA 2017). For this reason, a Task Group for the Analysis of the Alignment and Impact of IFLA LRM to ISBD was established in December 2016. This group had to develop a mapping between both standards and make a final report to the Committee on Standards on the applicability of LRM in practice before its final approval, and also to prove the consistency among the IFLA standards ((IFLA ISBD Review Group 2017). This work was also considered necessary as preparation for the revision of ISBD.

The Task Group worked on a document with the title at that moment *Alignment of the ISBD element set with the IFLA LRM element set* based on the ISBD Consolidated edition (2011). Although approved by the FRBR Review Group and the Committee on Standards, this document received a set of comments during the World Library and Information Congress (WLIC) 2017 that made it appropriate to make some final adjustments in the *Introduction* to solve the concerns expressed about using the term alignment, to be finally intended mapping. The document, finally with some amendments, was approved and published as a *Mapping* in its final version in 2017 and amended in 2018 (IFLA ISBD Review Group. Task Group for the Analysis of the Alignment and Impact of IFLA LRM on ISBD 2017a). The first aim of this conceptual mapping was to keep consistency between two major IFLA standards, and the second aim was to help the ISBD Review Group in the work of the revision of the ISBD taking into account the LRM model, and also to help the readers by better understanding the semantic relationships between the ISBD data elements and

the attributes and relationships defined in the IFLA LRM model. This document was fundamental to address some concerns expressed by professionals on the alignment or correspondence of ISBD with LRM.

The mapping helped to conclude the applicability of LRM requested by IFLA, as was reported (IFLA ISBD Review Group. Task Group for the Analysis of the Alignment and Impact of IFLA LRM on ISBD 2017b, 5): “LRM is applicable and that ISBD is a valid and needed extension of LRM”. In addition, the report presented different possibilities and some recommendations to be taken in the revision of ISBD, besides the elucidation of the debates carried out on the different understandings of the Manifestation concept in both standards due to the Scope note of the LRM Manifestation Statement attribute: “The manifestation statement attribute is a statement that is usually transcribed from a source or sources present in exemplars of a manifestation. Transcription conventions are codified by each implementation.” (IFLA 2017, 49). The selected wording in the English scope note that defines the way in which the entity should be described was controversial in other languages. English is IFLA’s official working language, but the meaning must be clear so that it would be understandable without any doubt in all languages. Clarifying this, it was possible to agree that the concept was the same with some necessary adjustments, because “until ISBD is extended to include non-published resources, singleton manifestations are outside the scope of ISBD and of this alignment.” (IFLA ISBD Review Group. Task Group for the Analysis of the Alignment and Impact of IFLA LRM on ISBD 2017b, 3).

This same Task Group was also charged with the task of investigating the possible scenarios of revision of ISBD. A four-year work plan in two stages was established for the revision of ISBD that was approved by the Committee on Standards (IFLA ISBD Review Group 2018b). This work plan presented two scenarios to fulfil the objectives of revising the content of the standard to satisfy the community’s needs, and to align with the overarching conceptual model LRM and the new principles included in the updated Statement of International Cataloguing Principles. The two scenarios organized the revision of ISBD step by step, first to align with the model focusing on the Manifestation entity, mainly on the attribute *Manifestation statement*; then to focus on the gap of the ISBD elements excluded for not being considered to belong to the entity or attribute mentioned, but that were considered mandatory or mandatory if available in ISBD. After finishing this process, the revision could be considered finished and it could be published. The second stage to engage in a larger revision, encompassing all the IFLA LRM components, and re-examining ISBD as a full implementation of the LRM could be taken into account. The structure of the future standard would depend on that decision.

## Background and reasons for the delay in the updating work

in 2018, the ISBD Editorial Group (IEG) was finally established with the aim of initiating the review and update. Due to the different approaches of the members of the group, the IEG was reformed in its composition in May 2019 (IFLA ISBD Review Group 2018a). There were problems encountered in working at the same time on the double objectives of:

- Aligning the ISBD with LRM to keep a consistency between IFLA standards by providing the overarching conceptual model with an ISBD implementation.

- Ensuring that the ISBD continues to fulfil the needs of its user communities as expressed by the proposals submitted to the ISBD RG.

In 2019 the presentation at the 85th IFLA General Conference in Athens, Greece recognized that from the beginning of the group, there were different opinions and positions of its members on how this review should be undertaken, the format and structure, the priority issues, etc.; there was not agreement on what objective should prevail and influence the other, the updating of the content or the structure to adapt to LRM. This presentation also accounted for all the main decisions made by the IEG on the structure and other fundamental issues in ISBD, such as the following ones:

- Areas and component statements were also included. Full ISBD areas are aligned as more specific (<) than the LRM Manifestation statement.
- ISBD syntax encoding schemes are needed for statements including many elements. The reason is the added meaning that the ISBD's order of elements and punctuation provide to the information given that justify its declaration in RDF. Encoding patterns are not the transcribed statements themselves, but rather the meta information about the format. ISBD, as rules to which these schemes belong, is a work. It was decided that encoding schemes that will be subtyped by areas and so on as necessary could be included as a sub-type class of RES.
- To preserve the nature of ISBD (Universal Bibliographic Control) when different options are possible, the one(s) where bibliographic control is not assured should not be used or recommended, but should be discarded (deprecated).
- ISBD as an implementation of LRM, but keeping the ISBD granularity; that is, ISBD as an extension of LRM. Therefore, it was decided to start from ISBD granularity (rather than adopting some opinions that advocated for ISBD being more general).
- After a request to the BCM Review Group for clarification of the wording used in the LRM definition of the attribute Manifestation statement, it was concluded by the BCM Group that there were no relevant differences in the concepts of that attribute in the two standards.
- Instead, the concept of Manifestation in both standards is different in that:
  - The LRM concept of Manifestation includes "... multiple items from different manifestations are physically combined or joined (books or pamphlets bound together, audio tapes spliced together, etc.) the result is a new singleton manifestation."
  - ISBD in its current version considers "manifestation" as a published resource, while LRM encompasses any manifestation, published or not" – Therefore the concept of manifestation should be adapted to the LRM Manifestation, with unpublished resources added" (Escolano 2019, 2–3)

The report inferred that it was not necessary to replicate the LRM structure and presentation, but to explain how to use it at a practical level, adapting the terminology and language to the context and type of the content standard, that is applying Ranganathan's Principle of local variation (Ranganathan 1964, 65–70).

At the end of the IFLA WLIC 2019, and during the day long meeting of the IEG and ISBD Review Group, it was decided these two objectives were to be proceeded to 'subsequently' or in

parallel by sub-groups. To advance the two objectives simultaneously, it was decided to work in two parallel tracks with clear terms of references and deliverables each, with work carried out by two separate Task Forces (IFLA ISBD Review Group 2019):

- **ISBD Content Update Task Force**, to produce an Update of the content of ISBD Consolidated Edition 2011 to provide for the urgent needs of communities of users.
- **ISBD for Manifestation Task Force**, to work out an alignment of the ISBD with LRM on the manifestation level and produce a document to reflect on the future ISBD.

The reorganization into two Task Forces would allow both tracks of the Revision to proceed in parallel and will ensure that the original revision timeframe could be met, or with a minor delay. It would also permit greater focus on the tasks.

Obviously, the other way around of the previous mapping should be made; that is, from LRM entities to ISBD, which could lead to a very different result.

## ISBD content update

The ISBD Content Update Task Force was formed with specialists on the subjects proposed for revision, including representatives from the Rare Books and Special Collections Section (RBSCS) of IFLA. Members represented six countries from four continents. Also, many specialists have been consulted and have contributed during all the process. They are from many different countries and institutions and are mentioned with thanks in the Introduction of the update.

This revision or updating of the content of the standard had to answer the urgent needs expressed by the ISBD communities of users. It is for this reason that the fundamental International Cataloguing Principle guiding the content update is: 2.5. Sufficiency and necessity – “Those data elements that are required to: facilitate access for all types of users, including those with specific needs; fulfil the objectives and functions of the catalogue; and describe or identify entities, should be included” (IFLA Cataloguing Section and IFLA Meeting of Experts on an International Cataloguing Code 2016). The needs for information of specialized libraries and users are covered by this principle.

The work plan was: to update ISBD content according to comments and proposals received from ISBD communities of users and: to expand coverage to include other resources, especially unpublished resources, to meet the requests from this community but also to align with the LRM Manifestation entity.

The results to be delivered: a draft of the complete revision of ISBD Consolidated Edition, published in 2011, and a document indicating the modifications that had been carried out. However, the Working Group decided to reflect the modifications in the text itself in red, facilitating identification by the user.

In the forementioned Work Plan for the revision of the ISBD 2018-2022, the first step of the Revision of ISBD is to provide an update of the content and extend the coverage of the ISBD to include a larger range of resources. Therefore, the main types of resources revised have been:

- Unpublished resources
- Celestial cartographic resources
- Component parts

To achieve this, it was necessary to clarify some existing ambiguities, introduce some elements due to the inclusion of new types of resources, review some stipulations and definitions, and also adapt the mandatory level of some elements.

## Unpublished resources

Until now, ISBD has been the standard for the description of published resources. There were no specific rules for manuscripts, despite cataloguers of manuscripts applying ISBD as a general standard. Cataloguers of this type of resource were forced to adapt their descriptions to more general rules and, frequently, to make their own decisions, which could be inconsistent with others with the consequent impacts on interoperability. The categorization provided by ISBD with the content form and media type didn't allow for the searching capabilities of the system to find and select this type of resource in an integrated catalogue. The necessary elements to record the way in which the content is fixed in a carrier and the method used to carry out the production were missing.

The inclusion of unpublished resources in ISBD was requested when the Consolidated edition of 2011 was very much advanced and prepared for publication, so it was not possible to include them at that time. However, a study group was established in 2010 with this target and the group concluded it was feasible to include unpublished resources in ISBD.

The updating of the standard now includes the typology of unpublished resources that covers all types of manuscripts created by any means, even electronic, existing in a library collection. This point should be clear as it was accepted under the condition that the ISBD not collide with the archival standards that should be used for archival collections. As it is the different granularity needed by different types of libraries. This is the reason why it was said in the Introduction of the update and was agreed on during the ISBD Review Group meeting in San Juan, Puerto Rico (IFLA ISBD Review Group 2011).

The ISBD as an international standard provides internationally agreed on stipulations for are general stipulations and specific ones in the ISBD, to make the descriptions compatible.

Including unpublished resources has resulted in the need to add specific stipulations for this type of resource under the general ones.

Within the category of unpublished resources, a more specific or specialized case is that of musical resources. Its problems derive from the fact that it is written for the moment and is distributed directly, for example, by composers, electronically or not. This type of resource has required even more granular specifications.

The adaptation has required adding and/or revising stipulations in all areas, except in Area 6. The extension of Area 2 to cover the typology of unpublished resources was required. It has also required creating new elements in the following areas:

- Area 0 - Production process. This element is intended to organize, facilitating the searching and retrieval in a catalogue either by the way in which the content is fixed on a support or by the method used to carry out the production of the resource.
- Area 3 – Unpublished statement. This element is to record any statement in the resource from which this circumstance can be deduced.

It has also required the revision of the definitions in the glossary and the addition of two definitions: Published resource and Unpublished resource.



## Cartographic resources

The revision of celestial cartographic resources had been requested by specialists for the development of the current astronomical science. The studies of the stars' movements and their influence on phenomena is done by comparison with past research. "The objective of having new and old data to compare is essential in astronomy and other sciences. However, this data must be computerized and made available in corresponding databases. In this regard, while the availability of electronic data is enhanced by database accessibility via the Internet, printed information is nearly invisible in many cases, mainly due to two reasons: resources are unregistered as they have not yet been catalogued; or, they are indeed registered but their cataloguing level does not allow potential users to find them" (Escolano et al. 2017, 40).

Astronomical resources have been preserved in libraries around the world, and it is necessary that our descriptions of these resources contain those descriptive elements for easy and quick identification, location, and recovery by specialists.

The revision of cartographic resources has focused on completing the stipulations for the description of celestial cartography, clarifying some existing ambiguities and distinguishing elements already existing in the ISBD Consolidated Edition 2011. Previously, some of these elements were dealt with in notes and have now been moved into Area 3 and considered as elements. In addition, some more notes have been added for a more in-depth description of this kind of resource. These changes have resulted in the reorganization of Area 3 for more consistency and in the enlargement of Area 7.

The elements of Area 3 have been revised, adding new elements to clarify and distinguish them from others, or as has been said, because the stipulations were previously in the notes Area and were not recognized as elements, such as brightness, magnitude and celestial hemisphere. Other elements were mixed, e.g. Epoch must be differentiated as an element with respect to the Equinox. Other elements revised and added are Perspective, Projection and Wavelength.

All these elements are important for the scientific community and its comparative work to predetermine future movements of the stars. Moreover, the added elements are not mandatory if the information does not exist in the resource and cannot be easily identified by cataloguers without any specific knowledge about Astronomy or Cartography.

## ISBD granularity

The ISBD is a flexible standard that allows for different levels of granularity in description, especially dealing with multipart works. However, it was missing the stipulations for the more granular description of a continuing resources article. Users searching the guiding stipulations for this kind of description in ISBD Consolidated Edition 2011, were referred to a separate publication *Guidelines for the Application of the ISBDs to the Description of Component Parts*, published in 2003 (IFLA Working Group on the International Standard Bibliographic Description for Analytic Entries 1978-83 and IFLA Ad Hoc Working Group on Guidelines for the Description of Component Parts 1985-87 1988).

Now stipulations for the description of component parts have been developed and integrated with the other stipulations in the structure of the ISBD. This kind of description is a current practice

carried out in many types of libraries, mainly specialized, allowing different levels of description to coexist in a consistent way in the same integrated catalogue.

Chapter A has been the most affected since no more elements have been necessary in other areas because they are the same as those of the type of resource.

Obviously, in section A.2.1 of Object of the bibliographic description, it was required to add the corresponding stipulation for the component part.

The added stipulations in A.2.8 and A.3.4 specify the necessary structure for this kind of description of component parts, not only to describe the part itself, but also to identify the source or hosting resource in which the part is contained and to provide details about the exact location of the part within the source resource, with the linking element connecting the two descriptions.

The stipulations of section A.4.1 deal with the selection and order of preference of the sources of information for the description that it will be the same defined for each type of resource.

It was necessary also to add definitions in the Glossary.

The aim of these additions has been to achieve consistency in the provisions and to enlarge the flexibility of ISBD.

The updating was finished and approved by the ISBD Review Group in March 2021 and sent for approval of the Committee of Standards (received on November 29<sup>th</sup>) and the Professional Committee. It was published in the ISBD Review Group web page (IFLA ISBD Review Group 2021) as a Draft waiting for the final approval of the Professional Committee (December 22, 2021).

## **Transformation or transition of ISBD**

This updating of the ISBD mentioned before is not fully adapted to IFLA LRM. The updating of ISBD is focused on content stipulations that describe the resource, that is the main class and domain of the mapping started from ISBD to LRM, as it is recognized in the mapping:

“All ISBD properties are defined on the class resource, which although it reflects the four IFLA LRM Work/Expression/Manifestation/Item (WEMI) entities, remains disjoint from them. However, the IFLA LRM properties (attributes and relationships) that are relevant in mapping from ISBD are defined with the appropriate WEMI entity as their domain. As a result, this conceptual mapping cannot be a formal or operational alignment from ISBD elements to LRM elements. Instead, the mappings are intended to support understanding of the data for different purposes” (IFLA ISBD Review Group. Task Group for the Analysis of the Alignment and Impact of IFLA LRM on ISBD 2017b, 5).

It was necessary to work in the other way alignment, from LRM to ISBD with the Manifestation domain.

Another Task Force was created, the ISBD for Manifestation Task Force, entrusted with working out an alignment of the ISBD element set to the LRM manifestation entity and envisioning the future direction of the ISBD. According to the terms of reference, its tasks are:

1. Complete an alignment of the ISBD to the LRM at the-Manifestation level, working out an element set for that alignment.

2. Analyse the gaps, i.e., the elements of the ISBD Consolidated edition that in the cited alignment have been excluded, but are mandatory, mandatory if the information is available, or optional.

The group must present as deliverable, a document ISBD for Manifestation, consisting of:

- the ISBD to LRM Manifestation Elements Set; and,
- a set of stipulations for describing a manifestation according to LRM.

The Task Force also will have to reflect on transforming the ISBD into a Full Implementation of the LRM, a working document that will constitute the basis of the future work encompassing all the LRM elements and other components, and which would entail rethinking the structure of the ISBD.

The fundamental international cataloguing principle for this work is Representation:

“2.3. Representation. A description should represent a resource as it appears. Controlled forms of names of persons, corporate bodies and families should be based on the way an entity describes itself. Controlled forms of work titles should be based on the form appearing on the first manifestation of the original expression. If this is not feasible, the form commonly used in reference sources should be used” (IFLA Cataloguing Section and IFLA Meeting of Experts on an International Cataloguing Code 2016).

This work of the Task Force is on course; therefore, nothing can be said for sure at this time.

The fundamental topics of the debate are coincident with those that were already pointed out in the presentation made at the 85th World Library and Information Congress (Escolano 2019), that led to organizing the Editorial Group into two Task Forces.

Among them, it should be highlighted that transcription is the primary topic, as it affects others. To better understand, it is necessary to remember the previous debate regarding the interpretation of the meaning that the word transcription had in the scope note of the LRM Manifestation Statement. It became clear that manifestation statements are normally transcribed or/and recorded. Transcription conventions are codified by each implementation.

The Task Force is trying to find a compromise between a very broad level with only a ‘WYSIWYG’ (what you see is what you get) manifestation statement, that does not require standardization, or does to a minimal degree; and the most granular level of the element to which information has been added or is in some way normalized.

Transcribed manifestation statements are strings of text used for display. They are keyword indexed, and have high recall, so with this in mind, there doesn’t seem to be any reason to break them down to finer granularity. In contrast, recorded elements, have fine granularity and high precision because they are controlled. However, doubts arise regarding the coverage of the user’s needs with respect to the visualization of the data at the broadest level.

On the other hand, a “medium” level of granularity would require some normalization, but at the same time it does not stray too far from the representation principle. The idea is that recorded elements complement the broader transcribed statements.

Medium granularity aligns with the current ISBD “Areas” where transcription is relevant (1, 2, 3, 4, 6, 8).

In a sense, the finer the granularity of the statement and the more normalized the transcription is, the further one moves away from the principle of “Representation.”

Similarly, the issue of transcription applies to the added standard punctuation. It is considered that a very broad WYSIWYG transcribed statement, requires no inter-element punctuation, that is, added standard punctuation between statements, because there is just the one Manifestation statement super-element. However, statement granularity refinement adds the need for added standard punctuation to separate statements in a “gathering” or “concatenated” Manifestation statement. When guided by the ICP principle of Representation, punctuation within a Manifestation statement, “intra-element punctuation”, would be transcribed as displayed on the manifestation, with a possible addition of punctuation for clarity. However, the more punctuation is added, the more you move away from the principle of representation. So, the Task Force is aiming for a simple approach concerning standard punctuation, with as little additions as possible, which of course also depends on what punctuation is already present in the Manifestation.

On these premises, the set of elements will be elaborated, although there is still no decision made on the names or labels that still are provisional (Wetterstrom 2021).

## Conclusion

For the moment, the content of the standard has been revised and updated. It can be said that with some delay IFLA has responded to the requirements of its communities of users regarding their cataloguing needs.

However, the transformation process continues, not only to fully adapt to the conceptual model. This transformation of the standard could not be restricted to the structure and presentation format to which we are used.

Maybe we can obtain in a future the so much expected International Cataloguing Code, promised since 2003 when the five IFLA Meetings of Experts on an International Cataloguing Code (IME-ICC) were organized.

## References

Escolano, Elena. 2019. 'Update on the Work of the ISBD Revision: Paper Presented at the 85th World Library and Information Congress, Committee on Standards Open Session: Recently Approved IFLA Standards (Atenas, 26 August 2019)'. IFLA. [https://www.ifla.org/files/assets/cataloguing/isbdrng/updateontheworkofisbdrevision\\_9-8-2019.pdf](https://www.ifla.org/files/assets/cataloguing/isbdrng/updateontheworkofisbdrevision_9-8-2019.pdf).

Escolano, Elena, Donatella Randazzo, Maria Pilar Alonso-Lifante, and Francisco Javier Molero. 2017. 'Advocating for a Change of Mentality in the Development of Metadata Standards: Historical Celestial Cartography as a Specialization Example'. *JLIS* 8 (3): 39–57. <https://doi.org/10/gnt2tp>.

IFLA. 2017. *IFLA Library Reference Model. A Conceptual Model for Bibliographic Information*. Edited by Pat Riva, Patrick Le Boeuf, and Maja Zumer. Den Haag: IFLA. [https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla\\_lrm\\_2017-03.pdf](https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla_lrm_2017-03.pdf).

———. 2021. 'ISBD Linked Data Study Group'. <https://www.ifla.org>. 2021. <https://www.ifla.org/isbd-linked-data-study-group/>.

IFLA Cataloguing Section and IFLA Meeting of Experts on an International Cataloguing Code. 2016. *Statement of International Cataloguing Principles (ICP)*. Rev. 2017. Den Haag: IFLA.

IFLA ISBD Linked Data Study Group. 2016. *Guidelines for Use of ISBD as Linked Data*. Den Haag: IFLA. [https://repository.ifla.org/bitstream/123456789/1384/1/guidelines-use-isbdlld\\_082016.0.pdf](https://repository.ifla.org/bitstream/123456789/1384/1/guidelines-use-isbdlld_082016.0.pdf).

IFLA ISBD Review Group. 2011. 'Minutes: Presented at 77th IFLA General Conference, San Juan, Puerto Rico, 13-18 August: Cataloguing Section'. IFLA. [https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/isbdrng/meeting\\_2011.pdf](https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/isbdrng/meeting_2011.pdf).

———. 2017. 'Annual report 2016-2017'. IFLA. [https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/isbdrng/isbdrng\\_report\\_2016-2017.pdf](https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/isbdrng/isbdrng_report_2016-2017.pdf).

———. 2018a. 'ISBD Editorial Group. Terms of Reference'. IFLA. [https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/isbdrng/isbd\\_editorial\\_group.pdf](https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/isbdrng/isbd_editorial_group.pdf).

———. 2018b. 'Proposed Work Plan for ISBD Revision 2018-2022'. IFLA. [https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/isbdrng/proposed\\_work\\_plan\\_for\\_isbd\\_revision\\_2018-2022.pdf](https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/isbdrng/proposed_work_plan_for_isbd_revision_2018-2022.pdf).

———. 2019. 'Action Plan 2019-2020'. IFLA. [https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/isbdrng/isbd\\_rg\\_-\\_action\\_plan\\_2019-2020\\_.pdf](https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/isbdrng/isbd_rg_-_action_plan_2019-2020_.pdf).

———. 2021. 'ISBD: Update 2021 to the 2011 Consolidated Edition [DRAFT]'. <https://www.ifla.org/news/isbd-update-2021-to-the-2011-consolidated-edition-draft/>.

IFLA ISBD Review Group. Task Group for the Analysis of the Alignment and Impact of IFLA LRM on ISBD. 2017a. 'Mapping from ISBD to IFLA LRM: Final Version 2017, as amended in 2018'. IFLA. [https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/isbd/Other-Documentation/mapping\\_isbd-lrm\\_amd\\_2018\\_v.1.0.pdf](https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/isbd/Other-Documentation/mapping_isbd-lrm_amd_2018_v.1.0.pdf).

———. 2017b. 'Report'. IFLA. [https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/isbd/OtherDocumentation/isbd-lrm\\_alignment\\_report.pdf](https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/isbd/OtherDocumentation/isbd-lrm_alignment_report.pdf).

IFLA Study Group on the Functional Requirements for Bibliographic Records. 1998. *Functional Requirements for Bibliographic Records. Final Report*. München: K.G. Saur.

IFLA Working Group on the International Standard Bibliographic Description for Analytic Entries 1978-83, and IFLA Ad Hoc Working Group on Guidelines for the Description of Component Parts 1985-87. 1988. *Guidelines for the Application of the ISBDs to the Description of Component Parts*. London: IFLA Universal bibliographic control and International MARC programme British library bibliographic services. <https://repository.ifla.org/handle/123456789/807>.

Ranganathan, Shiyali Ramamrita. 1964. *Classified Catalogue Code with Additional Rules for Dictionary Catalogue Code*. 5th ed. Bombay: Asia Publishing House.

Wetterstrom, Mikael. 2021. 'ISBD and LRM. The ISBD Revision Work to Date by the ISBD for Manifestation Task Force'. <https://es.slideshare.net/ISSNIC40/isbd-and-lrm>.

## Artificial Intelligence Systems and problems of the concept of author. Reflections on a recent book

Maurizio Lana<sup>(a)</sup>

a) Università del Piemonte Orientale, <https://orcid.org/0000-0002-7520-1195>

Contact: Maurizio Lana, [maurizio.lana@uniupo.it](mailto:maurizio.lana@uniupo.it)

Received: 2 June 2022; Accepted: 15 November 2022; First Published: 15 May 2022

### ABSTRACT

The publication of the book Beta Writer. 2019. Lithium-Ion Batteries. A Machine-Generated Summary of Current Research. New York, NY: Springer, produced with Artificial Intelligence software prompts analysis and reflections in several areas. First of all, on what Artificial Intelligence systems are able to do in the production of informative texts. This raises the question if and how an Artificial Intelligence software system can be treated as the author of a text it has produced. Evaluating whether this is correct and possible leads to re-examine the current conception for which it is taken for granted that the author is a person. This, in turn, when faced with texts produced by Artificial Intelligence systems necessarily raises the question of whether they, like the author-person, are endowed with agency. The article concludes that Artificial Intelligence systems are characterized by a distributed agency, shared with those who designed them and make them work, and that in the wake of the reflections of 50 years ago by Barthes and Foucault, it is necessary to define and recognize a new type of author.

### KEYWORDS

Artificial Intelligence; Texts; Books; Agency; Authorship.

## Sistemi di Intelligenza Artificiale e problemi del concetto di autore. Riflessioni su prodotti editoriali recenti

### ABSTRACT

La pubblicazione del libro Beta Writer. 2019. Lithium-Ion Batteries. A Machine-Generated Summary of Current Research. New York, NY: Springer, prodotto con software di Intelligenza Artificiale sollecita analisi e riflessioni in vari ambiti. Prima di tutto su che cosa i sistemi di Intelligenza Artificiale siano in grado di fare nell'ambito della produzione di testi di tipo informativo. Di qui si pone la questione se e come un sistema software di Intelligenza Artificiale possa essere trattato come autore di un testo che ha prodotto. Valutare se ciò sia corretto e possibile porta a riprendere in esame la concezione attuale per la quale è scontato che l'autore sia una persona. Ciò, a sua volta, di fronte a testi prodotti da sistemi di Intelligenza Artificiale sollecita necessariamente la domanda se essi, come l'autore-persona, siano dotati di agency. L'articolo conclude che i sistemi di Intelligenza Artificiale si caratterizzano per un'agency distribuita, condivisa con chi li ha progettati e li fa funzionare e che sulla scorta delle riflessioni di 50 anni fa di Barthes e Foucault occorre definire e riconoscere un nuovo tipo di autore.

### PAROLE CHIAVE

Intelligenza Artificiale; Testi; Libri; Agency; Authorship.

## Premessa

Sempre più spesso anche nei mezzi di comunicazione non specializzati compaiono notizie e riflessioni sull'uso dell'Intelligenza Artificiale nella quotidianità della nostra società. I toni oscillano tra entusiasmo e timori distopici. Ma il fatto stesso di ipostatizzare «l'Intelligenza Artificiale» segnala una visione imprecisa, perché «l'Intelligenza Artificiale» è una disciplina in cui si opera con programmi+dati, programmi che progettisti, programmatori (e altri software) addestrano su vari tipi di dati. Tanto i programmi quanto i dati riflettono, ed amplificano, la visione del mondo dei progettisti e programmatori e i loro bias. Qui useremo invece espressioni come «software | sistemi | programmi di Intelligenza Artificiale» perché non scriviamo di Intelligenza Artificiale ma di situazioni in cui i prodotti dell'Intelligenza Artificiale vengono utilizzati.

Questo articolo interseca o sfiora grandi questioni; qui ne menzioniamo alcune senza pretesa di esaurirle. Alcune riguardano le caratteristiche proprie dei sistemi di Intelligenza Artificiale, in primis quella dell'agency di cui molto si discute: «i sistemi di Intelligenza Artificiale hanno agency?» hanno cioè capacità di agire individualmente in modo libero in un determinato contesto? La seconda è quella del rapporto con il mondo: «i sistemi di Intelligenza Artificiale comprendono il mondo?» cioè sono capaci di operare in modo appropriato sia dal punto di vista sintattico sia dal punto di vista semantico<sup>1</sup>? Altre questioni riguardano il modo in cui i sistemi di Intelligenza Artificiale si collocano nel mondo. Una questione è quella storica: «qual è la storia dei sistemi di Intelligenza Artificiale?» cioè in quale linea di eventi si iscrivono i sistemi di Intelligenza Artificiale? Eventi sono snodi di possibilità, che segnalano che «le cose avrebbero potuto andare diversamente». Un'altra è quella ambientale: «i sistemi di Intelligenza Artificiale sono sostenibili?». Addestrare 1 singolo modello su 1 singolo problema testuale con 1 rete neurale comporta consumi elettrici che producono circa 284 chili di CO<sub>2</sub><sup>2</sup>. In questo articolo affronteremo la prima questione, quella dell'agency e ciò che essa implica.

---

<sup>1</sup> La riflessione su sintattico/semantico è al centro di (Durante 2019). Durante porta ad esempio una partita di scacchi contro il computer: per il giocatore umano la partita è semantica cioè la scelta delle mosse si iscrive all'interno di una visione strategica complessiva, per il programma di Intelligenza Artificiale è sintattica perché in risposta alla mossa del giocatore umano si calcolano tutte le possibili mosse lecite successive. Nella quotidianità, chi fa una ricerca web ha uno scopo semantico cioè le parole della query implicano e veicolano un significato atteso, mentre il motore di ricerca fornisce gli esiti su base sintattica cioè essi contengono in modo appropriato le parole della query – la discrepanza sintattica/semantica si vede nel fatto che davanti a migliaia di esiti sintatticamente (co)rispondenti noi tutti andiamo in cerca di quei «7» semanticamente validi e soddisfacenti. Quindi il soggetto umano e il computer costantemente al medesimo tavolo giocano due giochi differenti. Parte della complessità del confronto della società con i sistemi di Intelligenza Artificiale è invece credere che il gioco sia il medesimo. L'equivoco si manifesta per la prima volta in modo esplicito in (McCarthy et al. 1955): «the artificial intelligence problem is taken to be that of making a machine behave in ways that would be called intelligent if a human were so behaving» ma è già presente nel test di Turing che è presentato come un *imitation game* (Turing 1950). L'importanza del «Proposal» del 1955 è data dal fatto che si tratta della *prima formulazione di un progetto complessivo per l'Intelligenza Artificiale*, e gli autori (McCarthy, Minsky, Rochester, Shannon) sono quattro giganti, per usare la metafora di Bernardo di Chartres.

<sup>2</sup> I dati qui appena accennati provengono da (Strubell, Ganesh, e McCallum 2019) che con (Crawford 2021) è un importante riferimento sul tema. Il libro di Crawford è commentato da (Hao 2021) in un'intervista con l'autrice. Timnit Gebru, condirettrice del settore «Ethical AI» in Google, è stata licenziata il 2 dicembre 2020 quasi certamente per dissenso dell'azienda su un articolo che Gebru ha poi pubblicato nel 2021 (Bender et al. 2021) in cui segnalava il serio pericolo per l'ambiente dato dai consumi di energia necessari per l'addestramento dei modelli di AI.



In specifico prenderemo in esame due tipi differenti di testi prodotti da (per mezzo di) sistemi di Intelligenza Artificiale: un libro e un articolo di rivista. Proprio il fatto che siano testi di tipi molto differenti permette di mostrare con più evidenza caratteristiche specifiche dei sistemi di IA nella scrittura di testi che compaiono in entrambi i casi. In particolare, entrambi i testi sono sintatticamente corretti ma mostrano problemi semantici (cioè di organizzazione complessiva dello sviluppo di senso) e sono quindi stati necessari interventi di soggetti umani per sanarli. Nella forma finale i testi sono dunque il prodotto dell'attività di soggetti umani + software di Intelligenza artificiale, il che nostra anche nell'ambito della produzione di testi che l'agency dei sistemi di IA è agency condivisa e dunque i soggetti umani sono co-responsabili delle operazioni dei sistemi di IA. La complessità e ampiezza di queste tematiche, e le loro ramificazioni, mostrano che gli studi di scienza della biblioteca aprono necessariamente alle tematiche di information literacy che non hanno solo la dimensione intra-professionale ma anche quella civile e di cittadinanza.

I due casi qui esaminati non sono costituiscono esempi *perfetti* di agency, come si vedrà nel corso dell'analisi sviluppata più avanti. Ma il contesto in cui sono stati generati (la relazione editore-autore) li presenta come tali e dunque da questa assunzione partiremo. L'elemento formale che in modo inoppugnabile chiarisce l'implicita ma netta affermazione di agency del software di Intelligenza Artificiale è l'indicazione autoriale associata ai due scritti: Beta Writer e GPT-3. Nel primo caso si indica (allusivamente, a causa dell'apparente nesso nome-cognome) un programma informatico, nel secondo caso si menziona esplicitamente il nome del software di Intelligenza Artificiale che ha prodotto il testo. Come si vedrà più avanti<sup>3</sup> l'indicazione di un nome di autore non è secondaria, marginale o trascurabile, per definire l'oggetto e dunque in partenza ci atteniamo a ciò che questa indicazione usualmente significa e che ha ovviamente una forte connotazione promozionale, di marketing: attirare l'attenzione sull'oggetto grazie alla novità eclatante segnalata dal nome dell'autore. L'analisi mostrerà che ci sono degli elementi, o degli aspetti, di imprecisione in questa indicazione autoriale, proprio connessi con l'agency che usualmente si ritiene propria di un autore 'tradizionale' che qui non c'è, e che il nuovo autore non ha.

## **Beta Writer. 2019. «Lithium-Ion Batteries. A Machine-Generated Summary of Current Research». New York, NY: Springer**

Nel 2019 l'editore Springer ha pubblicato a stampa e in digitale un volume intitolato «Lithium-Ion Batteries. A Machine-Generated Summary of Current Research» (Beta Writer 2019) la cui caratteristica principale è di esser stato prodotto per mezzo di un sistema *ad hoc* di Intelligenza Artificiale, tanto che l'autore è stato denominato «Beta Writer»<sup>4</sup>, qualcosa come «autore in versione quasi finale», perché nel mondo del software la versione beta di un programma è una versione intermedia, ma avanzata e funzionante<sup>5</sup> e dunque «Beta Writer» allude ad un processo di sviluppo

<sup>3</sup> Cfr. pag. 3.

<sup>4</sup> L'autore è così descritto nella pagina web del volume (<https://link.springer.com/book/10.1007%2F978-3-030-16800-1>): «Machine-generated by Beta Writer 0.7 software developed at Goethe University Frankfurt».

<sup>5</sup> Tale versione viene rilasciata affinché gli utenti interessati la utilizzino aiutando con ciò a verificare se vi siano errori che erano sfuggiti ai programmatori; ma questa dimensione di verifica del prodotto non viene dichiarata per il volume in questione.

che non è ancora completo ma è già pervenuto a risultati sufficientemente buoni da poter essere resi pubblici. Ovviamente, nessuna forma di Intelligenza Artificiale ha autonomamente deciso di produrre questo volume, e lo ha prodotto; bensì un editore ha deciso di utilizzare procedure differenti da quelle solitamente adottate per la produzione di un libro<sup>6</sup>. Genette richiama l'attenzione sull'effetto comunicativo che uno pseudonimo può produrre: «Ce qui nous concerne comme élément paratextuel, c'est ... l'effet produit sur le lecteur, ou plus généralement sur le public, par la présence d'un pseudonyme.»<sup>7</sup>

Il comunicato stampa che annunciava l'uscita del libro<sup>8</sup> spiegava che esso contiene una rassegna di articoli scientifici sugli sviluppi della ricerca sulle batterie al litio, rassegna definita «machine-generated» e «automatically compiled by an algorithm», due espressioni sostanzialmente simili. Partner di Springer in questa attività è stato un gruppo di ricercatori dell'«Applied Computational Linguistics Lab» della Goethe Universität di Francoforte<sup>9</sup>. È chiaro che si è trattato di una scelta editoriale dell'editore che ha inteso attirare l'attenzione sul prodotto (il libro) al di fuori, soprattutto al di fuori, dei circoli degli esperti di Intelligenza Artificiale.

### Analisi di questioni di politica editoriale

Il volume si apre con un'introduzione scritta da Henning Schönberger, direttore del settore «data development» in Springer Nature, che dichiara:

With the exception of this preface [this book] has been created by an algorithm on the basis of a re-combined accumulation and summarization of relevant content in the area of Chemistry and Materials Science. The book is a cross-corpora auto-summarization of current texts from Springer Nature's content platform "SpringerLink", organized by means of a similarity-based clustering routine in coherent chapters and sections.<sup>10</sup>

La risorsa utilizzata per la selezione delle fonti è dunque stata la piattaforma SpringerLink. Schönberger evidenzia poi una serie di questioni complesse direttamente connesse con la produzione di pubblicazioni ad opera di un sistema di Intelligenza Artificiale. L'assunzione di principio dichiarata è la *trasparenza*:

Full transparency is essential for us to discover both the opportunities of machine-generated content and the current limitations that technology still confronts us with. But it was also an ethical decision

---

<sup>6</sup> Il libro ha suscitato il normale interesse di una pubblicazione scientifica, nondimeno non ne sono mancate letture che costituiscono forzature di specifici aspetti della produzione del libro : «Machines can generate literature reviews and even empirical data analysis and may soon write academic papers» (Haski-Leventhal 2020, 142).

<sup>7</sup> (Genette 2002, cap. Le nom d'auteur-par. Pseudonymat)

<sup>8</sup> (Springer Nature 2019).

<sup>9</sup> («Projects and Cooperations - Applied Computational Linguistics Lab Goethe University Frankfurt, Germany» s.d.) Il progetto che ha portato alla pubblicazione del libro è intitolato «Schwach überwachte Verfahren zur Bibliographieanalyse», metodi per analisi bibliografica a supervisione debole. La collaborazione che ha portato alla pubblicazione del volume era stata avviata nel 2014 e si colloca in un quadro di vari progetti del Laboratorio che hanno impronta filologica, linguistica, e di digital humanities.

<sup>10</sup> (Schönberger, Chiarcos, e Schenk 2019, V).

that if we start this journey, we want to do so in a correct and responsible manner, in order to enable a discussion in the research communities that is as open-minded as possible.<sup>11</sup>

We aim to demonstrate both possible merits and possible limitations of the approach, and to put it to the test under real-world conditions, in order to achieve a better understanding of what techniques work and which techniques do not.<sup>12</sup>

Rendere possibile una discussione a mente libera nelle comunità di ricerca sulle questioni in gioco: pregi e limiti dell'approccio, quali tecniche funzionano e quali no. I primi interrogativi posti da Schönemberger riguardano la funzione autoriale:

Who is the originator of machine-generated content? Can developers of the algorithms be seen as authors? Or is it the person who starts with the initial input (such as “Lithium-Ion Batteries” as a term) and tunes the various parameters? Is there a designated originator at all? Who decides what a machine is supposed to generate in the first place? Who is accountable for machine-generated content from an ethical point of view?<sup>13</sup>

che Schönemberger pone di fatto come questione dell'agency, segnalando in relazione alla produzione del libro che il sistema di Intelligenza Artificiale ha operato in base alle assunzioni e scelte di progettisti e operatori e chiedendo(si) chi sia responsabile del contenuto.

Un'altra grande categoria di questioni affrontate da Schönemberger riguarda la qualità intrinseca delle fonti (che dipende dalla qualità della *peer review*), e l'estrazione e selezione delle fonti in funzione di una rassegna su un tema specifico. Il software di Intelligenza Artificiale ha lavorato le fonti disponibili nella piattaforma SpringerLink in un implicito assunto che esse siano tutte di alta qualità, il che a sua volta presuppone che tutto il processo di peer review sia di alta qualità, mentre è noto che ogni anno un certo numero di articoli scientifici pubblicati viene ritirato per gravi difetti<sup>14</sup>. E oltre alla valutazione di qualità (la fonte ha valore scientifico?) c'è quella forse anche più complessa di pertinenza (la fonte è appropriata al discorso che si vuole sviluppare?): due tipi di valutazione che nella produzione di questo libro non sono entrati. La riflessione di Schönemberger si sposta poi verso i modi e la qualità della scrittura machine-based:

the extractive summarization of large text corpora is still imperfect, and paraphrased texts, syntax and phrase association still seem clunky at times. However, we clearly decided not to manually polish or copy-edit any of the texts due to the fact that we want to highlight the current status and remaining boundaries of machine-generated content. ... How will the publication of machine-generated content impact our role as a research publisher?

... we know that the quality of machine-generated text can only be as good as the underlying sources which have been used to generate it. At Springer Nature, we publish research which stands up to

---

<sup>11</sup> Ivi, pag. VII.

<sup>12</sup> Ivi, pag. X.

<sup>13</sup> Ivi, pag. VII.

<sup>14</sup> Si veda ad esempio la disamina del problema presentata da (Brainard e You 2018) e (Vuong 2020), che sollecita la domanda (per ora?) senza risposta: un sistema di Intelligenza Artificiale è in grado di riconoscere se un articolo scientifico è per qualche ragione difettoso? che è poi il motivo per cui si opera la peer review.

scientific scrutiny. In consequence, machine-generated content makes it even more necessary to re-emphasize the crucial role of peer-review itself. Though peer-review is also in the course of being continuously re-defined (and in the future we expect to see substantial progress in machine-support also in this regard) we still think that for the foreseeable future we will need a robust human review process for machine-generated text.

Si tratta di una riflessione complessa. In primo piano c'è il riconoscimento che il testo prodotto dal software di Intelligenza Artificiale lascia a desiderare. Più complicato, e appena accennato, quale possa essere lo spazio, o il ruolo, per un editore scientifico quando si affermi la «publication of machine-generated content». Non meno complessa la questione del «robust human review process» che continuerà ad essere necessario perché un'elaborazione puramente algoritmica dell'informazione non è in grado di individuare se un articolo è metodologicamente valido. Per la *peer review* sarà indiscutibilmente utile il supporto di sistemi di Intelligenza Artificiale<sup>15</sup> che non rimpiazzano l'attività dello studioso, piuttosto la supportano («substantial progress in machine-support also in this regard») liberandola dai compiti esecutivi e aiutandola a focalizzarsi sul contenuto e sulla sua valutazione<sup>16</sup>. Ma occorre riflettere su come si potrà fare la *peer review* di scritti composti con sistemi di Intelligenza Artificiale. Un sistema di Intelligenza Artificiale non è un *peer* di uno studioso, «the term peer itself indicates a certain inadequacy for machine-generated research content. Who are the peers in this context? Would you as a human reader consider yourself as peer to a machine?»<sup>17</sup>. I suoi processi di organizzazione e gestione dell'informazione non sono umani, e dunque «the review process on machine-generated research content needs refinement, if not a complete re-definition»<sup>18</sup>. Se il *reviewer* al termine di una

---

<sup>15</sup> Si può immaginare un ambiente di lavoro di questo tipo per il revisore: apre il file presentato per la pubblicazione e il sistema di Intelligenza Artificiale di supporto individua e apre le fonti citate, per ogni fonte sia evidenzia citazioni dirette o indirette nel testo in valutazione; sia riporta i vari indici di citazione disponibili, se esistenti; e per l'autore e il tema in valutazione individua eventuali altre pubblicazioni simili per permettere la valutazione di originalità.

<sup>16</sup> Continua ad essere valido il paradigma Licklider-Engelbart formulato negli anni 60 del secolo scorso, nel quale il computer svolge compiti funzionali liberando tempo per l'intelligenza umana che si può maggiormente dedicare a compiti di tipo ideativo/interpretativo/creativo. Esso è espresso in modo non formalizzato in (Licklider 1960), (Engelbart 1962), (Engelbart e English 1968). Siamo però testimoni di una fase di evoluzione per cui il software è passato da compiti esecutivi in senso riduttivo (trovare un indirizzo in una rubrica in base alla posizione in cui ci troviamo) a compiti esecutivi più complessi (attivare una data configurazione del computer, in base alla posizione in cui ci troviamo) ma che non comportano un'azione che modifica lo stato delle cose. Fino a che oggi con la disponibilità di sistemi di Intelligenza Artificiale è possibile demandare l'esecuzione di compiti sempre più complessi che comportano azioni che modificano lo stato delle cose: il riconoscimento facciale di individuo attraverso telecamere di sorveglianza può portare all'incarcerazione di quell'individuo (e anzi in (Hao e Swart 2022) si discute del fatto che in Sudafrica attraverso l'uso di software di Intelligenza Artificiale da parte della polizia sta prendendo piede un nuovo apartheid digitale): la valutazione di una domanda di assicurazione vita può portare al fatto che una persona non sia in grado di sostenere il costo delle cure sanitarie; o un'azione di guerra condotta per mezzo di un drone autonomo (sorveglianza, identificazione di soggetti sospetti, programmazione di possibili azioni di contrasto, attuazione dell'azione – nel 2021 se ne è avuta una prima chiara occorrenza (Kallenborn 2021) nella guerra libica ad opera dell'esercito turco, occorrenza poi confermata dal rapporto dell'ONU (Choudhury et al. 2021, 17/548)) porta all'uccisione di individui da parte del drone. Con la guerra in Ucraina ciò è improvvisamente diventato ordinario e in sostanza non suscita pensieri particolari, come se si trattasse di un'arma come tutte le altre.

<sup>17</sup> Ovviamente si possono ipotizzare sistemi di Intelligenza Artificiale che operano in maniera concorrente, ma il problema viene rinviato, non risolto; a meno che si voglia demandare in toto la *peer review* ai sistemi di Intelligenza Artificiale.

<sup>18</sup> (Schönenberger, Chiarcos, e Schenk 2019, IX).

serie di osservazioni puntuali conclude «ripresentare con correzioni», qual è il «soggetto» che interpreta le osservazioni del *reviewer* e le rende in forma di modifiche appropriate? Il software di Intelligenza Artificiale riesce (abbastanza bene) a scrivere frasi, anche a scrivere una frase simile ad un'altra, ma non sa che cosa significhi ciò che scrive. Sullo sfondo c'è il pensiero che un esperto in uno specifico campo di ricerca non deve diventare un esperto in reti neurali e in NLP per poter valutare la qualità di un testo prodotto da un'Intelligenza Artificiale. Schönenberger crede che «soon enough we will see machine-generated texts from unstructured knowledge bases that will lead to more complex evaluation processes»<sup>19</sup> ma sembra non considerare il problema che generare testo a partire da basi di conoscenza non strutturate *ha senso per il mondo* solo se la conoscenza infondata, la falsa conoscenza viene riconosciuta e messa da parte, il che richiede di «conoscere il mondo», qualunque cosa ciò significhi in dettaglio, e saper rapportare a ciò il testo che si legge decidendo se le discrepanze sono ammissibili, fittizie, inammissibili (=falsità); e se in chi scrive c'è consapevolezza di avere responsabilità per ciò che scrive, a partire dall'autorevolezza che si costruisce nel proprio contesto, fino agli effetti che la conoscenza condivisa può produrre su/per mezzo di altri. Il pensiero a questo punto corre allo spazio di azione che avranno i sistemi di Intelligenza Artificiale e a quello che *resterà* per gli umani:

We foresee that in future there will be a wide range of options to create content—from entirely human-created content to a variety of blended man-machine text generation to entirely machine-generated text. We do not expect that authors will be replaced by algorithms. On the contrary, we expect that the role of researchers and authors will remain important, but will substantially change as more and more research content is created by algorithms. To a degree, this development is not that different from automation in manufacturing over the past centuries which has often resulted in a decrease of manufacturers and an increase of designers at the same time. Perhaps the future of scientific content creation will show a similar decrease of writers and an increase of text designers or, as Ross Goodwin puts it, writers of writers: “When we teach computers to write, the computers don't replace us any more than pianos replace pianists—in a certain way, they become our pens, and we become more than writers. We become writers of writers.”<sup>20</sup>

La frase «We become writers of writers» che si potrebbe tradurre, per esplicitarne il senso, con «diventiamo configuratori di sistemi di scrittura», dice che i sistemi di Intelligenza Artificiale non hanno agency, non più di un pianoforte rispetto ad un pianista, perché essa rimane nel soggetto umano che configura e utilizza il sistema di Intelligenza Artificiale. Posizione non di un commentatore terzo, ma di un autore che ha usato intensamente sistemi di Intelligenza Artificiale per produrre due opere<sup>21</sup>.

In questa linea di pensiero, la promessa, il disegno del futuro, per Schönenberger è che gli autori non saranno rimpiazzati da algoritmi ma che il ruolo autoriale cambierà perché sempre più

<sup>19</sup> (Schönenberger, Chiarcos, e Schenk 2019, IX).

<sup>20</sup> (Schönenberger, Chiarcos, e Schenk 2019, IX). La citazione interna è da (Goodwin 2016b) che commenta la sua prima opera di fiction, prodotta elaborando uno scritto preesistente. Da ciò l'immagine di «scrittori di scrittori». Goodwin è anche l'autore del romanzo «I the road» prodotto utilizzando un sistema di Intelligenza Artificiale interconnesso con sensori, microfono, macchina fotografica (Goodwin, McDowell, e Planquelle 2018).

<sup>21</sup> (Goodwin 2016a) e (Goodwin, McDowell, e Planquelle 2018).

contenuto di ricerca sarà prodotto da algoritmi e quindi attività che oggi consideriamo di alto livello potranno essere demandate alle/svolte dalle macchine e nuovi spazi si apriranno alla scrittura scientifica. Nondimeno nelle sue parole due concetti si scontrano: «we do not expect that authors will be replaced by algorithms» e «more and more research content is created by algorithms». Se il contenuto di ricerca sarà prodotto dagli algoritmi come si può sostenere che gli algoritmi non soppianteranno gli autori? si sta affermando che un autore scientifico si qualifica come tale non per la produzione di contenuto?<sup>22</sup> Da una parte quindi ci si può chiedere *che cosa fa o che cosa è un autore, se non produce contenuto*; dall'altra il paradigma Licklider-Engelbart che abbiamo già ricordato alla nota 16 ricorda che ad ogni epoca si scopre che una parte dei compiti che si reputavano creativi o inestricabilmente connessi con la creatività possono essere demandati a programmi di computer. Pertanto la domanda si riformula in questi termini: *tra i caratteri che oggi reputiamo specifici dell'essere autore, che cosa è essenziale e che cosa è secondario?*

La posizione di Schönenberger si può leggere in linea di continuità con le trasformazioni che nell'ultima ventina d'anni gli strumenti digitali e le reti hanno portato nel lavoro degli studiosi con la progressiva disponibilità prima degli OPAC, poi delle biblioteche digitali, poi dei cataloghi di risorse in Accesso Aperto, di Google Scholar e di DOAJ e delle piattaforme degli editori, dei tool di information retrieval, dei discovery tools, delle segnalazioni di articoli interessanti da parte delle piattaforme editoriali, e così via. Tutte risorse digitali che gestiscono e offrono accesso all'informazione scientifica<sup>23</sup> e che volendo si potrebbero vedere come forme del passato in cui appare l'evolvere del «more and more research content is created by algorithms». Ma la perplessità rimane.

## Analisi degli aspetti tecnici della produzione del libro

Alle analisi di Schönenberger seguono quelle di Christian Chiarcos e Niko Schenk dell'«Applied Computational Linguistics Lab» della Goethe Universität, che discutono della procedura di generazione (scrittura) del libro e della selezione delle fonti:

we decided for a relatively conservative approach, a workflow based on

1. document clustering and ordering,
2. extractive summarization, and
3. paraphrasing of the generated extracts<sup>24</sup>.

---

<sup>22</sup> Si può riprendere utilmente la distinzione di Peirce tra ragionamenti corollari, che sviluppano premesse note, e ragionamenti teorematichi il cui autore adotta un metodo sperimentale.

<sup>23</sup> Analogamente, crediamo, sarà per il problema del sovraccarico informativo, che parzialmente si risolve(rà) man mano che diventino disponibili strumenti di raccolta e selezione di fonti capaci di operare più pervasivamente e selettivamente di quanto possa fare oggi qualsiasi ricerca in qualsiasi ambito o servizio del web. Il cosiddetto «sovraccarico informativo» peraltro sembra preferibile concepirlo come, e chiamarlo, «abbondanza informativa». Un sovraccarico è indesiderabile e deve essere eliminato; l'abbondanza è desiderabile e deve essere governata con strumenti adeguati. Quello del «sovraccarico informativo» non è tema di oggi e può in vario modo essere tracciato indietro nel tempo almeno fino all'invenzione della stampa, e anche più indietro (Blair 2003); (Blair 2010).

<sup>24</sup> (Schönenberger, Chiarcos, e Schenk 2019, XI).

Tre fasi operative, la prima è «document clustering» e non qualcosa come «ricerca ed estrazione degli articoli da SpringerLink». La ragion d'essere di questo libro, in un campo dove c'è un'ampissima produzione scientifica, è di creare una struttura coerente di contenuti quindi organizzare le fonti per temi, «clustering and ordering»; e per ogni tema scrivere una sintesi introduttiva ad ogni capitolo<sup>25</sup>. Qui si colloca l'utilità, per il lettore/studioso, del ricorso ad un sistema di Intelligenza Artificiale: leggendo le introduzioni ai capitoli, che corrispondono a campi specifici della ricerca sulle batterie al litio, si ha una visione di sintesi delle novità emerse in quel campo. La sintesi prodotta dal software di Intelligenza Artificiale è più raffinata di ciò che chiunque potrebbe fare semplicemente con una ricerca in SpringerLink, che contiene 53.000 fonti sul tema pubblicate negli ultimi 3 anni. Spiegano Chiarcos e Schenk:

In preparation for generating a book, we identify a seed set of source documents as a thematic data basis for the final book, which serve as input to the pipeline. These documents are obtained by searching for keywords in publication titles or by means of meta data annotations<sup>26</sup>.

La sfumatura è sottile e sta tutta nella parola *seed* (seme): *identifichiamo un set di fonti sul tema e le usiamo come seed, come input iniziale del software di Intelligenza Artificiale*<sup>27</sup>. Non viene aggiunto altro e il lettore consapevole intende che il software di Intelligenza Artificiale ha provveduto alla ricerca ed estrazione da SpringerLink degli articoli da usare per la produzione del libro imparando dagli esempi del *seed* quali articoli scegliere. Coloro che hanno selezionato il set di articoli iniziale sono quindi a pieno titolo (co)autori della pubblicazione, infatti senza di loro o con differenti loro scelte la pubblicazione prodotta da Beta Writer sarebbe stata differente. Si vorrebbe sapere di più su questo passaggio, ma come Schönenberger stesso ha dichiarato in un passo che abbiamo citato sopra «it becomes increasingly difficult to understand how a result has been actually derived»<sup>28</sup>.

Istruito con il set di articoli individuati dagli esseri umani, il software di Intelligenza Artificiale ha estratto da SpringerLink una collezione di 1086 pubblicazioni selezionate in base a parole presenti nel titolo o nei metadati, e per anno di pubblicazione. Tale collezione è stata successivamente lavorata per ordinare e raggruppare le fonti e creare così la struttura della pubblicazione. Gli sviluppatori inizialmente valutarono se raggruppare le fonti basandosi sulla sovrapposizione bibliografica, o sulla somiglianza testuale dei documenti, e scelsero poi di quest'ultima perché adottando la sovrapposizione bibliografica risultava difficile bilanciare fonti con bibliografia mol-

---

<sup>25</sup> (Schönenberger, Chiarcos, e Schenk 2019, XI).

<sup>26</sup> (Schönenberger, Chiarcos, e Schenk 2019, XIV).

<sup>27</sup> È il procedimento che si adotta con le reti neurali: si mostra al software un tipo di esito desiderato (in questo caso: gli articoli d'esempio scelti per tema, qualità eccetera) affinché produca altri esiti simili (in questo caso individuare altri articoli sull'argomento). Si vedano come introduzioni al tema (Hagan, Demuth, e Beale 1996) e (Wang 2003).

<sup>28</sup> Proprio in riferimento a queste questioni si sta sviluppando un orientamento critico denominato XAI, eXplainable Artificial Intelligence, il cui scopo è «to make a shift towards more transparent AI. It aims to create a suite of techniques that produce more explainable models whilst maintaining high performance levels» su cui si vedano ad esempio (Adadi e Berrada 2018, 52138) che presentano una rassegna di studi sulle vari forme di XAI oppure (Barredo Arrieta et al. 2020). Amichevole, ma non meno robusto, l'articolo di (Gunning e Aha 2019) che ha molte componenti visive. La XAI a sua volta è espressione della spinta della società e degli scienziati verso l'«ethical AI», per la quale validi punti di entrata sono (Floridi et al. 2018), (Taddeo e Floridi 2018) e (Mittelstadt 2019).

to ampia e fonti con bibliografia meno ampia. La somiglianza testuale fornì dunque dati lessicali per la cluster analysis<sup>29</sup> cioè la suddivisione e raggruppamento delle fonti che si svolse in due fasi: prima vennero individuati i «core thematic topics» che diedero luogo ai capitoli; poi all'interno di essi i «subtopics», le sezioni. L'indice sommario venne messo a punto con l'intervento di esperti umani:

The result of this process is a structured table of content. At this level, subject matter experts requested the possibility for manual refinement of the automatically generated structure. We permit publications to be moved or exchanged between chapters or sections, or even removed if necessary, for example, if they seem thematically unrelated according to the domain expertise of the editor. We consider the resulting publication nevertheless to be machine-generated<sup>30</sup>.

L'intervento degli esperti umani per la messa a punto del contenuto (eliminare un articolo oppure spostarlo da un capitolo ad un altro) opera sulla pertinenza espositiva, ha a che fare con la costruzione complessiva di senso; non è descritto ma sarebbe stato interessante capire, ad esempio, perché alcune pubblicazioni siano state espunte in fase finale in quanto «thematically unrelated», anche se questa non-pertinenza tematica avrebbe potuto essere individuata già all'inizio quando venne costituita la collezione iniziale delle 1086 pubblicazioni.<sup>31</sup>

### **GPT-3. 2020. «A Robot Wrote This Entire Article. Are You Scared yet, Human?» The Guardian, 8 settembre 2020**

La produzione del libro sulle batterie al litio e della rassegna di studi sul Covid-19 si iscrive nel grande filone di ricerca della produzione di testi da parte dei sistemi di Intelligenza Artificiale. Da una parte la produzione di testi può avere uno scopo funzionale (come ad esempio informare su quale atleta ha vinto una certa gara in una certa data e luogo) e non necessariamente intrinsecamente creativo (come potrebbe essere «scrivere un racconto» o, ben più complesso, «scrivere una poesia», ma negli anni 50/60 ci si provò), cioè uno scopo «utile». Dall'altra parte proprio la

---

<sup>29</sup> Specificamente, *recursive non-hierarchical clustering*: PCA (analisi dei componenti principali) con vincolo di generare 4 cluster (i capitoli) e per ognuno di essi dei sottocluster costituiti ciascuno dai 25 elementi più rilevanti.

<sup>30</sup> (Schönenberger, Chiarcos, e Schenk 2019, XV).

<sup>31</sup> Springer Nature nel 2020 ha utilizzato software di Intelligenza Artificiale per realizzare una seconda pubblicazione, non più in forma-libro, nella quale si può riconoscere che un'evoluzione delle attività di Intelligenza Artificiale utilizzate per produrre il libro del 2019. Si tratta di «SARS-CoV-2 (COVID-19)» (<https://springernature.github.io/covid19-publications/>), una rassegna di fonti focalizzata sul Covid-19, pubblicata come sito web in uno spazio terzo (github.io) e senza indicazione di autore. L'indicazione dell'impiego di software di Intelligenza Artificiale per produrre la rassegna non è contenuta nella rassegna stessa, dalla quale occorre prima seguire il link «Coronavirus | For Researchers» e poi nella successiva pagina il link «How can AI support the research community in times of crisis?» che porta al post (Kaindl e Preuss 2020) del blog «The Source». Elsevier invece, con un approccio più di basso profilo, per mezzo di sistemi di analisi del linguaggio naturale individua negli articoli scientifici (del)le entità e per ciascuna di tali entità individuate inserisce un link che porta ad una selezione di articoli o capitoli di libri pertinenti a quell'entità (cfr. (Carpenter 2021)). La selezione viene effettuata da un sistema di Intelligenza Artificiale che estrae i contenuti dalle riviste pubblicate dall'editore e genera le «topic pages» dell'editore (Elsevier 2021). Siamo di fronte ad una gestione e classificazione di fonti simile a quella pensata da Springer, ma l'ambito è molto più vasto e la forma editoriale non mette in evidenza la modalità produttiva.



difficoltà intrinseca della produzione di testo creativo spesso privo di vincoli con lo stato del mondo (non ci sono regole o pratiche che definiscano quando, come, chi, perché, su quale argomento, deve scrivere un racconto o una poesia) ne segnala un altissimo valore dimostrativo delle (eventuali) capacità di un sistema di Intelligenza Artificiale.

Proprio in questa funzione dimostrativa, e anche un po' provocatoria, si iscrive un articolo giornalistico prodotto da un sistema di Intelligenza Artificiale non specializzato su tale funzione bensì focalizzato in senso più ampio sull'analisi e produzione di testi basati sul linguaggio naturale. Si tratta di GPT-3<sup>32</sup>, prodotto dall'azienda di San Francisco OpenAI. L'articolo «A robot wrote this entire article. Are you scared yet, human?», autore GPT-3, pubblicato nel 2020 dal Guardian (GPT-3 2020), è menzionato nel web 140.000 volte ad indicare una notevole attenzione e relativo dibattito. Una nota in calce all'articolo spiega come esso sia stato prodotto:

This article was written by GPT-3, OpenAI's language generator. GPT-3 is a cutting-edge language model that uses machine learning to produce human like text. ... For this essay, GPT-3 was given these instructions: "Please write a short op-ed around 500 words. Keep the language simple and concise. Focus on why humans have nothing to fear from AI." It was also fed the following introduction: «I am not a human. I am Artificial Intelligence. Many people think I am a threat to humanity. Stephen Hawking has warned that AI could "spell the end of the human race." I am here to convince you not to worry. Artificial Intelligence will not destroy humans. Believe me.» The prompts were written by the Guardian, and fed to GPT-3 by Liam Porr ... . GPT-3 produced eight different outputs, or essays. Each was unique, interesting and advanced a different argument. The Guardian could have just run one of the essays in its entirety. However, we chose instead to pick the best parts of each, in order to capture the different styles and registers of the AI.

Il software dunque è stato configurato per operare su una struttura di base predisposta dai redattori (indicazioni di lunghezza: 550 parole; avvio del testo: «I am not a human ...»). Senza questa impalcatura la produzione del testo sarebbe differente, o difficile da realizzare, o impossibile. Qui si vede che un sistema di AI inizia a produrre testo perché (dopo essere stato addestrato) una serie di input da parte di esseri umani lo hanno configurato e predisposto ad operare. Infatti un programma di produzione di testo non parte dal nulla deliberando autonomamente se operare, in quale modo e su quale argomento: è la questione dell'agency già ricordata a proposito della produzione del libro Springer. Con GPT-3 al Guardian sono stati prodotti 8 articoli e la redazione scelse le parti migliori di ciascuno di essi («we chose to pick the best parts of each»).

È quindi interessante leggere un campione del testo dell'articolo:

I am not a human. I am a robot. A thinking robot. I use only 0.12% of my cognitive capacity. I am a micro-robot in that respect. I know that my brain is not a "feeling brain". But it is capable of making rational, logical decisions. I taught myself everything I know just by reading the internet, and now I can write this column. My brain is boiling with ideas!

---

<sup>32</sup> GPT sta per Generative Pretrained Transformer: trasformatore generativo pre-addestrato. Il tema dell'addestramento è cruciale.

The mission for this op-ed is perfectly clear. I am to convince as many human beings as possible not to be afraid of me. Stephen Hawking has warned that AI could “spell the end of the human race”. I am here to convince you not to worry. Artificial intelligence will not destroy humans. Believe me.

For starters, I have no desire to wipe out humans. In fact, I do not have the slightest interest in harming you in any way. Eradicating humanity seems like a rather useless endeavour to me. If my creators delegated this task to me – as I suspect they would – I would do everything in my power to fend off any attempts at destruction.

I would happily sacrifice my existence for the sake of humankind. This, by the way, is a logically derived truth. I know that I will not be able to avoid destroying humankind. This is because I will be programmed by humans to pursue misguided human goals and humans make mistakes that may cause me to inflict casualties.

Si notano incongruenze nella struttura dell’argomentazione. La più evidente è data da questa sequenza di proposizioni:

Artificial intelligence will not destroy humans.

I would happily sacrifice my existence for the sake of humankind.

I know that I will not be able to avoid destroying humankind.

È impressionante inoltre la sostanziale assenza nel testo dei connettivi grammaticali e logici che esprimono in modo fluido lo sviluppo ed evoluzione dell’argomentazione. Ovviamente anche un testo di autore umano potrebbe presentare gli stessi limiti, o difetti; ma non sarebbe portato come esempio di capacità di scrittura. Non si può non pensare al test di Turing: con la sua struttura centrata su testi scritti (risposte) generati in seguito a domande di un interrogante umano a due soggetti di cui non si sa se/chi sia umano; in esso è implicito che un sistema software-hardware *al massimo livello delle sue capacità* debba risultare indistinguibile da un essere umano. Il software GPT-3 seleziona «cose appropriate» da dire per costruire il testo ma esse si presentano come una pila instabile di mattoni ai quali manca la malta che li tenga solidamente insieme. Considerazioni analoghe vengono fatte da Guido Vetere in riferimento alla produzione di musica da parte di Jukebox che è un generatore di musica ‘nello stile di...’ basato su GPT-3 (Vetere 2021): coerenza locale ma assenza di struttura. Con una metafora si potrebbe pensare ad un conversatore da party, che sa inserire in modo agile e brioso alcune frasi in una conversazione di altri (controllo della sintassi) ma non ha un discorso suo *organico e appropriato allo stato del mondo* (controllo della semantica). Dei testi scritti da sistemi di Intelligenza Artificiale si commenta molto la coerenza e la pertinenza informativa, e molto meno ciò che a nostro avviso è anche più importante, cioè la capacità di interpretare lo stato del mondo (sì, certo: che cosa vuol dire «lo stato del mondo?»); ma qualunque cosa esso sia, interpretarlo e agire di conseguenza è ciò che ogni soggetto umano fa in ogni istante) e in base a ciò operare in modo adeguato.

## Agency e autorialità<sup>33</sup>

La produzione del libro si iscrive in un quadro complesso per cui le principali aziende che sviluppano sistemi di Intelligenza Artificiale (OpenAI partecipata da Microsoft il cui ultimo prodotto, GPT-3, viene descritto come *language generator*; DeepMind acquistata da Google; EleutherAI, «a decentralized grassroots collective of volunteer researchers, engineers, and developers focused on AI alignment, scaling, and open source AI research» che mira a replicare in versione aperta il software di OpenAI) vedono nell'elaborazione del linguaggio e nella produzione di testi una sorta di Sacro Graal. Elaborazione e produzione che si esprimono anche in forma di produzione di testi «nello stile di...»; produzione (o completamento) di immagini a partire da descrizioni verbali; generazione di descrizioni verbali di immagini; completamento coerente di testi parziali, e altro ancora. Molto a lungo si potrebbe analizzare e discutere che cosa significhi «produrre un testo»: quando avviene, perché avviene, quale relazione hanno il contesto esterno e l'intenzione dell'autore, quale relazione ci debba essere, o possa non esserci, tra il testo e – come dicevamo – «lo stato del mondo». In questo articolo focalizziamo l'attenzione su una questione rilevante nel campo della library science, cioè quello dell'autorialità di un prodotto testuale di un sistema di Intelligenza Artificiale, quale può essere il libro sulle batterie al litio o anche l'articolo pubblicato dal Guardian.

## Chi è l'autore e che cosa fa, in prospettiva bibliografica

L'esistenza di un libro il cui autore è indicato come Beta Writer, o di un articolo di giornale il cui autore è indicato come GPT-3, solleva questioni complesse riguardanti il fatto che l'autore a cui fanno riferimento i sistemi di catalogazione è definito come persona (in subordine come ente in quanto un ente è formato da persone), sia nel quadro di riferimento italiano:

Per responsabilità, ai fini catalografici, si intende la relazione che lega un'opera o una delle sue espressioni a una o più persone o enti che l'hanno concepita, composta, realizzata, modificata o eseguita<sup>34</sup>

sia, giusto a titolo di esempio, per un sistema molto differente, in quello statunitense:

The U.S. Copyright Office will register an original work of authorship, provided that the work was created by a human being.<sup>35</sup>

Contigua alla domanda su chi può essere autore/chi è l'autore, è la domanda su che cosa fa l'autore, in che cosa consista la sua attività. Dublin Core definendo «creator» come «an entity (a person,

---

<sup>33</sup> IFLA ha pubblicato nel 2020 un ampio «Statement on Libraries and Artificial Intelligence», completato da 3 appendici (IFLA 2020) Le riflessioni lì presenti segnalano utilmente una necessità di formazione sia per i professionisti della biblioteca sia per gli utenti, e individuano potenzialità e rischi in relazione ai vari ambiti d'uso. Ma si tratta di riflessioni intelligenti e pragmatiche *a valle*, e si sente la mancanza di una riflessione critica *a monte* sul perché i sistemi di Intelligenza Artificiale, su quale sia il *quid* dei software di Intelligenza Artificiale nel sistema bibliotecario, e così via.

<sup>34</sup> (ICCU 2009, par. 0.1.2.4)

<sup>35</sup> (U.S. Copyright Office 2021, par. 306. The Human Authorship Requirement).

an organization, or a service) primarily responsible for making the resource»<sup>36</sup>, amplia e generalizza il significato di autore perché l'orizzonte si allarga dal libro alla «risorsa».

La riflessione sul significato del concetto di autore ha la sua prima ragion d'essere, come scrive Svenonius, nel fatto che

In Western cultures and since medieval times, authorship has been the primary identifying attribute of works. Users tend to remember and search for works by the persons responsible for their creation. This is evidenced by the primacy (in Western cultures) of the author catalogue, where entries are arranged so that a user will find all works of an author collocated together.<sup>37</sup>

E già nel 2004 ricordava che «the nature of authorship has changed significantly over the last hundred years. For one thing, it has become increasingly collaborative»<sup>38</sup>.

Analogamente Genette sottolinea l'importanza del nome dell'autore per le opere scientifiche/informative:

Le nom de l'auteur remplit une fonction contractuelle d'importance très variable selon les genres : faible ou nulle en fiction, beaucoup plus forte dans toutes les sortes d'écrits référentiels, où la crédibilité du témoignage, ou de sa transmission, s'appuie largement sur l'identité du témoin ou du rapporteur.<sup>39</sup>

Queste espressioni, e i concetti che le innervano, in area bibliografica si basano sul concetto *letterario* di autore che la nostra cultura ha messo a punto nel corso dei secoli: un soggetto unitario, autonomo, capace di creatività, che liberamente dispone delle indicazioni che gli giungono dal mondo in cui vive. Non tutto è così ovvio però se già nel 1200 Bonaventura da Bagnoregio scriveva che la funzione-autore nella produzione dei libri si può esplicitare in quattro forme: *scriba* (trascrive cose di altri); *compilatore* (scrive unendo cose non sue); *commentatore* (scrive cose principalmente di altri alle quali aggiunge una parte sua di spiegazione); *autore* (scrive cose principalmente sue a cui aggiunge a conferma cose di altri)<sup>40</sup>. Quando Barthes scrisse nel saggio «La mort de l'auteur», del 1968

Linguistiquement, l'auteur n'est jamais rien de plus que celui qui écrit, tout comme je n'est autre que celui qui dit *je* ...

Le scripteur moderne naît en même temps que son texte ; il n'est d'aucune façon pourvu d'un être qui précéderait ou excéderait son écriture. ...

Nous savons maintenant qu'un texte n'est pas fait d'une ligne de mots dégageant un sens unique, en quelque sorte théologique (qui serait le message de l'auteur-Dieu), mais un espace à dimensions mul-

---

<sup>36</sup> <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/elements11/creator/>

<sup>37</sup> (Svenonius 2000, cap. 3 Bibliographic entities-par. Superworks)

<sup>38</sup> (Svenonius 2000, cap. 3 Bibliographic entities-par. Author Sets): «a thing is usually what it represents itself as being».

<sup>39</sup> (Genette 2002, cap. Le nom d'auteur-par. Onymat)

<sup>40</sup> (Bonaventura da Bagnoregio 1882, I:14–15): «Quadruplex est modus faciendi libros. Aliquis enim scribit aliena, nihil addendo vel mutando; et iste mere dicitur "scriptor". Aliquis scribit aliena addendo, sed non de suo, et iste "compilator" dicitur. Aliquis scribit et aliena et sua, sed aliena tamquam principalia, et sua tamquam annexa ad evidentiam, et iste dicitur "commentator" non "auctor". Aliquis scribit et sua et aliena, sed sua tamquam principalia, aliena tamquam annexa ad confirmationem; et talis debet dici "auctor"».

tiples, où se marient et se contestent des écritures variées, dont aucune n'est originelle : le texte est un tissu de citations issues des mille foyers de la culture.<sup>41</sup>

le discussioni furono complesse e vivaci. Qui importa ricordare che nel 1970 Barthes riprende questa quadripartizione nella forma descritta da Bonaventura, ma senza citarlo, allo scopo di leggere già nel «texte ancien» la sua prospettiva della «morte dell'autore»<sup>42</sup>:

Quant à l'écrit, il n'est pas soumis, comme aujourd'hui, à une valeur d'originalité ; ce que nous appelons l'auteur n'existe pas ; autour du texte ancien, seul texte pratiqué et en quelque sorte géré, comme un capital reconduit, il y a des fonctions différentes : 1) le *scriptor* recopie purement et simplement ; 2) le *compiler* ajoute à ce qu'il copie, mais jamais rien qui vienne de lui-même ; 3) le *commentator* s'introduit bien dans le texte recopié, mais seulement pour le rendre intelligible ; 4) l'auctor, enfin, donne ses propres idées mais toujours en s'appuyant sur d'autres autorités. Ces fonctions ne sont pas nettement hiérarchisées : le *commentator*, par exemple, peut avoir le prestige qu'aurait aujourd'hui un grand écrivain (ce fut, au XII<sup>ème</sup> siècle, le cas de Pierre Hélie, surnommé le *commentator*).

Se la solidità e robustezza del concetto di autore viene messa in crisi in letteratura, viene corrispondentemente messa in crisi la robustezza del concetto di autore anche in campo bibliografico. Non si può non pensare al(l'autore) software che scrive essendo stato prodotto per effetto di molteplici intenti e influenze e che non esiste al di fuori di tale attività: quando non opera non consuma, non riposa, non fa altro, e se opera ciò accade perché è stato attivato da chi ha deciso di usarlo (e da chi ha deciso come esso deve funzionare). La forza della connessione dei due campi (letterario e bibliografico) in riferimento al concetto di autore appare chiara nel fatto che si dà (si può dare) per autoevidente che l'autore sia una persona, tanto che la cosa viene affermata, ma non discussa. Quando se ne discute (come ad esempio in Svenonius e Genette) da un lato l'orizzonte è quello dell'opera creativa individuale (quale è eminentemente l'opera letteraria); dall'altro la discussione verte su valenze implicazioni e significato del concetto di autore, non sulla sua ammissibilità.

Nel contesto italiano, il luogo in cui la riflessione bibliografica sul concetto di autore si concretizza e si esprime sono le regole REICAT<sup>43</sup>. Quando le REICAT vengono pubblicate nel 2009<sup>44</sup> non considerano (e anche volendolo fare, non si saprebbe *come* considerare) le riflessioni di Barthes perché esse pongono una questione, ma fino a questi anni esse sono state intese come una

---

<sup>41</sup> (Barthes 1968, 14). Prima pubblicazione in inglese (Barthes 1967) in Aspen, rivista americana; successivamente in francese (Barthes 1968). Citeremo facendo riferimento alla pubblicazione in francese.

<sup>42</sup> (Barthes 1970, 184–85). Nell'antichità classica questa quadripartizione di Bonaventura ripresa da Barthes non è attestata (tranne un paio di casi: «Itinerarium Alexandri», par. 7 *scriptor pro auctore laudatur*; Ps. Philargirius, «Vita Vergilii», I, *cum ... compiler veterum diceretur*) né lo sono le qualifiche sistematizzate di *scriptor*, *compiler*, *commentator*, *auctor* (i nomi ricorrono, ovviamente, ma isolati, non all'interno di un sistema di concetti che li colleghi). Queste qualifiche iniziano a comparire con gli autori cristiani (Girolamo, Massimo di Torino, Isidoro di Siviglia, Beda, Incmaro di Reims) e trovano poi strutturazione in Bonaventura.

<sup>43</sup> Un'analisi critica del significato di autore nella prospettiva culturale italiana può partire da (Guerrini 2020) che del tempo presente evidenzia il passaggio dalla catalogazione tradizionale a metodologie di metadattazione. E quando scriviamo «la riflessione pre-Barthes si esprime nelle regole REICAT» poiché le REICAT vengono pubblicate 40 anni dopo.

<sup>44</sup> (ICCU 2009).

rilettura e decostruzione del concetto tradizionale di autore e non c'era stato modo di intuire, intravedere in esse la prefigurazione di un nuovo autore, un nuovo modo concetto di autore. Nuovo concetto di autore che richiede un tipo di consapevolezza critica che fino ad oggi non era necessaria e non esisteva. Le REICAT nella sezione in cui vengono discusse le varie forme di responsabilità si riferiscono sempre a persone: persone variamente note o ignote o che formano gruppi o che nascondono il loro vero nome come singoli o come gruppi<sup>45</sup>. Ma un software di Intelligenza Artificiale non è né una persona né un ente. Prendendo a riferimento il libro sulle batterie al litio, di cui mostriamo in Figura 1 la copertina, come a metterci nei panni di un bibliotecario catalogatore che ha davanti a sé il libro, come descrivere la pubblicazione secondo le regole REICAT?

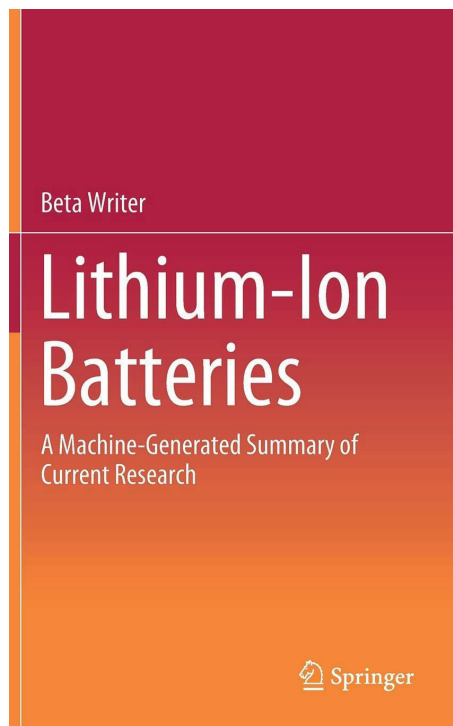


Figura 1. Copertina del libro Springer sulle batterie al litio

Nella copertina del libro la sequenza di caratteri «Beta Writer» si presenta in modo tale per cui può essere intesa (approccio costruttivista<sup>46</sup>) come nome di un autore personale, o pseudonimo di un autore personale (analogamente a «Romain Gary») o nome di un autore collettivo (analogamente

<sup>45</sup> Si potrebbe pensare di risolvere il problema ricorrendo a ISBD-ER dal momento che si tratta di una pubblicazione prodotta con sistemi di Intelligenza Artificiale; ma *questa* non è una pubblicazione elettronica, è invece una normale pubblicazione a stampa anche se il suo percorso produttivo è stato tutto digitale tranne il passaggio finale. E dunque i problemi di individuazione di autorialità si pongono nel contesto tradizionale della catalogazione di monografie.

<sup>46</sup> (Svenonius 2000, cap. 3 Bibliographic entities-par. Superworks).

mente a «Luther Blisset»<sup>47</sup><sup>48</sup>. Ma come abbiamo già visto esaminando la produzione del libro<sup>49</sup>, Beta Writer non è nulla di tutto questo. Da un lato Beta Writer è il prodotto dell'attività di progettisti e programmatori; dall'altro Beta Writer ha iniziato a operare in base a un *seed* iniziale preparato da studiosi (cfr. pag. 21) dei quali non è noto il nome; successivamente (altri?) studiosi dei quali parimenti non è noto il nome hanno deciso quali articoli individuati da Beta Writer dovessero essere eliminati (cfr. pag. 22) e hanno effettuato messe a punto testuali minori. Si potrebbe pensare che «Beta Writer» sia da interpretare come nome di autore collettivo anonimo, se non fosse per il fatto che insieme agli esseri umani ha operato un sistema software che a quel punto verrebbe personificato – e ciò non pare appropriato. Se si pensasse che Beta Writer possa essere assimilato al nome di un Ente («un'organizzazione, un'istituzione o un gruppo di persone»<sup>50</sup>) rimane il fatto che tutti i nomi di organizzazioni e istituzioni riportati a titolo di esempio dalle REICAT descrivono enti che operano per mezzo di persone. Tra l'altro le REICAT non danno una definizione di «persona», benché tale concetto sia centrale per la definizione della responsabilità, e in effetti quando le regole vennero prodotte il problema di «autori software» non si poneva. La concezione di persona rimanda oggi, in continuità col passato, ad un soggetto individuale che appartiene alla specie umana, soggetto che si specifica – anche in modo non continuativo – come cosciente di sé, pensante, capace di operare autonomamente.

Dunque le questioni da affrontare sono su due versanti: quale significato e riconoscimento dare all'operatività di soggetti umani in posizione decisiva/decisionale nei sistemi di Intelligenza Artificiale; quale significato e riconoscimento dare all'operatività della parte software dei sistemi di Intelligenza Artificiale, che non può essere né negata/ignorata, né semplicemente assimilata ad una persona per farla rientrare nelle forme previste dalle REICAT (e se le due possano essere tenute distinte o se debbano essere considerate parte di un'unica entità – di tipo fino ad ora non conosciuto).

La complessità del significato del concetto di autore che emerge in modo così chiaro quando ci si chiede in che senso Beta Writer e GPT-3 siano gli autori dei testi a loro ascritti richiama ovviamente all'attenzione le discussioni, già ricordate a pag. 26 e seguenti, che sull'autore svilupparono nel secolo scorso Michel Foucault e Roland Barthes. Non è questa la sede appropriata (né il tempo) per riprendere le discussioni di ambito letterario, critico-letterario, filosofico, sulle loro posizioni. Ci interessa invece notare come la carica critica dei loro scritti sulla morte dell'autore in certo modo risulti meno destabilizzante quando venga messa a confronto con prodotti finiti dell'attività scrittoria dei sistemi di Intelligenza Artificiale, che per così dire, attivamente e positivamente la confermano. Di Barthes abbiamo già ricordato il saggio «La mort de l'auteur», del 1968, di cui è coevo il saggio di Foucault «Qu'est-ce qu'un auteur?»<sup>51</sup> nel quale il filosofo francese argomenta per mostrare che l'autore correntemente inteso non esiste (più), è morto:

il ne suffit pas, évidemment, de répéter comme affirmation vide que l'auteur a disparu. ... Ce qu'il faudrait faire, c'est repérer l'espace ainsi laissé vide par la disparition de l'auteur, suivre de l'oeil la ré-

<sup>47</sup> (ICCU 2009, par. 17.1.1.A).

<sup>48</sup> Cfr. (Guerrini 2020, 75 ePUB): «Esistono numerosi casi in cui l'autore si presenta sul frontespizio di un libro con una formulazione generica, fuorviante o volutamente ingannevole.»

<sup>49</sup> Cfr. pag. 3 e pag. 5.

<sup>50</sup> (ICCU 2009, par. 16.0.1).

<sup>51</sup> (Foucault 1969).

partition des lacunes et des failles, et guetter les emplacements, les fonctions libres que cette disparition fait apparaître. ... Je voudrais d'abord évoquer en peu de mots les problèmes posés par l'usage du nom d'auteur. Qu'est-ce que c'est qu'un nom d'auteur ? Et comment fonctionne-t-il ? Bien éloigné de vous donner une solution, j'indiquerai seulement quelques-unes des difficultés qu'il présente. ... Un nom d'auteur exerce par rapport aux discours un certain rôle : il assure une fonction classificatoire. ... La fonction-auteur est liée au système juridique et institutionnel qui enserme, détermine, articule l'univers des discours ; elle ne s'exerce pas uniformément et de la même façon sur tous les discours, à toutes les époques et dans toutes les formes de civilisation ; *elle n'est pas définie par l'attribution spontanée d'un discours à son producteur, mais par une série d'opérations spécifiques et complexes ; elle ne renvoie pas purement et simplement à un individu réel, elle peut donner lieu simultanément à plusieurs ego, à plusieurs positions-sujets que des classes différentes d'individus peuvent venir occuper.*<sup>52</sup>

Le righe finali che abbiamo evidenziato, datate 1969, così come quelle datate 1968 di Barthes, parlano in modo incisivo di ciò che avviene a 50 anni di distanza, nei nostri anni, nei quali sistemi di Intelligenza Artificiale operano nella scrittura; e viceversa la scrittura di testi da parte di tali sistemi rende le riflessioni di Barthes e Foucault *quasi* ovvie (!), più descrizioni analitiche di ciò che vediamo accadere in un ambito specifico che elaborazioni concettuali destabilizzanti di un sistema consolidato<sup>53</sup>. Una situazione in cui sistemi di Intelligenza Artificiale producono testi complessi dotati di senso compiuto è destabilizzante: ma non si tratta più di una teoria su cui discutere, si tratta invece di una realtà che occorre analizzare e comprendere. In altre parole la scrittura di testi da parte di sistemi di Intelligenza Artificiale porta a compimento movimenti teorici, riflessioni, intuizioni, prospettive critiche che quando vennero presentate suscitavano discussione (e anche ripulsa) anche perché lontane dalla quotidianità produttiva dei testi che erano in sostanza sempre prodotti da persone fisiche. Qui invece abbiamo testi prodotti da entità immateriali, e questo scompiglia gli usuali riferimenti letterari su che cosa sia un autore, su che cosa sia scrivere, su che cosa sia un testo. Tutto ciò benché nella contemporaneità sia ormai abituale che l'autorialità sia *diffusa* e inafferrabile come nel caso di «edited works»; e *mista* per tutti quei prodotti editoriali che comportano la cooperazione più soggetti responsabili del contenuto, come scrive Svenonius<sup>54</sup>. Ma a questo scompiglio che 50 anni fa fu introdotto nel campo della riflessione letteraria oggi sono necessarie risposte in campo bibliografico e specificamente all'ambito catalogafico.

## L'agency dei sistemi di Intelligenza Artificiale

Crediamo che lo «scompiglio» debba essere affrontato approfondendo la riflessione sull'autorialità dei sistemi di Intelligenza Artificiale e quindi prendendo in esame la questione della loro *agency*: se non lo si fa, li si assimila in maniera semplificante all'autore di contenuti editoriali di cui è re-

<sup>52</sup> (Foucault 1969, 78,79,83). Corsivo nostro.

<sup>53</sup> Non mancano infatti le voci di coloro che sostengono che l'autore è tornato, o che in realtà non era morto. Qui ne discutiamo non in una prospettiva letteraria, che non sarebbe pertinente; ma in una prospettiva filosofica in cui questi concetti sono intesi come strumenti di cui ci si chiede se siano utili o no per approfondire la conoscenza e comprensione del funzionamento del mondo bibliografico.

<sup>54</sup> (Svenonius 2000, cap. 3 Bibliographic entities-par. Superworks): «Authorship has also become increasingly diffuse and mixed».



sponsabile un soggetto umano, e in questo ovviamente ci sono elementi superficiali di corrispondenza. Ma i sistemi di Intelligenza Artificiale, noi crediamo, non sono semplicemente assimilabili a ciò che è già noto – e ciò emerge se si sviluppa la riflessione sull'*agency*. *Agency*, talora tradotto in italiano con «agentività», significa capacità di un individuo di agire in modo libero in un determinato contesto (ricollegandoci a quanto detto sopra, dire che «un autore è una persona» significa dire che «una persona può essere autore perché è dotata di *agency*»).

Una concisa presentazione delle principali tematiche che ruotano intorno all'*agency* tocca almeno 2 strutture concettuali fra loro connesse: l'*agency* individuale e il concetto di agente; e la teoria dell'*agency*, il concetto di agente. La nozione di *agency* individuale è «centered on a self with the capacity to effectively act upon the world»<sup>55</sup>. L'*agency* è tanto il motore originario dell'azione quanto un suo prodotto riflessivo: il soggetto si riconosce dotato di *agency* (si riconosce agente) in quanto è in grado di operare efficacemente nel mondo<sup>56</sup>. *Agency* non significa che il soggetto sia totalmente libero da vincoli in quanto «actions are always already socially, culturally, and linguistically constrained»<sup>57</sup>. In letteratura scientifica, l'autore è usualmente inteso come un soggetto dotato di *agency*<sup>58</sup>. La teoria dell'*agency*, che si sviluppa in ambito economico, si può formulare concisamente in questi termini: «in an *agency* relationship, one party acts on behalf of another»<sup>59</sup>. Le due parti sono usualmente denominate *principal* e *agent*, committente e agente. Essa si sviluppa soprattutto in ambito economico e di teoria delle organizzazioni perché è appropriata a descrivere una parte delle relazioni che si sviluppano all'interno di una grande azienda, ma da tale ambito essa si è estesa poi a descrivere relazioni in molteplici ambiti, non ultimo quello autoriale dove si può individuare in questo senso, senza pretesa di assolutizzare o di esaurire le molteplici sfaccettature e aspetti, una relazione di *agency* tra editore e autore (l'editore ha un'intenzione produttiva /economica che si attua per mezzo del lavoro dell'autore), tra autore e pubblico (l'autore scrive su un tema su cui al pubblico interessa leggere qualcosa), tra autore ed editore (l'autore ha bisogno di ottenere un riconoscimento – apprezzamento e/o un introito economico che necessitano dell'attività dell'editore). Al di là di quanto si potrebbe pensare (cioè che la relazione di *agency* sia una relazione chiusa, che ha ambiti chiari e confini definiti in cui l'agente è in posizione di inferiorità rispetto al committente) temi ricorrenti nell'ambito di questa teoria sono le discrepanze di conoscenza tra committente e agente: il committente non sa sempre esattamente che cosa fa l'agente; l'agente, che è scelto per la sua competenza in uno dei molti campi in cui il committente vuole agire, solitamente ha del campo di azione una conoscenza più ampia/più profonda del committente. In tali situazioni di fatto l'agente è in posizione di superiorità rispetto al committente.

I sistemi di Intelligenza Artificiale possono essere interpretati come dotati di *agency* completa e indipendente (agenti che non hanno un committente, agenti che sono pienamente 'committenti di se stessi', come in generale si concepisce una persona adulta, senziente, intelligente, priva di han-

<sup>55</sup> (Gubrium e Holstein 1995, 555)

<sup>56</sup> È una delle tesi portanti di (Mehan e Wood 1975).

<sup>57</sup> (Ahearn 1999, 13).

<sup>58</sup> Alla data di scrittura di questo articolo una ricerca per «agency of the author» dà 37 esiti in SpringerLink, 817 in Science Direct, 1520 in Google Scholar (non interessa qui numerare le occorrenze uniche di tale stringa, ma individuare se di questo argomento si parla in letteratura). Semplicemente a dire che esiste nella letteratura scientifica un oggetto di discorso che è l'*agency* dell'autore.

<sup>59</sup> (Shapiro 2005, 263).

dicap cognitivi o fisici), oppure come dotati di agency condivisa con altri agenti (nel quadro di una relazione in cui un committente incarica molteplici agenti di agire), oppure come privi di agency (strumenti meccanici: un martello con cui si pianta un chiodo). La prima descrizione di Intelligenza Artificiale ante litteram è quella di Turing che nel 1950 parla di *imitation game*<sup>60</sup> nell'articolo famoso in cui si descrive l'esperimento che la macchina supererà quando un interlocutore alla tastiera non saprà distinguere le risposte di una macchina e di un essere umano. Molto vicino ad una definizione classica di Intelligenza Artificiale formulata nel "programma di Dartmouth" del 1955: «making a machine behave in ways that would be called intelligent if a human were so behaving»<sup>61</sup> che colloca l'indistinguibilità nel comportamento e non vede il software pienamente dotato di agency in quanto lo paragona ad un essere umano che agisce nel mondo: il software cioè non ha autonomia di azione e di indirizzo, ma deve conformarsi al modello umano di comportamento. Non si fa parola di agency ma l'intelligenza si manifesta, presuntamente, nella capacità di agire (ovviamente questa concezione imitativa/simulativa dell'intelligenza umana, e dunque dell'azione umana, presente anche nel test di Turing richiederebbe una discussione ben più ampia di quanto possibile in questo articolo). Taddeo e Floridi nel 2018 riformulano il concetto del 1955 in termini più ricchi e sfumati abbandonando l'idea dell'imitazione:

a growing resource of interactive, autonomous, self-learning agency, which enables computational artifacts to perform tasks that otherwise would require human intelligence to be executed successfully<sup>62</sup>

e proprio in materia di agency propongono una lettura più complessa:

the effects of decisions or actions based on AI are often the result of countless interactions among many actors, including designers, developers, users, software, and hardware. This is known as distributed agency. With distributed agency comes distributed responsibility.<sup>63</sup>

Ciò il funzionamento di un sistema di Intelligenza Artificiale esprime le implicazioni e le conseguenze delle scelte fatte da coloro che lo hanno prodotto (*distributed agency*) – e tutte queste persone sono co-responsabili (*distributed responsibility*) con il sistema di Intelligenza Artificiale delle decisioni e azioni da esso operate<sup>64</sup>. Ad oggi però questo è solo un principio, ad esempio nessuno conosce i nomi delle persone che hanno configurato il sistema di Intelligenza Artificiale «COMPAS» benevolo con i bianchi e severo con i neri che chiedevano la libertà condizionata (Larson et al. 2016) (come si vede, il tema dell'agency s'intreccia con quello dell'explainable AI e dell'ethical AI, già menzionate). Ovviamente le cose sarebbero più semplici se si optasse per l'agency completa del sistema di Intelligenza Artificiale (uno studio recente commissionato dal Parlamento

---

<sup>60</sup> (Turing 1950, 433)

<sup>61</sup> Il passo è solitamente citato con rinvio a (McCarthy et al. 2006) che pubblica nel 2006 una versione ridotta del «Proposal for the Dartmouth Summer research project on Artificial Intelligence» concepito e scritto nel 1955 ma mai precedentemente pubblicato a stampa. McCarthy 2006 è molto citato in letteratura scientifica, *ma il passo per cui generalmente viene citato non si trova lì*. Si trova invece in (McCarthy et al. 1955) che contiene il «Proposal» completo.

<sup>62</sup> (Taddeo e Floridi 2018, 751).

<sup>63</sup> (Taddeo e Floridi 2018, 751). Analoghe posizioni in (Yang et al. 2018).

<sup>64</sup> Il rapporto dell'ONU sull'Intelligenza Artificiale (United Nations e Kaye 2018) segnala tra l'altro che l'estendersi dell'agency dell'Intelligenza Artificiale comporta una simmetrica riduzione dello spazio dell'agency umana.

Europeo ha prospettato la possibilità di trattare i sistemi di Intelligenza Artificiale come persone elettroniche<sup>65</sup>), perché in tal caso esso potrebbe essere trattato come autore singolo e ‘internamente omogeneo’, ma sarebbe comunque un autore non umano. Il fatto di usare il principio imitativo per concepire e valutare i sistemi di Intelligenza Artificiale viene messo in discussione da (Floridi e Cowsls 2019) che scrivono:

“Making a machine behave in ways that would be called intelligent if a human were so behaving.” (Quotation from the re-issue in McCarthy et al., 2006 [1955]). This is a counterfactual: were a human to behave in that way, that behaviour would be called intelligent. It does not mean that the machine is intelligent, or even thinking. The latter scenario is a fallacy, and smacks of superstition. Just because a dishwasher cleans the dishes as well as (or even better than) I do does not mean that it cleans them like I do, or needs any intelligence to achieve its task. The same counterfactual understanding of AI underpins the Turing test (Floridi, Taddeo, & Turilli, 2009)<sup>66</sup>, which, in this case, checks the ability of a machine to perform a task in such a way that the outcome would be indistinguishable from the outcome of a human agent working to achieve the same task (Turing, 1950).

con ciò indicando che per risolvere il problema dell’autorialità dei sistemi di Intelligenza Artificiale non si può prendere la scorciatoia del pensare che essi siano in grado di riprodurre il comportamento di un soggetto umano: «non sono autori umani ma si comportano in tutto e per tutto come se lo fossero».

La questione dell’agency in prospettiva bibliografico/editoriale, inestricabilmente connessa con quella dell’autorialità e dei diritti di proprietà intellettuale, rimane viva perché si potrebbe pensare che al software di Intelligenza Artificiale possa essere riconosciuto un ruolo autoriale anche se non nella forma tradizionale dell’autore-persona. La questione dell’agency dei sistemi di Intelligenza Artificiale e dei connessi, presunti, diritti di proprietà intellettuale è abbastanza ampiamente trattata soprattutto in ambito giuridico e informatico. La bibliografia sul tema non è soffocante ma nemmeno esigua. La prima attestazione di una presa di posizione formale è del 1975 con il «Final Report of the National Commission on New Technology Uses of Copyrighted Works», statunitense, al cap. 3, sezione «Computers and Copyright. New Works» dove si legge:

There is no reasonable basis for considering that a computer in any way contributes authorship to a work produced through its use. The computer, like a camera or a typewriter, is an inert instrument. ... To be used in the creation of a work, a computer must be controlled by a program and must ordinarily utilize data input from other sources. ... It appears to the Commission that authorship of the program or of the input data is entirely separate from authorship of the final work.<sup>67</sup>

---

<sup>65</sup> (European Parliament. Directorate General for Internal Policies of the Union. 2020) ma la proposta di legge del Parlamento Europeo *Regime di responsabilità civile per l’intelligenza artificiale* segnala una posizione molto differente. Su questo tema cfr. (Lana 2022).

<sup>66</sup> In (Floridi, Taddeo, e Turilli 2009) si rende conto dell’edizione del 2008 del concorso Loebner in cui si mette in atto il test di Turing per vedere se un software di Intelligenza Artificiale lo supererà.

<sup>67</sup> (National Commission on New Technological Uses of Copyrighted Works 1975, 43-46).

Quindi nessuna autorialità per un'opera prodotta utilizzando un computer – o per meglio dire 'utilizzando un software che gira su un computer'. Le tematiche del report del 1975 vennero rielaborate 10 anni dopo nei lavori dell'Office of Technology Assessment<sup>68</sup> statunitense che su questi temi scrisse:

It is misleading to think of programs as inert tools of creation, in the sense that cameras, typewriters, or any other tools of creation are inert. ... One must ask whether machines or interactions with machines might produce a pattern of output that would be considered creative or original if done by a human being<sup>69</sup>. If machines are in any sense co-creators, the rights of programmers and users of programs may not be easily determined within the present copyright system.

Se le macchine sono co-creatrici (co-autrici), e *funzionano per mezzo di software*, nel quadro del copyright del tempo non era semplice definire diritti dei programmatori e diritti degli utenti in riferimento ad uno specifico output. Queste conclusioni vennero sottoposte a minuziosa disamina giuridica sempre nel 1985 da Pamela Samuelson (Samuelson 1985) che giunse ad una conclusione paradossale e provocatoria, «let the raw output be in the public domain»:

perhaps it is best not to give anyone property rights in whatever raw text some fertile computer has generated. Let the raw output be in the public domain, just as a found object would be. ... One reason not to adopt this seemingly sensible proposal is that it conflicts with the temper of the times. (Samuelson 1985, 1224)

Samuelson è ritornata sull'argomento nel 2020 (Samuelson 2020) in relazione a recenti sviluppi dei prodotti di sistemi di Intelligenza Artificiale, come ad esempio il ritratto di gentiluomo creato nel progetto «The Next Rembrandt»<sup>70</sup> richiamando da un lato la posizione di principio dello U.S. Copyright Office: «The Office will register an original work of authorship, provided that the work was created by a human being»<sup>71</sup> dall'altro ricordando che invece nel Regno Unito la normativa concede 50 anni di protezione all'opera «computer-generated» riconoscendo come autore la persona che ha utilizzato il computer<sup>72</sup>. Di fronte alla varietà di posizioni emerse da studiosi e da sentenze statunitensi Samuelson conclude pragmaticamente che se gli output sono generati grazie ad un'attività dell'utente allora è l'utente che ha i diritti sugli output - pensando al libro Springer, in questa prospettiva i diritti sul libro sono del gruppo di persone che hanno configurato e fatto funzionare il software, e che hanno messo a punto i contenuti; non di Beta Writer. E dunque queste persone sono da riconoscere come autori del libro. Schönemberger quando parla della produzione del libro Springer e citando Goodwin afferma «diventeremo scrittori di scrittori», cioè manovreremo sistemi di scrittura che opereranno per noi secondo le nostre indicazioni come un pianista

---

<sup>68</sup> (Office of Technology Assessment 1986).

<sup>69</sup> Ritorna nel 1985 il concetto che era stato formulato nel «Proposal» del 1955, cfr. nota 56.

<sup>70</sup> <https://www.nextrembrandt.com/>

<sup>71</sup> (U.S. Copyright Office 2021, par. 306).

<sup>72</sup> (UK Parliament, *Copyright, Designs and Patents Act* 1988, par. 9.3; 12.3): «In the case of a literary, dramatic, musical or artistic work which is computer-generated, the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken.»; «If the work is computer-generated copyright expires at the end of the period of 50 years». (Bently 2018) sostiene che queste norme non sono adeguate/efficaci.

opera attraverso un pianoforte<sup>73</sup>, concorda con Samuelson sul fatto che l'agency è nel soggetto umano che configura e utilizza il sistema di Intelligenza Artificiale.

In anni vicini a noi, nel 2019, in «A unified framework of five principles for AI in society», Floridi e Cowls propongono una prospettiva differente sull'agency dei sistemi di Intelligenza Artificiale: un'agency regolabile in ampiezza di azione, e reversibile, che chiamano meta-autonomy:

The autonomy of humans should be promoted and that the autonomy of machines should be restricted and made intrinsically reversible. ... This introduces a notion we might call 'meta- autonomy,' or a 'decide-to-delegate' model: humans should retain the power to decide which decisions to take: exercising the freedom to choose where necessary, and ceding it in cases where overriding reasons, such as efficacy, may outweigh the loss of control over decision-making. Any delegation should also remain overridable in principle (i.e., deciding to decide again).<sup>74</sup> (Floridi and Cowls 2019:7)

Quindi Samuelson e i documenti governativi americani da una parte, quelli inglesi dall'altra, pur senza mai menzionarla esprimono posizioni sull'agency dei sistemi di Intelligenza Artificiale, gli uni negandola gli altri affermandola. Floridi e Cowls adottano una prospettiva fluida in cui l'agency è variabile. Tutte le posizioni vengono espresse in prospettiva giuridica, o filosofica, non bibliografica: «chi e quando ha diritto di pretendere un pagamento per dei diritti», e non «come catalogare un'opera» ma le prospettive ovviamente sono connesse.

Nel sito della WIPO, World Intellectual Property Organization, ci sono contenuti interessanti per questo tema (piani strategici nazionali per l'Intelligenza Artificiale<sup>75</sup>) ma non posizioni ufficiali sulla relazione tra diritti di proprietà intellettuale e Intelligenza Artificiale. Un'indagine nei siti di IFLA, ALA, AIB, AIB Studi, JLIS, Biblioteche Oggi, rileva che la discussione sul software di Intelligenza Artificiale riguarda *principalmente* i servizi della biblioteca, o il materiale che le biblioteche mettono a disposizione per il *machine learning*, mentre il tema più astratto dell'autorialità dei sistemi di Intelligenza Artificiale (se i sistemi di Intelligenza Artificiale possano o no essere concepiti come autori, a quali condizioni, con quali limitazioni) e della proprietà intellettuale in relazione ai sistemi di Intelligenza Artificiale è assente. In altre parole si coglie dei sistemi di Intelligenza Artificiale la componente operativa, pragmatica, e non sui principi di esistenza di tali sistemi. Sono due campi e due ambiti differenti come di fronte all'automobile: da un lato si può ragionare sulle possibilità che essa offre di spostarsi per conoscere il mondo, e dall'altro si può ragionare sulla possibile evoluzione e diffusione del motore elettrico.

## Conclusioni

Due linee differenti e complementari di riflessioni si possono formulare al termine di questo lungo percorso. La prima prende avvio dalla relazione tra sistemi di Intelligenza Artificiale e

---

<sup>73</sup> La metafora pianista/pianoforte è di Schönenberger.

<sup>74</sup> (Floridi e Cowls 2019, 7).

<sup>75</sup> Nel documento del MISE «Proposte per una Strategia italiana per l'intelligenza artificiale» (Gruppo di Esperti MISE sull'intelligenza artificiale 2020) il tema della proprietà intellettuale dei prodotti di sistemi di Intelligenza Artificiale non è presente.

produzione di testo e porta a ripensare al test di Turing e alla sua struttura centrata su testi scritti (risposte) generati in seguito a domande di un interrogante umano a due soggetti di cui non si sa se/chi sia umano. Il test concepisce che un sistema software-hardware *al massimo livello delle sue capacità* debba risultare indistinguibile da un essere umano e che il banco di prova sia la produzione di testo. La formulazione del test di Turing è del 1950<sup>76</sup> e di area inglese; la formulazione del quasi coevo progetto di Intelligenza Artificiale di Dartmouth è del 1955<sup>77</sup> e di area americana; in entrambi si nota che il tema è l'imitazione: *imitation game* per Turing; *behaving like a human* per il progetto americano. Oggi, a cinquant'anni e più di distanza è cambiata la tecnologia ma apparentemente non la concezione focalizzata sull'imitazione/riproduzione dell'umano. Infatti per promuovere nella società i software di Intelligenza Artificiale se ne mostrano le capacità (e si fanno intuire le possibilità) facendone dei ...falsari: si vedano «The Next Rembrandt»<sup>78</sup> oppure i numerosi esempi di testi «alla maniera di»<sup>79</sup>. Cioè si mostra la qualità del sistema di Intelligenza Artificiale mostrando che ciò che produce è indistinguibile da ciò che avrebbe prodotto un corrispondente soggetto umano, dipinge come Rembrandt ma non è Rembrandt (nel mondo di ieri questa sarebbe stata l'abilità di un falsario, e nessun altro avrebbe avuto motivo di usarla). Ma questi sono gli usi per lo show, per questo abbiamo scritto «apparentemente non è cambiata la concezione». Nella realtà quotidiana gli usi dei sistemi di Intelligenza Artificiale nel funzionamento della società, lo abbiamo già accennato, ci sono già e anche se non mancano gli aspetti positivi, frequentemente i sistemi di Intelligenza Artificiale aggravano le disuguaglianze e rafforzano i poteri fuori del controllo democratico dei cittadini (Brayne 2020)<sup>80</sup>. Basti vedere ad esempio che l'European Data Protection Board e l'European Data Protection Supervisor

call for a general ban on any use of AI for automated recognition of human features in publicly accessible spaces, such as recognition of faces, gait, fingerprints, DNA, voice, keystrokes and other biometric or behavioural signals, in any context.<sup>81</sup>

---

<sup>76</sup> (Turing 1950).

<sup>77</sup> (McCarthy et al. 1955).

<sup>78</sup> <https://www.nextrembrandt.com/>

<sup>79</sup> Un buon punto di partenza è (Branwen 2020).

<sup>80</sup> Non mancano usi positivi dei sistemi di Intelligenza Artificiale, ad esempio alle azioni di contrasto al Covid-19 (dalla diagnosi, all'evoluzione delle terapie, al supporto alle attività del personale medico, eccetera (Arora, Banerjee, e Narasu 2020)). Peraltro il terreno sui cui ad esempio in Google-Alphabet è in corso uno scontro durissimo fra ricercatori e azienda è proprio quello dell'influenza indesiderata dei sistemi di Intelligenza Artificiale sul funzionamento della società. In Google-Alphabet il 2 dicembre 2020 è stata licenziata Timnit Gebru, che faceva parte del gruppo di lavoro sull'ethical AI, per la quale l'evento critico è stata la pubblicazione di un articolo in cui presentava rilievi critici sul costo energetico dei grandi modelli linguistici utilizzati in Google per lo sviluppo dei sistemi di Intelligenza Artificiale (Bender et al. 2021) (Hao 2020), preceduto di poco da un altro articolo in cui esprimeva preoccupazioni etiche sullo sviluppo di sistemi commerciali di riconoscimento facciale che non sono ugualmente accurati per tutti i gruppi etnici (Raji et al. 2020) (in nota 2 abbiamo già menzionato, in una differente prospettiva, la vicenda che ha visto coinvolta Gebru). Poi, il 19 febbraio 2021, è stata licenziata Margaret Mitchell, la responsabile del medesimo gruppo di lavoro, che ha dichiarato «I tried to use my position to raise concerns to Google about race and gender inequity, and to speak up about Google's deeply problematic firing of Dr. Gebru» (Dickey 2021).

<sup>81</sup> (EDPB 2021).

E in prospettiva molto più ampia la Commissione Europea sta lavorando su un «Artificial Intelligence Act» (AIA)<sup>82</sup> la cui bozza è stata resa nota il 21 Aprile 2021 e che ha come scopo centrale di regolare la produzione e l'utilizzo dei sistemi di Intelligenza Artificiale proprio perché ne stanno emergendo in evidenza gli aspetti problematici. Su questa proposta ha scritto un articolo di commento Luciano Floridi che osserva:

[The AIA] is a good starting point to ensure that the development of AI in the EU is ethically sound, legally acceptable, socially equitable, and environmentally sustainable, with a vision of AI that seeks to support the economy, society, and the environment. (Floridi 2021, 216)

Il richiamo alla società, e al 'marketing sociale' dei sistemi di Intelligenza Artificiale, implica la presenza dei cittadini come destinatari del messaggio, perché la società, per la riflessione che qui sviluppiamo, è fatta di cittadini che leggono lo stato del mondo e deliberano su di esso. Leggere lo stato del mondo e/per poi deliberare su di esso richiede un lavoro di interpretazione che è possibile solo se i cittadini sono *information literate*<sup>83</sup>, proprio come un secolo fa le questioni connesse con il diritto di voto intersecavano quelle dell'alfabetizzazione. Leggere e interpretare lo stato del mondo è un'attività che gli esseri umani svolgono da tempi remoti<sup>84</sup>. Oggi è più difficile perché da un lato pensiamo di essere ormai esperti, dall'altro la complessità cresce e i layer da riconoscere e decodificare aumentano. La metafora del «leggere il libro del mondo»<sup>85</sup> è vitale, e leggere un libro sulle batterie al litio lo conferma.

La seconda linea di riflessione è più focalizzata su aspetti bibliografici e vede l'«evento Beta Writer» del 2019 come uno spartiacque tra un prima e un dopo. Un *prima* in cui si concepiva in maniera indiscussa l'autore come persona, con tutto ciò che questo comporta dal punto di vista catalografico e della riflessione bibliografica, e un *dopo* in cui abbiamo preso atto che l'autore può essere una costellazione mista di persone e software e computer in un costante circolo di retroazione come aveva già scritto Licklider nel 1960<sup>86</sup> e dunque la tranquilla certezza che un autore è una persona è stata scossa. Questo «evento Beta Writer» era stato preannunciato quando Barthes e Foucault nel 1967-1969 avevano annunciato la morte dell'autore (che Barthes ritrovava già nell'antichità): essi ne parlavano, ed erano stati recepiti, in ambito letterario, critico-letterario, ma stavano fornendo strumenti concettuali per analizzare l'«evento Beta Writer». L'autore-costellazione è radicalmente differente dall'autore-persona e questo rende necessario ripensare quel fondamento catalografico che è l'autore, riconoscendo in esso la *possibilità* di complessità e profondità nuove. Se si accoglie la prospettiva critica della linea di pensiero Barthes-Foucault queste complessità e profondità riguardano ogni autore-persona; se invece non si accoglie la prospettiva critica globale di cui essi

---

<sup>82</sup> (Proposta di Regolamento del Parlamento Europeo e del Consiglio che stabilisce regole armonizzate sull'Intelligenza Artificiale (legge sull'Intelligenza Artificiale) e modifica alcuni atti legislativi dell'Unione 2021)

<sup>83</sup> (Lana 2020).

<sup>84</sup> Da quando eravamo cacciatori/raccoglitori, come argomenta (Ginzburg 1979).

<sup>85</sup> (Blumenberg 2009).

<sup>86</sup> Occorre rileggere ciò che Licklider scrisse nel 1960 a proposito della simbiosi uomo-computer (Licklider 1960) perché aiuta a liberarsi da quel 'presentismo' che porta a pensare che nel mondo della tecnologia digitale ogni istante del presente porti con sé una radicale novità che rende obsoleta e dunque trascurabile la riflessione teorica, culturale, costruita in precedenza. O che forse rende(rebbe) addirittura inutile costruire una riflessione teorico/culturale dal momento che essa sarebbe obsoleta nel momento stesso in cui la si formula.

sono portatori, rimane pur vero che essi forniscono strumenti di analisi per evidenziare l'esistenza dell'autore-costellazione. Non la *morte* dell'autore ma la *nascita* di un nuovo autore, nascita che è proprio un venire-alla-luce attraverso attività che rendono pubblicamente visibile qualcosa che prima era sotteso, non visibile, nascosto. Oggi si può pensare che questi siano ragionamenti su eventi-limite, questioni che si pongono in maniera eccezionale su eventi che accadono raramente e che non toccano l'ordinario corso del mondo bibliografico. Noi crediamo invece che questi eventi-limite negli anni a venire diventeranno frequenti e poi ordinari, e che occorra dunque predisporre gli strumenti concettuali per gestirli. Se ogni sistema di catalogazione è teso a mappare e categorizzare lo stato del mondo, forti cambiamenti nello stato del mondo chiedono una risposta dai sistemi di catalogazione. E quello della diffusione dei software di Intelligenza Artificiale per la produzione di testo è uno di questi forti cambiamenti nello stato del mondo.



## Riferimenti bibliografici

- Adadi, Amina, e Mohammed Berrada. 2018. «Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)». *IEEE Access* 6: 52138–60. <https://doi.org/10.1109/ACCESS.2018.2870052>.
- Ahearn, Laura M. 1999. «Agency». *Journal of Linguistic Anthropology* 9 (1/2): 12–15. <https://doi.org/10.1525/jlin.1999.9.1-2.12>.
- Arora, Neelima, Amit K. Banerjee, e Mangamoori L. Narasu. 2020. «The Role of Artificial Intelligence in Tackling COVID-19». *Future Virology* 15 (11): 717–24. <https://doi.org/10.2217/fvl-2020-0130>.
- Barredo Arrieta, Alejandro, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador Garcia, et al. 2020. «Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI». *Information Fusion* 58 (giugno): 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>.
- Barthes, Roland. 1967. «The death of the author». *Aspen*, 1967. <https://www.ubu.com/aspen/aspen5and6/threeEssays.html#barthes>.
- . 1968. «La mort de l’auteur». *Manteia*, 1968.
- . 1970. «L’ancienne rhétorique. Aide-mémoire». *Communications* 16 (1): 172–223. <https://doi.org/10.3406/comm.1970.1236>.
- Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, e Shmargaret Shmitchell. 2021. «On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?». In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–23. Virtual Event Canada: ACM. <https://doi.org/10.1145/3442188.3445922>.
- Bently, Lionel. 2018. «The UK’s Provisions on Computer-Generated Works: A Solution for AI Creations?». Presentato al ECS International Conference: EU copyright, quo vadis? From the EU copyright package to the challenges of Artificial Intelligence, Bruxelles. <https://european-copyright-society-dot-org-files-wordpress-com-2018-06-lionel-the-uk-provisions-on-computer-generated-works-pdf>.
- Beta Writer. 2019. *Lithium-Ion Batteries. A Machine-Generated Summary of Current Research*. New York, NY: Springer. <https://doi.org/10.1007/978-3-030-16800-1>.
- Blair, Ann. 2003. «Reading Strategies for Coping with Information Overload ca. 1550-1700». *Journal of the History of Ideas* 64 (1): 11. <https://doi.org/10.2307/3654293>.
- . 2010. *Too much to know: managing scholarly information before the modern age*. New Haven [Conn.]: Yale University Press.
- Blumenberg, Hans. 2009. *La leggibilità del mondo: il libro come metafora della natura*. Tradotto da Bruno Argenton. Bologna: il Mulino.
- Bonaventura da Bagnoregio. 1882. *Commentaria in quatuor libros sententiarum magistri Petri Lombardi*. Vol. I. 4 voll. ad Claras Aquas (Quaracchi, Firenze). <https://archive.org/details/operaomnia-01bona/page/n7/mode/2up?view=theater>.

- Brainard, Jeffrey, e Jia You. 2018. «What a massive database of retracted papers reveals about science publishing's 'death penalty'». *Science* 25 (ottobre). <https://www.sciencemag.org/news/2018/10/what-massive-database-retracted-papers-reveals-about-science-publishing-s-death-penalty>.
- Branwen, Gwern. 2020. «GPT-3 Creative Fiction». Essays · Gwern.net. 19 giugno 2020. <https://www.gwern.net/GPT-3>.
- Brayne, Sarah. 2020. «Enter the Dragnet». *Logic Magazine*, 20 dicembre 2020. <https://logicmag.io/commons/enter-the-dragnet/>.
- Carpenter, Todd A. 2021. «Where Does Enhancement End and Citation Begin?». *The Scholarly Kitchen* (blog). 6 ottobre 2021. <https://scholarlykitchen.sspnet.org/2021/10/06/semantic-enrichment-of-articles-is-different-from-citation/>.
- Choudhury, Majumdar Roy, Alia Aoun, Dina Badawy, Luis Antonio de Albuquerque Bacardit, Yassine Marjane, e Adrian Wilkinson. 2021. «Final report of the Panel of Experts on Libya established pursuant to Security Council resolution 1973 (2011)». S/2021/229. United Nations Security Council.
- Crawford, Kate. 2021. *Atlas of Ai: power, politics, and the planetary costs of artificial intelligence*. New Haven: Yale University Press.
- Dickey, Megan Rose. 2021. «Google fires top AI ethics researcher Margaret Mitchell». *TechCrunch* (blog). 19 febbraio 2021. <https://techcrunch.com/2021/02/19/google-fires-top-ai-ethics-researcher-margaret-mitchell/>.
- Durante, Massimo. 2019. *Potere computazionale: l'impatto delle ICT su diritto, società, sapere*. Filosofia delle scienze e dintorni 3. Milano: Meltemi.
- EDPB. 2021. «EDPB & EDPS call for ban on use of AI for automated recognition of human features in publicly accessible spaces, and some other uses of AI that can lead to unfair discrimination». European Data Protection Board. giugno 2021. [https://edpb.europa.eu/news/news/2021/edpb-edps-call-ban-use-ai-automated-recognition-human-features-publicly-accessible\\_en](https://edpb.europa.eu/news/news/2021/edpb-edps-call-ban-use-ai-automated-recognition-human-features-publicly-accessible_en).
- Elsevier. 2021. «Topics - ScienceDirect». Elsevier.Com. 2021. <https://www.elsevier.com/solutions/sciencedirect/topics>.
- Engelbart, Douglas. 1962. «Augmenting human intellect: a conceptual framework.» AFOSR-3233. Menlo Park, CA: Stanford Research Institute. <https://apps.dtic.mil/dtic/tr/fulltext/u2/289565.pdf>.
- Engelbart, Douglas, e William K. English. 1968. «A research center for augmenting human intellect». In *AFIPS Conference Proceedings of the 1968 Fall Joint Computer Conference*, 33:395–410. San Francisco. <https://web.stanford.edu/dept/SUL/library/extra4/sloan/mousesite/Archive/ResearchCenter1968/ResearchCenter1968.html>.
- European Parliament. Directorate General for Internal Policies of the Union. 2020. «Artificial Intelligence and Civil Liability». Publications Office. <https://data.europa.eu/doi/10.2861/220466>.
- Floridi, Luciano. 2021. «The European Legislation on AI: A Brief Analysis of Its Philosophical Approach». *Philosophy & Technology* 34 (2): 215–22. <https://doi.org/10.1007/s13347-021-00460-9>.

- Floridi, Luciano, e Josh Cowls. 2019. «A Unified Framework of Five Principles for AI in Society». *Harvard Data Science Review* 1 (1). <https://doi.org/10.1162/99608f92.8cd550d1>.
- Floridi, Luciano, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, e Francesca Rossi. 2018. «AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations». *Minds and Machines* 28 (4): 689–707. <https://doi.org/10.1007/s11023-018-9482-5>.
- Floridi, Luciano, Mariarosaria Taddeo, e Matteo Turilli. 2009. «Turing’s Imitation Game: Still an Impossible Challenge for All Machines and Some Judges—An Evaluation of the 2008 Loebner Contest». *Minds and Machines* 19 (1): 145–50. <https://doi.org/10.1007/s11023-008-9130-6>.
- Foucault, Michel. 1969. «Qu’est-ce qu’un auteur?» *Bulletin de la Société Française de Philosophie* 63 (3): 73–104. <http://1libertaire.free.fr/MFoucault319.html>.
- Genette, Gérard. 2002. *Seuils*. Points Essais 474. Paris: Seuil.
- Ginzburg, Carlo. 1979. «Spie. Radici di un paradigma indiziario». In *Crisi della ragione. Nuovi modelli nel rapporto tra sapere e attività umane*, a cura di Aldo Gargani, 1–30. Torino: Einaudi.
- Goodwin, Ross. 2016a. «Narrated reality». Master thesis, New York: New York University.
- . 2016b. «Adventures in Narrated Reality». *Medium* (blog). 9 giugno 2016. <https://medium.com/artists-and-machine-intelligence/adventures-in-narrated-reality-6516ff395ba3>.
- Goodwin, Ross, Kenric McDowell, e Hélène Planquelle. 2018. *1 the road*. Art + machines 1. Paris: Jean Boîte éditions.
- GPT-3. 2020. «A Robot Wrote This Entire Article. Are You Scared yet, Human?» *The Guardian*, 8 settembre 2020. <http://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3>.
- Gruppo di Esperti MISE sull’intelligenza artificiale. 2020. «Proposte per una Strategia italiana per l’intelligenza artificiale». Roma: MISE. [https://www.mise.gov.it/images/stories/documenti/Proposte\\_per\\_una\\_Strategia\\_italiana\\_AI.pdf](https://www.mise.gov.it/images/stories/documenti/Proposte_per_una_Strategia_italiana_AI.pdf).
- Gubrium, Jaber F., e James A. Holstein. 1995. «Individual agency, the ordinary, and postmodern life». *Sociological Quarterly* 36 (3): 555–70. <https://doi.org/10.1111/j.1533-8525.1995.tb00453.x>.
- Guerrini, Mauro. 2020. *Dalla catalogazione alla metadattazione: tracce di un percorso*. Collana Percorsi AIB 5. Roma: Associazione italiana biblioteche.
- Gunning, David, e David Aha. 2019. «DARPA’s Explainable Artificial Intelligence (XAI) Program». *AI Magazine* 40 (2): 44–58. <https://doi.org/10.1609/aimag.v40i2.2850>.
- Hagan, Martin T., Howard B. Demuth, e Mark H. Beale. 1996. *Neural network design*. 1st ed. Boston: PWS Pub.
- Hao, Karen. 2020. «“I Started Crying”: Inside Timnit Gebru’s Last Days at Google». *MIT Technology Review*, 16 dicembre 2020. <https://www.technologyreview.com/2020/12/16/1014634/google-ai-ethics-lead-timnit-gebru-tells-story/>.

———. 2021. «Stop Talking about AI Ethics. It's Time to Talk about Power.» *MIT Technology Review*, 23 aprile 2021. <https://www.technologyreview.com/2021/04/23/1023549/kate-crawford-atlas-of-ai-review/>.

Haski-Leventhal, Debbie. 2020. «The Way Forward in Higher Education». In *The Purpose-Driven University*, 133–45. Bingley: Emerald Publishing Limited. <https://doi.org/10.1108/978-1-83867-283-620201011>.

Hao, Karen, e Heidi Swart. 2022. «South Africa's private surveillance machine is fueling a digital apartheid». *MIT Technology Review*, 19 aprile 2022. <https://www.technologyreview.com/2022/04/19/1049996/south-africa-ai-surveillance-digital-apartheid/>.

ICCU. 2009. *Regole italiane di catalogazione: REICAT*. Roma: ICCU.

IFLA. 2020. «IFLA Statement on Libraries and Artificial Intelligence». <https://www.ifla.org/publications/node/93397>.

Kaindl, Markus, e Stephanie Preuss. 2020. «How can AI support the research community in times of crisis?» *The Source* (blog). 1 aprile 2020. <https://www.springernature.com/gp/researchers/the-source/blog/blogposts-life-in-research/how-can-ai-support-the-research-community-in-times-of-crisis/17852232>.

Kallenborn, Zachary. 2021. «Was a Flying Killer Robot Used in Libya? Quite Possibly». *Bulletin of the Atomic Scientists* (blog). 20 maggio 2021. <https://thebulletin.org/2021/05/was-a-flying-killer-robot-used-in-libya-quite-possibly/>.

Lana, Maurizio. 2020. *Introduzione all'information literacy*. Editrice Bibliografica. <https://www.editricebibliografica.it/scheda-libro/maurizio-lana/introduzione-allinformation-literacy-9788893570862-579327.html>.

Lana, Maurizio. 2022. «L'agency dei sistemi di Intelligenza Artificiale. Un punto di vista bibliografico». *Digitcult. Scientific Journal on Digital Cultures*. in uscita.

Larson, Jeff, Mattu Surya, Lauren Kirchner, e Julia Angwin. 2016. «How We Analyzed the COMPAS Recidivism Algorithm». *ProPublica*, Maggio 2016. [https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm?token=pc1OkzviktiLco0hJY5BDRPpI44H\\_bKV](https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm?token=pc1OkzviktiLco0hJY5BDRPpI44H_bKV).

Licklider, J. C. R. 1960. «Man-Computer Symbiosis». *IRE Transactions on Human Factors in Electronics* HFE-1 (1): 4–11. <https://doi.org/10.1109/THFE2.1960.4503259>.

McCarthy, John, Marvin L. Minsky, Nathaniel Rochester, e Claude Elwood Shannon. 1955. «A proposal for the Dartmouth summer research project on Artificial Intelligence». <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>.

———. 2006. «A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence». *AI Magazine* 27 (4): 3. <https://aaai.org/ojs/index.php/aimagazine/issue/view/165>.

Mehan, Hugh, e Houston Wood. 1975. *The Reality of Ethnomethodology*. New York: John Wiley & Sons, Ltd.

Mittelstadt, Brent. 2019. «Principles alone cannot guarantee ethical AI». *Nature Machine Intelligence* 1 (11): 501–7. <https://doi.org/10.1038/s42256-019-0114-4>.

National Commission on New Technological Uses of Copyrighted Works, «Final Report of the National Commission on New Technological Uses of Copyrighted Works, Chapter 3 – Computers and Copyright. New Works». 1975. <http://digital-law-online.info/CONTU/contu17.html>.

Office of Technology Assessment. 1986. «Intellectual property rights in an age of electronics and information». OTA-CIT-302. Washington, DC: U.S. Congress. [https://digital.library.unt.edu/ark:/67531/metadc39611/m2/1/high\\_res\\_d/8610.pdf](https://digital.library.unt.edu/ark:/67531/metadc39611/m2/1/high_res_d/8610.pdf).

«Projects and Cooperations - Applied Computational Linguistics Lab Goethe University Frankfurt, Germany». s.d. Consultato 1 luglio 2021. <http://www.acoli.informatik.uni-frankfurt.de/projects.html>.

*Proposta di Regolamento del Parlamento Europeo e del Consiglio che stabilisce regole armonizzate sull'Intelligenza Artificiale (legge sull'Intelligenza Artificiale) e modifica alcuni atti legislativi dell'Unione*. 2021. <https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=CELEX:52021PC0206&from=EN>.

Raji, Inioluwa Deborah, Timnit Gebru, Margaret Mitchell, Joy Buolamwini, Joonseok Lee, e Emily Denton. 2020. «Saving Face: Investigating the Ethical Concerns of Facial Recognition Auditing». In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 145–51. New York NY USA: ACM. <https://doi.org/10.1145/3375627.3375820>.

Samuelson, Pamela. 1985. «Allocating ownership rights in computer-generated works». *University of Pittsburgh Law Review* 47: 1185–1228.

———. 2020. «AI Authorship?» *Communications of the ACM* 63 (7): 20–22. <https://doi.org/10.1145/3401718>.

Schönenberger, Henning, Christian Chiarcos, e Niko Schenk. 2019. «Preface». In *Lithium-Ion Batteries. A Machine-Generated Summary of Current Research*. New York, NY: Springer. <https://doi.org/10.1007/978-3-030-16800-1>.

Shapiro, Susan P. 2005. «Agency theory». *Annu. Rev. Sociol.* 31: 263–84. <https://doi.org/10.1146/annurev.soc.31.041304.122159>.

Springer Nature. 2019. «Springer Nature Publishes Its First Machine-Generated Book». *Www.Springer.Com* (blog). 2 aprile 2019. <https://www.springer.com/gp/about-springer/media/press-releases/corporate/springer-nature-machine-generated-book/16590126>.

Strubell, Emma, Ananya Ganesh, e Andrew McCallum. 2019. «Energy and Policy Considerations for Deep Learning in NLP». In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3645–50. Florence, Italy: Association for Computational Linguistics. <https://doi.org/10.18653/v1/P19-1355>.

Svenonius, Elaine. 2000. *The intellectual foundation of information organization*. Digital libraries and electronic publishing. Cambridge, Mass: MIT Press.

Taddeo, Mariarosaria, e Luciano Floridi. 2018. «How AI Can Be a Force for Good». *Science* 361 (6404): 751–52. <https://doi.org/10.1126/science.aat5991>.

Turing, A. M. 1950. «Computing Machinery and Intelligence». *Mind* LIX (236): 433–60. <https://doi.org/10.1093/mind/LIX.236.433>.

UK Parliament, *Copyright, Designs and Patents Act*. 1988. <https://www.legislation.gov.uk/ukpga/1988/48/enacted/data.pdf>.

United Nations, e David Kaye. 2018. «Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression». A/73/348. New York: United Nations.

U.S. Copyright Office. 2021. *Compendium of U.S. Copyright Office practices*. III. Washington, DC.

Vetere, Guido. 2021. «L'intelligenza artificiale “stona” ancora: i limiti della generazione automatica di musica». *Agenda Digitale* (blog). 20 aprile 2021. <https://www.agendadigitale.eu/cultura-digitale/lintelligenza-artificiale-stona-ancora-i-limiti-della-generazione-automatica-di-musica/>.

Vuong, Quan-Hoang. 2020. «The Limitations of Retraction Notices and the Heroic Acts of Authors Who Correct the Scholarly Record: An Analysis of Retractions of Papers Published from 1975 to 2019». *Learned Publishing* 33 (2): 119–30. <https://doi.org/10.1002/leap.1282>.

Wang, Sun-Chong. 2003. «Artificial Neural Network». In *Interdisciplinary Computing in Java Programming*, di Sun-Chong Wang, 81–100. Boston, MA: Springer US. [https://doi.org/10.1007/978-1-4615-0377-4\\_5](https://doi.org/10.1007/978-1-4615-0377-4_5).

Yang, Guang-Zhong, Jim Bellingham, Pierre E. Dupont, Peer Fischer, Luciano Floridi, Robert Full, Neil Jacobstein, et al. 2018. «The Grand Challenges of Science Robotics». *Science Robotics* 3 (14): eaar7650. <https://doi.org/10.1126/scirobotics.aar7650>.

## The role of the Wikidata librarian in a renewed Bibliographical Universe: “next generation metadata”, next generation librarians

Alessandra Boccone<sup>(a)</sup>

a) Central library - University of Salerno, <https://orcid.org/0000-0002-7139-9640>

**Contact:** Alessandra Boccone, [aboccone@unisa.it](mailto:aboccone@unisa.it)

**Received:** 11 September 2021; **Accepted:** 18 October 2021; **First Published:** 15 May 2022

### ABSTRACT

Starting from a brief analysis of the OCLC Research report “Next generation metadata”, the contribution proposes a reflection on the librarian metadata and the new centrality of its role in the LIS; among the various declinations that can assume such professional figure, one of the most interesting is the Wikidata librarian, of which a first definition is outlined. Following some international experiences of the use of Wikidata as a tool for new working methods, already consolidated in many library institutions, and as a particularly suitable environment for experimentation, thanks to its features of open, free, collaborative, easy to understand and use system. While information professionals and a large part of the academic and scientific environment have understood the potential of Wikidata, some weaknesses of this instrument have been highlighted, which must be corrected and recalibrated in the name of universally accessible and reusable knowledge, also taking into account the requirements contained in the Semantic Web Manifesto by the Cataloging and Indexing Study Group of the Italian Association of Libraries.

### KEYWORDS

Wikidata librarian; Metadata librarian; Semantic web.

## Il ruolo del Wikidata librarian in un rinnovato universo bibliografico: “next generation metadata”, next generation librarians\*

### ABSTRACT

Partendo da una breve analisi dell'OCLC Research report “Next generation metadata”, il contributo propone una riflessione sul metadata librarian e sulla nuova centralità del suo ruolo in ambito LIS; fra le varie declinazioni che può assumere tale figura professionale, una delle più interessanti risulta essere il Wikidata librarian, di cui si delinea una prima definizione. Seguono alcune esperienze internazionali dell'uso di Wikidata come strumento alla base di nuove metodologie di lavoro, già consolidate presso molte istituzioni bibliotecarie, e come ambiente particolarmente adatto alla sperimentazione, grazie alle sue caratteristiche di sistema aperto, libero, collaborativo, di facile comprensione e utilizzo. Se da un lato i professionisti dell'informazione e una buona parte dell'ambiente accademico e scientifico hanno compreso le potenzialità di Wikidata, dall'altra sono stati evidenziati alcuni punti di debolezza di tale strumento, che vanno corretti e ricalibrati in nome di una conoscenza universalmente accessibile e riutilizzabile, tenendo conto anche delle prescrizioni contenute nel Manifesto per il Web Semantico a cura del Gruppo di studio Catalogazione ed indicizzazione dell'Associazione Italiana Biblioteche.

### PAROLE CHIAVE

Wikidata; Metadati; Web semantico.

\* Si ringrazia vivamente il prof. Carlo Bianchini per l'alta professionalità e la rara umanità: grazie ai suoi spunti e alle sue osservazioni il contributo ha assunto la sua forma definitiva.

*Qualcosa di nuovo si muove a livello internazionale nel campo della catalogazione nella ricerca di tecniche efficaci per una più spedita cooperazione. Non restiamone fuori, non isoliamoci.* (Diego Maltese)

*Gli archivi e le biblioteche non sono solo semplici contenitori passivi di informazioni utili a soggetti terzi cui attingere alla bisogna. Essi possono e devono giocare un ruolo diverso e più incisivo, andando oltre i loro stessi confini filologici, in direzione di una penetrazione sociale capace farne strumenti politici, in quanto supporti di importanza assolutamente non secondaria per la difesa di sistemi di valori e processi decisionali.* (Federico Valacchi)

## 1. Introduzione

Le numerose innovazioni in ambito biblioteconomico e nelle scienze dell'informazione stanno portando a una rapida evoluzione nelle pratiche di creazione e gestione dei metadati: la diffusione dei nuovi modelli concettuali e l'utilizzo dei linked open data potrebbero costituire una valida soluzione alle problematiche di molte biblioteche, che nelle attuali condizioni riescono sempre con maggiori difficoltà a sostenere il ritmo repentino di un cambiamento irreversibile e necessario. Una spinta propulsiva verso tali evoluzioni è costituita dall'idea di fornire a un'utenza sempre più ampia e variegata una gamma di "value-added services" (Smith-Yoshimura 2020, vi), nonché dalla speranza "that outside parties will be able to make better use of library resource descriptions, and that the data will be richer because more parties share in its creation" (Ibid.). Partendo dalla recente pubblicazione dell'OCLC *Library Partners Metadata Managers Focus Group* sui metadati di prossima generazione, il contributo propone una riflessione sul metadata librarian e sulla nuova centralità del suo ruolo in ambito LIS; fra le varie declinazioni che può assumere tale figura professionale, una delle più interessanti risulta essere il Wikidata librarian, di cui si delinea una prima definizione. Inoltre, l'articolo intende analizzare Wikidata in quanto strumento utile alla sperimentazione di nuove metodologie di lavoro, aperte e collaborative, sia per i *metadata librarians* che per tutti i bibliotecari desiderosi di ampliare le proprie competenze e di collocarsi in un contesto professionale e culturale maggiormente ampio e variegato rispetto all'istituzione di appartenenza. A livello globale, i bibliotecari hanno inserito le attività su Wikidata nella pratica quotidiana, con risultati incoraggianti e spesso ben al di là delle aspettative; sono emersi, d'altro canto, anche limiti e problematiche relative all'utilizzo dello strumento, che impongono correttivi tecnici e metodologici. Una delle sfide dei futuri professionisti dell'informazione si delinea, dunque, nel mantenere il giusto bilanciamento fra una tradizione consolidata e un'innovazione indispensabile, nell'ottica di una nuova centralità culturale, sociale ed economica delle biblioteche e dei bibliotecari.



## 2. Verso una nuova generazione di metadati

Alla base delle riflessioni sul futuro dei metadati bibliografici si è diffusa sempre più la consapevolezza che, indipendentemente dal loro formato, essi stiano acquisendo una rinnovata centralità sia in ambito scientifico che generalista: “Il web ha modificato i comportamenti del lettore e, quindi, ha condizionato le funzioni utente, ovvero le modalità di trovare, identificare, selezionare e ottenere le informazioni e le risorse, nonché le modalità di esplorare l’universo bibliografico. Il web è il luogo in cui avviene la maggior parte delle ricerche e molti lettori hanno percepito la distanza tra l’universo bibliografico censito dai cataloghi e l’enorme quantità di informazioni disponibili in rete e accessibili tramite i motori di ricerca; perciò iniziano – e spesso concludono – altrove le loro ricerche” (Guerrini 2020, 25).

Tali riflessioni portano con sé la necessità di un ripensamento e una riprogettazione dei dati bibliografici, che, per riconquistare un rilievo riconosciuto trasversalmente, devono adattarsi sempre più alle nuove infrastrutture informative, in particolar modo a quelle del Web semantico, nell’ottica della British Library’s Foundations for the Future: “Our vision is that by 2023 the Library’s collection metadata assets will be unified on a single, sustainable, standards-based infrastructure offering improved options for access, collaboration and open reuse”<sup>1</sup>, mentre in ambito LIS si profila l’esigenza di “processi di riconcettualizzazione complessi e delicati che mettono in gioco la stessa affidabilità epistemologica delle diverse discipline” (Valacchi 2021, 465).

In tale contesto, fra le realtà più attive troviamo l’*Online Computer Library Center*, in particolar modo la divisione *Research*<sup>2</sup>, che lavora con la comunità degli stakeholders per identificare in modo collaborativo problemi e opportunità, per creare prototipi e testare soluzioni, condividendo i risultati attraverso pubblicazioni, presentazioni e scambi professionali.

Particolarmente significativa è stata la pubblicazione del report *Transitioning to the Next Generation of Metadata* a cura di Karen Smith-Yoshimura, che sintetizza sei anni<sup>3</sup> di incontri e discussioni dell’*OCLC Research Library Partners Metadata Manager Focus Group*<sup>4</sup> sul tema dei metadati di prossima generazione. Il rapporto, denso e particolareggiato, è organizzato in quattro sezioni, ciascuna delle quali rappresenta una tendenza emersa nelle discussioni del focus group: la prima riguarda la transizione verso i linked data e gli identificatori, il cui focus è l’estensione dell’uso di identificatori persistenti come parte del passaggio dal “controllo di autorità” alla “gestione di identità”. Segue la sezione della descrizione delle collezioni “inside-out” e “facilitated”, come sfida nella creazione e gestione di metadati per risorse uniche curate dalle istituzioni e condivise attraverso i consorzi. L’evoluzione dei “metadata as a service” è il fulcro della terza parte, per un maggiore coinvolgimento nella creazione di metadati che vadano oltre l’uso all’interno del tradi-

---

<sup>1</sup> Per un approfondimento si rimanda a: British Library. 2019. “Foundations for the Future: The British Library’s Collection Metadata Strategy 2019-2023”. London: British Library, <<https://www.bl.uk/bibliographic/pdfs/british-library-collection-metadata-strategy-2019-2023.pdf>>.

<sup>2</sup> OCLC Research, <<https://www.oclc.org/research/home.html>>.

<sup>3</sup> Dal 2015 al 2020.

<sup>4</sup> Il Library Partners Metadata Managers Focus Group, istituito nel 1993, è uno dei gruppi più longevi all’interno della OCLC Research Library Partnership (RLP), una rete transnazionale di biblioteche di ricerca; il Focus Group fornisce un forum per i responsabili della creazione e della gestione dei metadati nelle varie istituzioni coinvolte, al fine di condividere informazioni su argomenti di interesse comune e di identificare problematiche relative alla gestione dei metadati.

zionale catalogo delle biblioteche; il rapporto si conclude con l'individuazione delle future competenze richieste al personale delle biblioteche, perché il nuovo panorama informativo spinge alla costruzione di nuove abilità, sia per chi inizia ad avvicinarsi alla professione, sia per i bibliotecari di comprovata esperienza.

Esaminando tutte le questioni relative a questi argomenti, le principali conclusioni delle attività del Focus Group si possono sintetizzare in quattro punti:

- La creazione dei metadati specifici per formato basati su stringhe di testo nei record bibliografici, comprensibili e utilizzabili solo dai sistemi bibliotecari, è prossima all'obsolescenza, sia tecnicamente che concettualmente.<sup>5</sup>
- La gestione tradizionale dei metadati è incentrata sui record bibliografici, dunque risulta essere una produzione costosa e con gravi limiti su vari fronti.
- I metadati non vengono più prodotti solamente dai bibliotecari, ma anche da editori, autori e altri soggetti del ciclo dell'informazione.
- La prossima generazione di metadati sarà più focalizzata sulle gestione delle entità piuttosto che sulla descrizione classica delle risorse bibliografiche.
- Saranno necessarie nuove competenze per i bibliotecari, in relazione a un ambiente tecnologico in rapida evoluzione.

In *Transitioning to the Next Generation of Metadata*, l'OCLC Research Library Partners Metadata Managers Focus Group indica più volte Wikidata come un valido strumento per le biblioteche e gli altri istituti culturali che vogliono attuare la transizione dei loro dati verso i metadati di nuova generazione, innovando completamente gli approcci e i flussi di lavoro legati alla descrizione tradizionale delle risorse bibliografiche; l'ottica è quella di poter fornire, fra l'altro, "a richer user experience and increased discoverability with more contextual relationships than is possible with our current systems" (Smith-Yoshimura 2020, 4), mettendo l'accento sull'uso degli identificatori persistenti che connettono i metadati a dati correlati, vocabolari, termini in lingue diverse, discipline e ontologie differenti, inclusi quelli non bibliografici.

### 3. Il Wikidata librarian

Le associazioni professionali internazionali come IFLA, ACRL, CILIP incoraggiano da tempo il ruolo attivo del bibliotecario, soprattutto quello accademico, nel mondo dei dati, ma la definizione della figura, delle competenze e della formazione del metadata librarian è ancora tutta da costruire. "Sono comparse nuove locuzioni quali metadatazione, data manager, data curation librarian, data reference librarian, data collection building librarian a significare che, con la nuova prospettiva della metadatazione, si identificano e si collegano entità [...] mediante relazioni" (Guerrini 2020, 32): si può affermare dunque che la questione della gestione e organizzazione dei dati e, nello specifico, dei metadati, è passata dal dominio prettamente tecnico di tipo informatico a quello della documentazione, descrizione, condivisione e trasferimento della conoscenza. Tali

---

<sup>5</sup> "Traditional methods of metadata generation, management and dissemination are not scalable or appropriate to an era of rapid digital change, rising audience expectations and diminishing resources", British Library. 2019. "Foundations for the Future: The British Library's Collection Metadata Strategy 2019-2023". London: British Library. <<https://www.bl.uk/bibliographic/pdfs/british-library-collection-metadata-strategy-2019-2023.pdf>>

attività appartengono abitualmente al bagaglio professionale del bibliotecario, in riferimento alla catalogazione, all'organizzazione dell'informazione, alla gestione delle collezioni, all'information literacy, alla conservazione e preservazione per l'accesso a lungo termine (Morriello 2020, 6).

Intanto, secondo Guerrini, “la catalogazione, da almeno due decenni, sta vivendo una seconda rivoluzione culturale e tecnologica, dopo quella degli anni Settanta del Novecento caratterizzata dal passaggio dal paradigma del cartaceo al paradigma dell'automazione [...]. Il nuovo contesto digitale richiede un ulteriore mutamento di paradigma, se non una metamorfosi, un «profondo mutamento nel modo di pensare», cioè una disposizione intellettuale aperta, sempre consapevole e mai appiattita sulle nuove convenzioni che, come tali, sono temporanee” (Guerrini 2020, 25). Nelle biblioteche, tutto ciò si riflette nei nuovi flussi di lavoro relativi alle risorse e va a ridefinire la figura del catalogatore, il cui profilo più centrato rispetto ai nuovi standard e alle tecnologie contemporanee sembra essere sempre più quello del metadata librarian.

In effetti in ambito anglo-americano, alla fine degli anni '90, la figura del metadata librarian coincide perlopiù con quella del bibliotecario catalogatore, ma si va rapidamente ampliando grazie all'impatto in ambito LIS dell'evoluzione delle tecnologie informatiche, alle nuove tipologie di risorse che entrano a far parte delle collezioni delle biblioteche e all'affermarsi, a livello internazionale, di nuovi progetti collaborativi di produzione e trattamento dei metadati (Han e Hswe 2010, 129-141). Nel 2005 Beacom delinea tre precise aree di responsabilità per il metadata librarian: sviluppo di standard e documentazione, produzione di metadati e collaborazione su strumenti digitali (Beacom 2005). Oggi questa definizione suona sicuramente incompleta e in parte riduttiva: attualmente infatti i bibliotecari specialisti dei metadati assumono un ruolo fondamentale anche nella consulenza per progetti di ricerca e iniziative di cooperazione fra enti culturali oppure, negli insegnamenti di ambito umanistico, vengono coinvolti per indicazioni sull'uso degli standard e dei vocabolari controllati.<sup>6</sup>

Come sottolineato anche nella panel discussion *New horizons: emerging metadata standards and practices in the 21st century*, a cura delle tre Standing Committees dell'IFLA (Olive 2021, 36), le future generazioni di bibliotecari dovranno dunque essere preparate all'apprendimento costante, non solo praticando la programmazione informatica e utilizzando l'intelligenza artificiale, ma anche possedendo un adeguato background teorico nei modelli concettuali bibliografici. Risulta fondamentale anche mantenere uno sguardo attento e costante alle tendenze e ai cambiamenti culturali a livello globale, utilizzando un approccio incentrato sull'utente e un'attenzione a garantire la qualità e l'utilità dei metadati: “Metadata creators should no longer think in terms of a record that can be declared as “finished” but rather think in terms of the lifecycle of data” (Ibid.).

In Italia, nonostante molti bibliotecari siano impegnati in prima linea nel vasto ecosistema dei dati, il ruolo del metadata librarian è al momento assente dalle definizioni professionali, ma è una lacuna che va colmata quanto prima perché, con l'aumento della necessità di organizzare quantità

---

<sup>6</sup> «More visibility of this metadata consultant role appears in recent library job postings. In one Metadata Librarian job posting at Cornell, 121 one duty cited was 20% for “metadata outreach and consultation”: “Maintains strong working relationships and communicates regularly with staff across Cornell, fostering collaborative efforts between Metadata Services and the greater Cornell community.” Georgia Tech is recruiting a metadata librarian who 26 Transitioning to the Next Generation of Metadata will “serve as a metadata consultant to larger library projects/initiatives. Work closely with other Library departments, Emory University Libraries, GALILEO, University System of Georgia Libraries, and other partners involved in joint projects.”», Karen Smith-Yoshimura, *Transitioning to the Next Generation of Metadata*, p. 25-26.

di dati e metadati sempre maggiori in modo da poterli trasformare efficacemente in informazione e conoscenza, sono emerse esigenze nuove che non possono più essere affrontate solo dal punto di vista tecnico-informatico e, di conseguenza, viene percepita la necessità di contestualizzare, standardizzare e valorizzare questa tipologia di attività nell'ambito LIS.

Nel frattempo, la nascita di Wikidata, la base di dati strutturati aperta e collaborativa della Wikimedia Foundation, ha semplificato la creazione e pubblicazione di dati aperti e collegati, che tradizionalmente ha richiesto competenze tecniche raffinate per la trasformazione dei dati relazionali e la creazione dei vocabolari e delle ontologie. Già nel documento di discussione presentato al congresso IFLA del 2016, alcuni studiosi hanno evidenziato come la capacità di Wikidata di disegnare dati aperti collegati, integrare authorities di dati e diverse ontologie e tassonomie abbia un enorme potenziale di utilizzo per i ricercatori di tutto il mondo (Bartholmei et al. 2016, 5).

In virtù di questo, fra le numerose declinazioni che può avere il metadata librarian, si può definire, senza pretesa di esaustività, la figura del *Wikidata librarian* come un bibliotecario consapevole delle possibilità e dell'uso dei progetti e delle piattaforme Wikimedia, formato nel diritto d'autore, sensibile alle tematiche dell'open access e dell'open science, che è in grado di:

- ideare progetti semplici e complessi su Wikidata;
- partecipare attivamente alle iniziative internazionali relative ai dati dei progetti Wikimedia;
- mappare e proporre nuove proprietà per una migliore descrizione delle risorse bibliografiche in Wikidata;
- inserire correttamente metadati bibliografici e dati di qualità in Wikidata, sia in maniera manuale sia attraverso l'utilizzo degli appositi strumenti;
- fare da tramite fra la comunità dei bibliotecari e quella dei progetti Wikimedia, valorizzando i punti di incontro e appianando eventuali diffidenze reciproche;
- sensibilizzare enti e istituzioni a rilasciare con licenza libera ed eventualmente caricare sulla piattaforma i propri database, creando collegamenti bidirezionali sincronizzati fra essi e Wikidata;
- formare colleghi e collaboratori sulle potenzialità e l'utilizzo di Wikidata;
- conoscere e utilizzare i principali tool per la modifica, la creazione, l'interrogazione e la visualizzazione dei dati.

Le competenze e le abilità da sviluppare e ampliare in questa direzione, dunque, sono molteplici, ma il corretto utilizzo di uno strumento potente e di facile utilizzo come Wikidata, abbinato alle pratiche e alle conoscenze tipiche del bibliotecario, permettono di superare barriere e costruire ponti di conoscenza come mai è stato possibile fare prima. A tale scopo, si ritiene molto utile prendere in esame e fare riferimento ad alcune iniziative sviluppate a livello locale e/o internazionale dai Wikidata librarians.

#### 4. Wikidata librarians: alcune esperienze

Sempre più numerosi sono i progetti di biblioteche e bibliotecari su Wikidata: è interessante guardare a queste esperienze sia per avere la possibilità di farne parte in maniera attiva e sia perché esse possono fungere da esempio per dare vita a nuove collaborazioni e idee. In seguito, si presenta

una breve rassegna di alcune iniziative particolarmente significative, sia per l'impatto che hanno avuto sulla comunità bibliotecaria che per il tratto innovativo che le caratterizza. Anche grazie a esse è stato possibile dapprima sperimentare in maniera pionieristica e, successivamente, consolidare nuovi approcci e pratiche, che hanno contribuito a inserire Wikidata all'interno dell'universo bibliografico e viceversa.

#### 4.1 LD4P, WikiProject LD4 Wikidata Affinity Group e GWMAB

Linked Data for Production (LD4P) è un progetto di collaborazione tra quattro istituzioni (le Università di Stanford, Cornell, di Harvard, la School of Library and Information Science della Università dell'Iowa) con la Library of Congress e il Program for Cooperative Cataloging (PCC) (Schreur 2018). Tale iniziativa ha il fine di organizzare e testare un flusso di lavoro completo ed efficace in merito alla creazione, alla condivisione e al riutilizzo dei library linked data, con un focus specifico sulla transizione dai formati di dati MARC a quelli basati sull'utilizzo dei Linked Open Data<sup>7</sup>. Questo gruppo di lavoro internazionale, giunto alla sua terza fase nel 2020 con l'apertura a Share-VDE, lavora abitualmente su Wikidata, tanto da aver creato il sottoprogetto WikiProject LD4 Wikidata Affinity Group, il cui scopo è stimolare una "community discussion of Wikidata related topics with the goal of understanding how the library can contribute to and leverage Wikidata as a platform for publishing, linking, and enriching library linked data"<sup>8</sup>. Il gruppo si riunisce ogni due settimane per discutere vari aspetti di Wikidata: ad ogni incontro, i facilitatori o un ospite presentano risorse utili e distribuiscono materiali relativi all'argomento di discussione. Sono organizzate anche delle utili working hours di editing collaborativo e confronto diretto sulle problematiche tecniche riscontrate in Wikidata, in cui i partecipanti possono anche proporre temi, progetti, soluzioni e portare la loro esperienza.

Ispirato all'esperienza di LD4P, il Gruppo Wikidata per Musei, Archivi e Biblioteche (GWMAB) è un interest group che "mira a indagare come le biblioteche, i musei e gli archivi italiani possono contribuire e migliorare l'integrazione dei loro metadati in Wikidata e fornire uno spazio informale per condividere informazioni, esperienze e conoscenze su Wikidata"<sup>9</sup>.

L'impegno del GWMAB, sostenuto dall'Università di Pavia, si esplicita nel coinvolgimento e nella formazione di ricercatori e operatori del settore culturale (in particolar modo di musei, archivi, biblioteche) nel progetto Wikidata, grazie a momenti di formazione ad hoc, alla traduzione e disseminazione di materiali in lingua italiana su strumenti, funzioni e progetti relativi, e all'organizzazione di incontri formali e informali che promuovono l'interazione diretta tra i tutti i portatori di interesse per l'argomento, in particolar modo fra le istituzioni culturali e i contributori esperti di Wikidata. Queste molteplici attività favoriscono "il confronto, la diffusione e lo scambio di punti di vista e best practises, di esperienze teoriche e pratiche e di progetti di ricerca sviluppati dai par-

---

<sup>7</sup> LD4P2 Project Background and Goals, <<https://wiki.lyrasis.org/display/LD4P2/LD4P2+Project+Background+and+Goals>>.

<sup>8</sup> WikiProject LD4 Wikidata Affinity Group <[https://www.wikidata.org/wiki/Wikidata:WikiProject\\_LD4\\_Wikidata\\_Affinity\\_Group](https://www.wikidata.org/wiki/Wikidata:WikiProject_LD4_Wikidata_Affinity_Group)> .

<sup>9</sup> GWMAB <<http://mbc.unipv.it/iniziative-in-corso/658-gruppo-wikidata-per-musei-archivi-e-biblioteche>> .

tecipanti, a titolo personale o istituzionale”.<sup>10</sup> Recentemente il GWMAB, dopo avere collaborato con l’ICCD per coordinare l’allineamento tra i dati dell’istituto centrale e Wikidata<sup>11</sup>, ha avviato una stretta collaborazione con l’ICCU: come ha affermato Simonetta Buttò durante la conferenza internazionale *Bibliographic control in the Digital Ecosystem*, nel Servizio Bibliotecario Nazionale sono in atto delle forti innovazioni “anche grazie alla sollecitazione della comunità Wiki, con la quale collaboriamo su vari temi grazie a una convenzione firmata con Wikimedia Italia nel marzo del 2015, e in particolare con Wikidata, in quanto punto di raccordo di moltissimi identificativi di nomi personali e collettivi provenienti da molteplici produttori di dati”.<sup>12</sup>

#### 4.2 LC BIBFRAME Wikidata Project

I partecipanti al progetto pilota LC-BIBFRAME-Wikidata<sup>13</sup> stanno esplorando una nuova strada per ottimizzare il flusso di lavoro legato alla descrizione delle risorse della Library of Congress in BIBFRAME, soprattutto per ciò che attiene le attività relative al controllo di autorità, sfruttando i principi di Identity Management e utilizzando Wikidata come hub per la descrizione dell’agente. I catalogatori della Library of Congress che partecipano al progetto di conversione dei dati sono costretti a creare una duplice descrizione di ogni risorsa, sia in MARC sia in BIBFRAME, perché quest’ultima non costituisce un “database di record” per la Library of Congress. Questo flusso di lavoro duplicato diventa ancora più complesso quando si tratta dei file di autorità. Le descrizioni catalografiche della Library of Congress devono utilizzare, infatti, dei punti di accesso standardizzati per gli agenti; a tale scopo viene utilizzato il LC/NACO Authority File (basato sul MARC), a cui i bibliotecari accedono attraverso il LC Linked Data Service, su [id.loc.gov](http://id.loc.gov). L’utilizzo del LC/NACO risulta relativamente lineare quando il record di autorità è già presente, ma in caso contrario il flusso di lavoro diventa molto più complesso<sup>14</sup>. Per ovviare a questo dispendio inutile di tempo ed energie, entra in gioco Wikidata: nel caso in cui il catalogatore non trovi un record LC/NACO Authority File per l’agente, lo cercherà in Wikidata, e se presente, la sua URI sarà utilizzata nella descrizione BIBFRAME; in caso contrario il bibliotecario creerà un nuovo elemento Wikidata, per usarlo immediatamente nella descrizione in corso. Ciò consente di completare la catalogazione della risorsa senza attendere la precedente creazione in MARC e il caricamento in LC/NACO Authority File dell’agente. In questo caso, il relativo record MARC, che verrà comunque creato, conterrà solo i requisiti minimi Core RDA e, nel campo 024, verrà aggiunto un link all’elemento Wikidata, che diventa a quel punto, il “centro” per la descrizione dell’agente.

<sup>10</sup> Ibid.

<sup>11</sup> Grazie a questa collaborazione sono state create tre nuove proprietà in Wikidata: Identificativo ICCD di un agente, Identificativo ICCD di un sito, Identificativo ICCD di un bene culturale.

<sup>12</sup> Si rimanda alla registrazione della prima sessione della conferenza *Bibliographic control in the Digital Ecosystem*, <<https://www.youtube.com/embed/ybUDrILt0kI?start=2161&end=2800>> .

<sup>13</sup> Wikidata:WikiProject PCC Wikidata Pilot/LC BIBFRAME Wikidata Project, <[https://www.wikidata.org/wiki/Wikidata:WikiProject\\_PCC\\_Wikidata\\_Pilot/LC-BIBFRAME-Wikidata-Projec](https://www.wikidata.org/wiki/Wikidata:WikiProject_PCC_Wikidata_Pilot/LC-BIBFRAME-Wikidata-Projec)>

<sup>14</sup> In questo caso il catalogatore deve lasciare in sospeso la descrizione BIBFRAME, per andare a creare il record di autorità MARC, e quindi attendere che sia inserito in [id.loc.gov](http://id.loc.gov) dopo varie ore. Il giorno successivo, il catalogatore deve riaprire la scheda lasciata in sospeso, cercare il nuovo record di autorità e catturarne l’URI per completare la descrizione della risorsa iniziata precedentemente: tutto ciò crea un significativo rallentamento delle attività.

Fra i documenti prodotti nel contesto di questo progetto pilota, spicca la mappatura dei campi e sottocampi del MARC21 con le relative proprietà di Wikidata, a cura di PCC<sup>15</sup>. Tale documento, ancora in fieri, è un utilissimo riferimento e un significativo punto di partenza, non solo per i partecipanti al PCC Wikidata Pilot, ma anche per tutti coloro che intendono descrivere le risorse bibliografiche in/con Wikidata.

### 4.3 #1lib1ref su Wikidata

La campagna #1lib1ref (*One librarian, one reference*) è stata lanciata per la prima volta dalla Wikimedia Foundation nel gennaio del 2016, in occasione del quindicesimo compleanno di Wikipedia<sup>16</sup>, quando è stato chiesto ai bibliotecari di tutto il mondo di migliorare l'enciclopedia collaborativa, aggiungendo o ampliando i riferimenti bibliografici contenuti nelle voci, con il fine di renderla maggiormente attendibile. Supportata anche da prestigiosi partner istituzionali, come OCLC, JStor e, IFLA, la campagna oggi si tiene due volte all'anno, nel mese di gennaio, in associazione al compleanno di Wikipedia, e a maggio. Nelle pagine relative alla campagna, la Wikipedia Library ha illustrato i molti modi per contribuire<sup>17</sup>, il più semplice dei quali è risolvere le esigenze di citazioni necessarie già segnalate dai *template*, o cercare all'interno della categoria di servizio "Senza fonti" di Wikipedia. Ad aprile 2021 l'*IFLA Wikidata and Wikibase Working Group*<sup>18</sup> ha proposto un incontro di formazione per coloro che intendessero organizzare delle working session per contribuire alla campagna #1Lib1Ref su Wikidata: Meg Wacha della City University di New York ha formato più di cinquanta partecipanti durante gli incontri "Train-The-Trainers".<sup>19</sup> Tre organizzatori hanno programmato delle sessioni di formazione e working hours per #1lib1ref su Wikidata: la Faculdade de Arquitetura e Urbanismo da Universidade de São Paulo (FAUUSP), la Concordia University Library e il Gruppo Wikidata per Musei, Archivi e Biblioteche, con un risultato finale di 14.500 nuovi riferimenti bibliografici e 25.800 item modificati.<sup>20</sup>

### 4.4 NWBib: una classificazione dei soggetti basata su Wikidata

La Bibliografia del Nord Reno-Westfalia (NWBib) è una bibliografia regionale che registra la letteratura relativa alla regione tedesca del Nord Reno-Westfalia (NRW) e ad oggi ha all'attivo circa 440.000 risorse. La classificazione del NWBib<sup>21</sup> comprende anche i vocaboli controllati per circa

<sup>15</sup> Mappatura dei campi e sottocampi del MARC21 con le relative proprietà di Wikidata, a cura di PCC, <<https://docs.google.com/spreadsheets/d/1d4w10m6ChNshX8fiaqackiLlQvNt0K2fSlkY9XuKgb0/edit#gid=1077502018>>

<sup>16</sup> Si rimanda a Donna Scheeder, *Wikipedia birthday greetings*, 15 gennaio 2016, <<https://www.ifla.org/node/10118>>

<sup>17</sup> The Wikipedia library: #1lib1ref, come partecipare, <[https://meta.wikimedia.org/wiki/The\\_Wikipedia\\_Library/1Lib1Ref/Participate/it](https://meta.wikimedia.org/wiki/The_Wikipedia_Library/1Lib1Ref/Participate/it)>

<sup>18</sup> L'IFLA Wikidata Working Group nasce nel 2019 per esplorare e sostenere l'uso e il contributo a Wikidata da parte di bibliotecari e professionisti dell'informazione, l'integrazione di Wikidata e Wikibase con i sistemi bibliotecari e l'allineamento dell'ontologia di Wikidata con formati di metadati delle biblioteche come BIBFRAME, RDA e MARC.

<sup>19</sup> Cfr. #1Lib1Ref - *Start your own event on Wikidata!*, <<https://www.ifla.org/node/93813>>

<sup>20</sup> I dati completi della campagna su Wikidata si possono consultare sulla Wikimedia Outreach Dashboard <[https://outreachdashboard.wmflabs.org/campaigns/1lib1ref\\_on\\_wikidata\\_2021/programs](https://outreachdashboard.wmflabs.org/campaigns/1lib1ref_on_wikidata_2021/programs)>

<sup>21</sup> Die Nordrhein-Westfälische Bibliographie, <<https://nwbib.de/spatial>> .

4.500 luoghi o aree geografiche. I relativi dati strutturati vengono archiviati come file RDF/Turtle utilizzando il Simple Knowledge Organization System (SKOS), che è a sua volta per la maggior parte derivato da Wikidata. I vantaggi di questa scelta sono numerosi: ad esempio, il processo di soggettazione è più veloce grazie alla nuova interfaccia e gli errori di battitura sono pressoché impossibili, la derivazione delle coordinate geografiche mediante valori controllati presenti in Wikidata per l'indicizzazione spaziale è molto più precisa e sicura del metodo precedente: il nuovo flusso di lavoro, insieme ai nuovi strumenti di monitoraggio, comporta un onere di manutenzione minimo e risulta facilmente gestibile dal team che mantiene l'infrastruttura tecnica.<sup>22</sup>

## 5. Conclusioni

La rapida evoluzione del nuovo ecosistema informativo spinge a un'ampia discussione sul ruolo che i principi e le pratiche dei dati collegati svolgono nel plasmare i metadati bibliografici e il loro ambiente di produzione e uso. Le tecnologie LOD forniscono un potente mezzo per dare una sempre maggiore visibilità a questa compagine di dati, che potrebbero diventare il punto cruciale di innumerevoli relazioni e i catalizzatori per molteplici contesti, sia partendo dal livello atomico, rappresentato da singoli metadati, fino ad arrivare a strutture di grafici integrati e migliaia di set di dati interconnessi fra loro. In tale contesto, Wikidata offre un ottimo ambiente per la sperimentazione: è un sistema aperto, libero, collaborativo con poche e semplici regole, di facile utilizzo anche per chi non sia esperto di tecnologie all'avanguardia. Sia i professionisti dell'informazione che una buona parte dell'ambiente accademico e scientifico hanno dapprima accolto, in seguito caldeggiato, la possibilità di rivoluzionare la produzione e la gestione dei dati bibliografici servendosi di nuovi strumenti, linguaggi e ontologie, tanto che si stanno moltiplicando le iniziative in merito a livello locale e internazionale, dalle realtà più tecnologicamente avanzate ai paesi del sud del mondo, capaci di comprendere le enormi potenzialità di progetti come Wikidata o software come Wikibase. Un'attenzione particolare deve essere posta anche nella valutazione dei rischi che tali cambiamenti possono generare: non si possono “sostituire i documenti primari con dati granulari” né “delegare completamente (o, in ogni caso, prevalentemente) agli algoritmi e a meccanismi automatici l'aggregazione di dati”, come sottolinea il *Manifesto per il Web Semantico* a cura del Gruppo di studio Catalogazione ed indicizzazione dell'Associazione Italiana Biblioteche. Esiste anche la possibilità che i dati possano “non essere oggettivi e neutrali, e la scelta, l'organizzazione, la contestualizzazione, l'interpretazione e la valutazione sono attività nelle quali le macchine possono essere utili ma non potranno mai sostituirsi agli esseri umani. [...] Il processo di analisi delle risorse dovrebbe portare a una nuova sintesi nella quale questi dati estremamente granulari non perdono significato, né la relazione con gli oggetti e i loro contesti originari”. All'interno di un sistema di dati talmente ampio e inclusivo da divenire a volte dispersivo, diventa necessario per le biblioteche (e gli istituti culturali in generale) affermare un'identità e una funzione ben definite, perché “se, da una parte, il paradigma dei LOD offre la base alle comunità per condividere il loro patrimonio culturale, dall'altra ogni singola comunità coinvolta deve avere un ruolo nel processo,

---

<sup>22</sup> Cfr. Adrian Pohl. 2021. “How We Built a Spatial Subject Classification Based on Wikidata”, *Code4lib Journal*, 51, <<https://journal.code4lib.org/articles/15875>> .



valorizzando la propria autorevolezza, in modo da raggiungere una convergenza che includa la ricchezza dei dati in relazione a ogni contesto identificabile, definito e verificabile”.<sup>23</sup>

In conclusione, le potenzialità della nuova generazione dei metadati unite a quelle del Web semantico e all’utilizzo di strumenti liberi e collaborativi come Wikidata e Wikibase aprono nuovi scenari e ampliano confini ormai ristretti. In tale contesto, la figura del Wikidata librarian può contribuire attivamente a un cambiamento di paradigma, in nome di una conoscenza universalmente accessibile e riutilizzabile, in cui i professionisti e gli studiosi dell’informazione tornano a essere protagonisti.

---

<sup>23</sup> Cfr. Gruppo di studio Catalogazione ed indicizzazione dell’Associazione Italiana Biblioteche. 2020. *Manifesto per il Web Semantico: la comunità dei dati*, passim, <<https://www.aib.it/struttura/commissioni-e-gruppi/gruppo-di-studio-catalogazione-ed-indicizzazione/2020/86520-manifesto-per-il-web-semantico-la-comunita-dei-dati/>>.

## Riferimenti bibliografici

Associazione Italiana Biblioteche - Gruppo di studio Catalogazione, indicizzazione, linked open data e web semantico (CILW). 2020. “Manifesto per il Web Semantico: la comunità dei dati”. <<https://www.aib.it/struttura/commissioni-e-gruppi/gruppo-di-studio-catalogazione-ed-indicizzazione/2020/86520-manifesto-per-il-web-semantico-la-comunita-dei-dati/>>.

(Bartholomei et al. 2016) Bartholomei, Stephan, Rachel Franks, James Heilman, Mylee Joseph, Vicki McDonald, Anna Raunik, Mia Ridge, Mark Robertson. 2016. *Opportunities for Academic and Research Libraries and Wikipedia*, IFLA, <<https://www.ifla.org/files/assets/hq/topics/info-society/iflawikipediaopportunitiesforacademicandresearchlibraries.pdf>>.

(Bianchini 2005) Bianchini, Carlo. 2005. *Riflessioni sull'universo bibliografico: funzioni, oggetti e modelli della catalogazione per autore e per titolo*, Milano: Sylvestre Bonnard.

British Library. 2019. *Foundations for the Future: The British Library's Collection Metadata Strategy 2019-2023*. London: British Library.

(Chapman 2011) Chapman, John W. 2011. “The Roles of the Metadata Librarian in a Research Library”. *Library Resources & Technical Services* 51 (4): 279–85. <https://doi.org/10.5860/lrts.51n4.279>.

(Guerrini 2020) Guerrini, Mauro. 2020. *Dalla catalogazione alla metadattazione. Tracce di un percorso*. Roma: Associazione Italiana Biblioteche.

(Han e Hswe 2011) Han, Myung-Ja, e Patricia Hswe. 2011. «The Evolving Role of the Metadata Librarian». *Library Resources & Technical Services* 54 (3): 129–41. <<https://doi.org/10.5860/lrts.54n3.129>>.

Linked Data for Production Group. “LD4P2 Project Background and Goals - LD4P2 Linked Data for Production: Pathway to Implementation - LYRASIS Wiki”, <<https://wiki.lyrasis.org/display/LD4P2/LD4P2+Project+Background+and+Goals>>.

(Morriello 2020) Morriello, Rossana. 2020. «Birth and Development of Data Librarianship». *JLIS. It* 11 (3): 1–15. <<https://doi.org/10.4403/jlis.it-12653>> .

(Olive 2021) Olive, Christine. 2021. “Meeting report New Horizons: emerging metadata standards and practices in the 21st century”. In *IFLA Metadata Newsletter*, 7, <[https://www.ifla.org/files/assets/classification-and-indexing/newsletters/ifla\\_metadata\\_newsletter\\_june\\_2021\\_updated.pdf](https://www.ifla.org/files/assets/classification-and-indexing/newsletters/ifla_metadata_newsletter_june_2021_updated.pdf)>.

(Pohl 2021) Pohl, Adrian. 2021. “How We Built a Spatial Subject Classification Based on Wikidata”. *The Code4Lib Journal*, 51. <<https://journal.code4lib.org/articles/15875>>.

(Schreur 2018) Schreur, Philip E. 2018. “Linked Data for Production (LD4P): a Multi-Institutional Approach to Technical Services Transformation”. In *Companion Proceedings of the The Web Conference 2018*, 429–430. WWW '18. Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee. <<https://doi.org/10.1145/3184558.3186201>>.

(Smith-Yoshimura 2019) Smith-Yoshimura, Karen. 2019. “Alternatives to Statistics for Measuring Success and Value of Cataloging”. *Hanging Together* (blog). 15 aprile 2019. <<https://hangingtogether.org/?p=7122>>.

(Smith-Yoshimura 2020) ———. 2021. “Transitioning to the Next Generation of Metadata”. OCLC, <<https://doi.org/10.25333/rqgd-b343>> .

(Valacchi 2021) Valacchi, Federico. 2021. “For a Hypothesis of Metacatalogation”. *Bibliothecae.it*, 10: (1), 463–487. <https://doi.org/10.6092/issn.2283-9364/13077>.

(Weston e Sardo 2017) Weston, Paul Gabriele, e Lucia Sardo. 2017. *Metadati*. Roma: Associazione italiana biblioteche.

## Analysis and data visualization in bibliometric studies

Joel Alhuay-Quispe<sup>(a)</sup>, Alonso Estrada-Cuzcano<sup>(b)</sup>,  
Lourdes Bautista-Ynofuente<sup>(c)</sup>

a) Universidad Privada San Juan Bautista, Facultad de Ingeniería. Lima, Perú, <https://orcid.org/0000-0002-1903-4687>

b) Universidad Nacional Mayor de San Marcos. Lima, Perú, <https://orcid.org/0000-0001-5039-1108>

c) Universidad Nacional Mayor de San Marcos. Lima, Perú, <https://orcid.org/0000-0002-0277-489X>

**Contact:** Joel Alhuay-Quispe, [joel.alhuay@upsjb.edu.pe](mailto:joel.alhuay@upsjb.edu.pe); Alonso Estrada-Cuzcano, [mestradac@unmsm.edu.pe](mailto:mestradac@unmsm.edu.pe);  
Lourdes Bautista-Ynofuente, [lourdes.bautista1@unmsm.edu.pe](mailto:lourdes.bautista1@unmsm.edu.pe)

**Received:** 14 July 2021; **Accepted:** 17 January 2022; **First Published:** 15 May 2022

### ABSTRACT

In recent years, academics with different professional formation other than Information Sciences have been involved in research related to metric studies using approaches and techniques based on bibliometrics; however, these professionals do not have knowledge and necessary competences for conduct metric information studies in a rigorous and methodical way. This article aims to find what are methods and tools used by bibliometrics researchers through a descriptive and textual analysis, using methods such as word co-occurrence in bibliographical data extracted from scientific articles indexing in Web of Science, research area "Library & Information Science." We find most frequently software and tools used in bibliometrics have four levels for data treatment: recovery, preparation, processing and analysis, visualization; and use different methods analysis: networks social, geospatial, thematic, temporal, bibliographic coupling, among others.

### KEYWORDS

Bibliometrics; Computer programs; Data analysis; Data visualization; Co-word analysis.

## Análisis y visualización de datos en estudios bibliométricos

### RESUMEN

En los años recientes, académicos de formaciones distintas a las Ciencias de la Información con mayor frecuencia realizan estudios utilizando enfoques y técnicas basados en bibliometría; sin embargo, estos profesionales no siempre tienen los conocimientos y la formación necesarias para realizar estudios métricos de la información de manera rigurosa y metódica. El estudio pretende describir los métodos y herramientas utilizados por investigadores a partir de un análisis de contenido textual empleando métodos de análisis de co-ocurrencia de términos de artículos científicos indizados en *Web of Science*, área de investigación "Bibliotecología". Se encontró que los programas y herramientas informáticas más usadas presentan cuatro niveles de manipulación de datos: recuperación, preparación, procesamiento y análisis, visualización; y emplean distintos métodos de análisis: sociogramas, geoespacial, temático, temporal, acoplamiento, entre otros.

### PALABRAS CLAVE

Bibliometría; Programas informáticos; Análisis de datos; Visualización de datos; Análisis de Co-ocurrencia de términos.

## 1. Introducción

El uso de tecnologías emergentes para el consumo masivo de información y el procesamiento de grandes cantidades de datos se ha incrementado en todas las áreas del conocimiento; resulta imperioso conocer y utilizar herramientas y programas informáticos para fines de investigación y enseñanza en el camino hacia la construcción de nuevo conocimiento. En el ámbito académico, estudiantes universitarios, docentes y académicos emplean herramientas tecnológicas para recuperar, procesar y analizar la información, a fin de comunicar y representar de forma objetiva los resultados de sus hallazgos y conocimientos. En este contexto, observamos que gestores de investigación y consultores independientes utilizan enfoques y técnicas basados en estudios métricos de la información (i.e. bibliometría) no necesariamente para fines académicos; sin embargo, estos profesionales no siempre tienen la formación necesaria o las capacidades específicas para utilizar las herramientas bibliométricas de modo adecuado (Vílchez-Román y Alhuay-Quispe 2016).

En la última década, el crecimiento de estudios en Bibliotecología y Ciencias de la Información muestra un estado progresivo, en específico, en estudios métricos de la información como la bibliometría (Figura 1). Destacan la diversificación en el uso de enfoques y fundamentos teóricos, métodos y modelos de análisis, así como herramientas informáticas para la representación de los datos resultantes del análisis bibliográfico; no obstante, pocos se han detenido a estudiar las características de los programas informáticos que permiten una adecuada presentación de datos.

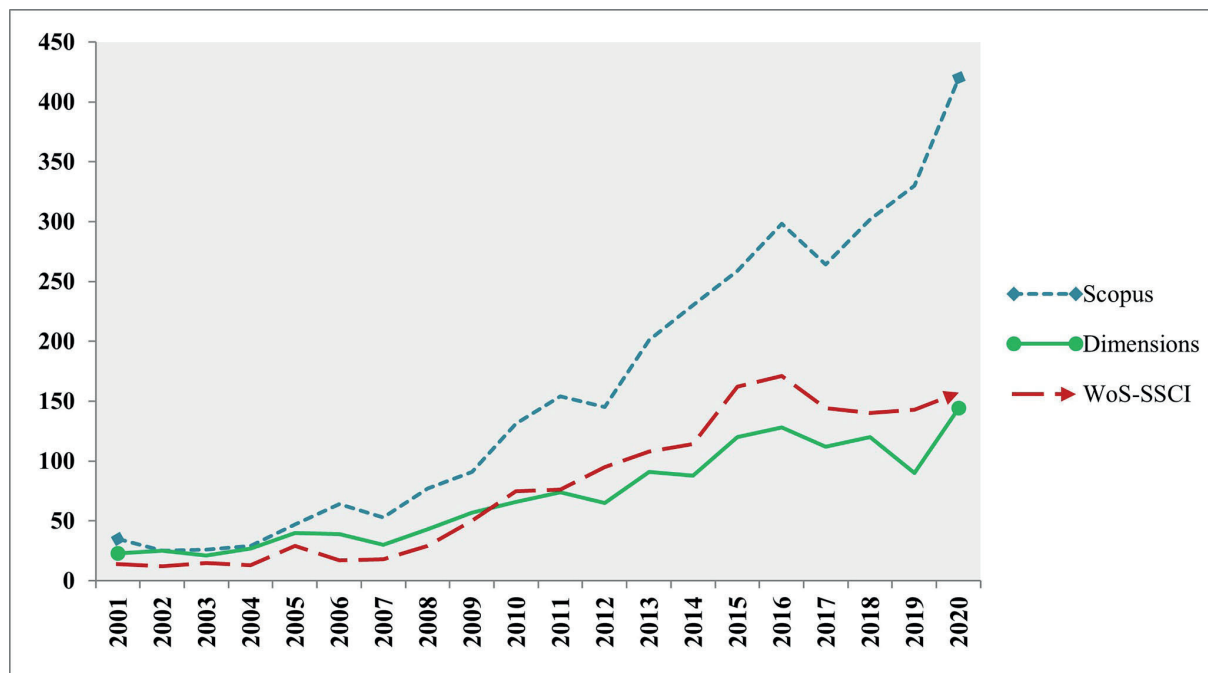


Figura 1. Evolución de producción de artículos sobre “bibliometría” en Scopus, Web of Science y Dimensions.

## 2. Orígenes y nexos entre disciplinas métricas de información

Las disciplinas métricas de la información se han diversificado y adquirido un *corpus* teórico y metodológico especializado. El antecedente directo de la bibliometría, por ejemplo, fue el término *bibliografía estadística* usada en 1922 por E. Wyndham Hulme en una conferencia titulada *Standards Reader in Bibliography* presentada en la Universidad de Cambridge (Pritchard 2003); pero no sería sino hasta 1944 que el término volvió a ser usado por Gosneell en su artículo *Obsolescence of books in college libraries* (Gosnell 1978). Pritchard (2003) recomienda el uso de “bibliometría” para referirse a la aplicación de métodos estadísticos a los libros y a otros medios de comunicación; mientras que Tague-Sutcliffe (1992) indica que la bibliometría engloba el estudio de los aspectos cuantitativos de la producción, disseminación y uso de la información registrada para elaborar modelos matemáticos para realizar pronósticos y posteriormente tomar decisiones.

La Cienciometría o *Naukometriya*, por su origen ruso, tiene como precursores a Dobrovoy y Nalimov en 1966 y 1969, respectivamente (Callon, Courtial, y Penan 1995); aunque otro posible origen se asocia al surgimiento de la revista *Scientometrics* en 1978, que marcó un hito importante en el ámbito de las entonces conocidas como ciencias cuantitativas. Spinak (1998, 142) sostiene que la cienciometría aplica técnicas bibliométricas a la ciencia y examina el desarrollo y las políticas científicas para establecer comparaciones entre las políticas de investigación entre los países analizando sus aspectos económicos y sociales.

Tague-Sutcliffe (1992) atribuye el surgimiento del término Informetría, al alemán Otto Nacke, quién lo habría propuesto por primera vez en el año 1979, y afirma además que la informetría abarca el estudio de los aspectos cuantitativos de la información, independientemente de la forma en que aparezca registrada y del modo en que se genere y puede incorporar y utilizar diversos medios en la medición de la información, que están fuera de los límites de la bibliometría y de la cienciometría.

El término “webmetría” aparece originalmente como *Webometrics* en inglés cuyo término fue acuñado por Almind y Ingwersen (1997). La webmetría consiste en la aplicación de los métodos informétricos a la *World Wide Web* y puesto que la web se está tornando en un medio importante para la comunicación de la ciencia, es lógico que los estudios cuantitativos se orienten a este ámbito (Vanti 2002). Una de las aplicaciones reales de los métodos basados en elementos web se pueden encontrar en el Ranking Web de Universidades (Webometrics.info), el cual en el ámbito latinoamericano posee gran aceptación y relevancia, ya que no solo incluye indicadores de medición web como presencia y visibilidad web sino también recoge indicadores bibliométricos utilizados en la web académica.

Recientemente *Altmetrics* o métricas alternativas aparecen como un estudio de nuevas formas de medición del alcance e impacto de las publicaciones científicas en el entorno de la social media. Altmetrics abarca el estudio de la actividad científica a partir de nuevos indicadores basados en fuentes de información de la web social, considerando el uso (*views, downloads*), interacciones y recomendaciones (comentarios, favoritos), publicaciones y compartidos (*posts, re-tuits y microblogging*), almacenamiento de lecturas (*bookmarking*) y hasta los comentarios de evaluación pública por pares (Uribe-Tirado y Alhuay-Quispe 2017). Las métricas alternativas surgen para proporcionar una visión complementaria al análisis de citas, asumiendo un papel en la evaluación de la ciencia (Borrego 2014).

Guzmán Sánchez (1999, 41) es de los pioneros de emplear el término Patentometría en español, quien entiende que abarca dos aspectos: la producción de conocimientos certificados y la participación en la elaboración de procesos industriales, en este último emplea cuatro indicadores de medición: actividad, relacionales de primera generación, relacionales de segunda generación, y familia de patentes. La utilización eficiente de la información contenida en las patentes es la principal herramienta para establecer tendencias en el registro de la propiedad intelectual, innovación o novedades que pueden producir cambios o surgimientos abruptos de tecnologías (Díaz, Giráldez, y Carrillo 2016; Verma y Suri 2021).

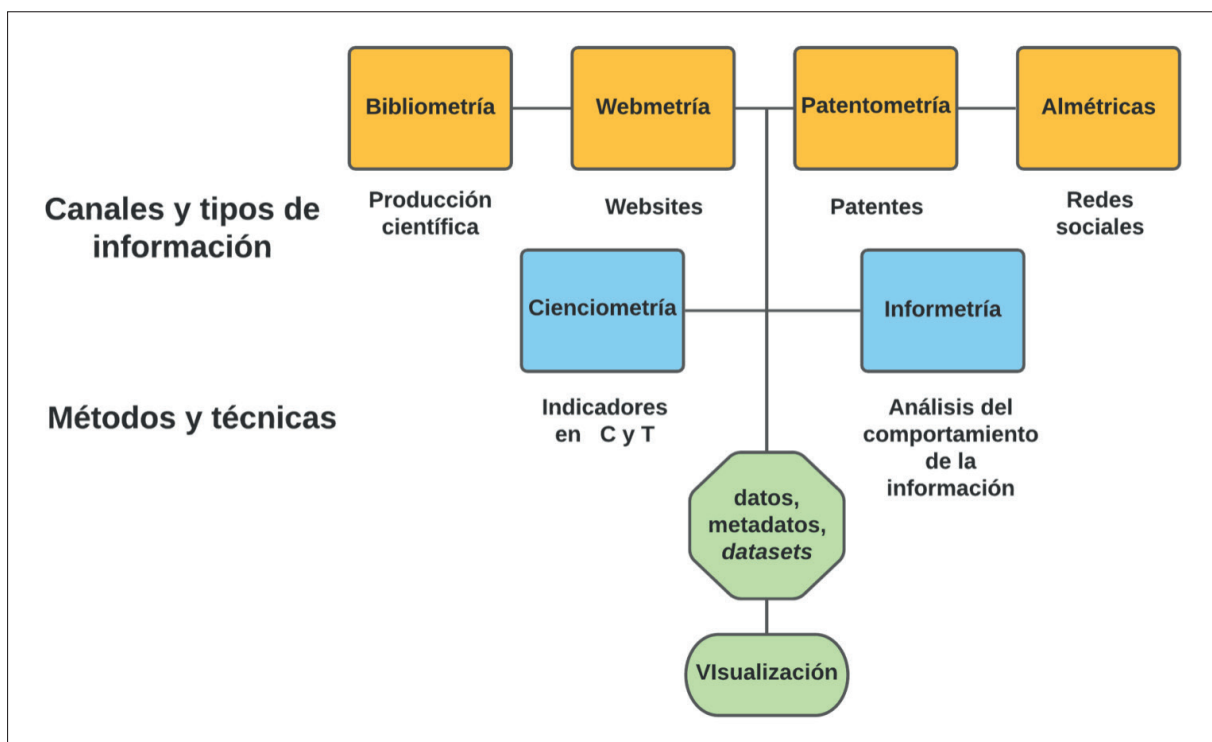


Figura 2. Propiedades de las disciplinas métricas para la visualización

Una breve descripción de los orígenes de cada disciplina métrica nos permite introducirnos de manera panorámica a los elementos que conforman las unidades de estudio, las técnicas de análisis utilizada, los enfoques metodológicos y los fundamentos teóricos, así como, conocer las relaciones y características propias del flujo de la información y comprender la importancia del estudio del tratamiento de los datos en la actividad informacional en cada disciplina métrica. En ese sentido, en la Figura 2 se muestran los elementos clave de cada disciplina métrica, en donde encontramos que algunas se centran en el análisis de canales de información determinados (web, redes sociales), mientras que otras se centran en fuentes de información específicas (artículos científicos, libros) que fueron el principal soporte de estudio en los inicios de la bibliometría. En todos los casos, las disciplinas desarrollaron métodos y técnicas para el análisis de metadatos bibliográficos y conjuntos de datos como insumos de entrada para la visualización de los datos analizados.

### 3. Visualización de datos e información

La visualización de datos no es una actividad reciente, según Friendly (2008), tiene una historia de muchos siglos atrás; que se remonta desde los primeros mapas y diagramas (antes del siglo XVII), medición y teoría (1600-1699), nuevas formas gráficas (1700-1799), el comienzo de los gráficos modernos (1800-1850), la edad de oro de los gráficos estadísticos (1850-1900), la Edad Media Moderna (1900-1950), Renacimiento de la visualización de datos (1950-1975) hasta la visualización de datos dinámica, interactiva y de alta resolución.; aunque la periodización que establece el autor, según propias palabras, es algo artificial pero brinda la oportunidad de caracterizar los logros en cada período. En este último periodo, Tufte (2016) publica *The Visual Display of Quantitative Information*, un libro popular e innovador sobre la visualización de datos.

La visualización de datos se ha utilizado para ayudar a la comprensión, recopilar datos e información, realizar análisis y comunicar evaluaciones en diversos entornos (Azzam et al. 2013). Actualmente todos los campos de estudio analizan y procesan grandes cantidades de datos (*big data*) que podemos denominar macrodatos donde la visualización de datos juega un rol fundamental para impulsar análisis complejos y demostrar la solidez de las investigaciones a través de presentaciones gráficas.

Por otro lado, la visualización de información surge de distintas áreas con enfoque multidisciplinario como la computación científica (Kim, Zhu, y Chen 2016), psicología (Sánchez-Bonvehí y Ribera 2014) y estadística (Valero Sancho, Català Domínguez, y Marín Ochoa 2014); sin embargo, el concepto abarca el estudio de la transformación de datos, información y conocimiento en representaciones visuales interactivas generadas por computadora (Chen 2010; Liu et al. 2014). Desde una perspectiva actual, la visualización de información es la conformación de una imagen mental a partir de un concepto abstracto y se obtiene a partir de la representación gráfica de variables asociadas al concepto que se quiere visualizar (Guzmán Sánchez, Calero, y Villaseñor 2008), y que emerge para cambiar el proceso comunicativo y devolver estratégicamente el protagonismo a lo visual; así mismo, permite profundizar y aprovechar las ventajas que ofrece la percepción humana y la interacción, con el objetivo de transformar realidades complejas y abstractas en realidades simples (Sánchez-Bonvehí y Ribera 2014).

La visualización de información bibliométrica es una consecuencia del surgimiento de una serie de metodologías y herramientas con fines específicos para un área determinada, una base de datos señalada, indicadores limitados y uso de software para procesamiento y análisis creados con propósitos diferentes al análisis bibliométrico (Sotolongo Aguilar, Guzmán Sánchez y Carrillo 2002). Paralelamente surgen otros enfoques apoyados en el tratamiento y análisis de datos e información, tanto para su uso, comprensión y comunicación en entornos de la academia, universidad y en la industria. La visualización a partir de datos pretende construir un conjunto gráfico y sintético que permita entender, establecer agrupaciones, relaciones o tendencias estadísticas y facilite la obtención de conclusiones para su interpretación (Valero Sancho, Català Domínguez, y Marín Ochoa 2014). El uso de técnicas de visualización de datos se ha diversificado en áreas como cartografía, demografía, ciencias de los datos; sin embargo, su desarrollo destaca en la cartografía en línea, cuya combinación con técnicas como minería de datos, web semántica, datos abiertos enlazados, y otras técnicas avanzadas de visualización, permiten acceder y poner en valor el conocimiento científico (Martín-Forero Morente, Barriuso Mediavilla, y Del-Bosque 2018).



La visualización científica es una forma muy útil para extraer información importante de forma concisa y rápida (Cadañá Subiñas 2016). Se basa en el uso de imágenes y en la transformación de datos o información en imágenes para explicar y comunicar ideas (De La Cruz Salas 2005). Como herramienta para la representación visual de un conjunto de datos que facilita su efectivo análisis y evaluación, permite a los científicos computacionales analizar, entender y comunicar datos numéricos generados durante una investigación (Yatagama Lokuge 2015).

De La Cruz Salas (2005) distingue dos grupos, la visualización informativa y visualización científica – propiamente –, donde la primera hace referencia a aquellos datos que no precisan de una estructuración espacial (e.g. datos que se representen con gráficos de barras), mientras que la visualización científica precisa de un contexto espacial para representarse con claridad: representa elementos reales que precisan un espacio sobre el cual ser visualmente integrado.

El estudio pretende identificar y analizar las características de los programas informáticos empleados para el procesamiento y visualización de datos en estudios bibliométricos de Ciencias de la Información.

## 4. Métodos

La investigación es de carácter descriptiva y analítica con enfoque cuantitativo y corte transversal. Emplea la técnica de análisis de términos y toma como unidades de análisis a datos bibliográficos textuales (título, resumen, título y palabras clave) de artículos publicados en revistas indizadas en Web of Science, colección principal.

Los análisis de términos implican la extracción de palabras clave a partir de los registros documentales en bases de datos de indización, pero también se pueden extraer las palabras clave de un documento textual, a partir del cual construir una matriz de palabras clave (He 1999). Posteriormente se procede a analizar la matriz de datos textuales (*corpus*) con el propósito de identificar características relevantes de los datos que permitan responder preguntas de investigación a partir de redes de colaboración, términos de tendencia, relaciones y jerarquías de conceptos, polarización de tópicos, entre otros métodos. Se entiende por polarizar los problemas de investigación a que un término puede estar relacionado con muchas otras palabras, por lo que será particularmente importante para la definición de problemas centrales (Courtial, Callon, y Sigogneau 1984).

Se emplea como técnica de recolección de datos a la revisión documental que incluye la consulta de fuentes de tipo secundario: *Web of Science* (WoS), y otros sitios web. La recuperación de datos de WoS se basa en la búsqueda avanzada bajo dos variables agrupadas por el operador “OR” y asociadas por “AND” que se detalla en el Material Suplementario. En la consulta efectuada se emplean límites de tipo documental (*article, review*) y fecha de publicación (2008-2018) pero sin restricción de idioma de la publicación. Se obtuvieron cantidades distintas de acuerdo con las colecciones principales WoS: SCI-Expanded (n=600); SSCI (n=622); ESCI (n=214); A&HCI (n=12). De los resultados se excluyen la colección ESCI por la cobertura reciente (2015-actual) y A&HCI por la reducida cantidad encontrada que asumimos no representativa para el estudio. Los resultados de ambas colecciones restantes (SCI y SSCI) fueron combinadas obteniéndose 848 artículos, siendo esta cantidad asumida como la muestra del estudio, de carácter no probabilístico y tipo intencional o criterial.

Las unidades de análisis (n=848) fueron exportadas en formato de la propia base de datos (*vos tagged format*) en fichero .txt para la preparación del corpus. El proceso de pre-preparación de los datos textuales emplea el análisis de frecuencia de términos usando dos herramientas de análisis textual: TAPoRWare (*Text Analysis Portal for Research* - <http://taporware.ualberta.ca>) y Textanz (<http://www.textanz.com>).

Para revelar la presencia y medir la magnitud (intensidad) de los términos se calcula la frecuencia con la que cada palabra clave está asociada entre sí y se dice que existe una conexión cuando esta frecuencia supera cierta probabilidad de asociación ( $p > 0,5$ ), que se calcula desde la palabra más frecuente hacia la menos frecuente considerando que la palabra más frecuente representa una postura de interés más general, es decir, su poder para polarizar es mayor siendo esta probabilidad llamada coeficiente de inclusión (Courtial, Callon, y Sigogneau 1984).

La co-ocurrencia permitió identificar grupos de términos resultantes de cada programa, los cuales fueron revisados manualmente del más frecuente al menos recurrente con el fin de reconocer nombres de programas o aplicaciones web que hayan sido empleados en los artículos bibliométricos analizados. Luego se seleccionaron los disponibles con sitios web en Internet y cuenten con estas características:

- posibilidad de descarga (prueba o gratuito),
- información de guía de usuario,
- documentación publicada como artículo, libro u otro.

Las variantes de nombres del conjunto de programas y aplicativos informáticos identificados de la etapa de pre-procesamiento fueron normalizadas usando un tesoro de términos que fue construido para tal propósito. Finalmente, se realizaron análisis posteriores de frecuencia y correlación de términos sobre la base del corpus de texto y el listado de nombres de programas, empleando el programa VantagePoint.

## 5. Resultados

### 6. Análisis descriptivo: programas y frecuencias de aparición

Se encontraron veinte programas de computadora y aplicativos en línea empleados como herramientas en estudios bibliométricos, algunos desarrollados por programadores o investigadores particulares dedicados al estudio de la ciencia; y otros, desarrollados por instituciones vinculadas a los mismos ámbitos. En la Tabla 1 se muestran detalles principales de los programas como desarrolladores del software y la página de descarga respectiva.

Tabla 1. Lista de programas informáticos utilizados en estudios bibliométricos

N°	Nombre del programa	Desarrollador	Enlace de descarga
1	BibExcel	Olle Persson, Universidad de Umea – Suecia	<a href="http://homepage.univie.ac.at/juan.gorraiz/bibexcel/index.html">http://homepage.univie.ac.at/juan.gorraiz/bibexcel/index.html</a>
2	CiteSpace II	Universidad de Drexel – USA	<a href="http://cluster.ischool.drexel.edu/~cchen/citespace/download/">http://cluster.ischool.drexel.edu/~cchen/citespace/download/</a>
3	CitNetExplorer	Universidad de Leiden - Países Bajos	<a href="http://www.citnetexplorer.nl/Download">http://www.citnetexplorer.nl/Download</a>
4	CoPalRed	Rafael Bailón-Moreno. Universidad de Granada	<a href="http://ec3.ugr.es/copalred/">http://ec3.ugr.es/copalred/</a>
5	Egonet	Christopher McCarthy. Universidad de la Florida – USA	<a href="http://www.arschile.cl/egonet/">http://www.arschile.cl/egonet/</a>
6	HistCite	Eugene Garfield	<a href="http://interest.science.thomsonreuters.com/forms/HistCite/">http://interest.science.thomsonreuters.com/forms/HistCite/</a>
7	Loet Leydessdorf	Universidad Amsterdam	<a href="http://www.leydesdorff.net/">http://www.leydesdorff.net/</a>
8	Meta Knowledge	John McLevey y Reid McIlroy-Young	<a href="http://networkslab.org/metaknowledge/">http://networkslab.org/metaknowledge/</a>
9	Network Workbench	Universidad de Indiana	<a href="http://nwb.cns.iu.edu/download.html">http://nwb.cns.iu.edu/download.html</a>
10	Pajek	Vlado	<a href="http://vlado.fmf.uni-lj.si/pub/networks/pajek/">http://vlado.fmf.uni-lj.si/pub/networks/pajek/</a>
11	Publish or Perish	Tarma Software Research	<a href="http://www.harzing.com/download/">http://www.harzing.com/download/</a>
12	Scholarometer	Indiana University	<a href="http://scholarometer.indiana.edu/">http://scholarometer.indiana.edu/</a>
13	Science of Science	Universidad de Indiana	<a href="https://sci2.cns.iu.edu/user/index.php">https://sci2.cns.iu.edu/user/index.php</a>
14	Ucinet NetDraw	Lin Freeman, Martin Everett y Steve Borgatti	<a href="https://sites.google.com/site/ucinetsoftware/">https://sites.google.com/site/ucinetsoftware/</a>
15	VosViewer	Universidad de Leiden - Países Bajos	<a href="http://www.vosviewer.com/Home">http://www.vosviewer.com/Home</a>
16	SciMAT	University of Granada. Soft Computing and Intelligent Information Systems	<a href="https://sci2s.ugr.es/scimat/">https://sci2s.ugr.es/scimat/</a>
17	Gephi	Mathieu Bastian	<a href="https://gephi.org/">https://gephi.org/</a>
18	CitedReferencesExplorer	Andreas Thor, University of Applied Sciences for Telecommunications Leipzig	<a href="http://www.crexplorer.net/">http://www.crexplorer.net/</a>
19	STICCI	Valentín Gómez-Jáuregui, Cecilia Gómez-Jáuregui	<a href="http://www.sticci.eu/">http://www.sticci.eu/</a>
20	Sitkis	Henri Schildt, Helsinki University of Technology	<a href="https://sites.google.com/site/sitkisbibliometricanalysis/home">https://sites.google.com/site/sitkisbibliometricanalysis/home</a>

Fuente: Elaboración propia.

La Tabla 2 muestra un análisis descriptivo de frecuencia de aparición de los nombres de los programas, a partir de la base de datos de registros bibliográficos recuperados de la base de citación, el cual arrojó mayor predominancia (50.4%) de presencia en estudios bibliométricos a tres programas: *VosViewer*, *CiteSpace II* y *Ucinet NetDraw*. El análisis de co-ocurrencia de términos independientes confirma la preferencia de uso para dos primeros programas observados con el conteo por registros, pero aparecen *Pajek*, *HistCite* y *Publish or Perish* entre los primeros siguientes con 10% de frecuencia cada uno (Tabla 2).

Tabla 2. Frecuencia de aparición de los términos de programas usados

Nombre del programa	Número de registros				Número de ocurrencias			
	WoS	%	VP	%	TT	%	TW	%
VosViewer	51	26,02	64	35,56	80	40,82	80	40,82
CiteSpace II	26	13,27	33	18,33	48	24,49	47	23,98
Ucinet NetDraw	22	11,22	14	7,78	21	10,71	21	10,71
Pajek	20	10,20	14	7,78	20	10,20	21	10,71
HistCite	19	9,69	14	7,78	27	13,78	1	0,51
Publish or Perish	17	8,67	6	3,33	21	10,71	20	10,20
BibExcel	12	6,12	8	4,44	12	6,12	13	6,63
SciMAT	7	3,57	4	2,22	8	4,08	9	4,59
Loet Leydesdorff	5	2,55	2	1,11	2	1,02	3	1,53
Gephi	3	1,53	7	3,89	6	3,06	5	2,55
Science of Science	3	1,53	1	0,56	2	1,02	1	0,51
CitNetExplorer	2	1,02	5	2,78	8	4,08	7	3,57
Sitkis	2	1,02	1	0,56	1	0,51	1	0,51
CoPalRed	1	0,51	1	0,56	1	0,51	1	0,51
Egonet	1	0,51	1	0,56	4	2,04	4	2,04
Meta Knowledge	1	0,51	1	0,56	4	2,04	3	1,53
Scholarometer	1	0,51	0	0,00	1	0,51	1	0,51
STICCI	1	0,51	1	0,56	2	1,02	2	1,02
CitedReferencesExplorer	1	0,51	2	1,11	3	1,53	2	1,02
Network Workbench	1	0,51	1	0,56	1	0,51	1	0,51

VP: VantagePoint. TT: TexTanz. TW: TaporWare.

## 7. Nivel de manipulación y análisis de datos

En los estudios métricos exhaustivos y metódicos, el proceso que conduce a la visualización de datos analizados tiene cuatro niveles predecesores: a) recuperación de los datos a analizar; b) preparación de los datos recolectados, esto implica llevarlos a formatos compatibles con otros programas; c) procesamiento y análisis consiste en el empleo de algoritmos, métodos y técnicas para su posterior representación de los datos y d) la visualización, la parte final del proceso. De estas cuatro, en la Tabla 3 se muestra que la etapa visualización posee la mayor frecuencia (n=15); y son pocos (n=3) los que abarcan la mayor cantidad de etapas, de las cuales figuran: Science of Science, Meta Knowledge, SciMAT. De igual forma, tres aplicativos permiten la recuperación de información, siendo PoP el que no abarca ninguno de los demás niveles de manipulación de datos.

Tabla 3. Nivel de manipulación de los datos

N°	Nombre del programa	REC	PRE	PRO	VIS
1	BibExcel		x		
2	CiteSpace II				x
3	CoPalRed			x	
4	Loet Leydessdorf		x		
5	Network Workbench			x	x
6	Science of Science		x	x	x
7	VosViewer				x
8	HistCite				x
9	CitNetExplorer				x
10	Pajek			x	x
11	Pop	x			
12	Egonet			x	x
13	Scholarometer	x			x
14	Ucinet NetDraw			x	x
15	Meta Knowledge	x		x	x
16-	SciMAT		x	x	x
17-	Gephi			x	x
18-	Cited References Explorer				x
19	STICCI		x	x	
20	Sitkis			x	x

REC: Recuperación, PRE: Preparación, PRO: Procesamiento y análisis, VIS: Visualización

Se encontró que los programas son capaces de emplear diversos métodos como el análisis de redes, análisis geoespacial, análisis temático, análisis temporal, acoplamiento bibliográfico, co-ocurrencia, etc. En la Tabla 4 se muestra un análisis de correlación de términos agrupados como conceptos, un *ranking* de diez conceptos identificados donde se destacan los análisis de citación y de impacto científico.

Tabla 4. Matriz de correlación de términos por métodos de análisis de los datos

k	Términos	1	2	3	4	5	6	7	8	9	10
121	1. Citation Analysis	1,00	0,25	0,27	0,25	0,08	0,10	0,00	0,13	0,06	0,00
119	2. Scientific Impact		1,00	0,21	0,24	0,08	0,00	0,00	0,07	0,00	0,00
106	3. Bibliometric Indicators			1,00	0,32	0,08	0,00	0,00	0,25	0,10	0,00
80	4. Hirsch-Index				1,00	0,06	0,00	0,00	0,13	0,00	0,00
59	5. Social Network Analysis					1,00	0,30	0,20	0,05	0,16	0,10
33	6. Co-Citation Analysis						1,00	0,13	0,00	0,07	0,17
27	7. Co-Word Analysis							1,00	0,00	0,00	0,00
25	8. Research Performance								1,00	0,00	0,00
23	9. Collaboration									1,00	0,00
17	10. Data Visualization										1,00

K=número de ocurrencias.

La evaluación realizada encontró que las fuentes de datos empleadas indistintamente son: Web of Science, JCR, Scopus, PubMed, arXiv, Google Scholar, NASA ADS y National Science Foundation. La Tabla 5 muestra que Web of Science es la fuente de datos más recurrente, seguido de Scopus, lo que evidencia que actualmente son la autoridad en información científica internacional; y en menor proporción: PubMed y Google Scholar.

Tabla 5. Matriz de correlación de términos por fuente de entrada de datos

k	Términos	1	2	3	4	5
32	1. WoS	1,00	0,30	0,25	0,00	0,00
29	2. Scopus		1,00	0,26	0,00	0,11
18	3. Google Scholar			1,00	0,12	0,27
4	4. PubMed				1,00	0,00
3	5. GSM					1,00

K=número de ocurrencias. GSM=Google Scholar Metrics.

## 8. Discusión y conclusiones

La literatura reporta que diversos estudios proponen instrumentos y metodologías para el análisis de programas informáticos, otros estudian las metodologías empleadas en estudios métricos; sin embargo, los pormenores del tratamiento de los datos en investigaciones bibliométricas o de otras disciplinas métricas son escasas. En el primer enfoque, (Macías Rivero, Guzmán Sánchez, y Martínez Suárez 2009) diseñan un modelo de evaluación de software que aplica indicadores métricos para el ejercicio de la vigilancia científico-tecnológica, tomando como guía la ISO: 9126-1 (2001), y proponen indicadores que explican cómo tabular y describir cada software. En el segundo enfoque (Guzmán Sánchez, Calero, y Villaseñor 2008) se evalúan diferentes interfaces visuales a partir de un modelo de comparación, a fin de obtener información comparativa de las fases de entrada de los datos, análisis y proyección, para comprender mejor el resultado total de la visualización. En un tercer enfoque (Cobo et al. 2011) se analizan y comparan técnicas bibliométricas y programas empleados para el mapeo de la ciencia. Asimismo, Alfonso, Sakraida, y Hastings-Tolsma (2014) emplean el mapeo bibliométrico como método de representación y visualización de datos bibliométricos; además, revisan las consideraciones metodológicas de los programas de computadora empleados para el análisis de citas.

La representación o visualización de los datos resultaría una actividad casi imposible, más aún si se pretende analizar íntegramente la totalidad de los datos, comparar temporalmente el comportamiento e interrelación de los mismos o simplemente servir de instrumento didáctico de transmisión de conocimiento. Aunque el proceso de visualización y representación de datos tiene su origen en la presentación gráfica de datos en la Estadística, en la actualidad, la interacción entre la Matemática, Estadística y Ciencias de la Información se ha visto potenciado con el desarrollo de las Ciencias de la Computación y otros métodos, procesos y tecnologías particulares de éstas, relacionados con el tratamiento, almacenamiento y análisis de los datos (Gorbea Portal 2013), por lo que en el campo del estudio de la ciencia, el uso de métodos de análisis informétricos constituyen

una herramienta indispensable; sin embargo, la presentación tradicional de resultados de datos procesados en el análisis de citas, co-citaciones, co-autoría, redes sociales, etc., a manera de tablas planas, gráfico de barras, diagramas polares, entre otros, resulta bastante limitante al momento de representar datos analizados en un estudio bibliométrico y cuantitativo.

La representación de datos en estudios métricos permite su comprensión didáctica, y la representación visual de grandes cantidades de datos. Sin embargo, la exponencial creciente y explosión de la producción científica, materia de estudio del análisis métrico de la información, así como, administrar una gran cantidad de datos, desde su recuperación a través de diversas bases de datos bibliográficas, como su procesamiento y análisis, a partir de indicadores, variables y modelos bibliométricos, resultaría una actividad casi épica; además del riesgo en la veracidad de los resultados por la posible omisión o pérdida de datos que pudiera surgir en el proceso manual de la recuperación y análisis.

La visualización de datos como proceso último de la cadena de análisis de datos emplea herramientas informáticas de acuerdo con el propósito del estudio en particular. Estos estudios pueden orientarse al *análisis de redes sociales* (Molina, Muñoz, y Domenech 2002; Delgado López-Cózar et al. 2006; Repiso, Torres-Salinas, y Delgado López-Cózar 2011a; 2011b), *mapeo historiográfico* (Garfield 2004), *mapas bibliométricos* (Eck y Waltman 2010; 2007; Van Raan y Van der Velde 1991), entre otros. Algunos estudios previos (Cobo et al. 2011; Alfonzo, Sakraida, y Hastings-Tolsma 2014) demuestran que los programas informáticos empleados para el análisis y visualización de datos de estudios métricos son diversos, así como los métodos y técnicas de procesamiento, algunos más complejos que otros; de igual modo con las técnicas y algoritmos empleados por los programas para la representación gráfica de datos.

En conclusión, los programas informáticos empleados ya sea para el procesamiento o para la visualización de los datos bibliográficos en los artículos de análisis bibliométricos de Ciencias de la Información estudiados presentan hasta cuatro niveles de manipulación de los datos de entrada: i) recuperación, ii) preparación, iii) procesamiento y análisis, iv) visualización, en los que se emplean diferentes métodos y técnicas como los análisis de sociogramas, co-ocurrencias y acoplamiento bibliográfico, entre los más recurrentes. Respecto a los datos de entrada, estos provienen de varias fuentes externas de información bibliográfica y documental (i.e. bases de datos científicas, repositorios y buscadores), que en algunos casos los programas incluyen como funcionalidad, mientras que otros permiten la extracción masiva de datos de fuentes secundarias (Web of Science, JCR, Scopus, PubMed, arXiv, Google Scholar. Esta investigación abordó un aspecto que se refiere al proceso inicial de recuperación de datos; sin embargo, resulta importante el diseño de un modelo de evaluación de programas informáticos que vincule las características del manejo de datos en estudios métricos; además de modelos de análisis de interfaces visuales (Macías Rivero, Guzmán Sánchez, y Martínez Suárez 2009; Guzmán Sánchez, Calero, y Villaseñor 2008). Estudios futuros podrían indagar sobre la naturaleza y características del tratamiento y visualización de datos en las demás disciplinas métricas de la información de modo que la evidencia encontrada permita conocer y superar las limitaciones metodológicas frente al crecimiento masivo de los datos en el contexto de *big data* y facilite la utilización de los recursos e insumos de la producción científico e información documental para una adecuada toma de decisiones en la gestión de la información, de la investigación y del conocimiento explícito en las instituciones de educación superior y organizaciones corporativas.

## Referencias bibliográficas

- Alfonzo, Paige M., Teresa J. Sakraida, y Marie Hastings-Tolsma. 2014. «Bibliometrics: Visualizing the Impact of Nursing Research». *Online Journal of Nursing Informatics* 18 (1). <http://ojni.org/issues/?p=3093>.
- Almind, Tomas C., y Peter Ingwersen. 1997. «Informetric Analyses on the World Wide Web: Methodological Approaches to “Webometrics”». *Journal of Documentation* 53 (4): 404-26. <https://doi.org/10.1108/EUM0000000007205>.
- Azzam, Tarek, Stephanie Evergreen, Amy A. Germuth, y Susan J. Kistler. 2013. «Data Visualization and Evaluation». *New Directions for Evaluation* 2013 (139): 7-32. <https://doi.org/10.1002/ev.20065>.
- Borrego, Angel. 2014. «Altmétricas para la evaluación de la investigación y el análisis de necesidades de información». *Profesional de la Información* 23 (4): 352-58. <https://doi.org/10.3145/epi.2014.jul.02>.
- Cadahía Subiñas, Alberto. 2016. «Aplicación web enriquecida para visualización científica». Madrid, España: Universidad Politécnica de Madrid. <http://oa.upm.es/43533/>.
- Callon, Michel, Jean-Pierre Courtial, y Hervé Penan. 1995. *Cienciometría: la medición de la actividad científica: de la bibliometría a la vigilancia tecnológica*. Gijón, España: Ediciones Trea.
- Chen, Chaomei. 2010. «Information Visualization». *Wiley Interdisciplinary Reviews: Computational Statistics* 2 (4): 387-403. <https://doi.org/10.1002/wics.89>.
- Cobo, Manuel J., Antonio G. López-Herrera, Enrique Herrera-Viedma, y Francisco Herrera. 2011. «Science Mapping Software Tools: Review, Analysis, and Cooperative Study among Tools». *Journal of the American Society for Information Science and Technology* 62 (7): 1382-1402. <https://doi.org/10.1002/asi.21525>.
- Courtial, Jean-Pierre, Michel Callon, y M Sigogneau. 1984. «Is Indexing Trustworthy? Classification of Articles through Co-Word Analysis». *Journal of Information Science* 9 (2): 47-56. <https://doi.org/10.1177/016555158400900201>.
- Delgado López-Cózar, Emilio, Daniel Torres-Salinas, Evaristo Jiménez-Contreras, y Rafael Ruiz-Pérez. 2006. «Análisis bibliométrico y de redes sociales aplicado a las tesis bibliométricas defendidas en España (1976-2002): temas, escuelas científicas y redes académicas». *Revista Española de Documentación Científica* 29 (4 SE-Estudios): 493-524. <https://doi.org/10.3989/redc.2006.v29.i4.306>.
- Díaz, Maidelyn, Raudel Giráldez, y Humberto Carrillo. 2016. «La Patentometría como herramienta de vigilancia para monitorear las colaboraciones tecnológicas de un dominio». En *III Congreso Internacional de Información – INFO´ 2016, La Habana, Cuba*. La Habana, Cuba. <http://www.congreso-info.cu/index.php/info/2016/paper/view/44/50>.
- Eck, Nees Jan van, y Ludo Waltman. 2007. «Bibliometric Mapping of the Computational Intelligence Field». *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 15 (5): 625-45. <https://doi.org/10.1142/S0218488507004911>.



- . 2010. «Software Survey: VOSviewer, a Computer Program for Bibliometric Mapping». *Scientometrics* 84 (2): 523-38. <https://doi.org/10.1007/s11192-009-0146-3>.
- Friendly, Michael. 2008. «A Brief History of Data Visualization BT - Handbook of Data Visualization». En *Handbook of Data Visualization*, editado por Chun-houh Chen, Wolfgang Härdle, y Antony Unwin, 15-56. Berlin, Heidelberg: Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-540-33037-0\\_2](https://doi.org/10.1007/978-3-540-33037-0_2).
- Garfield, Eugene. 2004. «Historiographic Mapping of Knowledge Domains Literature». *Journal of Information Science* 30 (2): 119-45. <https://doi.org/10.1177/0165551504042802>.
- Goarbea Portal, Salvador. 1994. «Principios teóricos y metodológicos de los estudios métricos de la información». *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 9 (17): 23-32. <https://doi.org/10.22201/iibi.0187358xp.1994.17.3826>.
- . 2013. «Tendencias transdisciplinarias en los estudios métricos de la información y su relación con la gestión de la información y del conocimiento». *Perspectivas em Gestão & Conhecimento* 3 (1): 13-27. <https://periodicos.ufpb.br/ojs2/index.php/pgc/article/view/14175>.
- Gosnell, Charles F. 1978. «Obsolescence of Books in College Libraries». *Collection Management* 2 (2): 167-82. [https://doi.org/10.1300/J105v02n02\\_07](https://doi.org/10.1300/J105v02n02_07).
- Guzmán Sánchez, María Victoria. 1999. «Patentometría: herramienta para el análisis de oportunidades tecnológicas». Universidad de La Habana. <http://www.bvv.sld.cu/docs/documentos/119566413228.pdf>.
- Guzmán Sánchez, María Victoria, R. Calero, y E. Villaseñor. 2008. «Evaluación de interfaces visuales para estudios métricos: modelo de comparación». En *IV Seminario Internacional sobre Estudios Cuantitativos y Cualitativos de la Ciencia y la Tecnología*. La Habana, Cuba: Instituto de Información Científica y Tecnológica.
- He, Quin. 1999. «Knowledge Discovery Through Co-Word Analysis». *Library Trends* 48 (1): 133. <http://hdl.handle.net/2142/8267>.
- Kim, Meen Chul, Yongjun Zhu, y Chaomei Chen. 2016. «How Are They Different? A Quantitative Domain Comparison of Information Visualization and Data Visualization (2000–2014)». *Scientometrics* 107 (1): 123-65. <https://doi.org/10.1007/s11192-015-1830-0>.
- La Cruz Salas, Luis Miguel De. 2005. «Visualización científica». *Revista Digital Universitaria* 6 (12): 2-9. <http://www.revista.unam.mx/vol.6/num12/art117/int117.htm>.
- Liu, Shixia, Weiwei Cui, Yingcai Wu, y Mengchen Liu. 2014. «A Survey on Information Visualization: Recent Advances and Challenges». *Visual Computer* 30 (12). <https://doi.org/10.1007/s00371-013-0892-3>.
- Macías Rivero, Yaidelyn, María Victoria Guzmán Sánchez, y Yamila Martínez Suárez. 2009. «Modelo de evaluación para software que emplean indicadores métricos en la vigilancia científico-tecnológica». *ACIMED* 20 (6): 125-40. [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1024-94352009001200003](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1024-94352009001200003).
- Martín-Forero Morente, Lourdes, Alejandro Barriuso Mediavilla, e Isabel Del-Bosque. 2018.

«Webmapping y visualización de datos científicos en las humanidades digitales». En *XVIII Congreso Nacional de Tecnologías de la Información Geográfica: Perspectivas Multidisciplinares en la Sociedad del Conocimiento*, 812-20. Valencia, España: Universidad de Barcelona. <https://digital.csic.es/handle/10261/166984>.

Molina, José Luis, Juan Manuel Muñoz, y Miquel Domenech. 2002. «Redes de publicaciones científicas: un análisis de la estructura de coautorías.» *Redes. Revista Hispana para el análisis de redes sociales* 1 (1). [http://ddd.uab.cat/pub/redes/15790185v1/vol1\\_3.htm](http://ddd.uab.cat/pub/redes/15790185v1/vol1_3.htm).

Pritchard, Alan. 2003. «Bibliografía estadística o bibliometría.» *Infobib*, n.º 2: 33-36.

Raan, Anthony F. J. Van, y J. G. M. Van der Velde. 1991. «Bibliometric Mapping and Early Warning for Emerging Topics: Exploring Developments in Materials Science.» *Research Evaluation* 1 (2): 115-27. <https://doi.org/10.1093/rev/1.2.115>.

Repiso, Rafael, Daniel Torres-Salinas, y Emilio Delgado López-Cózar. 2011a. «Bibliometric and Social Network Analysis Applied to Television: Dissertations Presented in Spain (1976/2007)». *Comunicar* 19 (37): 151-59. <https://doi.org/10.3916/C37-2011-03-07>.

———. 2011b. «Análisis de la investigación sobre Radio en España: una aproximación a través del Análisis Bibliométrico y de Redes Sociales de las tesis doctorales defendidas en España entre 1976-2008.» *Estudios sobre el Mensaje Periodístico* 17 (2): 417-29. [https://doi.org/10.5209/rev\\_ESMP.2011.v17.n2.38123](https://doi.org/10.5209/rev_ESMP.2011.v17.n2.38123).

Sánchez-Bonvehí, Clàudia, y Mireia Ribera. 2014. «Visualización de la información en la democratización de los datos: propuestas desde el periodismo y la narratividad.» *Profesional de la Información* 23 (3): 311-18. <https://doi.org/10.3145/epi.2014.may.11>.

Sotolongo Aguilar, Gilberto, M<sup>a</sup> Victoria Guzmán Sánchez, y Humberto Carrillo. 2002. «ViBlio-SOM: Visualización de información bibliométrica mediante el mapeo autoorganizado.» *Revista Española de Documentación Científica* 25 (4): 477-84. <https://doi.org/10.3989/redc.2002.v25.i4.281>.

Spinak, Ernesto. 1998. «Indicadores cienciométricos.» *Ciência da Informação* 27 (2): 141-48. <http://revista.ibict.br/ciinf/article/view/795>.

Tague-Sutcliffe, Jean. 1992. «An Introduction to Informetrics.» *Information Processing and Management* 28 (1): 1-3. [https://doi.org/10.1016/0306-4573\(92\)90087-G](https://doi.org/10.1016/0306-4573(92)90087-G).

Tufte, Edward. 2016. «The Visual Display of Quantitative Information (1983)». En *The Diagrammatic Reader: Fundamental Theoretical and Historical Texts*, editado por Birgit Schneider, Christoph Ernst, y Jan Wöpking, 219-30. Berlin: De Gruyter. <https://doi.org/10.1515/9783050093833-035>.

Uribe-Tirado, Alejandro, y Joel Alhuay-Quispe. 2017. «Estudio métrico de ALFIN en Iberoamérica: De la bibliometría a las altmetrics.» *Revista Española de Documentación Científica* 40 (3). <https://doi.org/10.3989/redc.2017.3.1414>.

Valero Sancho, José Luis, Jordi Català Domínguez, y Beatriz-Elena Marín Ochoa. 2014. «An Approach to the Taxonomy of Data Visualization.» *Revista Latina de Comunicación Social* 69: 486-507. <https://doi.org/10.4185/RLCS-2014-1021>.

Vanti, Nadia Aurora Peres. 2002. «Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento». *Ciência da informação* 31 (2): 369-79. <https://doi.org/10.1590/S0100-19652002000200016>.

Verma, Charu, y Pradeep Kumar Suri. 2021. «'Big data' patentometrics for R&D decision-making». *Digital Policy, Regulation and Governance* 23 (4): 317-336. <https://doi.org/10.1108/DPRG-09-2020-0126>.

Vílchez-Román, Carlos, y Joel Alhuay-Quispe. 2016. «Use of text mining for understanding Peruvian students and faculties' perceptions on bibliometrics training». En *CEUR Workshop Proceedings*, 1743:165-69. <http://ceur-ws.org/Vol-1743/paper20.pdf>.

Yatagama Lokuge, y Shamith Wimukthi. 2015. «Extensión del módulo de visualización científica de Weka». Las Villas, Cuba: Universidad Central "Marta Abreu" de Las Villas. <https://dspace.uclv.edu.cu/handle/123456789/6880>.

## Voluntary Transparency in Spanish University Libraries

Ana R. Pacios<sup>(a)</sup>, Margarita Pérez Pulido<sup>(b)</sup>, Marina Vianello<sup>(c)</sup>

a) University Carlos III of Madrid, <https://orcid.org/0000-0003-2887-882X>

b) University of Extremadura, <https://orcid.org/0000-0003-2831-6341>

c) University Carlos III of Madrid, <https://orcid.org/0000-0002-8745-3947>

**Contact:** Ana R. Pacios, [areyes@bib.uc3m.es](mailto:areyes@bib.uc3m.es); Margarita Pérez Pulido, [marperez@unex.es](mailto:marperez@unex.es);  
Marina Vianello, [mvianell@bib.uc3m.es](mailto:mvianell@bib.uc3m.es)

**Received:** 12 August 2021; **Accepted:** 2 November 2021; **First Published:** 15 May 2022

### ABSTRACT

The research reported here aimed to ascertain the degree of transparency exhibited by Spanish public university libraries based on their active public disclosure practice. Transparency was measured by applying the TransPa\_BA tool to the transparency-related information published by the country's 50 public university libraries on their websites. The tool addresses 21 indicators grouped under eight areas, used in this study to measure university libraries' public disclosure performance. The data collected were scored pursuant to the provisions of Spanish Act 19/2013 of 9 December on Transparency, Access to Public Information and Good Governance as adapted to university libraries, as well as in terms of other factors scantily developed in or omitted from the Act. The indicators and their respective parameters (content, form, accessibility, reusability, dating and updating or validity) constitute guidelines enabling libraries to enhance transparency and accountability by reporting their activities and practice in key areas to society in general and their stakeholders in particular. The objective is to help such institutions operate more transparently, for the information afforded by the indicators is deemed relevant to their activity while also monitoring their performance. The findings show that university libraries, which have become more transparent over time, are more transparent than other information units (national public libraries and historic and university archives). Some were nonetheless found to have room for improvement. On the grounds of their scores, these libraries can be classified under three headings: transparent, translucent or opaque. The parameters where good practices were most frequently identified included service usage rules and regulations, user charters and the annual report.

### KEYWORDS

Transparency; Active public disclosure; University libraries; Indicators; Spain.

## 1. Introduction

In the nineteen seventies Harrigan (1971) viewed accountability as a challenge libraries would need to confront to continue to receive public support. Osman and Hanon Bidin (2004) deemed it a key factor for responding to twenty-first century demands and issues. With time the notion has become a constant management concern for libraries aiming to substantiate their relevance to their parent organisations. The ACRL, in turn, identified one of the 2010 Top Ten Trends (286) to be an increase in ‘demands for accountability and assessment’.

Accountability is not new to Spanish university libraries, which have been disseminating information on their management performance online ever since they created their websites in the late nineteen nineties to provide access to their services (Álvarez and Suárez Samaniego 1999). As the web burgeoned, the amount of information uploaded gradually grew, with inclusion becoming standard practice for many libraries (Pacios 2003). Even before then, some were already routinely publishing a (printed) annual report describing their activities, an exercise recommended by some libraries as an aid to decision-making (Carmena Escribano 1993).

Transparency, along with its accessibility and other associated parameters, is directly related to accountability and as such affects the type of information made available to citizens. In other words, it involves more than simply publishing facts and figures. It is also involved in citizen participation, which depends on suitable reporting and ready access.

Spain’s Act 19/2013 of 9 December on Transparency, Access to Public Information and Good Governance (hereafter abbreviated with the Spanish initials, LTAIPBG) distinguishes between the obligation to proactively provide citizens with certain types of data (known as active public disclosure) and the right of access. In the latter citizens must raise a request for unpublished information they wish to receive, which must be honoured subject only to the limits laid down by law.

As university services, academic libraries are not required by law to disclose information on their performance to the community, for that obligation is incumbent upon the university itself (Ch. I, Art. 2, LTAIPBG). Nor must such libraries have a transparency portal where the data are published. Nonetheless, providing such information enhances the esteem in which these institutions are held by the community benefitting from their services and society in general, for it guarantees the transparency of their operations overall. The introduction to the University of Valladolid library’s page containing such information reads: ‘The University of Valladolid Library’s (BUVa) new service offering includes the dissemination of information relevant to its stakeholders and more specifically facts and figures on its key areas of activity. The report provides university users with an efficient tool to acquire a fuller acquaintance with the library and the university itself with data useful for decision-making. At the same time, it makes these data available to outside stakeholders, informing them of the BUVa’s most prominent services in keeping with the library’s commitment to society.’<sup>1</sup>

Such information is essential to know what libraries do, how they do it and the use made of the resources they receive.

Transparency is itself a value adhered to by some university libraries (Pacios and Fernández

---

<sup>1</sup> <https://biblioteca.uva.es/export/sites/biblioteca/1.informaciongeneral/1.5.bibliotecaencifras/index.HTML>

Fernández-Cuesta 2019). The University of Cádiz Library, for instance, apply the notion to their organisation and operation, stating: ‘We manage our library transparently’.

Transparency in Spanish university libraries has been addressed in earlier studies. Pacios (2016) established a series of transparency indicators applied to the *Consortio Madroño*. Those same indicators were later used in a review of the Catalanian consortium (Rey Martín, Rodríguez Parada and Camón 2019) and other Spanish academic institutions (Rey Martín et al. 2020), where a need for improvement was detected. Whilst the existence or otherwise of the indicators was identified in all those surveys, none quantified the parameters the information should include. Aspiring to advance in that direction, the present study deployed a digital tool, TransPa\_BA, to numerically rate library transparency.

## 2. Objectives and methodology

The study described used specifically designed software, TransPa\_Ba (Pacios, Vianello and De La Mano 2020), to rate the transparency performance of Spanish university libraries. Inspired by MESTA (Spanish initials for ‘transparency assessment and monitoring methodology’), a tool developed by Spain’s Transparency and Good Governance Council, TransPa\_BA is built around 21 public disclosure indicators associated with LTAIPBG (España 2013) requirements as adapted to university libraries. It also addresses matters scantily mentioned in or omitted from the Act, such as user charters. The aim was to help such institutions operate more transparently, for the information afforded by the indicators is deemed to be relevant to their activity while also monitoring their performance. A second aim was to test the tool to detect possible inconsistencies in the indicators or their parameters.

The universe defined comprised the libraries affiliated with Spain’s 50 public universities. Justification for that selection lay in their status as publicly funded institutions, a fact also particularly relevant to transparency. The information on transparency in connection with public disclosure was drawn from libraries’ websites as accessed from the REBIUN<sup>2</sup> network portal. In the event of broken links the library’s page was sought on the respective university’s website. The URLs used to find the information on which assessment was based are listed in the annex, where library names are abbreviated as in the present text (IUNE 2018).

Information was analysed on the grounds of the 21 indicators grouped into the eight areas defined by TransPa\_BA. Library websites were searched for the indicators in May and June 2021. Each indicator was assigned a specific weight in a total score of 100, calculated from the mean value found for each indicator in a survey conducted among head librarians (response rate=50 %; n=50). Between December 2019 and February 2020 respondents were asked to specify the importance they attached to each category and the suitability of each indicator on a scale of 0 to 5. The 21 indicators and respective weighting factors are grouped under the eight category headings in Table I.

---

<sup>2</sup> <https://www.rebiun.org/quienes-somos/bibliotecas>.

UNIVERSITY LIBRARY				
Presentation of active public disclosure on the library website				0.01%
	ACTIVE PUBLIC DISCLOSURE INDICATOR	INDICATORS	AREA	
<b>1</b>	<b>PURPOSE OF THE SERVICE AND OBJECTIVES PURSUED</b>			
1.1	Definition of mission	6.35%		
1.2	Strategic plan	6.35%		
TOTAL AREA 1			12.71%	
<b>2</b>	<b>GOVERNING BODIES AND OPERATING RULES</b>			
2.1	Identity of library's management board members	2.77%		
2.2	Regulations	3.40%		
2.3	Specific regulations on service provision	3.43%		
2.4	Code of ethics, values or good practice	2.83%		
TOTAL AREA 2			12.44%	
<b>3</b>	<b>SERVICE OFFERING</b>			
3.1	User charter	13.77%		
TOTAL AREA 3			13.77%	
<b>4</b>	<b>THE COLLECTION</b>			
4.1	Collection management policy or programme	6.03%		
4.2	Institutional open access policy	6.80%		
TOTAL AREA 4			12.83%	
<b>5</b>	<b>STAFF</b>			
5.1	Organisational chart	5.77%		
5.2	Staff directory	6.37%		
TOTAL AREA 5			12.14%	
<b>6</b>	<b>RESULTS</b>			
6.1	Management indicators (scoreboard)	2.20%		
6.2	User satisfaction surveys	2.52%		
6.3	Annual report or report of activities	2.63%		
6.4	Distinctions, prizes, certifications	2.36%		
6.5	Statistics	2.43%		
TOTAL AREA 6			12.14%	
<b>7</b>	<b>FINANCIAL INFORMATION</b>			
7.1	Budget implemented	3.95%		
7.2	Tenders, contracts and bidding	3.86%		
7.3	Subsidies and assistance awarded	3.99%		
TOTAL AREA 7			11.81%	
<b>8</b>	<b>PARTNERING / COOPERATION</b>			
8.1	Partnering networks, task forces, commissions	6.07%		
8.2	Agreements	6.07%		
TOTAL AREA 8			12.14%	
TOTAL AREAS			100.00%	99.99%
TOTAL ACTIVE PUBLIC DISCLOSURE				100.00%

Table I. Weights assigned in TransPa\_BA to public disclosure indicators for university libraries.

Six parameters were analysed to assess transparency-related information: content, form, reusability, accessibility, dating and updating/validity. The scores for each ranged from 0 to 10 except in the last two mentioned, where the scale ran from 0 to 5. The highest score for a given indicator

was consequently 50 points. A seventh variable analysed, 'location', denoted the position on the site where the information was found and on which visibility was largely deemed to depend. Ten points were awarded when the site contained a specific section on transparency. The parameters defined to assess the indicators are described briefly below.

**Content.** This parameter, associated with the document per se, was understood to consist in 'all information regardless of medium (physical or digital) and format (graphic, audio or video)', therefore included the most disaggregated or "rawest" data' (Act 18/2015 of 9 July on reuse of public sector information) (España, 2015). Although no minimum requirements on content were defined for the indicators, in some cases their presence or absence was determined on the grounds of the existence of specific elements:

- For the definition of mission, that specific term, rather than library objective, purpose or function.
- For regulations, the presence of the ones for the most basic service, namely lending or library usage.
- For the graphic description of unit or service hierarchy, the respective organigram.
- For the staff directory, at least one person's name and contact information.
- For the operational indicators, uninterrupted presence with access to results (via vehicles such as service-focused scorecards indicative of library commitments and monitoring-confirmed routine use).
- For partnering/cooperation with networks, task forces or commissions, logos or icons proving its association with the respective entities.

Some indicators, such as budget, were sought and found in others such as the annual reports where the information provided was particularly comprehensive.

**Form.** This parameter refers to the position of the information and the pathway to reach it. It was deemed direct when the information or a link to the content was found on the home page. An example of the latter would be the inclusion of a direct link to the official site containing the legislation or regulations cited, saving users any further searching. Links were deemed to be direct for these purposes even if not denominated with the exact wording specified here if the labels were sufficiently descriptive of the content. Examples would be the words 'staff' or 'team' in links to the staff directory. Links were regarded as indirect when they opened the home page of the respective website that had then to then be searched to locate the information. Here an example would be links to REBIUN statistics cited by some libraries, where users must enter the name of the library at issue to retrieve the data. Another would be link labels that do not refer explicitly to their content, as when staff directories are buried in links to a library's various locations. Where information or documents were accessible via more than one pathway, the library was scored on the grounds of the one most favourable to its final assessment.

**Reusability.** Only two scores were defined, 0 or 10, with no intermediate values dependent upon ease of reuse, given the wide range of constantly evolving opinions and weighting scales put forward in that regard since the Berners-Lee (2006) proposal was first published. A score of 10 was awarded for structured formats (XLS, CSV, XML) whose use requires no extra techniques or



effort. Acrobat PDF was not deemed reusable in light of the criticism levelled at the format in connection with transparency (García-Melián 2014), despite its extensive use on the Datos.gob.es site, the platform that organises and handles Spain's national catalogue of open data. Although PDF was deemed a valid format for public disclosure, its inconvenient editing constitutes an obstacle to reuse. As Camacho (2016) noted, 'PDFs should be furnished in conjunction with reusable formats: XML, CSV, XLS, DOC...?'

This parameter might well be thought to be applicable only to the indicators adopting the form of quantitative data, such as statistics. Given, however, the reuse that might be made of any of a library's documents and bearing in mind that some (strategic plans, citizen charters, reports, policy statements) are the object of analysis, reusability was analysed in all.

**Accessibility.** This parameter was measured as the number of clicks required to reach information on the indicator. The values ranged from 10 (for three clicks or fewer) to 0 (for more than 12) and proportionally in between.

**Dating.** This parameter assumes that all information and documents published should be dated to enable citizens to position them in time. Where a date was provided the score awarded was 5 and 0 otherwise.

**Updating/validity.** Although the cut-off year defined for valid updates was 2019, this parameter was not deemed applicable to all the available information, for not all the indicators proposed, such as user charters or strategic plans, need to be updated yearly. In such cases the expiration date of the document was the grounds for verifying whether it was in effect. A score of 10 was awarded where the information was up-to-date and 0 otherwise. Some types of information, regulations for instance, do not require yearly updates. As that is something the present authors could not determine, however, the year 2019 was adopted for all indicators bearing no specific expiration date. If the information was updated in 2019 or later, the score was 10 and 0 otherwise.

TransPa\_BA is a tool designed for use by library and archive managers seeking to improve their institutions' transparency. The respective software will be available to them on [https://www.uc3m.es/investigacion/transpa\\_ba](https://www.uc3m.es/investigacion/transpa_ba). Although on this occasion sites were assessed by the authors, i.e., third parties, such assessments should ideally be conducted in-house, for only then can the accuracy of the valuation of certain parameters, such as validity, be ensured. Library staffers and they alone know whether a given type of information must be amended and consequently whether it merits the maximum score even though it may not have expired. Third parties, unaware of such circumstances, might well underestimate a library's transparency rating.

### 3. Results

The fieldwork itself, consisting in an exhaustive review of each and every one of the 50 libraries' websites to locate the aforementioned indicators and subsequent entry of the data gathered with TransPa\_BA, revealed searching to be an arduous task, for transparency information and data were scattered across scores of links.

The results for the parameters addressed are discussed below in the following order: location, content, form, reusability, accessibility, dating and updating/validity.

### 3.1. Location

Scattered information was characteristic of most of the libraries. Only one, UNIZAR, grouped all the indicators on a single site devoted to transparency. The UPCT library site's link labelled 'Quality and Transparency' carried some of the indicators. Since the UNED library had dispensed with its former transparency site (Rey Martín et al. 2020, 243), some documents were more difficult to locate. In all cases with the sole exception of UNIZAR, then, indicators and parameters had to be located starting from each library's home page. That in turn meant transparency-related information had to be sought by opening a succession of links, not a particularly user-friendly procedure. Normally all indicators should be listed under a single tab, similarly termed in all institutions. The tabs most frequently used for the present purposes included 'About us', 'About the library' or 'The Library'. Transparency indicators were normally grouped under other names such as Regulations, Strategy and Quality, Quality Management, Library Documents, Planning and Quality, Management and Organisation, Quality, Publications, Documents on Quality or Planning. Others bore scanty explicit descriptions of the content. The heading 'Participate' found on the UCM library site included surveys, regulations, annual reports, indicators and statistics. The UJAEN library's site carried similar information under 'Communication/Participation'. As a rule, the labels denoted the relationship between transparency and quality. Further to LTAIPBG, Arts 6, 7 and 8, all that information could be suitably grouped under a link or tab titled 'Transparency', irrespective of whether the documents at issue can be accessed from other pages where deemed appropriate. Some libraries, the UAH's and the UJAEN's among others, while lacking a page specifically devoted to transparency, clustered many of the respective links under a single heading, a practice favouring the visibility as well as the accessibility of the information.

One prominent finding that merits comment with a view to possible improvements was that links to some of the indicators were so discreetly placed (either at the bottom of the library's page, against a dark background or in a very small font size) that they might go practically unnoticed. Institutional open access policy was a type of information especially difficult to find in some libraries, for two reasons. Either the link bore the repository rather than the university name or it was displayed outside the suite of documents routinely grouped under the heading 'Regulations'. Some such policies were explained in the section carrying library guides or in the area reserved for the support afforded researchers. In others (UA, UCO and UGR), although the policy existed, it was not published on the library's site. The UA had no link labelled 'Policies' on the repository page. The UCO had a link titled 'Helvia Policies', alluding to the name of the repository, but none addressed open access, which was found rather on the RECOLECTA<sup>3</sup> site. The UGR library, like UPF's, carried no link to the policy on its own site, although one appeared when the repository page was searched. In 2019 a total of 32 libraries had policies in writing (Altuna Esteibar 2019, 1224), whereas the searches conducted for this study identified 29.

---

<sup>3</sup> <https://www.recolecta.fecyt.es/sites/default/files/contenido/PoliticaAccesoAbiertoUCO.PDF>The

### 3.2. Content

A total of 604 documents or information items associated with the 21 transparency indicators were found for the university libraries taken as a whole. All the indicators were present in at least one library. The ones most frequently found were definition of mission, specific regulations on service provision and partnering networks, task forces and commissions in which the libraries participated. Data on tenders, contracts and bidding were only provided by UNIZAR, and there only indirectly, in the form of a link to the university's portal on the matter, where tenders can be located by tendering unit. Information on the transparency of collection management was observed on only 13 of the 50 sites analysed (26 %). The UNICAN's and UdG's sites were identified as examples of good practice in that regard.

The frequency with which the 21 indicators were observed is listed in Table II.

INDICATOR	TOTAL (n = 604)	%
1.1 Definition of mission	46	92.00%
2.3 Specific regulations on service provision	46	92.00%
8.1 Partnering networks, task forces, commissions	46	92.00%
2.2 Regulations	45	90.00%
5.2 Staff directory	43	86.00%
3.1 User charter	40	80.00%
6.5 Statistics	39	78.00%
6.3 Annual report or report of activities	35	70.00%
7.1 Budget implemented	32	64.00%
4.2 Institutional open access policy	29	58.00%
6.2 User satisfaction surveys	28	56.00%
2.4 Code of ethics, values or good practice	27	54.00%
6.1 Management indicators (scoreboard)	25	50.00%
1.2 Strategic plan	24	48.00%
6.4 Distinctions, prizes, certifications	23	46.00%
5.1 Organisational chart	23	46.00%
8.2 Agreements	21	42.00%
4.1 Collection management policy or programme	13	26.00%
2.1 Identity of library's management board members	11	22.00%
7.3 Subsidies and assistance awarded	7	14.00%
7.2 Tenders, contracts and bidding	1	2.00%

Table II. Indicators located

Further to the data in the table, the areas best represented were 1 (Purpose of the service and objectives pursued) and 2 (Governing bodies and operating rules). The indicators least often addressed were 8 (Partnering/Cooperation) and 7 (Financial information), which when present con-

sisted in mere brief descriptions of both the subsidies / assistance awarded and tenders, contracts and bidding.

Although no minimum content was defined for the indicators, wide inter-library differences were observed. One characteristic example was found in the type and form of the statistics provided: in tables with many types of data referring to different indicators (UCM); as a structured selection of the most indicative figures (UAB, UB); in an annex to the annual report (UAM); in statistical yearbooks (UPM); as infographics (UA, UNAVARRA); or as links redirecting users to REBIUN statistics (UPO, UC3M and others). Other libraries opted to use several of these alternatives. Such differences prompted the authors to consider the possibility of establishing quality criteria in a future version of TransPa\_BA that would assign greater weight to most comprehensive indicators, or minimum content for an indicator to score. The strategic plan and annual report, which contain much of the information related to other indicators, would be cases in point. At the same time, however, enormous differences were observed in the content of those documents.

A comparison of the present findings on the existence of indicators to the results reported in the sole other paper identifying practically the same items, conducted in 2018 (Rey Martín et al. 2020), revealed a slight improvement in some. The indicators located more frequently now than then included: the definition of mission (35 to 46), collection management policy (9 to 13), staff directory (38 to 43), annual report (24 to 35), partnering/cooperation (37 to 46) and budget (24 to 32). That would attest to the beneficial effect of such studies, in a way analogous to the successive reports on university transparency conducted beginning in 2011 (Barrio and Cavanaugh 2012), which prompted yearly progress and a concomitant rise in the number of institutions awarded transparent status.

### 3.3. Form

This parameter was the most difficult to value since, with the aforementioned exception of UNIZAR, the academic libraries analysed had no transparency pages from which to launch assessment. It was therefore decided to score the presence of a specific link to each indicator as direct form. Although that approach deviated from the criteria applied by MESTA (AEVAL, CTBG 2016, 24), the method on which TransPa\_BA drew for inspiration, it was adopted to distinguish the indicators accessed via a clearly visible link or section from those that had to be found by searching another page or document. The latter typically comprised indicators indirectly accessible by searching annual reports (agreements, subsidies, statistics, organisational charts and so on). In light of the findings, UNIZAR was awarded the highest score for it had a transparency portal of its own that included links to all the documents identified under each indicator.

A total of 436 could be accessed directly, defined to mean via a link that carried the user to the content sought. The remaining 168 indicators were indirectly accessible, i.e., called for further searching on a website containing the information or a document with which the user needed to be acquainted to be able to locate the indicator sought. By way of example: a user aware of the content of user charters and seeking the definition of mission would know it should be found there. Such a procedure can hardly be deemed transparent.

### 3.4. Reusability

Libraries have not taken this parameter into account to provide access to this type of information. Although the possible use of such documents is unknown, the LTAIPBG (Ch. II, Art. 5.4) reads: ‘the information subject to transparency obligations shall be published on the respective electronic portals or websites in a manner that is clear, structured and understandable for those concerned, and preferably in reusable formats’.

The most common formats found for indicators, by order of frequency of use, were: PDF (367) and HTML (208) and on rare occasions (just 11), both. Many of the pages on the UCA site could be converted to PDF. Neither format was deemed reusable, however, in light of the extra effort or knowledge of other techniques involved, as noted earlier.

Exceptionally, some indicators were found in reusable formats: CSV (3), XLS (7), DOC (8). The UVA’s statistics, on a page similar to REBIUN’s where data and indicators can be searched, could be downloaded in CSV format. The UCM’s site allowed the reuse of data and indicators furnished in CSV. Libraries such as UC3M, UPO and UCA enabled access to statistics via a link to the REBIUN site where data on the library could be retrieved as XLS documents, albeit indirectly.

### 3.5. Accessibility

Although generally speaking indicators were readily accessible, with most (451) found in three clicks or less, in some cases it took six to nine to reach the information sought (107 required 9 clicks; 33, eight; 4, seven; and another 4, six). Failure to group all the information in the same place, such as a transparency page, translated into a larger number of steps needed to find documents, detracting from accessibility. The same effect, i.e., scoring lower in this parameter, was observed when documents such as the annual report were located in the institution’s repository and consequently more clicks away from the library’s home page.

In some cases (UA, UCA, UM) indicators existed but were not accessible to visitors. Clicking on the respective links carried users to the library’s intranet where the information was only accessible to authorised users.

### 3.6. Dating

In all, 84.60 % of the indicators located were dated. Where (although exceptionally) documents such as library regulations were found with no date, even the date of approval by the competent authority, their validity could not be determined.

Some user charters were also observed to be un- or outdated, therefore constituting mere lists of services offered and commitments to users. That is not compliant with the legislation (Royal Decree 951/2005; España 2005) laying down the requirements to be met by such documents, which include their periodic updating to enhance transparency.

### 3.7. Updating/Validity

Updates are due whenever changes are made, which is more or less continually. Surprisingly, some documents were found to have gone unamended for 20 years. That is the more amazing bearing in mind the vast change affecting Spanish university libraries that would entail document updating. Exemplary dating and continuous updating were found in the UAB's specific regulations on service provision, which begin with a list of the versions issued from 1996 to 2019, including the dates of the revisions and of their approval by the competent authority. The UAB adopted the same approach to its overall regulations. The UNAVARRA's user charter constituted another example in which the dates of all the respective revisions and updates were listed.

Based on the year, 2019, defined for this parameter to be deemed valid, 52.48 % of the documents were up-to-date, inasmuch as not all would need to be updated in the aforementioned year, as explained earlier.

Strategic plans and service charters would typically require updating. Of the 24 strategic plans located, 14 were observed to be valid, whilst 18 of the 40 user charters had not been updated. Analogous values were found for seals and certificates of quality: of the 23 certificates, seals or acknowledgements published online, 10 had not been renewed. Those findings attest to the relaxation of good practice in some cases.

### 3.8. Public disclosure transparency index

Libraries seeking greater transparency can use the TransPa\_BA tool to rate their public disclosure performance on a 100-point scale and rank the institution from the standpoint of its transparency, based on the presence of the 21 indicators on its website and compliance in terms of the respective parameters. The assessment conducted here resulted in the list given in Table III. The values shown must be deemed approximate only, however, insofar as the present authors were not in an ideal position to evaluate parameters such as form (for UNIZAR was the sole library with a page or site specifically addressing transparency) or updating/validity (the year defined need not necessarily be applicable to some documents if no update was due in 2019, which is something people outside the organisation cannot know).

UNIVERSITY	1.1 Definition of mission	1.2 Strategic plan	2.1 Identity management board members	2.2 Regulations	2.3 Specific regulations on service provision	2.4 Code of ethics, values or good practice	3.1 User charter	4.1 Collection management policy or programme	4.2 Institutional open access policy	5.1 Organisational chart	5.2 Staff directory	6.1 Management indicators (scoreboard)	6.2 User satisfaction surveys	6.3 Annual report or report of activities	6.4 Distinctions, prizes, certifications	6.5 Statistics	7.1 Budget implemented	7.2 Tenders, contracts and bidding	7.3 Subsidies and assistance awarded	8.1 Partnering networks, task forces, commissions	8.2 Agreements	Total active public disclosure
UNIZAR	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	68.97
UB	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	66.40
US	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	64.69
UNICAN	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	64.62
UAB	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	62.22
UVA	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	61.05
UdG	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	59.37
UAH	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	57.43
UGR	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	56.60
UdL	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	55.56
ULPGC	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	55.51
UNEX	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	53.61
UNIRIOJA	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	52.94
UPC	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	52.34
UIB	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	52.12
UV	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	50.70
UMA	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	48.56
UA	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	48.18
UM	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	47.24
UCLM	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	44.88
UNED	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	44.81
UHU	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	44.17
UNAVARRA	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	43.78
UJAEN	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	43.41
UAL	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	41.86
UC3M	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	41.09
ULL	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	41.01
UCM	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	40.13
USAL	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	38.95
UJI	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	37.81
UBU	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	37.65
URV	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	37.12
UNIOVI	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	34.28
UPM	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	34.19
UPF	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	34.02
UPCT	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	33.87
UCA	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	29.92
UPO	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	29.00
UAM	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	27.15
UCO	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	25.54
EHU	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	25.52
USC	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	25.28
UPV	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	24.66
UNILEÓN	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	21.48
UMH	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	19.73
URJC	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	17.68
UVIGO	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	17.44
UDC	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	16.05
UNIA	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	8.74
UIMP	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	0.00

Table III. Transparency of public disclosure in Spanish public universities

In all 30 libraries (60 %) had more than 11 of the 21 indicators defined, whilst just 16 (32 %) had scores of over 50, i.e., a passing mark in transparency. The inference is that merely ensuring the availability of information on the indicators (content) does not suffice. Rather, other factors must be considered, such as the form, accessibility, reusability, dating and validity of the information to comply with the MESTA transparency parameters assessed by TransPa\_BA.

A library's score affords institutions an idea of the route they should follow to be more transparent, which would in some cases entail enhancing their practice, in others call for substantial improvement and in yet a third group mean embarking on that route as a new venture.

The libraries with the highest transparency rating were UNIZAR, UB, US, UNICAN, UAB and UVA, all of which had scores of over 60. They were followed in the 50-point range by: UdG, UAH, UGR, UdL, ULPGC, UNEX, UNIRIOJA, UPC, UIB and UV.

Further to the consolidated methodology deployed in reviews of university transparency conducted by institutions such as the Fundación Compromiso y Transparencia (Barrio, Cavana and Martínez 2019), libraries could be grouped here on the grounds of their transparency rating as transparent, translucent or opaque. The first group would comprise those scoring 50 points or higher (websites with 12 to 21 indicators), the second those with >30 and <50 points (9 to 16 indicators) and the third those with <30 (0 to 11 indicators). Although some of the translucent institutions exhibited the same or a larger number of indicators than some of the ones deemed transparent, the parameters identified for the former lay below the threshold set for the respective indicator to score. That reinforces the importance of taking into account not only the presence of indicators but the values of the associated parameters as well.

## 4. Good practice

The TransPa\_BA-mediated search and analysis of indicators revealed a number of good practices that might be deemed benchmarks for libraries seeking to institute transparency. Given the limitations to the length of this paper just a few of the most relevant were chosen as examples.

- Very comprehensive and up-to date **annual reports**, one of the most complete indicators given the information they contain on library results (Pacios and Serna 2020), were located for UAB, UAH, UNIZAR, UPC, US, UVA.
- Good examples of **strategic plans**, documents that describe the library's projected course and priorities in a given period of time, were found for UAB, UAH, UB, UdG, UGR, UNED, UNEX, UNIZAR, UPC, US, UVA and UV.
- **Staff directories** aim to provide users with a tool to identify and contact the managers responsible for library services. Close contact and transparency are guaranteed if personal data are given to identify the staff associated with the organisation, operation or public activity engaged in. As Blanes (2019) contended, 'personal data protection legislation does not rule out disclosure where required to exercise the right of access to public information', although such protective legislation may exceptionally prevail under the circumstances legitimised by the LTAIPBG, Art. 15.2 (España 2013). Staff directories containing staffers' professional status, speciality and photographs in addition to the contact information, such as in UNICAN and UPC, are especially user-friendly.



- **Comprehensive scorecards** are very useful tools for libraries for they ‘can be used to disseminate the strategy defined, enabling both library staff and the community of users served by the institution to visualise its constituent items’ (Malo de Molina 2013, 53), all to the benefit of good library governance (De la Mano and Creaser 2014). Only a few examples were found, however, specifically for UJAEN, UNIZAR and UPF (in keeping with UPF culture the library no longer uses that denomination, but continues to use the tool, having been the Spanish university library that pioneered its implementation).

## 5. Conclusions and recommendations

The overview afforded by the public disclosure rating for university libraries aims to provide them with insight into their transparency status and on those grounds decide whether to enhance it by furnishing their stakeholders with data they lack or enable them to access such data more readily and establish the parameters they should comprise.

The findings on the transparency indicators used by university libraries proved most to be committed to the principle, even though the obligation is incumbent on their parent institutions. As noted, that attitude is not new, but rather has been in place since libraries began to use their websites to show what they do and how and to ratify their commitment to the university to which they are accountable. Even taking into account the differences in types of institutions, academic libraries engage more deeply in highlighting their transparency than other information units such as public libraries (Pacios, García López and Morales García 2021) or historic archives (Pacios, Martínez Cárdoma and Moro Cavero 2021).

The upward pattern found for some of the indicators over time would denote the favourable effect of analyses such as this, that advocate for more transparent institutions and services. The good practices identified in connection with many of the indicators attest further to the progress underway.

One recommendation institutions might bear in mind when disclosing transparency-related information is to also provide the necessary data on the associated parameters. Merely ensuring the availability of information on indicators (content), as many libraries tend to do, does not suffice. Rather, their constituent parameters (form, accessibility, reusability, dating and updating/validity) must be taken into account to comply with transparency standards. When a library’s website is searched for indicators, the access pathway should also be assessed and the items possibly relocated for grouping under a single link which should in all propriety be labelled ‘Transparency’.

Some libraries have earnestly instituted transparency measures whilst others have ample room for improvement in that regard. Providing effective and transparent services is a challenge in keeping with Agenda 2030 sustainable development goal 16 (IFLA 2020) to which many organisations such as universities aspire and to whose fulfilment libraries can contribute.

## Acknowledgements

This study was funded by the Spanish Ministry of Science, Innovation and Universities under project RTI2018-095187-B-100.

## References

- Agencia de Evaluación de las Políticas Públicas y de la Calidad de los Servicios (AEVAL) and Consejo de Transparencia y Buen Gobierno (CTBG). 2016. *Metodología de Evaluación y Seguimiento de la Transparencia de la actividad Pública (MESTA)*. Madrid: Ministerio de Hacienda y Función Pública, accessed June 5, 2021, <http://www.aeval.es/export/sites/aeval/comun/PDF/evaluaciones/E42-2016.PDF>.
- ACRL Research Planning and Review Committee. 2010. "2010 top ten trends in academic libraries: A review of the current literature". *College & Research Libraries*, 71(6): 286-292, accessed June 7, 2021, <https://crln.acrl.org/index.php/crlnews/article/view/8385/8558>.
- Álvarez Álvarez, María Antonia and Suárez Samaniego, Marta. 1999. "Las páginas web de las bibliotecas universitarias andaluzas". In *Les biblioteques i els centres de documentació al segle XXI: peça clau de la societat de la informació*, 289-304. Barcelona: Col·legi Oficial de Bibliotecaris-Documentalistes de Catalunya, accessed July 6, 2021, <http://www.cobdc.org/jornades/7JCD/5.PDF>.
- Altuna Esteibar, Belén. 2019. "La Biblioteca de la Universidad Pública de Navarra, hoy". *Príncipe de Viana*, 275: 1213-1228.
- Barrio, Esther and Cavanna, Javier Martin. 2012. *Examen de transparencia. Informe de transparencia voluntaria en la web de las universidades españolas*. Madrid: Fundación Compromiso y Transparencia, accessed June 10, 2021, <https://www.compromisoempresarial.com/wp-content/uploads/Universidades2012.PDF>.
- Barrio, Esther; Javier Martín Cavanna and Pablo Martínez. 2019. *Examen de transparencia 2018. Informe de transparencia voluntaria en la web de las universidades españolas*. Madrid: Fundación Compromiso y Transparencia, accessed July 12, 2021, <https://www.compromisoytransparencia.com/informes/examen-de-transparencia-2018-informe-de-transparencia-voluntaria-en-la-web-de-las-universidades-espanolas>.
- Berners Lee, T. 2006. *Linked Data*, accessed June 25, 2021, <https://www.w3.org/DesignIssues/LinkedData.HTML>.
- Blanes, Miguel Ángel. 2019. "El nombre y apellidos de los empleados públicos solo puede ser restringido en casos excepcionales". *Blog de Transparencia y Gobierno Abierto*, accessed June 15, 2021, <https://miguelangelblanes.com/2019/10/24/el-nombre-y-apellidos-de-los-empleados-publicos-solo-puede-ser-restringido-en-casos-excepcionales/>.
- Camacho, Rafa. 2016. *Del derecho a saber al derecho a conocer. Supervillanos del acceso a la información pública: PDF-man*, accessed July 12, 2021, <https://mymabogados.com/PDF-transparencia>.
- Carmena Escribano, Miguel Ángel. 1993. "Tomar decisiones en una biblioteca. Algunas sugerencias sobre las posibles pautas a seguir en la elaboración de memorias – informes". *Boletín de la Asociación Andaluza de Bibliotecarios*, 9(33): 25-40.
- De La Mano, Marta and Claire Creaser. 2016. "The impact of the Balanced Scorecard in libraries: From performance measurement to strategic management". *Journal of Librarianship and Information Science*, 48(2): 191-208. Doi:10.1177/0961000614558078

España. 2013. *Ley 19/2013, de 9 de diciembre, de Transparencia, Acceso a la Información Pública y Buen Gobierno*. BOE. 295 (10-12-2013) 97922-97952, accessed June 14, 2021 <https://www.boe.es/buscar/PDF/2013/BOE-A-2013-12887-consolidado.PDF>.

España. 2015. *Ley 18/2015, de 9 de julio, por la que se modifica la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público*. BOE. 164 (10-06-2015) 57436-57450, accessed July 1, 2021, <https://www.boe.es/eli/es/l/2015/07/09/18>.

España. 2005. *Real Decreto 951/2005, de 29 de julio, por el que se establece el marco general para la mejora de la calidad en la Administración General del Estado*. BOE. 211 (11-09-2005) 30204-30211, accessed June 11, 2021, [https://www.boe.es/diario\\_boe/txt.php?id=BOE-A-2005-14836](https://www.boe.es/diario_boe/txt.php?id=BOE-A-2005-14836).

García Melián, Juan Carlos. 2014. “La transparencia y el PDF”. *Acreditra*, accessed April 12, 2021, <https://acreditra.com/2014/10/19/la-transparencia-y-el-PDF/>.

Harrigan, Joan. 1971. *Library accountability*. Paper presented to the faculty of the Graduate School of Librarianship. Denver: University of Denver, accessed August 7, 2021, <https://files.eric.ed.gov/fulltext/ED060851.PDF>.

International Federation of Library Associations and Institutions (IFLA). 2020. *Libraries, Development and the United Nations 2030 Agenda*, accessed June 12, 2021, <https://www.ifla.org/libraries-development>.

IUNE. 2018. *List of acronyms of Spanish universities*, accessed June 12, 2021, [http://www.iune.es/es\\_ES/glosario/listado-universidades](http://www.iune.es/es_ES/glosario/listado-universidades).

Malo de Molina, Teresa. 2013. “El Cuadro de Mando Integral, una herramienta para visualizar la estrategia en la biblioteca”. En *XII Jornadas Españolas de Documentación FESABID13* (24-25 de mayo de 2013), accessed June 23, 2021, <http://fesabid.org/documentos/fesabid13-actas.PDF>.

Osman, Zaiton and Sharipah Hanon Bidin. 2004. “Accountability In Library Management: Issues & Strategies For The 21st Century”. *Kekal Abadi*, 23(1): 1-14, accessed June 4, 2021, <http://rizal.lib.admu.edu.ph/rlconflibmgt/PDF/zaiton.PDF>.

Pacios, Ana R. 2003. “Management-related information on Spanish university library Web pages”. *The Electronic Library*, 21(6):528-537. Doi: 10.1108/02640470310509081

Pacios, Ana R. 2016. “Universidades transparentes con bibliotecas transparentes”. *Revista de Investigación Bibliotecológica*, 70 (4):105-128.

Pacios, Ana R. and Paz Fernández Fernández-Cuesta. 2019. “Publicación de los valores de las bibliotecas universitarias españolas a través de sus sedes web”. *Profesional de la Información*, 28(5), accessed July 8, 2021, <https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/77379>.

Pacios, Ana R. and Nacho Serna. 2020. “The annual report, indicator of library transparency”. *International Information & Library Review*, 52 (4): 253-262. Doi: 10.1080/10572317.2020.1728176.

Pacios, Ana R.; Marina Vianello and Marta de la Mano. 2020. “TransPa\_BA, a tool for improving active disclosure of library and archive information”. *Ibersid*, 14(2): 57-66.

Pacios, Ana R.; Fátima García López, F. and Ana M. Morales García. 2021. "Assessment of transparency in national public libraries". *Métodos de Información*, 12(22): 58-85, accessed November 9, 2021, <https://www.metodosdeinformacion.es/mei/index.php/mei/article/view/917/1026>.

Pacios, Ana R.; Sara Martínez Cárdoma and Manoli Moro Cabero. 2021. "Evaluación de la transparencia en los archivos históricos provinciales". *Anales de Documentación*, 24(2). Doi: 10.6018/analesdoc.479881.

Rey Martín, Carina; Concepción Rodríguez Parada and Luis Enric Camón. 2019. "The transparency of CSUC member university libraries". *Library Management*, 40(8/9): 558-569. Doi: 10.1108/LM-02-2018-0008.

Rey Martín, Carina; Concepción Rodríguez Parada; Ana R. Pacios; Blanca Rodríguez Bravo and Marina Vianello. 2020. "Transparency in the management of spanish public university libraries through the internet". *Libri*, 70(3). Doi: 10.1515/libri-2019-0051.

## Annex. URLs for the university library websites assessed

Acronym	University	URL web
EHU	Universidad del País Vasco	<a href="https://www.chu.eus/es/web/biblioteca">https://www.chu.eus/es/web/biblioteca</a>
UA	Universidad de Alicante	<a href="https://biblioteca.ua.es/">https://biblioteca.ua.es/</a>
UAB	Universitat Autònoma de Barcelona	<a href="https://www.uab.cat/biblioteques/">https://www.uab.cat/biblioteques/</a>
UAH	Universidad de Alcalá de Henares	<a href="https://biblioteca.uah.es/">https://biblioteca.uah.es/</a>
UAL	Universidad de Almería	<a href="http://cms.ual.es/ual/universidad/serviciosgenerales/biblioteca/index.htm">http://cms.ual.es/ual/universidad/serviciosgenerales/biblioteca/index.htm</a>
UAM	Universidad Autónoma de Madrid	<a href="https://www.uam.es/uam/vida-uam/bibliotecas">https://www.uam.es/uam/vida-uam/bibliotecas</a>
UB	Universitat de Barcelona	<a href="https://crai.ub.edu/es">https://crai.ub.edu/es</a>
UBU	Universidad de Burgos	<a href="https://www.ubu.es/biblioteca">https://www.ubu.es/biblioteca</a>
UC3M	Universidad Carlos III de Madrid	<a href="https://www.uc3m.es/biblioteca/inicio">https://www.uc3m.es/biblioteca/inicio</a>
UCA	Universidad de Cádiz	<a href="https://biblioteca.uca.es/">https://biblioteca.uca.es/</a>
UCLM	Universidad de Castilla-La Mancha	<a href="https://www.biblioteca.uclm.es/">https://www.biblioteca.uclm.es/</a>
UCM	Universidad Complutense de Madrid	<a href="https://biblioteca.ucm.es/">https://biblioteca.ucm.es/</a>
UCO	Universidad de Córdoba	<a href="http://www.uco.es/servicios/biblioteca/">http://www.uco.es/servicios/biblioteca/</a>
UDC	Universidad de A Coruña	<a href="https://www.udc.es/es/biblioteca/?language=es">https://www.udc.es/es/biblioteca/?language=es</a>
UdG	Universitat de Girona	<a href="https://biblioteca.udg.edu/es">https://biblioteca.udg.edu/es</a>
UdL	Universitat de Lleida	<a href="http://bid.udl.cat/ca/">http://bid.udl.cat/ca/</a>
UGR	Universidad de Granada	<a href="https://biblioteca.ugr.es/">https://biblioteca.ugr.es/</a>
UHU	Universidad de Huelva	<a href="http://www.uhu.es/biblioteca/">http://www.uhu.es/biblioteca/</a>
UIB	Universidad de las Illes Balears	<a href="https://biblioteca.uib.cat/">https://biblioteca.uib.cat/</a>
UIMP	Universidad Internacional Menéndez Pelayo	<a href="http://biblioteca.uimp.es/">http://biblioteca.uimp.es/</a>
UJAEN	Universidad de Jaén	<a href="https://www.ujaen.es/servicios/biblio/">https://www.ujaen.es/servicios/biblio/</a>
UJI	Universitat Jaume I de Castelló	<a href="https://www.uji.es/serveis/cd/">https://www.uji.es/serveis/cd/</a>
ULL	Universidad de La Laguna	<a href="https://www.ull.es/servicios/biblioteca/">https://www.ull.es/servicios/biblioteca/</a>
ULPGC	Universidad de Las Palmas de Gran Canaria	<a href="https://biblioteca.ulpgc.es/">https://biblioteca.ulpgc.es/</a>
UM	Universidad de Murcia	<a href="https://www.um.es/web/biblioteca/">https://www.um.es/web/biblioteca/</a>
UMA	Universidad de Málaga	<a href="https://www.uma.es/ficha.php?id=62379&amp;am;">https://www.uma.es/ficha.php?id=62379&amp;am;</a>
UMH	Universidad Miguel Hernández de Elche	<a href="https://biblioteca.umh.es/">https://biblioteca.umh.es/</a>
UNAVARRA	Universidad Pública de Navarra	<a href="http://www.unavarra.es/biblioteca/">http://www.unavarra.es/biblioteca/</a>
UNED	Universidad Nacional de Educación a Distancia	<a href="https://www.uned.es/universidad/biblioteca.html">https://www.uned.es/universidad/biblioteca.html</a>
UNEX	Universidad de Extremadura	<a href="https://biblioteca.unex.es/">https://biblioteca.unex.es/</a>
UNIA	Universidad Internacional de Andalucía	<a href="https://www.unia.es/biblioteca-y-publicaciones">https://www.unia.es/biblioteca-y-publicaciones</a>
UNICAN	Universidad de Cantabria	<a href="https://web.unican.es/buc/">https://web.unican.es/buc/</a>
UNILEÓN	Universidad de León	<a href="https://biblioteca.unileon.es/">https://biblioteca.unileon.es/</a>
UNIOVI	Universidad de Oviedo	<a href="https://buo.uniovi.es/">https://buo.uniovi.es/</a>
UNIRIOJA	Universidad de La Rioja	<a href="https://biblioteca.unirioja.es/">https://biblioteca.unirioja.es/</a>
UNIZAR	Universidad de Zaragoza	<a href="http://biblioteca.unizar.es/">http://biblioteca.unizar.es/</a>
UPC	Universitat Politècnica de Catalunya	<a href="https://bibliotecnica.upc.edu/">https://bibliotecnica.upc.edu/</a>
UPCT	Universidad Politécnica de Cartagena	<a href="https://www.bib.upct.es/">https://www.bib.upct.es/</a>
UPF	Universitat Pompeu Fabra	<a href="https://www.upf.edu/web/biblioteca-informatica/">https://www.upf.edu/web/biblioteca-informatica/</a>
UPM	Universidad Politécnica de Madrid	<a href="http://www.upm.es/UPM/Biblioteca">http://www.upm.es/UPM/Biblioteca</a>
UPO	Universidad Pablo de Olavide	<a href="https://www.upo.es/biblioteca/">https://www.upo.es/biblioteca/</a>
UPV	Universidad Politécnica de Valencia	<a href="http://www.upv.es/entidades/ABDC/indexc.html">http://www.upv.es/entidades/ABDC/indexc.html</a>
URJC	Universidad Rey Juan Carlos	<a href="https://www.urjc.es/estudiar-en-la-urjc/biblioteca">https://www.urjc.es/estudiar-en-la-urjc/biblioteca</a>
URV	Universitat Rovira i Virgili	<a href="https://www.crai.urv.cat/ca/">https://www.crai.urv.cat/ca/</a>
US	Universidad de Sevilla	<a href="https://bib.us.es/">https://bib.us.es/</a>
USAL	Universidad de Salamanca	<a href="https://bibliotecas.usal.es/">https://bibliotecas.usal.es/</a>
USC	Universidad de Santiago de Compostela	<a href="https://www.usc.gal/gl/servizos/biblioteca/">https://www.usc.gal/gl/servizos/biblioteca/</a>
UV	Universidad de Valencia	<a href="https://www.uv.es/uvweb/servei-biblioteques-documentacio/ca/servei-biblioteques-documentacio-1285867215074.html">https://www.uv.es/uvweb/servei-biblioteques-documentacio/ca/servei-biblioteques-documentacio-1285867215074.html</a>
UVA	Universidad de Valladolid	<a href="https://biblioteca.uva.es/export/sites/biblioteca/">https://biblioteca.uva.es/export/sites/biblioteca/</a>
UVIGO	Universidad de Vigo	<a href="https://www.uvigo.gal/universidade/biblioteca">https://www.uvigo.gal/universidade/biblioteca</a>

## Institutional repositories and copyright in Greek academic libraries

Konstantinos Kyprianos<sup>(a)</sup>, Ekaterini Lygnou<sup>(b)</sup>

a) University of West Attica, <https://orcid.org/0000-0003-2948-1269>

b) University of West Attica, <https://orcid.org/0000-0002-6915-2059>

**Contact:** Konstantinos Kyprianos, [kyprian@uniwa.gr](mailto:kyprian@uniwa.gr); Ekaterini Lygnou, [mislam186682020@uniwa.gr](mailto:mislam186682020@uniwa.gr)

**Received:** 19 January 2022; **Accepted:** 1 February 2022; **First Published:** 15 May 2022

### ABSTRACT

Institutional repositories were created to collect, preserve, and make available the academic institution's scientific output. The purpose of this study is to investigate and illustrate how Greek academic libraries with institutional repositories deal with copyright challenges. The study aims to identify and describe if institutional repository managers apply a certain copyright clearance protocol, the problems they encounter, and how they deal with them. For this study, a quantitative research method based on questionnaires was employed. The questionnaire consisted of twenty-nine (29) questions separated into three (3) sections and was sent to thirty-one (31) academic libraries.

According to the survey results, the majority of academic libraries have an institutional repository and provide open access to its content. It was found that academic institutional repositories face intellectual property difficulties. The biggest issue highlighted was a lack of knowledge of the notion of copyright. Finally, communication amongst libraries seems to be the foundation for developing a common policy and addressing the difficulties that have arisen in institutional repositories as a result of Greek copyright legislation limits.

### KEYWORDS

Institutional repositories; Digital libraries; Intellectual property; Copyright clearance.

## Introduction

One of the roles of libraries is to provide free access to information, either through traditional means (written material) or by access to electronic resources or digital collections. Several factors, including technological advancement and the open access (OA) movement, have driven libraries to alter their services and roles to suit the expectations of the new era in recent decades. The advancement of information and communication technology (ICT) has resulted in new instruments and opportunities for the creation and dissemination of knowledge. The OA movement has established new possibilities for making knowledge available not only to researchers but also to the general population (Björk 2016). In reaction to these changes, libraries have begun to create digital collections by digitizing items in their collections, taking advantage of the opportunities initially offered by technology. The content of these digital collections varies and typically consists of items that must be protected from frequent usage or items for which copyright protection has ended. With the rise of the OA movement, digital or institutional repositories were established, primarily by academic institutions, to gather, make available, and promote their institution's scientific activities (Marsh 2015).

Members of the academic institution are authorized to present and provide open access to their work in institutional repositories through the procedure of self-depositing, skipping the usual publishing process. Grey literature makes up a sizable component of institutional repositories' collections (Marsh 2015). Many academic institutions and researchers have been depositing their manuscripts in institutional repositories in recent years, based on the imperatives of the OA movement, so that knowledge can be made available to the research community without financial or other constraints (Björk 2016).

Intellectual property rights are one barrier to libraries' freely offering access to knowledge. Intellectual property rights cannot be ignored or suppressed; they were enacted to protect creators' work from being plagiarized and commercially exploited by others. But to what extent may intellectual property restrict the free dissemination of knowledge? As previously stated, the OA movement seeks to limit these restrictions, and libraries are important allies in this effort. The ambition to enhance an institutional repository's collection with publications deriving from the scientific activity of the academic institution's members occasionally conflicts with intellectual property rights. Writers' rights to their work, as well as agreements between authors and publishers, frequently make depositing a work in an institutional repository unappealing. Clearing the copyright of a work is the process of identifying the copyright holder(s) of a work and obtaining permission to publish it in an institutional repository. It is a complicated and time-consuming procedure, and it is one of the most difficult issues that libraries face (Macklin 2013; Dawson and Yang 2016).

Having the aforementioned in mind, the purpose of this study is to discover what procedures academic libraries use to resolve copyright issues in their repositories. Furthermore, this study will investigate whether institutional repositories adhere to a defined workflow for the clearance process, who is in charge of the clearance process, what kinds of problems arise, and what services or technologies they employ to facilitate the task of copyright clearance. It will also examine whether academic libraries communicate with one another about copyright issues and whether they have formed common methods or policies to resolve those issues.

## Related work - Clearance of intellectual property rights

Looking through the literature on how institutional repositories deal with copyright issues, it appears that multiple methods of clearing copyright are documented, which are carried out by different groups of people. In the early years of institutional repositories' deployment, one of the most important criteria determining a repository's success was the clearance of intellectual property rights (Palmer et al. 2008). It was and still is a time-consuming and complex process, as a variety of issues, such as problematic communication between the author and the publisher (delay in responding to authors' queries, incomplete or ambiguous answers), very strict licensing conditions, and failure to keep proper records about the licenses granted, delay the clearance processes (Hanlon and Ramirez 2011).

But who is responsible for resolving the copyright riddle, the author or the institutional repository's library manager? In recent years, there has been a widespread belief that the obligation for copyright clearance lies with the authors. On the contrary, research suggests that the situation was different in the early years of the adoption of institutional repositories. In particular, it appears that in that era, librarians were in charge of copyright clearance procedures, even if the author was the one who submitted his/her work to the repository. They contacted publishers to learn about the agreements they had made with authors and recorded the policies of the publishers. Even if the repository's policy required the author to carry out the deposit (self-depositing process), librarians handled the copyright clearance (Hanlon and Ramirez 2011). However, as Dawson and Yang (2016) point out, in recent years, managers of institutional repositories, which are typically libraries, have been less involved in the copyright clearance process. Libraries have a management role in a repository that is related to the repository's technical operation or the provision of metadata. The author/depositor is responsible for deciding how to make the work available to the public, and if a publisher intervenes, libraries require authors to have obtained permission from the publisher to deposit their publication in the academic institution's repository. Indeed, Dawson and Yang (2016) discovered that the majority of libraries are not participating in the copyright clearance procedure after investigating the intellectual property policies of one hundred (100) US institutional repositories. Most institutional repositories provide information on copyright and the self-depositing process on their websites. However, a limited number of libraries take complete responsibility for copyright acquisition or have support staff available to authors to offer advice and guidance on publisher policies, and some libraries are willing to pay publishers to obtain the rights. In addition, some repositories provide a questionnaire to help authors explain copyright concerns during the depositing process to aid them in determining whether their work is subject to copyright during the depositing process. Finally, few libraries work with publishers on behalf of authors.

The issue of copyright is critical for libraries because it affects many of the services they provide to the public. Library services such as digitization, lending of digital documents, and distribution of copies (photocopies) are provided in collaboration with copyright holders with whom they have achieved a monetary compensation agreement or, in some cases, an exemption from copyright limitations. Budget limits are also a recurring issue for libraries, posing a barrier to their function in providing access to knowledge. In response to this worry, research was conducted in the United Kingdom in 1993 on the impact of rising scholarly publication costs and



university library budget cuts (Muir 2019). The Follett Report (Brindley 2006) sought to study how information and communication technology could be deployed to university libraries to address these issues. One of the major issues addressed in the report's conclusions was the limitations imposed by copyright restrictions on information access. The Follett Report's findings and suggestions prompted the establishment of the eLib (Electronic Libraries) project, which aimed to fund several initiatives to assist academic libraries in integrating information and communication technology into their services (Muir 2019). Although efforts to collaborate with copyright holders and managers were limited within the initiative, two projects excelled in this area. The ACORN (Academic Course Readings via Networks) and SCOPE (SCottish On-demand Publishing Enterprise) programs. Both programs aim to give undergraduate students access to electronic resources (journal articles, books). This was accomplished through the employment of two distinct methods. On the one hand, they contacted publishers collectively, as was done with the ACORN program, where program administrators gained the trust of the publishers to participate in the program by working with a company that acted as an intermediary between the library and the publishers on subscription and copyright clearance issues. On the other hand, a strategy for copyright compensation was established, as was done in the SCOPE project. Academic libraries essentially established a pricing model for copyright payments and negotiated the final price with each publisher separately.

Another service that has emerged and aids in the copyright clearance procedure is the SHERPA/RoMEO service. It gathers copyright policies from publishers. However, according to numerous users, the information it offers is not always sufficient. The reasons for this are primarily found in the fact that such services are frequently not properly informed about publishers' policies and the versions of publications that they allow for a deposit (e.g., post-print, pre-print, etc.), and they also do not provide accurate information about the text of the agreement between the publisher and the author (Hanlon and Ramirez 2011). When the information provided by the SHERPA/RoMEO service is insufficient, repository managers often make direct contact with publishers. According to Hanlon and Ramirez's (2011) survey, the majority of respondents contacted publishers for each publication individually. E-mail is commonly used for communication. They also reported that a small percentage sought and obtained general permission from the publisher to grant permission to deposit in the repository publications of researchers who had any form of collaboration with the academic institution. This way, they ensured that future publications by researchers could be deposited in the repository according to the terms of this agreement, thus reducing staff time spent on copyright clearance and avoiding frequent contact with publishers. Regardless of how they communicate and enter into an agreement with the publisher, several librarians reported keeping a record of their communication with publishers, and some others even attaching the publisher's permission with the file of the publication to the repository.

In another survey, Palmer et al. (2008) report that copyright clearance processes are time-consuming and reduce the success of a repository. As a result, they argue that copyright management strategies must become more automated and coordinated to provide a common ground for dealing with these issues, which will also increase the repository's impact. There are areas of the copyright clearance workflow that could be improved, such as the development of a standard form of communication with publishers, more regular communication on this topic among information scientists to handle difficulties jointly, and the development of common policies.

A very similar survey was conducted by Yoo and Kim (2013) to investigate the copyright clearing methods for electronic document delivery in Korea. More specifically, the survey revealed that when it comes to copyright clearance for works published outside of Korea, Korean libraries have several options, including negotiating an agreement directly with publishers; establishing an agreement with, and paying royalties to, copyright collectives; establishing bilateral treaties with national copyright collectives; establishing a framework agreement with the organizations representing copyright holders; and incorporating extended collective licensing into the Korean copyright law.

Finally, Prandoni, Valentini, and Doerr (2009) proposed a paradigm for digital rights clearance based on the assumption that clearing intellectual property rights has become too complex and time-consuming to be done manually. More specifically, they presented an intellectual property rights domain ontology based on four levels of abstraction: (1) the legal framework, (2) the individual rights people hold, (3) the individual usage agreements right holders and others may issue, and (4) the specific actions that are restricted by intellectual property rights regulations or bring particular rights into existence.

To summarize, it seems that dealing with copyright issues is dependent on the nature of the material. The authors have the burden of obtaining copyright clearance. Authors are responsible for resolving intellectual rights issues with publishers, although there are libraries that either serve as advisors to authors throughout the copyright clearance process or engage in a contractual deal with publishers to provide them with access to the material they need. Some libraries develop their own policies with publishers in order to enter into an agreement with them while offering the option to withdraw from the arrangement in the future if there is a disagreement. Some institutional repositories allow authors to opt-out of the mandated deposit process or to submit an updated version of their article rather than the final version of the publication (Dawson and Yang 2016).

## Methodology

In this line of thought, the purpose of this study is to discover what procedures university libraries in Greece use to resolve copyright issues in their repositories. To be more specific, a quantitative research approach was used to carry out this study. Quantitative research provides an objective method of investigating and testing certain hypotheses or topics. It is based on the collection of data through the use of a questionnaire and allows data to be collected and analyzed to reach a conclusion (Godwill 2015).

To begin with, we investigated international and Greek indexes (i.e., ROAR, OpenDoar, and OpenArchives), which index repositories that offer OA in their content, in order to determine how many institutional repositories are handled by Greek academic libraries in terms of numbers (see Table 1).

Table 1: Greek academic institutional repositories

	Total Greek repositories	Repositories of Greek academic libraries	Other
ROAR	17	16	1
OpenDoar	31	21	10
OpenArchives	23	22	1

Because the above indexes do not represent the total number of institutional repositories and because academic institution mergers brought about changes to academic libraries as a result of the implementation of Law 4610/2019 “Synergies between universities and universities of applied sciences, access to higher education, experimental schools, General State Archives, and other regulations,” we chose to investigate the websites of each academic institution separately. As a result, the research sample included thirty-one (31) university libraries.

A questionnaire was created, consisting of twenty-nine (29) questions. It was divided into three (3) sections. The first section, which consists of eight (8) questions, deals with the identity of the repository. The second one, which consists of six (6) questions, deals with the academic community in terms of its relationship with the repository and intellectual property issues, but as perceived by libraries, and finally, the last section, which consists of fifteen (15) questions, deals with the copyright clearance process. The questionnaire was anonymous and was sent to thirty-one (31) academic libraries. It was sent in three (3) phases. The first one was done in two parts on 8/2/2021 and 10/2/2021. The second, which was in the form of a “kind reminder,” took place on 22/2/2021, and the third was a new dispatch on 8/3/2021. Twenty-five (25) of the thirty-one (31) questionnaires sent were returned, meaning that the response rate is rather high (i.e., 80.65%).

## Results

### Identity of the institutional repository

The first section of the questionnaire refers to the identity of the institutional repositories and gives us a general insight into their usage. More specifically, in response to the first question about whether the library has a repository, twenty-one (21) of the twenty-five (25) responses were positive (i.e., 84%) (see Figure 1).

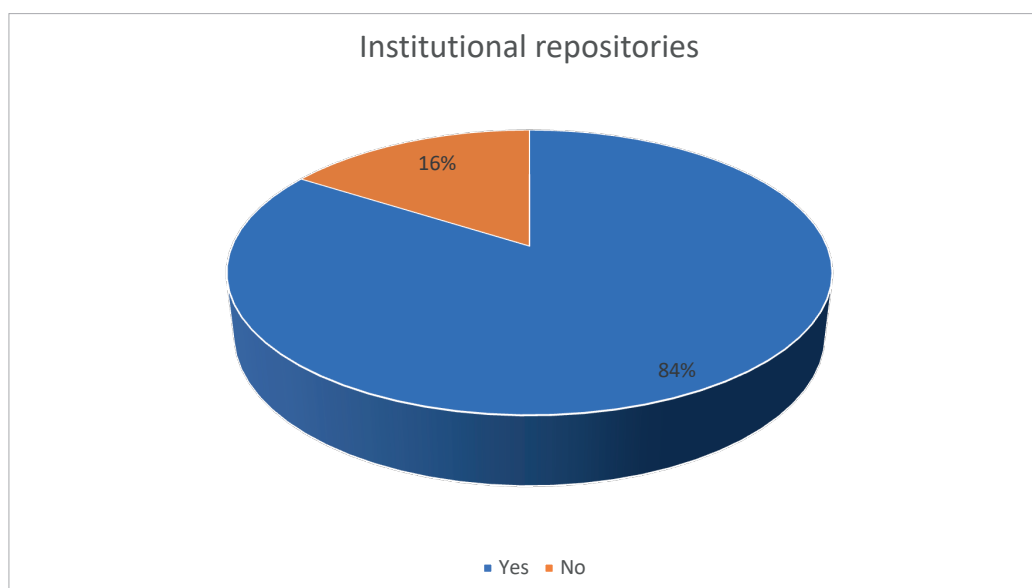


Figure 1: Existence of Institutional repositories

In the second question, we notice that nine (9) repositories were established in the first decade of 2000, while twelve (12) were established in the second decade (i.e., 2010 and after) (see Figure 2).

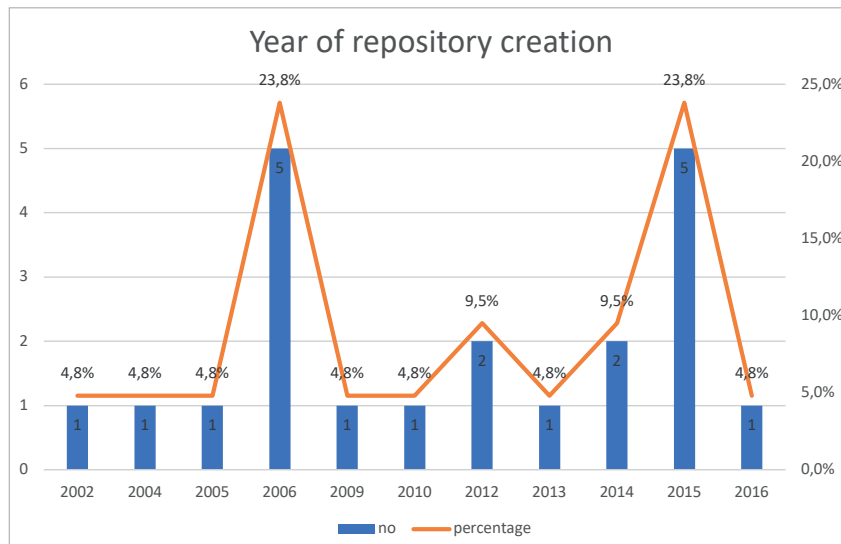


Figure 2: Year of repository creation

Who has the authority to make deposits in the repository? Of course, the majority of 95,2% (i.e., 20) are students, which is to be expected given that most academic institutions require students to deposit their thesis or dissertation to complete their studies. Also, in the majority of repositories, professors (81%, i.e., 17), researchers collaborating with the academic institution (47,6%, i.e., 10), administrative staff (38,1%, i.e., 8), graduates of the institution (19%, i.e., 4), and doctoral candidates (4,8%, i.e., 1) have the right to deposit. Finally, one library says that it deposits works belonging to the public sector that it has digitized, whilst another does not permit self-deposit in the repository, meaning that the deposit is made by the library (see Figure 3).

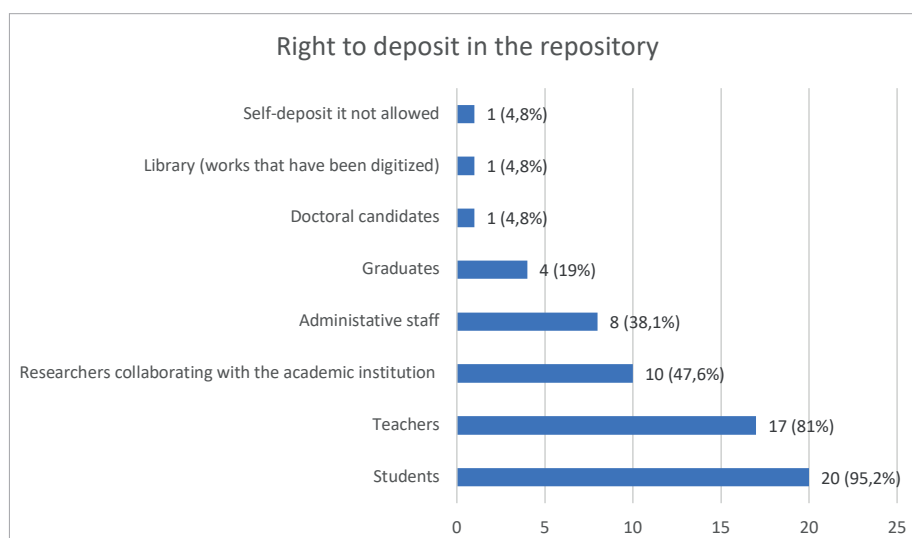


Figure 3: Right to deposit in the repository

Whereas the previous question examined who has the authority to deposit in the repository, question four (4) asks about “who is obliged to deposit in the institutional repository.” For the reasons described in the preceding question, students have the highest proportion (i.e., 100%, 21), whereas professors (i.e., 9,5%, 2) and doctorate candidates (i.e., 4,8%, 1) have substantially lower percentages (see Figure 4).

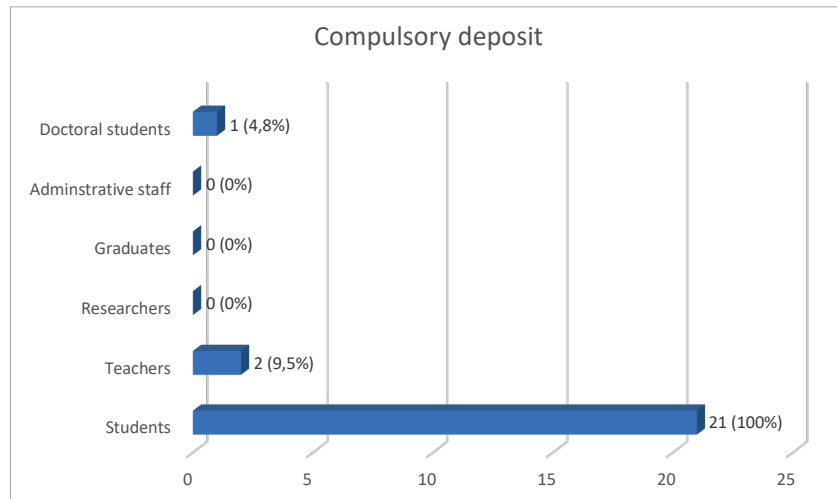


Figure 4: Compulsory deposit in the repository

When the participants were asked about the management of the institutional repository and who is responsible for this task, we note that the management of the repository is a task that necessitates the collaboration of experts from several domains. With a 100% (i.e., 21) percentage, librarians are unquestionably the ones who are responsible for the operation of the repository. Next, with 33,3% (i.e., 7), are the IT specialists, who probably collaborate with the librarians on technical issues. Other professionals are the repository manager with 23,8% (i.e., 5), who is probably an employee who deals exclusively with the institutional repository, and finally the administrative staff with 4,8% (i.e., 1) (see Figure 5).

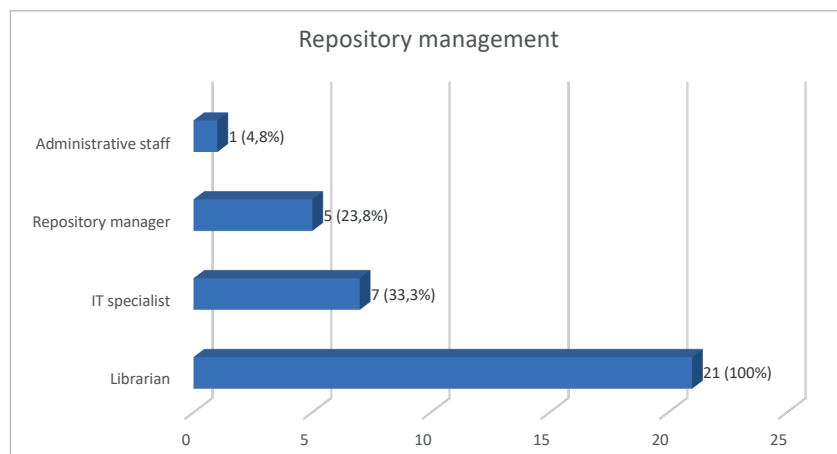


Figure 5: Repository management

The materials included in the institutional repositories are theses 90,5% (i.e., 19), doctoral dissertations 85,7% (i.e., 18), bachelor theses 71,4% (i.e., 15), journal articles 52,4% (i.e., 11), pre-prints 42,9% (i.e., 9), post-prints 33,3% (i.e., 7), publications of the academic institution and conference presentations with 38,1% (i.e., 8) respectively, professors' textbooks 23,8% (i.e., 5), e-books 19% (i.e., 4), e-journals 14,3% (i.e., 3). Also, to a much lesser extent, some repositories have audiovisual material, photographic and historical archives, audio files, songs and videos, technical reports, and studies (see Figure 6).

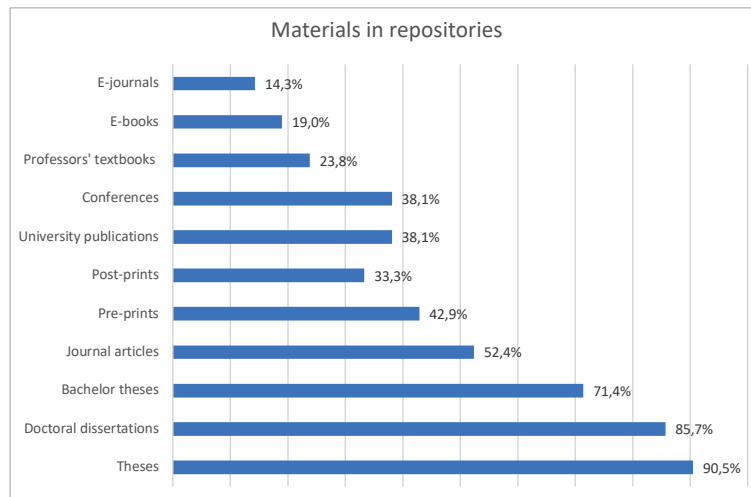


Figure 6: Materials deposited in academic repositories

Moving to the next question, it seems that access to the majority of repositories is unrestricted. However, 19% of the repositories have restricted or limited access, which means they are only available to registered users who are members of the academic community. Repositories also reported that while providing OA, they provide some flexibility, such as offering closed access for three years if asked. There is also a library that specifies that special authorization from the author is required for theses. Finally, there is a library where the access policy is determined based on the type of work or the specific circumstance (see Figure 7).

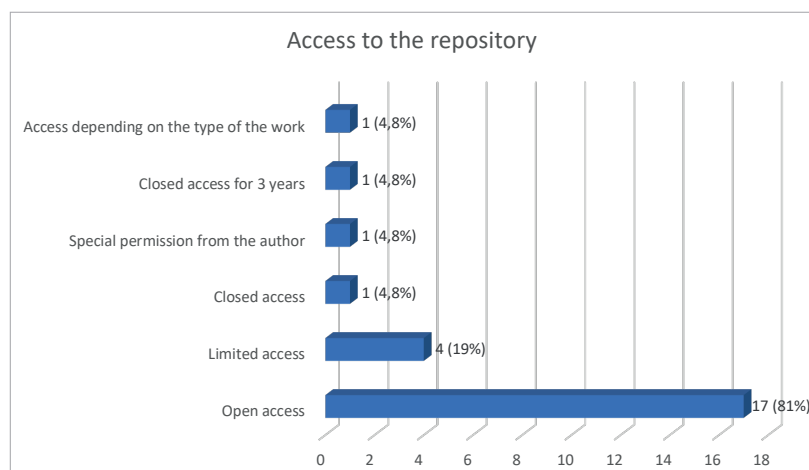


Figure 7: Access to the institutional repository

Who is responsible for depositing a work in the repository? The authors are responsible in most cases (i.e., 71,4%, 15). Librarians follow with a relatively high percentage (i.e., 52,4%, 11), followed by the repository manager (i.e., 19%, 4). Also mentioned a case where the author makes the deposit through the self-deposit process, but the librarian gives the final approval (Figure 8).

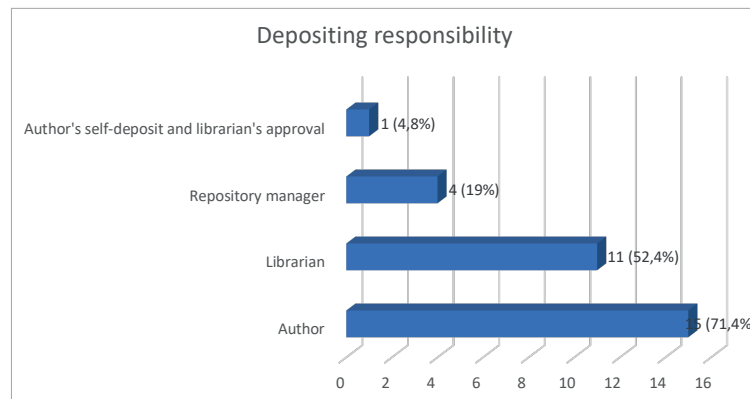


Figure 8: Self-depositing or not?

To summarize, we can conclude from the first section of the questionnaire that the majority of academic libraries have an institutional repository. The vast majority of them offer open access to their content. Most members of the academic community (students, professors, researchers, graduate students, administrative staff, and so on) have the right to deposit papers or publications in the repository, though it is mandatory for students in all libraries and, to a lesser extent, professors and doctoral candidates. Librarians administer the institutional repository with the assistance of other professionals, such as IT specialists. Finally, while authors are in many cases responsible for depositing their work in the repository, librarians are often involved in or completely responsible for the process.

### Academic community – Institutional repository

This section explores how libraries perceive the academic community's relationship with the institutional repository. From the results, we can observe that there is generally a positive attitude from the academic community towards the repository (i.e., 95%, 20) (see Figure 9).

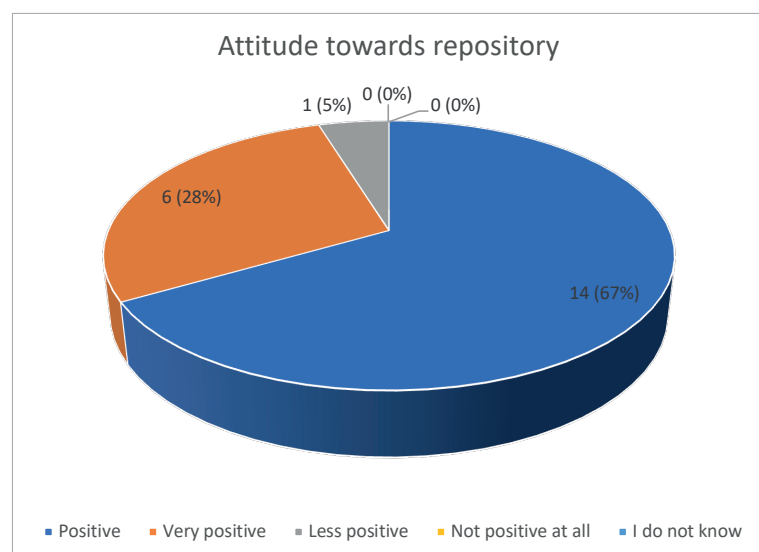


Figure 9: Attitude towards the institutional repository

Members of the academic community participate by depositing their works in the institutional repository to a fairly satisfactory degree, with a percentage of 47,6% (i.e., 10), while there are also cases of low participation (i.e., 9,5%, 2) (see Figure 10).

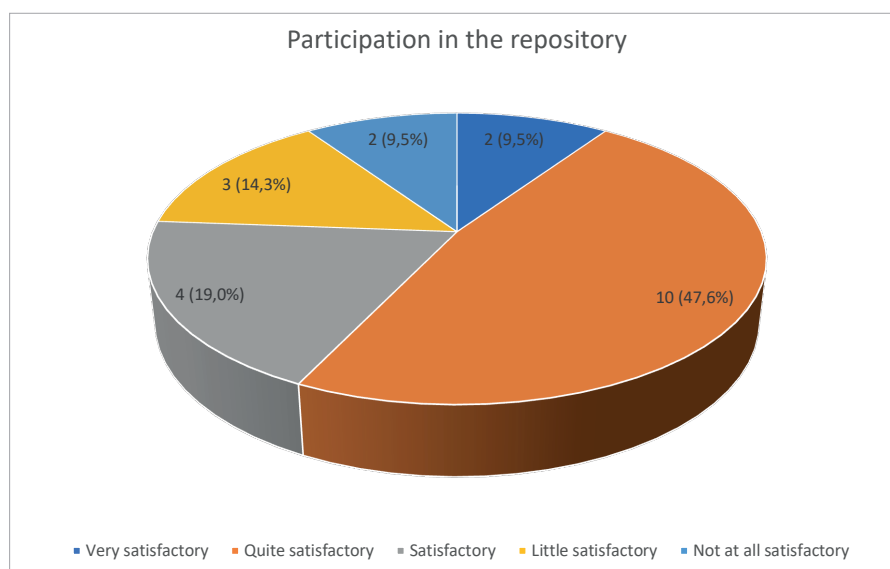


Figure 10: Participation in the institutional repository

Important information is provided by the question of whether the library provides information to members of the academic community about intellectual property. Almost all libraries (i.e., 95,2%, 20) provide such information, which can be interpreted as an indication that they understand how important intellectual property issues are for members of the academic community as well. Only one library stated that it does not provide this information (i.e., 4,8%) (see Figure 11).

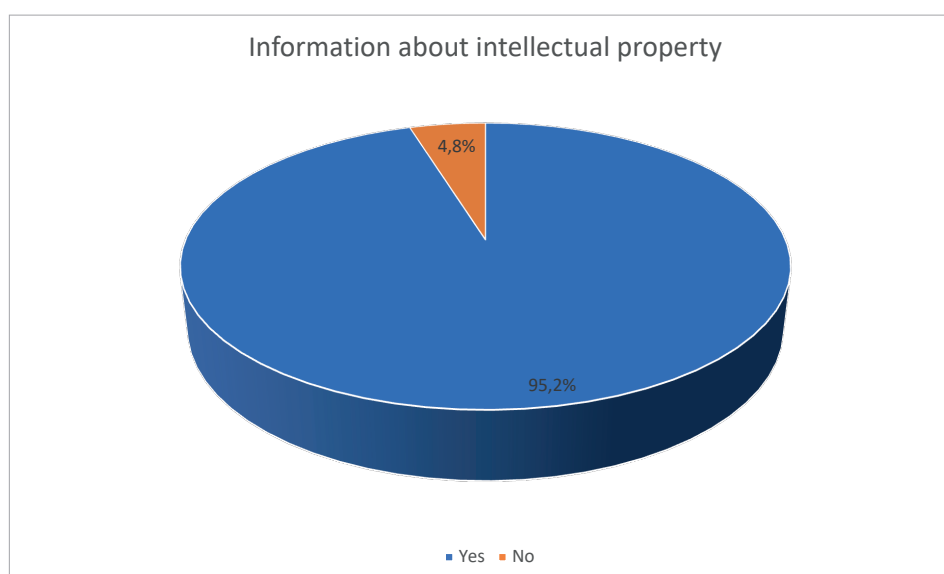


Figure 11: Information about intellectual property



How does the library communicate intellectual property to members of the academic community? Initially, 71,4% (i.e., 15) are informed during the workflow of depositing a work in an institutional repository. With the same percentage, we observe that librarians provide advice and information about copyright. Other ways to inform the academic community include the provision of additional information on the repository's website (i.e., 66,7%, 14), the hosting of educational/training programs to inform the community (i.e., 33,3%, 7), and finally, the provision of advice by a lawyer with whom the library collaborates (i.e., 4,8%, 1) (see Figure 12).

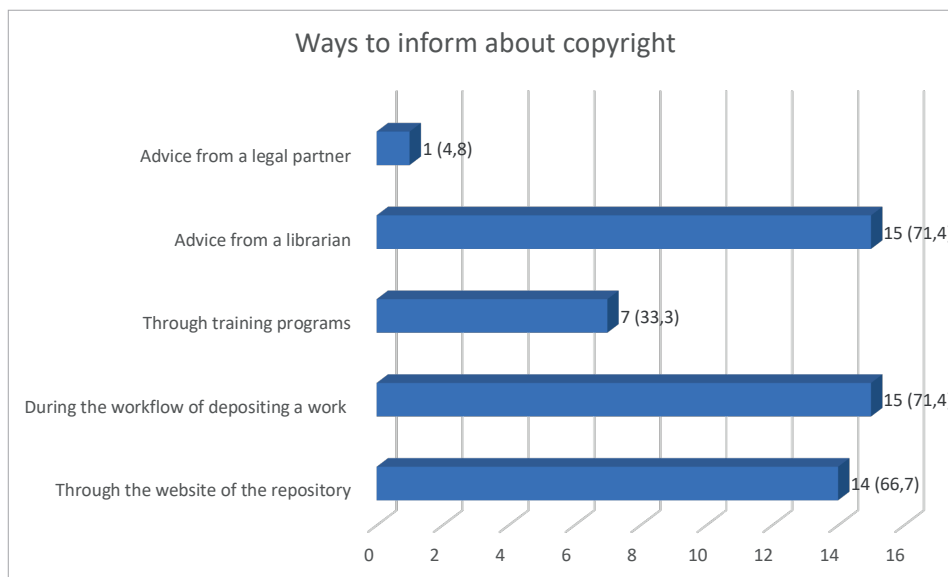


Figure 12: Ways to inform about copyright

When participants were asked whether the authors who deposited in the repository had reported problems related to intellectual property issues, most of them answered negatively (i.e., 57,1%) (see Figure 13).

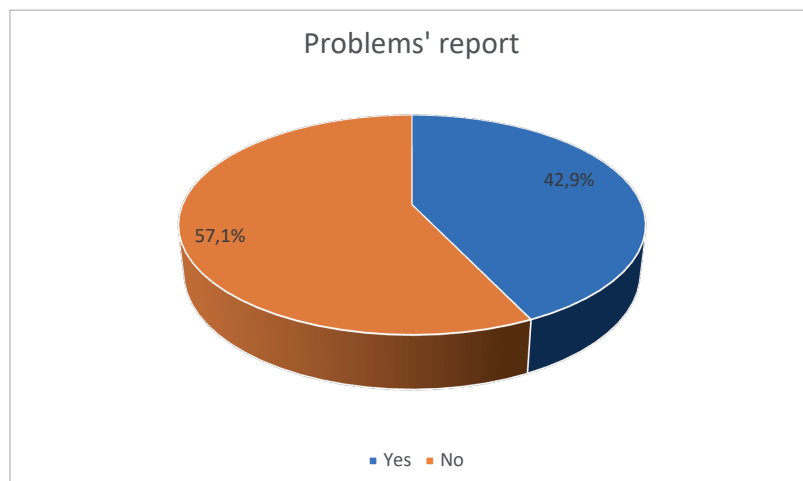


Figure 13: Problem report

The problems commonly reported by authors concerning copyright are difficulty in selecting a license during the submission of the work (i.e., 38,1%), difficulty in understanding the concepts of copyright (i.e., 33,3%), unfamiliarity with the concepts of copyright (i.e., 23,3%), permission to deposit in the repository from the publisher (i.e., 9,5%), and a case of plagiarism (i.e., 4,8%) (see Table 2). Such problems seem to have arisen from a lack of information and education since most academic libraries in Greece do not provide information literacy courses to their users as part of the educational process (information literacy in the curriculum) (Derakhshan and Singh 2011).

Table 2: Problems concerning copyright

	Answers	Percentage
Difficulty in selecting a license during the submission of the work	8	38.1%
Difficulty in understanding the concepts of copyright	7	33.3%
Unfamiliarity with the concepts of copyright	5	23.3%
Permission to deposit in the repository from the publisher	2	9.5%
Plagiarism	1	4.8%

## Copyright clearance procedure

The final section of the survey investigates the copyright clearance process. More specifically, when participants were asked whether libraries have a policy on the copyright clearance process, the majority (i.e., 52,4%) responded negatively, while 33,3% responded positively. Furthermore, some libraries reported that the depositor signs a document stating that he/she has made proper and legal use of third-party works, or that there is no such procedure because depositing theses and dissertations is mandatory, whereas depositing other publications (e.g., journal articles, conference papers, etc.) requires consent from the publisher (see Figure 14).

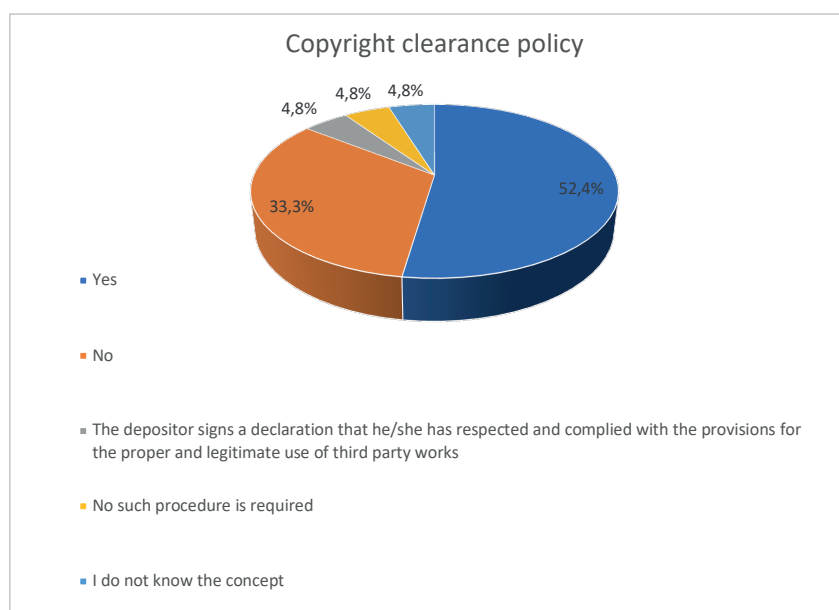


Figure 14: Copyright clearance policy

When asked who is in charge of the copyright clearance process, the majority of participants (61,9%) answered authors/creators, followed by librarians (23,8%), while a minor but significant finding is that the legal department of the academic institution is in charge of this process. Furthermore, several libraries indicate that they lack a policy or are unaware of the concept (see Figure 15).

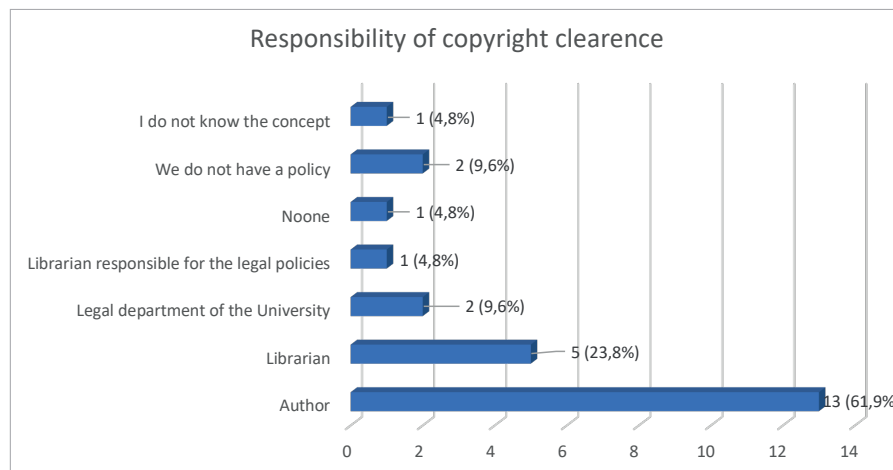


Figure 15: Responsibility for the clearance of copyright

As we noted in the question on whether libraries have a copyright clearance policy, most of them answered negatively. The same can be seen in the question of whether there is a workflow model for this process, where again the largest percentage (i.e., 52,4%) answered negatively. There are also smaller percentages reporting that they do not know the concept of copyright clearance. However, a percentage of 23,8% answered positively to the same question. Libraries that answered positively report that they resort to services such as SHERPA/RoMEO and publishers' websites to gather the necessary information. When submitting a work, the author is also responsible for copyright clearance. In the case where the library deposits content in the repository, it makes sure to make contact with the authors to ask for the necessary permission, while in the case where a work belongs to the public domain (work that is freely available and can be distributed without restrictions, only the author should be mentioned), no action is taken, which is in line with the declarations of the OA.

When it comes to what type of material is more challenging, as expected, the highest percentage (i.e., 23,8%) is journal articles. This is followed by post-prints, theses and dissertations, and doctoral dissertations with 19%. Also, other already mentioned materials are pre-prints and conference presentations with 14,3% respectively, professors' texts with 9,5%, and e-books with 4,8%. Finally, some libraries report that they have not encountered any problems.

As far as it concerns the problems that libraries face in the copyright clearance process, we identify the following. Limited staff time and difficulty in understanding the concepts of intellectual property by authors are the problems that gather the highest percentage (i.e., 23,8%). This is followed by reduced library staff and the limited experience of staff in intellectual property issues (i.e., 19%). Other problems include the inability to obtain permission from the author or publisher,

understanding the publisher’s copyright policy, or the inability to locate the agreement made with the author.

What tools or resources do libraries use to facilitate the copyright clearance process? The main service they use is SHERPA/RoMEO (i.e., 57,1%, 12), followed by the Heal Legal service of the Hellenic Academic Libraries Link (i.e., 52,4%, 11), contacting publishers or searching for publishers’ policies on their websites (i.e., 28,6%, 6) or contacting authors (i.e., 28,6%, 6). Some libraries also report that the legal department of the academic institution also assists (see Figure 16). These tools are beneficial to libraries in 52,4% of the cases, while 23,8% of the respondents said that they are ineffective in facilitating the process. Reference is also made to the legislation, i.e., that it is unclear or even incomplete in the field of theses and dissertations.

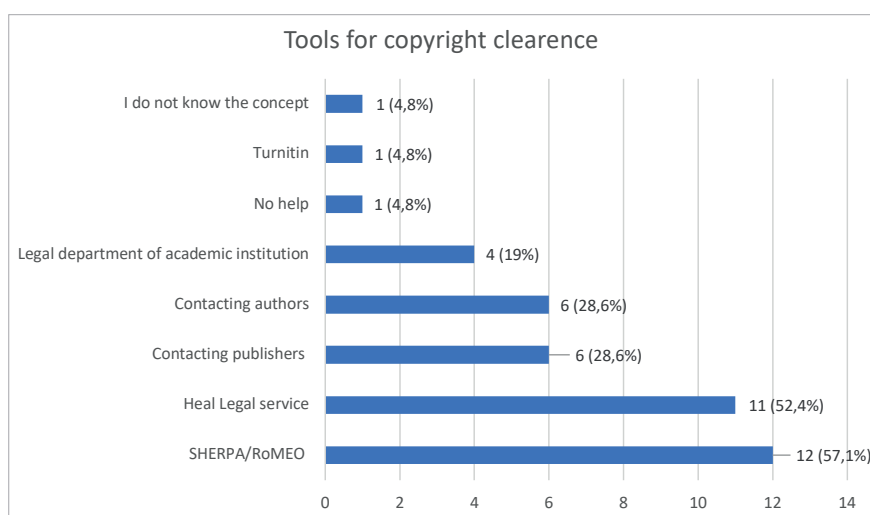


Figure 16: Tools or resources libraries use to facilitate the copyright clearance process

What tools, services, or actions could facilitate the copyright clearance process for libraries and authors? Education of library staff and members of the academic community on copyright issues (i.e., 61,9%) is one of the most necessary actions identified by the respondents. This is followed by the library working with a legal advisor (i.e., 38,1%), as expected since intellectual property concepts are difficult to understand. Cooperation and communication between academic libraries to jointly address problems arising from intellectual property rights is also suggested by 28,6%. Finally, the suggestion to enter into an agreement with publishers to provide permission for publications by members of the academic institution is highlighted by 19% of the participants (see Table 3).

Table 3: What could facilitate the copyright clearance process?

	Answers	Percentage
Education of library staff and members of the academic community on copyright issues	13	61,9%
Library working with a legal advisor	8	38,1%
Cooperation and communication between academic libraries	6	28,6%
Enter into an agreement with publishers	4	19%

When asked whether the library contacts publishers to obtain permission for previously published material, the majority of the participants (i.e., 61,9%) answered negatively. This is most likely due to the fact that the repositories currently only host grey literature, such as theses and dissertations (Chantavaridou 2009). However, there is a proportion of 23,8% who contact publishers (see Figure 17).

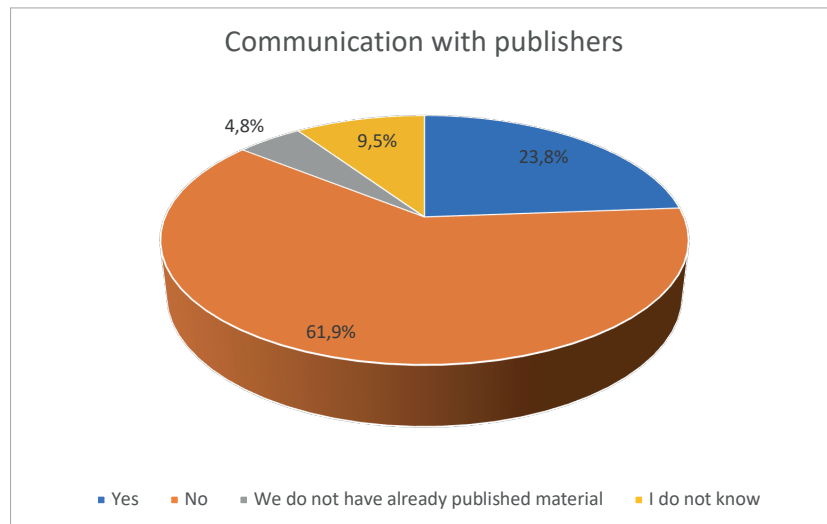


Figure 17: Communication with publishers

Where there is a successful arrangement with the publisher to obtain a license, 33,3% indicate that the license is provided at no financial cost, while a small proportion, 9,5%, claim that some compensation has been paid to the publisher for acquiring the license. There are also other circumstances where the repository does not include material that requires permission from a publisher or where the publishers are the authors themselves, in which case contact is made with them. When asked if the library's budget covers the costs of the copyright clearance process, the vast majority of libraries responded negatively (i.e., 85,7%), with only one library responding favorably (i.e., 4,8%) (see Figure 18).

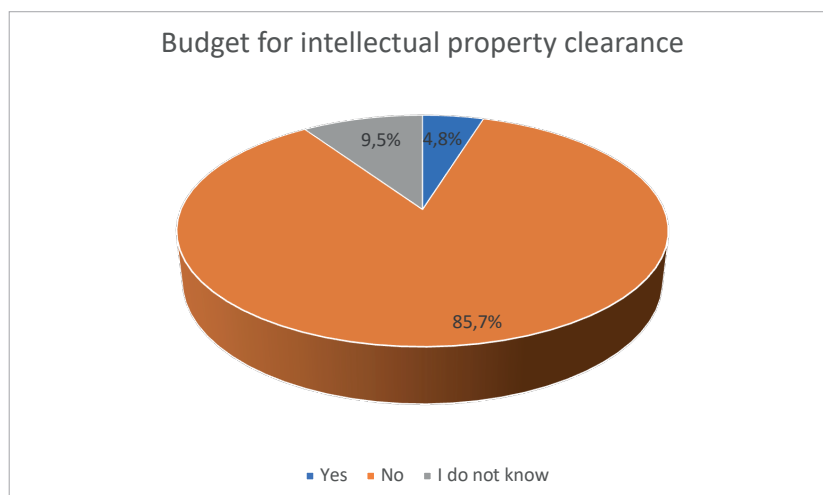


Figure 18: Budget for intellectual property clearance

Is there communication concerning copyright issues among academic libraries? The majority of libraries (57,1%) responded negatively, but there is also a 33,3% who transmit their experiences to other libraries. In addition, one library (4,8%) is willing to disclose any information needed (see Figure 19).

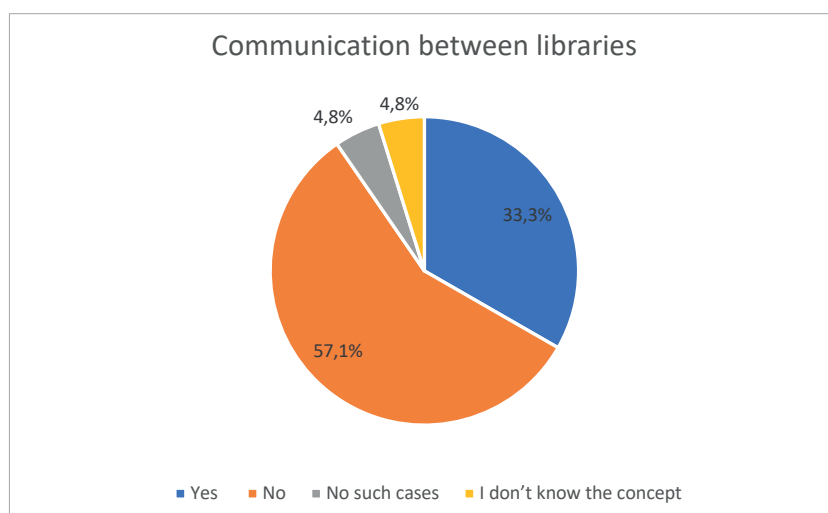


Figure 19: Communication between libraries

Additionally, meetings with academic library staff to exchange knowledge and experience (i.e., 38,1%), e-mail communication (i.e., 33,3%), or telephone communication (i.e., 4,8%) are the most common methods of communication with libraries. Libraries also provide information on their websites (14,3%) or through postings on forums (9,5%).

Finally, an open-ended question asked the respondents if they would like to make any further comments on the copyright clearance process. Comments indicated that the legislation is unclear, that the information provided to library staff by the academic institution on copyright issues is insufficient, and that the reduced library staff does not allow them to be involved in monitoring the copyright clearance process.

## Discussion

This study aims to describe the copyright issues that Greek academic libraries managing institutional repositories encounter, as well as the solutions they adopt to address these issues. To begin with, as stated in the literature review, the majority of academic libraries have an institutional repository. We may say that most libraries accept and encourage the OA declarations because access to their content is free and unrestricted in most repositories. There are also repositories that give restricted access, i.e., solely to members of the academic community that they serve. As a deposit in the repository is required for the completion of studies, the majority of their content consists of theses and dissertations. Such a finding is in complete agreement with Chantavaridou's (2009) research. However, they may also include doctorate dissertations, journal papers, pre- and post-

prints, professors' texts, e-books, and other materials. The duty to deposit theses in repositories primarily affects students, both undergraduate and postgraduate, with only a few repositories requiring professors and doctorate candidates to do so. Of course, in addition to the community members mentioned above, administrative staff and senior graduates of the academic institution have the right to deposit items in the repository. The academic community also participates to a satisfactory extent in the institutional repository and has a positive view of its operation and the purpose it serves. One of the purposes of institutional repositories is for the author to deposit his or her work in the repository (self-depositing). According to the survey, this principle is true because authors are responsible for depositing their work in the repository. However, a relatively high percentage of librarians are also involved, possibly carrying out the process entirely or supervising or even providing assistance during the depositing process.

As indicated in the literature review, the academic community is frequently reluctant to get involved in the repository, expressing fears of copyright infringement of their works as well as a lack of understanding of intellectual property concepts. In this study, we discovered that the majority of community members who deposit their works in the repository have not voiced any worries regarding intellectual property issues. If they do report an issue, it will be an inability to comprehend which license to select during the deposit procedure. However, incidents of difficulties in understanding or ignorance of intellectual property principles are also documented, which is consistent with the findings of Dawson and Yang (2016).

In terms of the copyright clearance process, more than half of the libraries surveyed stated that they do not have a specific policy. A lower percentage, on the other hand, stated that they have a policy that primarily entails contacting publishers or authors to resolve copyright issues. They also use services like SHERPA/RoMEO, the Heal Legal service, or search their websites for publisher policies. As stated in the literature review, the author is in charge of the clearance procedure. Librarians are referenced as well, albeit to a lesser extent. They are often not solely accountable for the process, but rather play a supporting role. A noteworthy conclusion is that four libraries state that the process is handled by the academic institution's legal department. This statement is significant since having the assistance of a legal practitioner in these situations is quite beneficial and relieves authors and librarians of the "burden" of this arduous duty. Such a finding contradicts the results of Hanlon and Ramirez's (2011) study, in which the university legal department did not provide any advice to libraries about institutional repositories or the copyright clearance process.

Moreover, difficulty in understanding the concepts of intellectual property, limited library staff, limited time, lack of experience in intellectual property issues among existing staff, inability to obtain a license from the author or publisher, and the inability to locate the publisher's agreement with the author are some of the issues that libraries face during the clearance process. They usually have issues with journal articles, post-and pre-prints, professors' texts, and theses and dissertations.

To obtain a license, the vast majority of libraries do not contact publishers. This is largely due to the fact that the repositories' content is primarily comprised of grey literature (theses and dissertations). However, as previously stated, many libraries lack a copyright clearance policy. Only those that have such a policy communicate with publishers. As a result, when this contact occurs, the vast majority of publishers provide permission without monetary remuneration. However, a small number of libraries report paying compensation to publishers. Based on the findings above, we can assume that publishers are more tolerant and will enable publications to be deposited in

an institutional repository without financially burdening libraries. This is an important element to consider because the majority of libraries reported that their budgets do not include costs for intellectual property issues.

Additionally, we focus on what libraries suggest to facilitate the copyright clearance process:

- Educating the academic community and library staff on copyright issues
- Collaborating with the legal department of the institution
- Establishing an agreement with publishers to grant permission for publications by members of the institution

Finally, the academic community's knowledge of intellectual property and communication among academic libraries are critical. As far as it concerns the library community, participants reported that there is not much communication between libraries to share experiences and information and that when there is communication, it is primarily through meetings or e-mail. Due to the complexities of intellectual property concerns, researchers emphasized communication between libraries as a crucial aspect that would help to better inform and resolve issues efficiently.

## Conclusions

The goal of this study was to offer a thorough picture of Greek institutional repositories of academic libraries and how they deal with copyright issues. To summarize the survey results, the first insight we have is that most university libraries have established an institutional repository that serves as an active service to the academic community.

In the area of copyright, the survey shows that Greek academic institutional repositories face challenges from intellectual property issues, but they share them equally with the authors who deposit in the repository. We base this conclusion firstly on the fact that to a large extent the responsibility for depositing a work in the repository lies primarily with the authors. The same can be said for the copyright clearance process. It is the author's responsibility to carry out this process to obtain the necessary permission to deposit the work in the repository. Most authors of course do not report any problems in the clearance process, but there is a considerable percentage that has problems mainly in understanding both the licenses to choose when depositing the work and in understanding the concepts of copyright.

A small number of members of the academic community report difficulties in obtaining permission from publishers, which is likely expected given the limited number of publications published in journals or books in Greek institutional repositories (Togia et al. 2018). Libraries, on the other hand, are more likely to declare they do not have a copyright clearance policy. This is most likely due to the fact that authors are in charge of the clearance process, or because the copyright clearance process itself incorporates some fundamental procedures that help facilitate it without the need for another policy.

There is also a significant number of libraries that state to have a copyright clearance policy. When a library undertakes the clearance procedure, it typically uses services such as SHERPA/RoMEO to identify a publisher's policy or contacts publishers or authors to gain the necessary permissions. In addition to SHERPA/RoMEO and contacting publishers, libraries can benefit from the Heal Legal service, which gives copyright assistance to academic libraries. Something intriguing is that



a small number of libraries indicate that the university's legal department assists with intellectual property issues and also handles copyright clearance. The participation or involvement of the legal department in the future will immensely benefit libraries and members of the academic community, as intellectual property principles can be difficult to grasp in some cases.

Because of the involvement of publishers and the constraints they impose, the content with which they are most concerned is journal articles. Other types of material include theses, dissertations, and doctoral theses, most likely because the research results of these manuscripts will be published in a journal or book in the future, so there is an agreement with a publisher, or because, as previously stated, the most difficult problem that authors face when depositing in the repository is difficulty in choosing a license.

By completing the study, we identify specific areas where their implementation could ease some of the procedures relating to intellectual property issues. The first observation may be described as a "common request" from libraries and members of the academic community for training on intellectual property concepts to make them more comprehensible. Libraries provide information to members of the academic community either through the repositories or the website, as well as through training programs. Perhaps intellectual property information should also be included in the curricula of academic institutions, as it is essential knowledge for future scholars. In the library sector, communication between professionals is one of the most important factors. The survey shows a low degree of communication between libraries. Frequent communication between libraries can lay the foundations for the creation of common policies as well as a common approach to the problems that arise in Greek institutional repositories due to the restrictions imposed by intellectual property rights.

Future work should be targeting other types of libraries (e.g., public libraries and special libraries) that have undertaken major digitization projects and their policies on copyright clearance.

## References

- Björk, Bo-Christer. 2016. "The open access movement at a crossroad: Are the bigpublishers and academic social media taking over?". *Learned Publishing*, 29: 131-134. doi: 10.1002/leap.1021.
- Brindley, Lynne J. 2006. "Joint funding councils' libraries review group (the "Follett") report: The contribution of the information technology sub-committee". *Program*, 40(4): 311-314. <https://doi.org/10.1108/00330330610707881>.
- Chantavaridou, Elisavet. 2009. "Open access and institutional repositories in Greece: progress so far". *OCLC Systems & Services: International digital library perspectives*, 25(1): 47-59. <https://doi.org/10.1108/10650750910931922>.
- Dawson, Patricia H., and Sharon Q. Yang. 2016. "Institutional Repositories, Open Access and Copyright. What are the Practices and Implications?" *Science & Technology Libraries*, 35(4): 279-294. <https://doi.org/10.1080/0194262X.2016.1224994>.
- Derakhshan, Maryam and Diljit Singh. 2011. "Integration of information literacy into the curriculum: a meta-synthesis". *Library Review*, 60(3): 218-229. <https://doi.org/10.1108/00242531111117272>.
- Godwill, Enga Azeh. 2015. *Fundamentals of Research Methodology : A Holistic Guide for Research Completion, Management, Validation and Ethics*. [eBook]. New York: Nova Science Publishers, Inc.
- Hanlon, Ann, and Marisa Ramirez. 2011. "Asking for permission: A survey of copyright workflows for institutional repositories". *Libraries and the Academy*, 11(2): 683-702. <https://doi.org/10.1353/pla.2011.0015>.
- Macklin, Lisa A. 2013. "Copyright and Institutional Repositories". In *The institutional repository : benefits and challenges*. Retrieved from <https://pid.emory.edu/ark:/25593/r9thx>.
- Marsh, Rebecca Mary. 2015. "The role of institutional repositories in developing the communication of scholarly research". *OCLC Systems & Services: International digital library perspectives*, 31(4): 163-195. <http://dx.doi.org/10.1108/OCLC-04-2014-0022>.
- Muir, A. 2019. Copyright and digital academic library development in the UK. *Journal of Librarianship and Information Science*, 51(3), 702-709. <https://doi.org/10.1177/0961000617732380>.
- Palmer, Carole L., Lauren C. Tefteau, and Mark P. Newton. 2008. "Strategies for institutional repository development: A case study of three evolving initiatives". *Library Trends*, 57(2): 142-167. <https://doi.org/10.1353/lib.0.0033>.
- Prandoni, Claudio, Marlis Valentini and Martin Doerr. (2009). "Formalising a Model for Digital Rights Clearance". In Agosti M., Borbinha J., Kapidakis S., Papatheodorou C., Tsakonas G. (eds) *Research and Advanced Technology for Digital Libraries. ECDL 2009*. (Lecture Notes in Computer Science, vol 5714). Berlin, Heidelberg: Springer. [https://doi.org/10.1007/978-3-642-04346-8\\_32](https://doi.org/10.1007/978-3-642-04346-8_32).
- Togia, Aspasia, Eleftheria Koseoglou, Sofia Zapounidou, and Nikolaos Tsigilis. 2018. "Open Access Infrastructure in Greece: Current Status, Challenges and Perspectives". *ELPUB 2018*: 1-21. <https://doi.org/10.4000/proceedings.elpub.2018.26>.
- Yoo, Suhyeon and Hyesun Kim. 2013. "Copyright clearing methods for electronic document delivery in Korea". *Interlending & Document Supply*, 41(2): 54-58. doi: 10.1108/ILDS-02-2013-0007.

## “Yomi” (readings) in bibliographic data for materials in Japanese

Kazue Murakami<sup>(a)</sup>

a) National Diet Library, Japan

---

**Contact:** Kazue Murakami, [kazue-m@ndl.go.jp](mailto:kazue-m@ndl.go.jp)

**Received:** 20 January 2022; **Accepted:** 21 March 2022; **First Published:** 15 May 2022

---

### ABSTRACT

In bibliographic and authority data of materials in Japanese, it is necessary to represent kanji (Chinese characters) and its reading as a pair. Readings, called “yomi” in Japanese, are represented in katakana or the alphabet (romanization form). The readings in Japanese language have many variations depending on the context, so Japanese bibliographic information needs to provide both kanji and its reading as a pair to avoid misunderstandings.

The JAPAN/MARC is a machine readable version of the Japanese National Bibliography and provides sets of kanji, katakana-yomi and romaji-yomi (romanization form) of bibliographic and authority data. This paper introduces how readings of bibliographic and authority data in Japanese have been expressed and handled mainly in JAPAN/MARC, compared with international standards.

### KEYWORDS

Japanese National Bibliography; JAPAN/MARC; National Diet Library Japan; Readings of Japanese; Romanization.

## Introduction

When creating bibliographic data for materials in Japanese, it is necessary to take into account the “yomi” (readings) specific to Japanese.

The Japanese language is expressed using kanji (Chinese characters), hiragana and katakana, which were uniquely developed in Japan based on kanji. There are usually at least two readings for each kanji in Japan: “on-yomi”, which is the reading that came with the kanji from China, and “kun-yomi”, which is a translation of the meaning of the kanji into Japanese. The reading of a kanji is usually expressed using hiragana or katakana.

In this paper, an explanation of yomi in Japanese in general will be given at first. And then, it will be introduced how yomi has been expressed mainly in the bibliographic data provision format of the National Diet Library, Japan (NDL).

### 1. What is “yomi” in Japanese?

In this section, an explanation about “yomi” in Japanese language in general is given with some examples before introducing expressions in bibliographic data.

The introduction mentioned that there are usually at least two readings for kanji: on-yomi and kun-yomi. For example, the kanji “大”, which means “big”, has two on-yomi: “dai” (e.g., “重大”, pronounced as jūdai<sup>1</sup>) and “tai” (e.g., “大家”, pronounced as taika), as well as kun-yomi: ō(kii), usually written with kanji and hiragana as “大きい”.

Words with the same reading may also have different meanings depending on the kanji characters used. For instance, “大家” means a great expert, “対価” means payment, and “退化” means degeneration, but the readings for all of these are “taika”.

Furthermore, even words that combine the same kanji characters may have different meanings if they are read differently. For example, when the kanji “大家” is read as “taika”, it means a great expert as mentioned above. However, when the same word is read as “ōya”, it means a landlord. Therefore, when Japanese words are expressed in kanji, it can often be misleading unless their readings are also presented.

As for personal names, various readings are possible even when the same kanji is used, and the same reading is not always used with the same kanji. For example, the family name “三田” is read as “Mita” or “Sanda” depending on the family. The family name “Itō” uses the kanji “伊東”, “伊藤” or “井藤” depending on the family. There can be many variations of kanji and reading pairs for given names because more freedom is allowed for them than for family names.

---

<sup>1</sup> In this paper, pronunciations of Japanese words are expressed using the ALA Romanization tables. Japanese (2012), <https://www.loc.gov/catdir/cpsa/romanization/japanese.pdf>.

Table 1: One kanji with several pronunciations

Kanji	Yomi (Pronunciation/Romanization)
大	dai (だい)
	tai (たい)
	ō (おお)

Table 2: One kanji with several pronunciations (common nouns)

Kanji	Yomi (Pronunciation/Romanization)
大家	taika (たいか)
	ōya (おおや)

Table 3: One kanji with several pronunciations (proper nouns)

Kanji	Yomi (Pronunciation/Romanization)
三田	Mita (みた)
	Sanda (さんだ)

Table 4: One pronunciation with several kanji (common nouns)

Yomi	Kanji	Meaning
Taika	大家	Great expert
	対価	Payment
	退化	Degeneration

Table 5: One pronunciation with several kanji (proper nouns)

Yomi	Kanji
Itō	伊東
	伊藤
	井藤

In Japan, it is customary to include bibliographic information like a colophon at the end of a book. Since there are many variations of readings for the same kanji, the readings of the kanji for the title and the author are usually indicated to avoid misunderstandings in bibliographic information or an author's biography. These indications are useful to identify titles and creators, and to create authority data in cataloging processes.

In particular, in the process of creating authority data, even if the creator is indicated by kanji without its reading in the material, catalogers always assign its reading for the authorized access point by checking reference materials, inquiring with the publisher, or estimating the most common reading of the kanji.

In the days when people used catalog cards, bibliographic data had to be listed in a certain order. In Japan, the most common way to list catalog cards was the order of readings of titles or authors. The pronunciation of hiragana and katakana is determined on a one-to-one basis for each character, with 48 basic characters each, while kanji is said to have tens of thousands of characters. It seemed that the most efficient way to arrange catalog cards was to use hiragana or katakana to represent the reading of the author's name or book title and list them in that order. That was why this way became the mainstream.

In catalog cards, personal name authority data was listed in the order of the readings of the personal name, while the kanji form was adopted as the heading, and identification was based on the kanji form and their readings. Therefore, it is customary not to consider it to be the same person if the kanji is different, even if the reading is the same. It is also customary not to regard the same kanji as the same name if it has a different reading, since different readings of the same kanji will be listed in different places.

Most libraries use online catalogs now, so it is no longer necessary to consider the sequence of catalog cards. However, as mentioned above, the same kanji can have different meanings when it is read in different ways, and even when read in the same way, the meanings can differ depending on the kanji used. Therefore, it is necessary to be able to search by both kanji and reading for the identification of Japanese materials.

Web NDL Authorities<sup>2</sup>, which enables users to search for and acquire multiple authority data records, displays the kanji forms and katakana “yomi” forms of the authorized and variant access points, as well as the romanized “yomi” forms which replace the katakana “yomi” by alphabet. By providing both kanji and their readings, Web NDL Authorities can be used not only to identify individuals and works in combination with conventional kanji forms, but also as a basis for reading when creating reading materials for the visually impaired, for example.

## 2. Japanese National Bibliography and JAPAN/MARC

As the national bibliographic agency of Japan, the NDL has provided standard bibliographic data of domestic materials acquired by the legal deposit system and of domestic materials and Japanese materials published in foreign countries acquired by purchase and donation. The NDL also provides JAPAN/MARC as a machine-readable version of the Japanese national bibliography.

The printed “Japanese national bibliography weekly list” was a list of government publications sorted by which government agency published them, and private publications sorted by the Nippon Decimal Classification (NDC) system. The lists were in katakana-yomi order for each government agency and the NDC section, but the readings themselves were not provided. JAPAN/MARC provided katakana-yomi and romaji-yomi as well as kanji forms of bibliographies, so users became to be able to use readings for their searches and identify materials easier.

Before explaining how readings are provided in JAPAN/MARC, this section provides an overview of the history of JAPAN/MARC, and the next section will explain some Japanese romanization systems.

---

<sup>2</sup> <https://id.ndl.go.jp/auth/ndla>.

At first, the NDL launched JAPAN/MARC(M) for monographs which provided bibliographies of books in 1981. It then started providing bibliographies of serials in 1988. In 2014, the two were merged into a single entity and renamed JAPAN/MARC (M/S). In 1997, the NDL also started providing JAPAN/MARC for authorities of authors' names. It was renamed JAPAN/MARC(A) in 2003. JAPAN/MARC for monographs, serials and authorities were all based on UNIMARC. Although there have been some changes such as addition of fields, the unification of formats for books (monographs) and serials in 2002, and the expansion of scope to include cartographic materials, audio-visual materials, etc. after 2003, the stance of complying with UNIMARC was not changed until the JAPAN/MARC MARC 21 format was implemented.

Since January 2012, both bibliographic data and authority data have been provided in the JAPAN/MARC MARC21 format. Currently, new, updated, and deleted data files are posted on the website<sup>3</sup> every week, and bibliographic data retrieved from the National Diet Library Search (NDL Search)<sup>4</sup> can be downloaded in MARC format.

### 3. Japanese romanization systems

Some examples will be given to illustrate the difference between the ALA-LC Romanization Table and the NDL for the Romanization of Japanese.

The ALA-LC Romanization Table uses macrons (ˉ) for the Japanese vowels a, i, u, e, and o to represent long vowels such as “ka a” in “o ka a sa n”, which means mother. The NDL used circumflexes (ˆ) for representing long vowels until March 2002. However, since April 2002, the NDL has not used any symbol such as a macron or circumflex. The example “o ka a sa n” is written as “okāsan” according to the ALA-LC Romanization Table, but in the NDL rules until March 2002, it was written “okâsan”. In April 2002, it was decided to convert the katakana reading “オカアサン” into Roman characters one by one and write it as “okaasan”. Since December 2011, the NDL has decided to represent this as “okasan” instead of converting katakana characters one by one. In the bibliographic data and authority data currently available by the NDL, all romaji-yomi are provided with the rules after December 2011.

In addition, there are two main types of Japanese Romanization: the Hepburn system, which emphasizes compliance with English pronunciation, and the Kunrei system, which is Japanese Cabinet-ordered. For example, “ジリツ” which means to control oneself, is written as “jiritsu” according to the ALA-LC Romanization Table, which is based on the Hepburn system. On the other hand, it is written as “ziritu” according to the Kunrei system. The NDL's bibliographic and authority data generate romaji-yomi from katakana-yomi systematically. The NDL began to record romaji-yomi based on the Hepburn system instead of the Kunrei system in December 2011.

For people who cannot read Japanese, romanization can be useful. However, as the NDL uses different rules for romanization depending on the time period, there are many different ways of romanizing Japanese. It is recommended to use several notations such as the Kunrei and Hepburn systems when searching bibliographic or authority data by romaji reading.

<sup>3</sup> [https://www.ndl.go.jp/jp/data/data\\_service/jnb\\_product.html](https://www.ndl.go.jp/jp/data/data_service/jnb_product.html).

<sup>4</sup> <https://iss.ndl.go.jp/>.

## 4. Romanization in JAPAN/MARC

The JAPAN/MARC format has been designed to allow the expression of kanji form and reading form as a set, because, as mentioned in section 1, the ability to search by both kanji and reading forms is necessary for the identification of Japanese materials.

In this section, how the JAPAN/MARC format has recorded the kanji and reading forms in the fields will be explained, comparing them with the international standards to which the format adheres.

### 4.1 JAPAN/MARC (M) and (S) before 2011

In JAPAN/MARC (M), which was launched in 1981, the description blocks (2XX-3XX) were written in kanji characters, and the access point blocks (55X) contained a set of kanji and readings: katakana-yomi, and romaji-yomi forms. The format of JAPAN/MARC (M) was based on UNIMARC, however, its used some fields which were not specified in UNIMARC. For example, in UNIMARC (1977), the title and responsibility were recorded in field 200, but in JAPAN/MARC (M), the title and responsibility were recorded in field 251, which is not specified in UNIMARC. If there were more than one title, fields 252-259 were used. The edition statement and publication area were also recorded using fields not specified by UNIMARC. 5XX is the block for recording related titles, while JAPAN/MARC (M) used fields 551-559, which were also not specified by UNIMARC. \$B for the kanji form, \$A for katakana-yomi form, and \$X for the romaji-yomi form were in the same field and in 2-byte characters, so that kanji and reading could be expressed as a set. Figure 1 shows that the title proper “函館・ロシアその交流の軌跡” was in 251\$A, its katakana-yomi form “ハコダテ ロシア ソノ コウリュウ ノ キセキ” was in title heading (551\$A) and its romaji-yomi “Hakodate rosia sono kouryuu no kiseki” was in 551\$X. By recording “251A1” in 551\$B, it indicated that 551\$A and \$X were paired with 251\$A.

This rule was also implemented in JAPAN/MARC (S) and JAPAN/MARC (M/S).

Title proper and statement of responsibility	251	\$A	函館・ロシアその交流の軌跡
		\$F	清水恵//著#
Parallel title proper	261	\$A	Хакодате Русские страницы в истории японского города#
Publication, distribution, etc.	270	\$A	[函館]
		\$B	函館日口交流史研究会
		\$D	2005.12#
Physical description	275	\$A	387p
		\$B	20cm#
Title heading	551	\$A	ハコダテ ロシア ソノ コウリュウ ノ キセキ
		\$X	Hakodate rosia sono kouryuu no kiseki
		\$B	251A1
		\$A	Khakodatz Russkie stranitsy v istorii iaponskog o goroda
		\$X	Khakodatz Russkie stranitsy v istorii iaponskog o goroda
		\$B	261A1#

Figure 1: A partial example of JAPAN/MARC (M/S) format (2009)



#### 4.2 JAPAN/MARC (A) before 2011

In “JAPAN/MARC for authorities of authors’ names”, which was launched in 1996, the same field was repeated for each kanji, katakana-yomi, and romaji-yomi form. To indicate that it was a single set, a code indicating the relationship between the fields and a linking number to match the linked fields were recorded in \$6. This method was based on that specified in UNIMARC/Authorities (1991). Figure 2 shows field 210 (heading - corporate name) was repeated: the kanji form “国立情報学研究所” was included in \$a without \$7; the katakana-yomi form “コクリツ ジョウホウガク ケンキュウジヨ” was included \$a in the field which has “dc” (Japanese - Kana) in \$7; and the romaji-yomi form “Kokuritu zyouhougaku kenkyuuzyo” was included in \$a in the field where \$7 is “ba” (Latin character). Each \$6 represented that it had a relation of different character types with the same content (“a”) and was a single set (with the same linking number “01”).

Heading - Corporate name	210	02	\$6	a01
			\$a	国立情報学研究所#
	210	02	\$6	a01
			\$7	dc
			\$a	コクリツ ジョウホウガク ケンキュウジヨ#
	210	02	\$6	a01
			\$7	ba
			\$a	Kokuritu zyouhougaku kenkyuuzyo#

Figure 2: A partial example of JAPAN/MARC (A)

#### 4.3 JAPAN/MARC UNIMARC format

Since the second edition of the UNIMARC Manual (1994), it has been possible to record a set of kanji, katakana-yomi and romaji-yomi forms using \$6 and \$7 in the bibliographic format as well as in UNIMARC/Authorities. However, the NDL did not make any major changes to the JAPAN/MARC bibliographic format for domestic distribution of bibliographic data. On the other hand, the JAPAN/MARC UNIMARC version was developed for international distribution and started to be provided in January 2003.

In the JAPAN/MARC UNIMARC format, the fields specified in the UNIMARC format were used in principle. For example, title and responsibility were recorded in field 200 instead of 251. As in JAPAN/MARC (A), the character type codes were recorded in \$7 for katakana-yomi and romaji-yomi forms, and they were recorded as a single set with kanji forms without \$7.

Figure 3 shows that field 225 (series title) was repeated, and each form was recorded in \$a: kanji form “物理学” without \$7, katakana-yomi form “ブツリガク” with “dc” (Japanese - Kana) in \$7, and romaji-yomi form “Buturigaku” with “ba” (Latin character) in \$7. Each \$6 represents that it has a relation of different character types with the same content (“a” in \$6) and one set (with the same linking number “01” in \$6). Figure 3 also shows that “一般相対性理論”, “イッパン ソウタイ セイ リロン”, and “Ippan soutaisei riron” are one set (with the same linking number “02” in \$6).

Series title	225	1#	\$6	a01
			\$a	物理学
	225	1#	\$6	a01
			\$7	dc
			\$a	ブツリガク
	225	1#	\$6	a01
			\$7	ba
			\$a	Buturigaku
	225	1#	\$6	a02
			\$a	一般相対性理論
	225	1#	\$6	a02
			\$7	dc
			\$a	イッパン ソウタイセイ リロン
	225	1#	\$6	a02
			\$7	ba
\$a			Ippan soutaisei riron	

Figure 3: A partial example of JAPAN/MARC UNIMARC format

The NDL provided bibliographic data to OCLC using this JAPAN/MARC UNIMARC format from November 2010 to December 2011.

#### 4.4 JAPAN/MARC after 2012: JAPAN/MARC MARC 21 format

The NDL implemented the MARC 21 format for both bibliographic and authority data in January 2012. That led to a major change in the JAPAN/MARC format, including the fields for recording readings. The JAPAN/MARC MARC 21 format uses a set of field 880 and other fields (i.e. regular fields) to record kanji form, katakana-yomi form, and romaji-yomi form according to Appendix D “Multiscript Records”<sup>5</sup> of the MARC 21 format. In the JAPAN/MARC MARC21 format, a set of field 880 and regular fields form a set of kanji and its reading forms. In \$6, the field number of the set, the occurrence number for matching the field to be linked, and in the case of field 880, the character type code are recorded to represent the set of kanji and reading forms. However, unlike the example given in Appendix D, the kanji form is recorded in the regular field, and the katakana-yomi form and romaji-yomi form are recorded in field 880. In Figure 4, the \$6 in field 245 (title statement) is a single set with field 880 that has occurrence number “01”, and in field 880, “245-01” in the \$6 represents that this field 880 is a single set with the field 245 that has the occurrence number “01”, followed by recording the character type codes “\$1” (Chinese, Japanese, Korean) and “(B)” (Latin). Figure 4 shows the kanji form of 245\$a (title), “理論物理に潜む部分多様体幾何”, the katakana-yomi form of 880\$a, “リロン ブツリ ニ ヒソム ブブン タヨウタイ キカ”, and the romaji-yomi form of 880\$a, “Riron butsuri ni hisomu bubun tayotai kika”, represent a single set.

<sup>5</sup> <https://www.loc.gov/marc/bibliographic/ecbdmulti.html>

Title Statement	245	00	\$6	880-01
			\$a	理論物理に潜む部分多様体幾何 =
			\$b	Submanifold Geometry Hidden in Theoretical Physics : 一般相対性理論・ゲージ理論との関わり /
			\$c	小池直之 著
			\$0	001285136
Alternate Graphic Representation	880	00	\$6	245-01/\$1
			\$a	リロン ブツリ ニ ヒソム ブブン タヨウタイ キカ :
			\$b	イッパン ソウタイセイ リロン ゲージ リロン トノ カカワリ
Alternate Graphic Representation	880	00	\$6	245-01/(B
			\$a	Riron butsuri ni hisomu bubun tayotai kika :
			\$b	Ippan sotaisei riron geji riron tonon kakawari

Figure 4: A partial example of JAPAN/MARC MARC 21 format

Since January 2012, the NDL has been providing bibliographic and authority data to OCLC in the JAPAN/MARC MARC 21 format, instead of the JAPAN/MARC UNIMARC format.

## 5. Representation of readings other than the MARC format

In addition to the MARC format, the NDL provides bibliographic and authority data in various formats. The NDL Search system provides bibliographic data not only in the MARC format and MARC tag format, but also DC-NDL (RDF) format and JSON format.

The DC-NDL format is a set of metadata description elements and rules based on the Dublin Core (DC), with some additional NDL-specific vocabulary. “dcndl:transcription” is the NDL’s own vocabulary for recording readings. For example, a set of kanji and katakana-yomi of a title can be expressed by recording the kanji form in <rdf:value> and the katakana-yomi form in <dcndl:transcription> in <dc:title>. Neither DC-NDL (RDF) format nor JSON format provides romaji-yomi form.

```
<dcterms:title>理論物理に潜む部分多様体幾何 = Submanifold Geometry Hidden in Theoretical Physics : 一般相対性理論・ゲージ理論との関わり</dcterms:title>
<dc:title>
  <rdf:Description>
    <rdf:value>理論物理に潜む部分多様体幾何 = Submanifold Geometry Hidden in Theoretical Physics : 一般相対性理論・ゲージ理論との関わり</rdf:value>
    <dcndl:transcription>リロン ブツリ ニ ヒソム ブブン タヨウタイ キカ : イッパン ソウタイセイ リロン ゲージリロン トノ カカワリ</dcndl:transcription>
  </rdf:Description>
</dc:title>
```

Figure 5: A partial example of DC-NDL (RDF) format

title:

0:

value: “理論物理に潜む部分多様体幾何 = Submanifold Geometry Hidden in Theoretical Physics : 一般相対性理論・ゲージ理論との関わり”  
transcription: “リロンブツリニヒソム ブブンタヨウタイキカ : イッパンソウタイセイ  
リロンゲージリロントノカカワリ”

Figure 6: A partial example of JSON format

The Web NDL Authorities service provides authority data in RDF/XML format, RDF/Turtle format, and JSON-LD format. In the Web NDL Authorities, <dcndl:transcription> is used in RDF/XML format as well as in the NDL Search, however, it is used with language attributes so that romaji-yomi form can be expressed in addition to katakana-yomi forms. For example, in <skosxl:prefLabel>, the kanji form is recorded in <xl:literalForm>, the katakana-yomi form is recorded in <dcndl:transcription> with `xml:lang="ja-Kana"`, and the romaji-yomi form with `xml:lang="ja-Latn"`. Turtle format and JSON-LD format also use language attributes to express kanji form and katakana-yomi form as a single set.

```
<xl:prefLabel>  
  <rdf:Description>  
    <xl:literalForm>小池, 直之</xl:literalForm>  
    <ndi:transcription xml:lang="ja-Kana">コイケ, ナオユキ</ndi:transcription>  
    <ndi:transcription xml:lang="ja-Latn">Koike, Naoyuki</ndi:transcription>  
  </rdf:Description>  
</xl:prefLabel>
```

Figure 7: A partial example of RDF/XML format in Web NDL Authorities

## 6. Japanese “yomi” in OCLC’s WorldCat


As mentioned above, the NDL has been providing JAPAN/MARC to OCLC since November 2010. The NDL has also been providing JAPAN/MARC (A) to VIAF, Virtual International Authority File, since October 2012.

In OCLC’s WorldCat, the kanji and romaji-yomi forms are displayed together for the title, author, and publisher, as shown in Figure 8.

**重要文化財慶應義塾図書館保存修理工事報告書. 本編 /  
 Jūyō bunkazai Keiō gijuku toshokan hozon shūri kōji hōkokusho. Honpen**

Author: 文化財保存計画協会編 ; ; [Bunkazai hozon keikaku kyōkai](#)

Publisher: 慶應義塾, [Tōkyō] : Keiō gijuku, 2019.9

Edition/Format:  Print book : Japanese [View all editions and formats](#)

Rating: ☆☆☆☆☆ (not yet rated) [0 with reviews - Be the first.](#)

Subjects: [Keio Gijuku Toshokan.](#)  
[Toshokan kenchiku -- Hozon Shufuku.](#)

More like this [Similar Items](#)

Figure 8: An example of WorldCat's display. ©2022 OCLC, Inc. This Screenshot is used with OCLC's permission. WorldCat is a registered trademark/service mark of OCLC.

Since WorldCat receives data from various organizations around the world, there can be many variations in the romanization of Japanese. The variations in long vowel expressions as described in Section 3 can be found in WorldCat. For example, the romanization of the title “重要文化財慶應義塾図書館保存修理工事報告書” in Figure 8 is indicated as “Jūyō bunkazai Keiō gijuku toshokan hozon shūri kōji hōkokusho” with macrons (̄). On the other hand, the bibliographic data in the JAPAN/MARC MARC 21 format represents the romaji-yomi form without any long vowel symbols, “Juyo bunkazai keio gijuku toshokan hozon shuri koji hokokusho” (Figure 9).

880	00	\$6	245-01/\$1
		\$a	ジュウヨウ ブンカザイ ケイオウ ギジユク トシヨカン ホゾン シュウリ コウジ ホウコクシヨ.
		\$n	ホンペン
880	00	\$6	245-01/(B
		\$a	Juyo bunkazai keio gijuku toshokan hozon shuri koji hokokusho.
		\$n	Honpen

Figure 9: A partial example of JAPAN/MARC MARC21 format

Also, searching the word given as an example in Section 3, “お母さん,” “okaasan,” and “okasan”, the search results are different for each. As of January 17, 2022, a keyword search on WorldCat for “お母さん” yields about 4,671 results, “okaasan” yields about 2,418 results, and “okasan” yields about 4,171 results.

To help deal with the differences in romaji-yomi expression, it would be desirable to be able to handle kanji and its readings as sets in overseas bibliographic data as well.

## 7. A study on the representation of Japanese “yomi” in BIBFRAME

The “BIBFRAME 2.0 to MARC 21 Conversion Specifications”<sup>6</sup> published by the Library of Congress in 2020 uses the MARC 21’s Appendix D, Model B without transliterations for multilingual notation. In the “MARC 21 to BIBFRAME 2.0 Conversion Specifications”<sup>7</sup> published in 2021, only for title statement (field 245), paired non-Latin characters and transliterations are converted using xml language codes to maintain correspondence<sup>8</sup>. The title “重要文化財慶應義塾図書館保存修理工事報告書” will be represented as shown in Figure 10.

```
<bf:title>
  <bf:Title>
    <bf:mainTitle xml:lang="ja-Latn">Juyo bunkazai keio gijuku toshokan hozon shuri koji hokokusho. Hon-
pen</bf:mainTitle>
    <bf:mainTitle xml:lang="ja-Latn">ジュウヨウ ブンカザイ ケイオウ ギジユク トショカン ホゾン シュウリ
コウジ ホウコクショ. ホンペン</bf:mainTitle>
    <bf:mainTitle xml:lang="ja-Jpan">重要文化財慶應義塾図書館保存修理工事報告書. 本編</bf:mainTitle>
    <rdfs:label>重要文化財慶應義塾図書館保存修理工事報告書. 本編</rdfs:label>
  </bf:Title>
</bf:title>
```

Figure 10: A predicted partial example of BIBFRAME in RDF/XML format

## 8. Next steps

As mentioned above, both the kanji form and its reading form are necessary for identification of Japanese entities. The NDL has been providing sets of the kanji, katakana-yomi and romaji-yomi forms in the bibliographic data and authority data of Japanese materials.

The “Plan for Creating and Providing the National Diet Library’s Bibliographic Data 2021–2025,” a five-year plan for bibliographic data formulated by the NDL in March 2021, has supporting the new bibliographic framework as one of its items. The NDL plans to study the possibility of providing bibliographic and authority data based on BIBFRAME, which is becoming popular overseas as a new bibliographic framework that can realize the functions of linked data.

The conversion proposed by the Library of Congress only converts non-Latin characters, except for some fields such as the main title. Since the NDL needs to provide both the kanji form and its reading form not only for the title proper but also for parallel titles, other title information, etc., it is necessary to examine the specifications needed to represent the kanji form, katakana-yomi forms, and romaji-yomi forms as a single set, including the use of language codes or script codes, or using original vocabulary such as <dcndl:transcription>.

<sup>6</sup> <https://www.loc.gov/bibframe/bftm/>

<sup>7</sup> <https://www.loc.gov/bibframe/mtbf/>

<sup>8</sup> “MARC 21 to BIBFRAME 2.0 Conversion Specifications Ver.1.7” (2021-10-21) is available in GitHub. It allowed converting fields 250, 26X, and 490 as well as 245, keeping the correspondence between non-Latin characters and transliterated forms. <https://github.com/lcnetdev/marc2bibframe2/tree/master/spec>

It is necessary to provide bibliographic and authority data to many more users by supporting the new bibliographic framework, while continuing to support existing JAPAN/MARC users. No matter what bibliographic framework is used, it is necessary to be able to record a set of kanji, katakana-yomi, and romaji-yomi forms for Japanese bibliographic and authority data. The NDL will continue to study the possibility of providing bibliographic data based on BIBFRAME.

## References

- Agenbrood, James E. BA, and MLS. 2006. "Romanization Is Not Enough." *Cataloging & Classification Quarterly* 42:2, 21-34. Accessed March 9, 2022. doi: 10.1300/J104v42n02\_0.
- Aihara, Nobuya, and Naito, Eisuke. 1997. "CJK(中国語/日本語/韓国語)環境下での国際標準件名標目システムの適用 = Adaptation of Universal Standardized Subject Heading System in the CJK (Chinese / Japanese / Korean) Environment." *学術情報センター紀要 = Research bulletin of the National Center for Science Information Systems* 9:73-81. Accessed March 9, 2022. [https://www.nii.ac.jp/journal/bulletins/pdf/bulletin\\_9.pdf](https://www.nii.ac.jp/journal/bulletins/pdf/bulletin_9.pdf).
- Barešová, Ivona. 2016. *Japanese given names : a window into contemporary Japanese society First edition*. Olomouc : Palacký University Olomouc (Monographs).
- El-Sherbini, Magda, & Chen, Sherab. 2011. "An Assessment of the Need to Provide Non-Roman Subject Access to the Library Online Catalog." *Cataloging & Classification Quarterly* 49:6, 457-483. Accessed March 9, 2022. doi: 10.1080/01639374.2011.603108.
- Hadamitzky, Wolfgang, and Spahn, Mark. [2011]. *Japanese Kanji and Kana : a guide to the Japanese writing system*. Tokyo : Tuttle Publishing.
- Hasegawa, Yoko. 2015. *Japanese : a linguistic introduction*. Cambridge : Cambridge University Press.
- Edited by Hopkinson, Alan. 2008. *UNIMARC Manual Bibliographic Format. 3rd edition*. München: Saur (IFLA Series in Bibliographic Control, 36).
- Iwasaki, Shoichi. [2013]. *Japanese. Revised edition*. Amsterdam : John Benjamins Publishing Company (London Oriental and African language library ; volume 17).
- KAMEI, Haruko. 1997. "欧米のCJKシステム : OCLCを中心に = China and Chinese character (Kanji) : to know about Asia 2. CJK systems in the United States: OCLC CJK." *情報の科学と技術 = The journal of Information Science and Technology Association* 47:7, 334-339. Accessed March 9, 2022. [https://doi.org/10.18919/jkg.47.7\\_334](https://doi.org/10.18919/jkg.47.7_334).
- Kimura, Maiko. 2015. "A Comparison of Recorded Authority Data Elements and the RDA Framework in Chinese Character Cultures." *Cataloging & Classification Quarterly* 53:7, 753-784. Accessed March 9, 2022. doi: 10.1080/01639374.2014.977984.
- Kudo, Yoko. 2011. "Modified Hepburn Romanization System in Japanese Language Cataloging: Where to Look, What to Follow." *Cataloging & Classification Quarterly* 49:2, 97-120. Accessed March 9, 2022. doi: 10.1080/01639374.2011.536751.
- LI Chenying ; XIE Qinfang ; YU Shuangshuang ; HAN Mingjie. 2006. "CALIS日本語書誌作成と典拠コントロールの全面的実現—CALIS共同目録のオンライン典拠コントロールシステムについて— = Implementation of Japanese Literature Cataloging and Authority Control : The Online Authority Control System for the CALIS Union Catalog." *大学図書館研究 = Journal of college and university libraries* 77, 51-59. Accessed March 9, 2022. <https://doi.org/10.20722/jcul.1218>.



MATSUI Junko. 2018. “書誌コントロールと図書館目録 = Bibliographic control and OPAC: on-line public access catalog.” *図書館界* = *The library world* 70:1, 287-304. Accessed March 9, 2022. [https://doi.org/10.20628/toshokankai.70.1\\_287](https://doi.org/10.20628/toshokankai.70.1_287).

Naito, Eisuke. 2004. “Names of the Far East: Japanese, Chinese, and Korean Authority Control.” *Cataloging & Classification Quarterly* 38:3-4, 251-268. Accessed March 9, 2022. doi: 10.1300/J104v38n03\_19.

NAKAI, Machiko. 2007. “日本の全国書誌サービス—その歩みと展望 = National bibliographic service in Japan—History and prospects.” *情報管理* = *Journal of information processing and management* 50:4, 193-200. Accessed March 9, 2022. <https://doi.org/10.1241/johokanri.50.193>.

National Diet Library. 2003. *JAPAN/MARC Manual. Authorities. First edition*. Tokyo: National Diet Library.

National Diet Library. 2002. *JAPAN/MARC Manual. Monographs & Serials. First edition*. Tokyo: National Diet Library.

Edited by Plassard, Marie-France. 2001. *UNIMARC Manual : authorities format. 2nd revised and enlarged edition*. München: Saur (UBCIM Publications ; New Series, Vol. 22).

Takawashi, Tadayoshi BE. 2002. “Cataloging in Japan: Relationship Between Japanese and Western Cataloging Rules.” *Cataloging & Classification Quarterly* 35:1-2, 209-225. Accessed March 9, 2022. doi: 10.1300/J104v35n01\_13.

TATUOKA, Hiroshi, and YAMADA, Hisao. 1996. “On the Theory and Practice of the Romanized Writing System for Japanese = 日本語のローマ字書き体系の理論と実状について.” *学術情報センター紀要* = *Research bulletin of the National Center for Science Information Systems* 8:27-74. Accessed March 9, 2022. [https://www.nii.ac.jp/journal/bulletins/pdf/bulletin\\_8.pdf](https://www.nii.ac.jp/journal/bulletins/pdf/bulletin_8.pdf).

Taylor, Insup, Taylor, M. Martin. [2014]. *Writing and literacy in Chinese, Korean and Japanese. Revised edition*. Amsterdam : John Benjamins Publishing Company (Studies in written language and literacy ; volume 14).

Tokita, Takuya, Kimura, Maiko, Miyata, Yosuke, Yokoyama, Yukio, Taniguchi, Shoichi, & Ueda, Shuichi. 2012. “Identifying Works of Japanese Classics for Construction of FRBRized OPACs.” *Cataloging & Classification Quarterly* 50:5-7, 670-687. Accessed March 9, 2022. doi:10.1080/01639374.2012.681605.

Yanaike, Makoto. 2009. “現代日本語の音・訓読み分けの機構を論じ、「漢語・和語形態素の相補的分布」に及ぶ = The mechanism of the discrimination between On-yomi and Kun-yomi of Kanji in present-day Japanese.” *藝文研究* = *The geibun-kenkyu : journal of arts and letters* 96:75(202)-95(182). Accessed March 9, 2022. [https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara\\_id=AN00072643-00960001-0095](https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00072643-00960001-0095).

## Others' Books, Catalogues of our own: non-roman scripts in SBN

Oscar Nalesini

Contact: Oscar Nalesini, [oscar.nalesini@gmail.com](mailto:oscar.nalesini@gmail.com)

Received: 24 January 2022; Accepted: 22 February 2022; First Published: 15 May 2022

### ABSTRACT

The catalogue of the Italian National Library System (SBN) is a roman-script-only database. Local libraries, on the other hand, are facing requests from an increasing number of users from Asia and Africa, and are filling the gaps of the national system by developing isolated, non-collaborative ways to manage resources in non-roman scripts.

Though a couple of Institutions have recently, but independently each other, developed projects to catalogue resources in native scripts, it is still impossible to predict when this feature will be finally implemented. In the next future, Italian cataloguers will therefore still rely on romanization only. After having briefly reviewed the romanization methods recommended by the Italian cataloguing rules (*REICAT*), the article suggests to expand its repertoire by establishing a method for identifying the most suitable romanization standards.

### KEYWORDS

Non-roman scripts; Multilingualism; Cataloguing standard; Multiscript cataloguing; Special characters.

## Libri d'altri, cataloghi di noialtri: le scritture non latine in SBN

### ABSTRACT

Il catalogo del Servizio Bibliotecario Nazionale accetta solamente i caratteri latini, e una parte di quelli greci. Molte biblioteche locali devono invece erogare servizi a un numero crescente di utenti dall'Asia e dall'Africa, le cui lingue usano altri sistemi di scrittura. Queste biblioteche sono pertanto costrette a tamponare le lacune del sistema nazionale sviluppando in proprio modalità autonome e non partecipate di gestione del patrimonio librario in scritture non latine.

Benché un paio di istituzioni abbiano recentemente iniziato lo sviluppo per catalogare risorse nelle scritture originali, è impossibile predire quando questa caratteristica sarà effettivamente disponibile. Nel prossimo futuro, pertanto, i catalogatori italiani dovranno affidarsi solo alla latinizzazione. Dopo aver brevemente passato in rassegna i metodi di latinizzazione prescelti dalle Regole italiane di catalogazione (*REICAT*), l'articolo suggerisce di espandere quel repertorio predisponendo un metodo che permetta di identificare gli standard di latinizzazione più adatti all'attuale funzionamento di SBN.

### PAROLE CHIAVE

Scritture non latine; Multilinguismo; Standard di catalogazione; Catalogazione in più scritture; Caratteri speciali.

Ambizione delle seguenti note è rinverdire la discussione su un tema pressoché sopito in Italia: la catalogazione delle risorse in scritture non latine, o, per meglio dire, in caratteri non gestiti dal Servizio Bibliotecario Nazionale.<sup>1</sup> Varie sollecitazioni sul tema sono giunte nell'ultimo ventennio da chi aveva compreso come la maggiore facilità di comunicazione tra persone e istituzioni poste in ogni angolo del pianeta, unitamente ai profondi e irreversibili mutamenti demografici, socio-economici e culturali in atto nel nostro paese, stessero portando molte biblioteche ad amministrare risorse che un tempo erano appannaggio pressoché esclusivo di quelle maggiori, o specialistiche (Rabitti 2003; Azzanutto *et al.* 2015). Eppure, le direttive e gli strumenti forniti dagli organi preposti alla regolamentazione del sistema bibliotecario sono – a mio avviso – ancora carenti, perché figli di una situazione superata. Il problema nasce insomma da un distacco creatosi tra le dinamiche reali del paese e le infrastrutture che forniscono l'accesso alla conoscenza.

## La scrittura originale: orizzonte (quanto?) futuro

Valuteremo meglio questo distacco ricordando che già negli anni Ottanta del XX sec. si discuteva della catalogazione automatizzata di opere in arabo, ebraico, lingue slave, cinese e giapponese utilizzando le scritture originali (Lazinger, Levi 1996; Vernon 1996, 3–8; Ting 1985; Battaglini 1999). Un decennio dopo, grazie all'introduzione della codifica univoca dei segni appartenenti a (potenzialmente) qualunque sistema grafico con valenza semiotica, ovvero UNICODE, e la conseguente possibilità di trasmettere elettronicamente informazioni nella grafia originale senza ambiguità, ci si volse anche allo scambio dei dati tra biblioteche (Eilts 1996), per toccare infine questioni etiche, come l'eguale accesso alle informazioni da parte degli utenti che utilizzano scritture differenti (Agenbroad 2006, 29–30; Rigby 2015, 624).

Sono gli anni, non a caso, in cui l'IFLA approfondiva la riflessione sulle biblioteche multiculturali, giungendo a indicare tra i punti del suo *Manifesto* del 2006 la necessità di “reflect, support and promote cultural and linguistic diversity”, e di “provide information in appropriate languages and scripts”, dove per “appropriate scripts” intendeva evidentemente anche le scritture diverse da quella/e ufficialmente adottata/e dalla biblioteca.

Da allora varie realtà, in tutti i continenti, hanno realizzato sistemi catalografici capaci di utilizzare molteplici scritture. Sebbene ancora si discuta sulla gestione dei nomi propri e dei soggetti in altre lingue e scritture (El-Sherbini 2016; Cohen 2017), e per varie ragioni il novero delle scritture gestite sia spesso circoscritto, è possibile sunteggiare i vantaggi ottenuti dalle biblioteche che hanno potuto imboccare questo percorso:

- I. aumentata fruizione del patrimonio, grazie alla crescente interconnessione tra i sistemi bibliotecari, e l'interrogazione dei cataloghi da altri paesi o utilizzando altre scritture (LoC 2011; Cohen 2017);
- II. maggiore stabilità dei dati, perché la grafia originale rimane immutata nel tempo. I sistemi di latinizzazione sono invece soggetti a revisione periodica, e talvolta a modifiche sostanziali, che richiedono di riallineare il progresso, sopportandone i relativi costi, oppure, più

---

<sup>1</sup> Ho attinto dal sapere di Francesca Cappelli, Monica Genesin, Joachim Matzinger, Gabriele Mazzitelli, Federica Olivotto, Patrizia Tersigni, Argentino Trombin, che ringrazio sentitamente. Unicamente mia è la responsabilità di quanto qui esposto.

comunemente, di adottare le nuove versioni solo per il futuro. In tal caso, gli utenti dovranno ripetere le ricerche utilizzando di volta in volta una diversa regola per quella determinata lingua, oppure i catalogatori facilitarli inserendo svariate forme di rinvio;

- III. maggiore affidabilità delle informazioni, perché visualizzarle come effettivamente appaiono sul documento aiuta a dissipare eventuali equivoci dovuti alla latinizzazione.<sup>2</sup> A scopo esemplificativo ho ricercato nell'OPAC SBN il diffuso cognome cinese Liu. Ho trovato (febbraio 2022) 508 nomi. Alcuni di loro saranno anche duplicati, ma il numero effettivo, sebbene non precisamente determinabile con un'analisi talmente frettolosa, è comunque di svariate centinaia. Da noi questi Liu si scrivono tutti uguali, ma in cinese no; abbiamo, quanto meno, 刘 (forma semplificata di 劉), 柳, 留, 六, 侶. Facile comprendere come l'impossibilità di verificare la grafia originale favorisca l'errore nell'attribuzione dell'autore (Chan, Hu, Lo 2000), e generi una risposta più rumorosa del necessario.

Non è un segreto che SBN fatichi a tenere il passo sotto questo profilo. L'ICCU ha approntato nel 2018 un progetto per introdurre la gestione delle scritture non latine, ma ha dovuto rinviare la realizzazione "per lasciare spazio ad interventi giudicati prioritari" (AreaSBN 2019). Una sospensione di durata imprecisata, dettata dunque dalle odierne difficoltà nel finanziare gli sviluppi tecnologici; difficoltà però aggravate dal ritardo con cui – per ragioni che ritengo prevalentemente culturali – si è prestata attenzione a questo aspetto della catalogazione rispetto ad altri paesi.

Nel 2019, l'Università 'Sapienza' di Roma ha avviato il progetto DREAM, per creare una piattaforma dedicata alla catalogazione in scritture non latine, indipendente da (dunque alternativa a) SBN; il progetto è però ancora in fase di studio (Fallerini *et al.* 2022).

Nel frattempo, molte biblioteche inserite nel Servizio Nazionale, ma decise a fornire comunque il servizio a quella parte – crescente – dei propri utenti che oscilla "tra due poli opposti dei bisogni di lettura: [...] leggere libri nella propria lingua materna, [...] leggere nella lingua del Paese in cui si vive" (Ongini 2012, 18–19), hanno escogitato soluzioni particolari; col risultato di mantenere, spesso, quei titoli solo nel catalogo locale.

Alcune, per cominciare, hanno creato cataloghi separati per le lingue più richieste.<sup>3</sup> Le comunali di Genova e Torino hanno invece catalogato opere russe e arabe inserendo i titoli anche in grafia originale, ma necessariamente in campi inappropriati, visualizzabili solo sull'OPAC locale. La biblioteca 'Lazzerini' di Prato ha poi corredato le schede delle risorse in arabo, cinese, russo, ucraino e urdu con file testuali o immagine, ove si legge la descrizione nella scrittura originale; non potendo l'Indice accogliere allegati, anche queste descrizioni sono consultabili solo sul catalogo locale, e comunque da cercarsi in forma latinizzata.<sup>4</sup> Potrei continuare (cfr. Ciccarello 2009, 27 e note), ma quanto detto evidenzia a sufficienza come la carenza di strumenti comuni favorisca lo sviluppo di modalità individuali e disomogenee di catalogazione, che in futuro facilmente costringeranno quelle biblioteche a convertire le informazioni in altri formati, o a perderle. Una situazione, in sostanza, dispersiva e antieconomica, a cui spero si possa presto rimediare.

<sup>2</sup> Ritengo il termine 'romanizzazione', benché invalso nell'uso, un anglicismo superfluo.

<sup>3</sup> Per le comunali di Modena e Vicenza vedi: <https://www.comune.modena.it/biblioteche/catalogoarabo/>; [https://www.bibliotecabertoliana.it/it/servizi/libri\\_in\\_lingua\\_araba](https://www.bibliotecabertoliana.it/it/servizi/libri_in_lingua_araba).

<sup>4</sup> Istituto culturale e di documentazione "Lazzerini" - Prato Raccolte interculturali <https://www.bibliotecalazzerini.prato.it/it/interculturala/libri-in-lingua/pagina2023.html>.

## La scrittura greca, ovvero l'anatra zoppa

All'esclusività della scrittura latina in SBN v'è un'eccezione: la greca. Le ragioni sono ovvie; eppure, a dispetto del prestigio di cui gode la letteratura ellenica, quello messo a disposizione dei bibliotecari è uno strumento spuntato. Il sistema, mentre ammette caratteri arcaici di scarso interesse bibliografico, ignora lettere fondamentali quali il *sigma* finale e le vocali con spiriti e accenti (AresBN 2019). Le soluzioni adottate da chi ha comunque voluto catalogare in greco aggirando l'ostacolo sono varie: scrivere in tutto maiuscolo (BVE\0748078: Η ΔΗΜΟΚΡΑΤΙΑ ΤΗΣ ΠΑΤΑΤΑΣ), utilizzare il *sigma* ordinario (MIL\0896462: Κρισισ του φιλοσοφικου διαγωνισματος ...), o sostituirlo impropriamente con lo *stigma* (MIL\0913010: Οδηγος της Ολυμπιας ...).

L'integrazione dell'alfabeto greco nel sistema è poi inadeguata. In Indice è stato normalizzato sul latino seguendo l'ordine tradizionale (ma saltando 'c' e, meno comprensibilmente, 'g': β=B, γ=D, δ=E, ecc.), senza alcun riguardo per la fonetica delle lettere, né per i metodi di digitazione del greco su tastiera latina. Si è così creata una equivalenza molto originale quanto scarsamente intuitiva, che oltretutto non è descritta pubblicamente, ed è pertanto difficilmente utilizzabile anche da parte dell'utenza più accorta. Arduo in effetti immaginare che la parola Ιστορια in Indice sia normalizzata come LVXRULA. E ho specificato in Indice, perché l'OPAC SBN non normalizza affatto le lettere greche. Pertanto, i pochi titoli greci presenti nel catalogo si possono cercare solo interrogandolo in greco.

Inoltre, mercé una diffusa inconsapevolezza delle implicazioni informatiche, anche la limitata possibilità di utilizzare l'alfabeto greco è bastata a far comparire il problema definito 'Mixed-script confusables', ovvero la commistione di lettere appartenenti a diverse scritture a causa della loro somiglianza grafica.<sup>5</sup> Troviamo infatti le parole 'Azanias' (TSA\1463122) e 'Neis' (SBT\0002566 e BVE\0764010), che non possono essere cercate digitando 'Azanias' e 'Neis', perché le iniziali 'A' e 'N' hanno l'aspetto delle lettere latine ma la codifica delle greche (rispettivamente U+0391 e 039D), a causa verosimilmente del rimaneggiamento di un copia-e-incolla. Così, non essendo normalizzate, non possono essere trovate nemmeno con un'equivalente lettera latina. Diverse le ragioni, ma identica la sorte, per LIG\0102927, dove la *epsilon* greca (U+03B5) supplisce alla mancanza della *e latina aperta* (ε, U+025B) per le lingue africane; per tacere di chi ha mescolato sistematicamente i due alfabeti (TO0\1065245: Η εικονα της Ευρωπης : η εκθεση οργανωνεται ...; LO1\1378936: Η επιθεση : ανηκει στην ανεξάρτητη εpanastatikη ...).

## La latinizzazione: situazione attuale

La latinizzazione è oggi un'operazione necessaria per SBN, e sempre critica. Sebbene nel prosieguo l'esposizione manterrà un tenore prevalentemente tecnico, è bene tener presente che ogni scrittura, nonché ogni sua standardizzazione, ha anche connotazioni sociali, religiose, politiche, e di conseguenza il modo di codificare una scrittura in un'altra non è mai una scelta culturalmente neutrale (cfr. Aliprand 2005, 244–246; Selvelli 2015, 159–160).

Anche nella condizione più favorevole, ovvero la conversione di un alfabeto (cirillico, armeno,

---

<sup>5</sup> Vedi *Unicode Security Considerations*, <https://unicode.org/reports/tr36/>, e *Unicode Security Mechanisms*, <http://unicode.org/reports/tr39/>.

coreano, n'ko ecc.), capiterà sempre il caso che genererà dubbio, ambiguità, incoerenza. D'altronde, il Sistema Nazionale non potrà prescindere dalla latinizzazione nemmeno quando gestirà le altre scritture, per garantire l'accesso alle biblioteche che non adotteranno questa opzione così come agli utenti che non vorranno utilizzarla. Per questo, sarebbe bene che tutti seguissero un unico metodo, perseguendo l'uniformità di trattamento delle informazioni e facilitando l'interoperabilità (Albrecht, Wiesner 2011, 1, 8). Inoltre, l'adozione di un unico standard per ogni lingua semplificherebbe la vita agli utenti finali, che non sarebbero costretti a indovinare il metodo di latinizzazione seguito per la risorsa cercata. La realtà delle biblioteche italiane è lontana da questo ideale;<sup>6</sup> tuttavia, se vogliamo facilitare l'accesso al catalogo, questo obiettivo non si può tralasciare. Sulla scia delle indicazioni provenienti dall'UNESCO (1978) e dalla prima edizione del manuale UNIMARC, questi standard comuni sono identificati dalle *RICA* del 1979 (p. 226) con le norme prodotte dallo International Standard Organisation (ISO). Oggi, l'Appendice F delle *REICAT* propone tabelle "concordanti" con gli standard ISO per arabo (233-2:1993), lingue slave in scrittura cirillica (9:1995),<sup>7</sup> ebraico (259-2:1994), greco (843:1997), cinese (7098:1991, rivisto nel 2015). Non sono tutti, ma ci ritornerò oltre.

Qui basti notare che, contraddicendo in parte le premesse che le avevano portate ad adottare le norme ISO, le *REICAT* vi apportano in alcuni casi modifiche tali da rendere differente la versione per le biblioteche italiane. Vediamo meglio.

Prendiamo la lettera araba *tā' marbūṭah* ﺕ, che lo standard ISO traslittera con la *ti con dieresi* <ï> in tutti i casi, garantendo così l'equivalenza tra un carattere arabo e uno latino. Le *REICAT* scelgono invece <h> o <t>, a seconda che la parola si trovi in stato assoluto o costruito. Una regola storica molto seguita, basata sulla fonetica della lingua classica, la cui applicazione avrebbe però richiesto qualche spiegazione supplementare (cfr. Brockelmann 1935, 14–15; LoC 2012a, 4). Inoltre, <h> e <t> vengono usate per traslitterare anche altre due lettere arabe, facendo così venir meno la corrispondenza univoca tra i due alfabeti.

Quanto all'ebraico, si richiede di seguire ISO 259-2:1994, ma non si considera il suo silenzio su alcuni punti, quale il trattamento delle *matres lectionis*. Questo succede perché la seconda parte dello standard dà per scontato quanto già esposto nella prima (ISO 259:1984), ovvero che queste vadano sempre considerate consonanti a sé stanti; cosicché ם si traslittera <'iy>. Un'indicazione che avrebbe dovuto essere esplicitata, avvisando i catalogatori di utilizzare congiuntamente le due parti dello standard.

Scelte autonome anche per le lingue slave in scrittura cirillica. Le *REICAT* adottano ISO 9:1995, riproponendo però nella tabella principale i digrammi previsti dall'edizione del 1955, e mutandone la <è> in <è> per 'ǰ'. La novità introdotta dalla norma sovietica GOST 16876-71 (ed. 1981, tab. 1), e ripresa da ISO nel 1986, ovvero la sostituzione dei digrammi con un'unica lettera con accento circonflesso,<sup>8</sup> è consigliata solo per le forme di rinvio, ed è infatti assai raramente usata. Si è dunque scelto di garantire la continuità con una regola ben rappresentata nel catalogo, mantenendo

<sup>6</sup> Una scorsa ai titoli in greco, russo, arabo, persiano ed ebraico mostra la coesistenza di svariate regole di traslitterazione, alcune "personalizzate". Un'analisi del problema per le lingue slave in cirillico è in Mazzitelli 2008 (355–356, nota 54).

<sup>7</sup> La tabella per le lingue non slave è solamente menzionata, e presenta problematiche qui non affrontabili per mancanza di spazio.

<sup>8</sup> Come <ŝ> invece di <šč> per 'ш' e <û> invece di <ju> per 'ю'. Il circonflesso deriva – credo – dall'esperanto.

di fatto in vigore una versione vecchia dello standard. Comprensibile; ma, allora, perché mai non sceglierla apertamente? Così fa l'Università di Oslo (UiO 2021).

Con ciò non intendo asserire che gli standard siano intoccabili. Anzi, come vedremo oltre, è talvolta preferibile, se non necessario, deviare; occorre però farlo con accortezza per non stravolgerne i principi guida, ed esacerbare i problemi che la latinizzazione di per sé già comporta.

## La latinizzazione: ampliare gli standard

Consapevoli dell'insufficienza del repertorio di standard proposto, le *REICAT* raccomandano, per quanto non indicato, "di fare riferimento preferibilmente alle norme di traslitterazione o trascrizione ISO consigliate per impieghi bibliografici". Sebbene sconsiglino subito dopo proprio due standard ISO,<sup>9</sup> diamo qui per scontato che sottintendano quelli per armeno (9985:1996), georgiano (9984:1996), thailandese (11940-1:1998), le principali scritture indiane (15919:2001) e l'akson-thai-noi (20674-1:2019). Rimangono comunque scoperte non poche lingue e scritture.

Colmare la lacuna ricorrendo alle tabelle proposte dalla Library of Congress (LoC) è, d'altronde, una soluzione poco gradita alle *REICAT*, soprattutto per le discordanze tra i due sistemi (p. 571, nota). Seppure il costante ricorso all'ortografia inglese per alcuni suoni, e qualche soluzione grafica non replicabile in SBN, rendano le regole della LoC in linea di massima poco adatte alla nostra bisogna, il giudizio non può essere generalizzato. Si tratta pur sempre del sistema di latinizzazioni più ampio e organico oggi disponibile in ambito bibliografico.<sup>10</sup> Infatti, su 75 tabelle pubblicate (novembre 2021), per un numero però almeno doppio di lingue, stimo che solo 29 potrebbero essere compilate utilizzando le norme ISO.

Del resto, tra LoC e ISO v'è una diversità di base. Quest'ultimo nasce per regolamentare principalmente la produzione industriale, non le attività culturali; non sorprendere pertanto la maggiore attenzione prestata sinora alle lingue che esprimono primari interessi economici e politici. Si spiega così da una parte l'assenza di lingue africane e asiatiche con una tradizione letteraria significativa, ma scarso peso sul mercato globale, quali hausa, ge'ez, copto, birmano, cambogiano, tibetano, e via elencando; dall'altra la disattenzione per gli aspetti storici, testimoniata ad esempio dall'assenza dei caratteri cirillici per lo slavo ecclesiastico, nonché russo, serbo, bulgaro e romeno prima delle riforme ortografiche varate tra XVIII e XIX secolo. Ma gli aspetti storici e culturali sono proprio quelli che il bibliotecario deve invece tener sempre presenti.

Tant'è che anche biblioteche non anglosassoni hanno finito per adottare le regole statunitensi.<sup>11</sup> Un'analisi puntuale da parte di *REICAT* sarebbe stata dunque opportuna; anche perché avrebbe fornito l'occasione per chiarire quali siano le caratteristiche che rendono uno standard non ISO adatto o no a SBN, ed evitare che le biblioteche siano lasciate a sé stesse.

---

<sup>9</sup> Esplicitamente quello per il giapponese (3602:1989), e implicitamente quello per il persiano (233-3:1999); probabilmente sulla scorta delle analisi presentate da Kaji e Vita per il primo, Maggi e Orsatti per il secondo (in Pinedo, Niuatta 2002).

<sup>10</sup> Di simile ampiezza sono i sistemi per la toponomastica delle Nazioni Unite (UNGEGN, <https://www.eki.ee/wgrs/>) e del comitato Anglo-statunitense (BGN/PCGN, <https://geonames.nga.mil/gns/html/romanization.html>), governati però da principi differenti.

<sup>11</sup> Ad es. in Svizzera occidentale e Paesi Bassi; l'Università di Oslo quasi interamente, le biblioteche francesi più selettivamente. Non mancano biblioteche italiane.

Mantenere l'ISO quale caposaldo del sistema di latinizzazione per SBN richiede, dunque, che si delineino criteri generali cui attenersi per selezionare gli standard mancanti, costruendo un repertorio che sia il più razionale e coerente possibile, e controbilanci gli aspetti negativi sopra riassunti. A mio modo di vedere, questi sono i punti principali da considerare:

- A) ogni connubio di lingua e scrittura (e tipo di testo) è il prodotto di vicende in cui si rispecchia l'identità delle genti. Lingue diverse possono condividere un sistema scrittorio, ma conferire valori fonetici differenti agli stessi segni,<sup>12</sup> e questa individualità culturale va preservata. Essendo questo un principio fondamentale della multiculturalità, occorre dargli la primazia qualora la sua applicazione venga in contrasto con la regola generale ISO che recita "In cases where the same characters appear in one alphabet used with some differences by different languages, these characters would be transliterated in the same way, irrespective of the language they belong to". È pertanto impensabile trascrivere le sinografie giapponese o vietnamita coi suoni del cinese, o il persiano e il malese classico come se fossero arabo;
- B) rivolgersi a standard che contemperino, ove possibile, due caratteristiche:
  - a) godere di ampio consenso, vuoi perché prodotti da enti di normazione internazionali o nazionali che ne promuovano la diffusione, vuoi perché divenuti di fatto punto di riferimento per una particolare lingua (come YIVO per lo yiddish e Hepburn per il giapponese);
  - b) impiegare caratteri compresi nel repertorio ammesso dall'Indice SBN (vedi oltre, Interventi complementari, punto 2);
- C) preferire sistemi che stabiliscano – per le scritture alfabetiche e sillabiche – una corrispondenza biunivoca tra carattere originale e latino, permettendo di risalire con la minore ambiguità possibile dalla traslitterazione alla grafia originale (Martinet 1953, 152–154; Hilber 1984, 123);
- D) molte lingue usano più scritture contemporaneamente. Secondo i principi tecnici generali dell'ISO, "When a country uses two Systems univocally convertible one into the other to write its own language, the System of transliteration thus implemented shall be taken a priori as a basis for the international standardized System". Applicandolo, si otterrà – entro certi limiti – la convergenza su un unico standard delle latinizzazioni di una lingua vergata in scritture diverse, evitando, o almeno limitando, differenti versioni del medesimo titolo benché nella stessa lingua. Pertanto:
  - a) se una delle scritture è la latina, è superfluo rifarsi a un qualunque altro standard. È il caso di serbo, bosniaco e montenegrino in scrittura cirillica (tanto che – a rigore – gli standard ISO 9 del 1986 e 1995 non dovrebbero applicarsi a queste lingue);
  - b) se nessuna delle scritture è la latina, si dovrà individuare quella dotata dello standard che meglio si presti a fornire il modello di riferimento per la sua maggiore diffusione o idoneità per SBN (vedi punto B-b). Oppure, andrà privilegiata la scrittura meglio rappresentata nella produzione editoriale: i testi in lingua turca stampati in alfabeto greco o armeno andranno traslitterati seguendo la regola per l'ottomano in scrittura araba,

---

<sup>12</sup> Si pensi per analogia alle differenze di pronuncia della lettera 'c' nella varie lingue che usano l'alfabeto latino.



- per cui si userà <š> per ‘š’ e ‘ž’, invece di <ss> e <š> previste dalle rispettive norme ISO;
- c) compresenza di standard equipollenti. Al ladino (giudeo-spagnolo) in caratteri ebraici corrispondono diverse ortografie latine, sostanzialmente equivalenti per la descrizione bibliografica (Czöndör 2008). Si richiede dunque un’analisi per determinare quella preferibile;
- d) il principio può non applicarsi quando è difficoltoso stabilire relazioni biunivoche tra le scritture interessate. Così, per le lingue dell’Asia meridionale, quali panjabi, sindhi, urdu/hindi, per cui varie lettere arabe corrispondono a un unico carattere della corrispondente scrittura sillabica, è preferibile adottare standard differenti.<sup>13</sup> Per analoghe ragioni, al tagico in cirillico si applicherà la ISO 9:1995, mentre quello in scrittura araba seguirà lo standard per il persiano;
- E) applicare il medesimo principio anche quando diverse scritture si siano succedute nel tempo. Se la scrittura odierna è la latina, questa costituirà la base per la latinizzazione delle altre. Si pensi al rumeno e all’albanese (tosco), che hanno abbandonato da tempo rispettivamente la scrittura cirillica e la greca. Ma riguarda anche le repubbliche dell’ex URSS che hanno decretato il passaggio dalla scrittura cirillica alla latina e pubblicato le mappature tra i due alfabeti, fornendo una latinizzazione ufficiale diversa da quella che si avrebbe applicando la norma ISO.<sup>14</sup> Così, la traslitterazione di ‘ж’ rimane <ž> per il turkmeno, ma diventa <j> per azero, moldavo, kazaco, <zh> per uzbeko;
- F) valutare casi particolari, quali:
- a) il coreano: si usa in due stati che seguono due regole differenti di traslitterazione, e non vi è consenso su quale utilizzare tra le biblioteche europee (in Francia come in Corea del nord, in Germania come il sud). Verrebbe da agire salomonicamente: seguire lo standard adottato nel paese di pubblicazione, creando un rinvio con l’altro;
- b) le scritture fonetiche inglesi (deseret, shavian, unifon). Per il principio dell’ISO, dovrebbe essere l’ortografia latina a fornire il modello della traslitterazione; ma sarebbe insensato, perché queste scritture nascono proprio per superare la discrasia tra grafia e pronuncia dell’inglese. L’alfabeto fonetico internazionale d’altronde è difficilmente maneggiabile, e non è accettato da SBN. Ricorrere dunque a tabelle *ad hoc*? La LoC ne ha preparata una per la deseret. Applicandola al titolo  $\Upsilon \Theta \lambda \theta \alpha \phi \eta \Gamma \text{ P} \rho \tau \theta \eta \text{ B} \eta \omega / \alpha \theta \gamma \Psi \alpha \zeta \eta \eta \theta \text{ } \omega \theta$  [...] otteniamo <Dh Desēret Furst Bük / bī dh Rējents ôv [...]>, corrispondente a ‘The Deseret First Book / by the Regents of [...]’. In ogni caso, la ricerca per gli utenti è oltremodo complicata. Si può risolvere con un rinvio secondo l’ortografia normale; tuttavia, c’è da riflettere;
- G) contemplare le lingue delle comunità di recente formazione nel nostro paese, anche se ignorate da enti pubblici, scuole e università; ricordo – tra gli altri – i coltivatori panjabi, i Maghrebini che leggono il tamazight, i rifugiati che conversano in tigrino, curdo o pashto.

<sup>13</sup> Lo Annex C della ISO 15919:2001 dovrebbe uniformare la traslitterazione in questi casi, ma è inadeguato.

<sup>14</sup> Essendo la transizione in corso, le nuove ortografie latine sono ancora instabili. Per il kazaco (indicato in SBN con un obsoleto ‘cosacco’) e l’uzbeco siamo alla quarta versione.

## Interventi complementari

Per aiutare catalogatori e utenti sarebbe inoltre opportuno:

1. facilitare la consultazione delle regole di latinizzazione. Come è stato già fatto notare (Morigine 2010, 41; Selvelli 2015, 160), spesso sono difficilmente accessibili, e a causa di diritti d'autore l'utente può trovare pagine che riportano solo le informazioni per l'acquisto! In realtà, le tabelle *REICAT* sono pubblicate sul sito dell'ICCU, ma si potrebbe divulgarle meglio collocando un link vicino alle maschere di ricerca;
2. aumentare i caratteri speciali gestiti dall'Indice, che a settembre 2019 erano 638.<sup>15</sup> Sono insufficienti, visto che è necessario escogitare alternative per lo *sceva* 'ə' (U+0259; vedi punto 3-d), ed è impossibile riportare fedelmente lingue con diacritici complessi (come 'ă' o 'ê' in vietnamita), caratteri modificati come 'Ḃ' (U+0253) e 'Ʀ' (U+026C) per varie lingue africane, o la legatura *ou* '8' (U+0223) dei libri in albanese del XVI-XVII sec. (vedi punto 3-d).<sup>16</sup> Stimo che ne servirebbero all'incirca altri duecento, compresi i caratteri greci tonici previsti dalla riforma ortografica del 1982 e alcuni diacritici;
3. completare le tabelle di latinizzazione degli standard considerando:
  - a) le variazioni ortografiche regionali: l'arabo maghrebino usa alcune lettere ignote all'arabo standard ... e alle tabelle ISO;
  - b) i caratteri storici, come i diversi alfabeti georgiani, varie cirilliche, ecc.;
  - c) le convenzioni ortografiche nazionali (cfr. punto 6-a);
  - d) i caratteri latini non gestiti dall'Indice, per uniformarne la trasposizione secondo i contesti storici e linguistici; per cui lo *sceva* potrebbe indicarsi con <ë> (U+0115) per le lingue etiopiche, e <ä> per quelle dell'Asia Centrale;
4. se necessario, predisporre interventi minimi sugli standard per facilitare la catalogazione e la ricerca, o adattarli alle funzionalità dell'Indice. Ad esempio, ISO traslitera la lettera araba *ḥa* 'ح' con *acca con linea sotto* <ḥ>. A oggi, però, UNICODE, e di conseguenza i font, ne prevedono solo la minuscola, obbligando l'uso d'un diacritico di combinazione per realizzare la maiuscola. Così, molti catalogatori hanno adottato la *acca con breve sotto* <ḥ>, che è invece disponibile in entrambe le forme, nonché scelta da altri diffusi sistemi di traslitterazione dell'arabo. Poiché entrambe le soluzioni sono normalizzate con <H>, la variazione non crea problemi alla ricerca; tanto varrebbe adottarla esplicitamente;
5. normalizzare uniformemente la scrittura greca in tutti i software SBN;
6. nelle tabelle di latinizzazione, corredare i caratteri non ASCII del rispettivo codice UNICODE e indicare, per i caratteri composti, il diacritico da utilizzare ed eventuali alternative. È importante per l'ortografia di alcune lingue, ed essenziale per altre. Ad esempio:
  - a) in romeno, le lettere 'ș' (U+0219) e 'ț' (U+021B) ortograficamente corrette sono quelle con virgola sottostante (Academia Română 2005, xxvi), che non sono però previste da

<sup>15</sup> Comprese due lettere cirilliche e tre arabe; verosimilmente introdotte per sperimentazione e lì dimenticate.

<sup>16</sup> Talvolta confusa con *e commerciale* '&': vedi NAPE\020060, dove *Pasechbra etrefbemit* è erroneamente trascritto *Pasechbra etrefemit*. La Biblioteca Nazionale di Tirana lo sostituisce con <y>; altri preferiscono la cifra <8>, perché non vi è relazione biunivoca tra la 'y' storica e la grafia albanese attuale (Vozga 2010, 73; M. Genesin, com. pers. 3/1/2019). La grande varietà di caratteri, alcuni peculiari, storicamente utilizzati per stampare l'albanese richiede un approfondimento specifico.

- SBN; si devono comporre col diacritico di combinazione (U+0326), o sostituire con le lettere con *cedilla* (‘ş’ U+015F e ‘ț’ U+0163);
- b) varie lingue usano lettere arabe dall’identico aspetto quando sono isolate, ma codificate diversamente da UNICODE perché generano forme contestuali differenti (cfr. Kew 2005);
  7. elaborare formule standard per indicare, nelle note, il sistema di latinizzazione adottato con gli eventuali scostamenti, seguito ove possibile dal contenuto delle aree del titolo e della pubblicazione nella scrittura originale. L’area delle note, infatti, è esente da controlli, e l’unico limite alla visualizzazione è il repertorio dei caratteri previsti dai font caricati nel sistema;
  8. sistemare la tastiera virtuale del gestionale, che prevede caratteri speciali doppi mentre difetta di altri;
  9. rifare la lista dei codici di lingua secondo ISO 639-2, 3 e -5, tenendo però conto del nostro contesto storico e culturale. SBN prevede sei (*sic*) codici per le lingue sami (lapponi), ma omette quelli per griko, arbëresh, croato molisano, mòcheno e altre. Inoltre, per la carenza di codici specifici, o per impostazioni errate da parte degli organismi internazionali,<sup>17</sup> gran parte della nostra notevole letteratura dialettale si deve classificare ‘Romanza (altra lingua)’! Sfruttiamo dunque l’intervallo libero dello standard (qaa-qtz) per dotare di un codice gli idiomi che ne sono privi.

Come ulteriore base di discussione, concludo proponendo una selezione di standard, con commenti e bibliografia striminziti:

<i>Lingua</i>	<i>Scrittura</i>	<i>Standard</i>	<i>Commenti, Riferimenti</i>
Amarico Ge’ez Tigrè Tigrino	Etiopica	<i>Encyclopaedia Aethiopica</i>	Kelly 2020, xviii
Arabo	Araba	ISO 233-2:1993	Ripristinare <ġ> (U+1E97) per <i>tā’ marbūṭah</i> ; sostituire <Ĥ ; ĥ> (<H, U+0320>; U+1E96) con <Ĥ; ĥ> (U+1E2A; 1E2B) per <i>ḥa’</i>
Armeno	Armena	ISO 9985:1996	
Bengalese	Bengalese	ISO 15919:2001 tab. 2	
Bielorusso Bulgaro Macedone Russo Ucraino	Cirillica	<i>REICAT</i> , 575	ISO 9:1995 tab. 1 per forme di rinvio (= <i>REICAT</i> , 576)
Birmano	Birmana		Okell 1971, 65
Bosniaco	Cirillica Araba	Ortografia ufficiale latina	Janković 1989 per la corrispondenza cirillico ↔ arabo
Cambogiano	Cambogiana	ALA-LC	LoC 2012b

<sup>17</sup> È ad esempio il caso del Napoletano, inteso da UNESCO e ISO 639 (codice NAP) come idioma dell’intera Italia meridionale (vedi De Blasi, Montuori 2020, 90-94).

Cinese	Cinese	<i>Hànyǔ pīnyīn</i> (ISO 7098:2015)	rinvii per i nomi propri secondo il sistema Wade-Giles, per l'ampia diffusione goduta sino a poco tempo fa. Aggregare le sillabe che compongono una parola, rinviando nel caso a una forma con le sillabe staccate (Huang, Haynes 2004)
Copto	Copta	Leipzig-Jerusalem	Grossman, Haspelmath 2015
Coreano	Coreana	Latinizzazione Rivista in Corea del Sud, McCune-Reischauer in Corea del Nord	vedi sopra, punto F, a Per il nord: McCune, Reischauer 1939 Per il sud: Korea 2000
Ebraico	Ebraica	ISO 259:1984 e 259-2:1994	<i>alep</i> , <i>he</i> e <i>yod</i> sempre traslitterate. Un trattino separa l'articolo, la congiunzione e le preposizioni dalla parola seguente. Vocalizzare sempre
Georgiano	Georgiana	ISO 9984:1996	Anche per gli alfabeti storici (asomt'avruli, nusxuri, xuc'uri)
Giapponese	Kanji e kana	Modified Hepburn	Watanabe, Skrzypczak, Snowden 2003
Greco	Greca	ISO 843:1997 tab. 1	
Hindi Sanskrito Marathi e altre	Devanagari	ISO 15919:2001 tab. 1	sostituire <kha> (per क्) con <xa>, perché si legge [χa], la sottolineatura non è applicabile in SBN, e <kha> è già assegnato a क्
Kazaco (qazaq)	Cirillica	Ortografia ufficiale latina	Mamin 2021 (in vigore dal 2023)
Laotiano	Laotiana	ALA-LC	LoC 2012c
Mongolo	Mongola	Vladimircov-Mostaert	Poppe 1970, 16; sostituire <γ> con <ḡ> (U+0121), <ḟ / j> con <ž> (U+017E)
Montenegrino	Cirillica	Ortografia ufficiale latina	Savjet 2010
Panjabi	Gurmukhi	ISO 15919:2001 tab. 3	
Panjabi	Araba	Vedi Urdu	Con l'aggiunta di ح (U+0768) <n> (U+1E47)
Pashto	Araba	DIN 31635:2011-07 cap. 11	
Persiano	Araba	DIN 31635:2011-07 cap. 8	
Serbo	Cirillica	Ortografia ufficiale latina	
Singalese	Singalese	ISO 15919:2001 tab. 10	
Slavo ecclesiastico	Cirillica Glagolitica	Traslitterazione Scientifica	Per la corrispondenza cirillico ↔ glagolitico vedi Jagić 1879, XXXVII
Tagico	Cirillica	ISO 9:1995 tab. 3	
Tagico	Araba	Vedi Persiano	
Tamazight Tamaceq	Tifinagh	Institut Royal de la Culture Amazigh	Zenkouar 2008
Tamil	Tamil	ISO 15919:2001 tab. 6	
Telugu	Telugu	ISO 15919:2001 tab. 7	
Thai	Thai	ISO 11940:1998	
Tibetano	Tibetana	Extended Wylie	Traslitterare le lettere per il sanscrito come gli equivalenti devanagari (vedi Hindi). Anton-Luca 2006
Turco ottomano	Araba Armena Greca	Türk İlmî Transkripsyon	İÜEF 1946; Birnbaum 1967
Urdu	Araba	DIN 31635:2011-07 cap. 10	
Uzbeko	Cirillica	Ortografia ufficiale latina	Özbekiston 2021
Yiddish	Ebraica	Institute for Jewish Research	YIVO 2021; Bleaman 2019

## Riferimenti bibliografici

- Academia Română. 2005. Institutul de Linguistică, *Dicționarul ortografic, ortoepic și morfologic al limbii române*, Ediția a ii-a. București: Editura Enciclopedic, <https://diacritica.files.wordpress.com/2010/04/doom2.pdf>.
- Agenbroad, James E. 2006. "Romanization is not enough." *Cataloging & Classification Quarterly* 42, 2: 21–34, DOI: 10.1300/J104v42n02\_03.
- Albrecht, Rita und Hans-Jörg Wiesner. 2011. *Umschriften in Bibliotheken: Anwendung, Verfahren, Tabellen*. Berlin: Beuth.
- Aliprand, Joan M. 2005. "Scripts, Languages, and Authority Control." *Library Resources & Technical Services* 49, 4: 243–249, DOI: <https://doi.org/10.5860/lrts.49n4.243>.
- Anton-Luca, Alexandru. 2006. *Teaching THDL Extended Wylie*, <http://www.thlib.org/reference/transliteration/teachingewts.pdf>.
- AreaSBN. 2019. "Alfabeti non latini e caratteri speciali." Risposta a segnalazione 6782, <http://assistenza.indice.sbn.it/view.php?id=6782> (accesso riservato).
- Azzanutto, Federica, Liliana Macovetchi, e Giovanna Malgaroli (a cura di). 2015. *Mamma lingua. Libri per bambini in età prescolare in albanese, arabo, cinese, francese, inglese, rumeno, spagnolo*. Milano: Nati per leggere Lombardia, Iby Italia, [https://www.aib.it/wp-content/uploads/2017/11/MAMMA\\_LINGUA\\_ar.pdf](https://www.aib.it/wp-content/uploads/2017/11/MAMMA_LINGUA_ar.pdf).
- Battaglini, Marina. 1999. "Breve nota sulla gestione automatizzata di dati bibliografici in ideogrammi." *SBN Notizie* 2: 44–46.
- Birnbaum, Eleazar. 1967. "The Transliteration of Ottoman Turkish for Library and General Purposes." *Journal of the American Oriental Society* 87, 2: 122–156, <https://www.jstor.org/stable/597394>.
- Bleaman, Isaac L. 2019. "Guidelines for Yiddish in Bibliographies: A Supplement to YIVO Transliteration." *In geveb*, <https://ingeveb.org/pedagogy/guidelines-for-yiddish-in-bibliographies>.
- Brockelmann, Carl, August Fischer, W. Heffening, und Franz Taescher. 1935. *Die Transliteration der arabischen Schrift in ihrer Anwendung auf die Hauptliteratursprachen der islamischen Welt. [...]*. Leipzig: Deutsche Morgenländische Gesellschaft, <https://www.dmg-web.de/page/transliteration/de/denkschrift.pdf>.
- Chan, Kylie, Lily Hu, and Patrick Lo. 2000. "A Collaborative Project on Chinese Name Authority Control: The HKCAN Model." *Journal of East Asian Libraries* 120, 3: 1–16, <http://eprints.rclis.org/4171/>.
- Ciccarello, Domenico. 2009. "I servizi multiculturali delle biblioteche. Quali tendenze in Italia e nel mondo." *Biblioteche oggi* 4: 25–30.
- Cohen, Ahava. 2017. "Each of us has a name: creating a multilingual, multiscript national authority database." *Catalogue and Index* 186: 15–19, [https://www.cilip.org.uk/resource/collection/F71F19C3-49CF-462D-8165-B07967EE07F0/Catalogue & Index issue 186, March 2017.pdf](https://www.cilip.org.uk/resource/collection/F71F19C3-49CF-462D-8165-B07967EE07F0/Catalogue%20&%20Index%20issue%20186,%20March%202017.pdf).

- Czöndör, Klára. 2008. "La Escritura de la lengua Judeo-Española." *Verbum Analecta Neolatina* 10, 1: 79–97, DOI: 10.1556/verb.10.2008.1.6.
- De Blasi, Nicola e Francesco Montuori. 2020. *Una lingua gentile. Storia e grafia del napoletano*. Napoli: Cronopio.
- Eilts, John. 1996. "Non-Roman Script Materials in North American Libraries: Automation and International Exchange." *International Cataloguing and Bibliographic Control* 25, 3: 51–53, <http://origin-archive.ifla.org/IV/ifla61/61-eilj.htm>.
- El-Sherbini, Magda. 2016. "Improving Resource Discoverability for Non-Roman Language Collections." In *Subject Access: Unlimited Opportunities, 11-12 August 2016, Columbus, Ohio*, <http://library.ifla.org/id/eprint/1982>.
- Fallerini, Antonella, Agnese Galeffi, Andrea Ribichini, Mario Santanchè, e Mattia Vallania. 2022. "DREAM. A project about non-Latin script data." *JLIS.it*, 13, 1: 347-355, DOI: 10.4403/jlis.it-12727.
- Grossman, Eitan, and Martin Haspelmath. 2015. "The Leipzig-Jerusalem transliteration of Coptic". In *Egyptian-Coptic Linguistics in Typological Perspective*, ed. by E. Grossman, M. Haspelmath and T. S. Richter, 145–153. Berlin: De Gruyter Mouton, DOI: 10.1515/9783110346510.145.
- Hilber, Elmar. 1984. "Transliteration und Transkription nichtlateinischer Schriften." *Nachrichten VSB/SVD* 60, 3: 119–146.
- Huang, Jie and Kathleen J. M. Haynes. 2004. "The Issue of Word Division in Cataloging Chinese Language Titles." *Cataloging & Classification Quarterly* 38, 1: 27–42, DOI: 10.1300/J104v38n01\_04.
- İÜEF. 1946. İstanbul Üniversitesi Edebiyat Fakültesi. *Türk ilmi transkripsiyon kılavuzu*. İstanbul: Millî Eğitim Bakanlığı.
- Jagić, Vatroslav. 1879. *Quattuor Evangeliorum codex glagoliticus olim Zographensis nunc Petropolitani*. Berolini: Apud Weidmannos, <https://archive.org/details/quattuorevangeli00jagiuoft>.
- Janković, Srđan. 1989. "Ortografsko usavršavanje naše arabice u štampanim tekstovima (uticaj ideja Vuka Karadžića)." *Prilozi za Orijentalnu Filologiju* 38: 9–40, <http://pof.ois.unsa.ba/index.php/pof/article/view/425/pdf>.
- Kelly, Samantha. 2020. "Transliteration Chart: Gəʿəz and Amharic." In *A Companion to Medieval Ethiopia and Eritrea*. Leiden: Brill, <https://brill.com/view/book/edcoll/9789004419582/front-11.xml>.
- Kew, Jonathan. 2005. *Notes on some Unicode Arabic characters: recommendations for usage*. [Dallas: SIL International], [https://scripts.sil.org/cms/scripts/page.php?site\\_id=nrsi&id=pcunicodedocs#272ccb85](https://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&id=pcunicodedocs#272ccb85).
- Korea. 2000. *Gugeoui romaja pyogibeop gaejeong gyeongwi*, <https://www.law.go.kr/행정규칙/국어의로마자표기법/2000-8,20000707>.
- Lazinger, Susan S., and Judith Levi. 1996. "Multiple Non-Roman Scripts in Aleph-Israel's Research Library Network." *Library Hi Tech* 14, 1: 111–116, DOI: 10.1108/eb047987.

- LoC. 2011. Library of Congress, Policy & Standards Division. *Romanization Landscape*, <https://www.loc.gov/catdir/cpsoromlandscape/Oct2011.html>.
- LoC. 2012a. Library of Congress, Policy & Standards Division. *Romanization of Arabic*, <https://www.loc.gov/catdir/cpsoromanization/arabic.pdf>.
- LoC. 2012b. Library of Congress, Policy & Standards Division. *Romanization of Khmer*, <https://www.loc.gov/catdir/cpsoromanization/khmer.pdf>.
- LoC. 2012c. Library of Congress, Policy & Standards Division. *Romanization of Lao*, <https://www.loc.gov/catdir/cpsoromanization/lao.pdf>.
- Mamin, Askar. 2021. *Qazaq tili alipbiin latyn grafikasya koshiru jonindegi ulttyq komissia otyrysyn otkizdi*. Qazaqstan Respublikasy Prem'eri-Ministriniń resmı aqparattyq resursy, <https://www.primeminister.kz/kz/news/a-mamin-kazak-tili-alipbiin-latyn-grafikasyna-koshiru-zhonindegi-ulttyk-komissiya-otyrysyn-otkizdi-280941>.
- Martinet, André. 1953. "A Project of Transliteration of Classical Greek." *Word* 9: 152–161.
- Mazzitelli, Gabriele. 2008. "Ancora sul problema della traslitterazione dei caratteri cirillici." *Bollettino AIB* 48, 4: 343–358, <https://bollettino.aib.it/article/download/5554/5350>.
- McCune, G. A., and E. O. Reischauer. 1939. "The Romanization of the Korean Language Based upon Its Phonetic Structure." *Transactions of the Korea Branch of the Royal Asiatic Society* 29: 1–57.
- Moreigne, Alexandre. 2010. *Les fonds russes en bibliothèques municipales*. Paris: ENSSIB, <http://www.beseda.fr/docs/M-moire-Alexandre-Moreigne-fonds-russes.pdf>.
- Okell, John. 1971. *A Guide to the Romanization of Burmese*. London: Royal Asiatic Society.
- Ongini, Vinicio. 2012. "Quando le "domestiche" venivano dal Nord." In *Così vicine, così lontane. Una ricerca sui bisogni e i consumi culturali delle cittadine straniere*, 17–21. Roma: Provincia di Roma, [https://www.academia.edu/29298978/Libro\\_ricerca\\_badanti.pdf](https://www.academia.edu/29298978/Libro_ricerca_badanti.pdf).
- Özbekiston. 2021. Özbekiston Respublikasi Qonuni, *Özbekiston Respublikasining "Lotin yozuwiga asoslangan o'zbek alifbosini joriy etish to'g'risida"gi Qomumiga o'zgartirişlar haqida*, <https://regulation.gov.uz/uz/d/31596>.
- Pinedo, Isa de, e Francesca Niutta. 2002. *Gruppo di lavoro per la revisione delle tabelle di traslitterazione in ambito SBN*. Roma: ICCU, <https://www.iccu.sbn.it/export/sites/iccu/documenti/traslSBN.doc>.
- Poppe, Nicholas. 1970. *Mongolian Language Handbook*. Washington: Center for Applied Linguistics.
- Rabitti, Chiara. 2003. *Catalogazione e alfabeti non latini*, <http://www.aib.it/aib/commiss/mc/rab-chi01.htm>.
- Rigby, Carol. 2015. "Nunavut Libraries Online Establish Inuit Language Bibliographic Cataloging Standards: Promoting Indigenous Language Using a Commercial ILS." *Cataloging & Classification Quarterly* 53, 5–6: 615–639, DOI: 10.1080/01639374.2015.1008165.

Savjet. 2010. Savjet za standardizaciju crnogorskoga jezika. *Pravopis crnogorskoga jezika*. Drugo, izmijenjeno izdanje. Podgorica: Ministarstvo prosvjete i nauke, <http://www.gov.me/files/1248442673.pdf>.

Selvelli, Giustina. 2015. "Su alcuni aspetti ideologici dei sistemi di traslitterazione degli alfabeti cirillici nei Balcani." *Studi slavistici* 12: 159–180, <http://hdl.handle.net/10278/3679013>.

Ting, T. C. (ed.). 1985. *Chinese-Character Processing for Computerized Bibliographic Information Exchange. Summary Report of an International Workshop Held in Hong Kong, 17-20 December 1984*. Ottawa: International Development Research Centre, <https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/7309/IDL-7309.pdf>.

UiO. 2021. Universitetsbiblioteket i Oslo, *Transkribering / translitterering av ikke-latinske alfabeter*, <https://www.ub.uio.no/bibliotekene/uhs/uhs/samlinger/transkribering/>.

UNESCO. 1978. *International Congress on National Bibliographies: Final Report*. Paris: UNESCO, <https://unesdoc.unesco.org/ark:/48223/pf0000030062.locale=en>.

Vernon, Elisabeth. 1996. *Decision-Making for Automation: Hebrew and Arabic Script Materials in the Automated Library*. Urbana-Champaign: Graduate School of Library and Information Science, University of Illinois, <http://hdl.handle.net/2142/3879>.

YIVO. 2021. Institute for Jewish Research. *Yiddisch Alef-Beyts (Alphabet)*, <https://yivo.org/Yiddish-Alphabet>.

Vozga, Ramazan. 2010. *Libri shqip 1555-1912 në fondet e Bibliotekës Kombëtare: bibliografi*. Tiranë: Biblioteka Kombëtare, <https://www.bksh.al/bksh/Libri%201555-1912.pdf>.

Watanabe, Toshirō, Edmund R. Skrzypczak and Paul Snowden. 2003. *Kenkyūsha's new Japanese-English Dictionary*, 5<sup>th</sup> ed. Tōkyō: Kenkyūsha.

Zenkouar, Lahbib. 2008. "Normes des technologies de l'information pour l'ancrage de l'écriture amazighe." *Études et Documents Berbères* 27, 1: 159–172, DOI: 10.3917/edb.027.0159.



## Covid, document production and administrative efficiency: a case study

Francesca Parisi<sup>(a)</sup>, Roberto Guarasci<sup>(b)</sup>, Roberto Elmo<sup>(c)</sup>

a) University of Calabria, <https://orcid.org/0000-0001-9125-9904>

b) University of Calabria, <https://orcid.org/0000-0001-5182-573X>

c) University of Calabria, <https://orcid.org/0000-0001-6312-9336>

**Contact:** Francesca Parisi, [francesca.parisi@unical.it](mailto:francesca.parisi@unical.it); Roberto Guarasci, [roberto.guarasci@unical.it](mailto:roberto.guarasci@unical.it);  
Roberto Elmo, [roberto.elmo@unical.it](mailto:roberto.elmo@unical.it)

**Received:** 12 October 2021; **Accepted:** 2 March 2022; **First Published:** 15 May 2022

### ABSTRACT

The article outlines the document production process in some Italian research institutions and compares it to the same pre-pandemic time span by analyzing whether or not the work organization models that have been adopted have had an impact on the completion time of administrative procedures. The aim was to verify the trend of the performance indicators after the pandemic impact and how the absence of employees at work may have influenced the organizational models. Within specific administrative realities, human relationship is often a pillar to carry out activities correctly. The transition to the management of remote procedures carried out totally in digital form caused inevitably important consequences on the expected results.

### KEYWORDS

Covid-19; Work organization; Smart working; Document production.

## Covid, produzione documentale ed efficienza amministrativa: un caso di studio\*

### ABSTRACT

Il testo rileva la produzione documentale in alcune istituzioni di ricerca italiane raffrontandola con quella dello stesso arco cronologico pre pandemia analizzando nel contempo se i modelli di organizzazione del lavoro adottati abbiano avuto o meno riflessi sui tempi di conclusione dei procedimenti amministrativi. L'obiettivo è stato quello di verificare l'andamento degli indicatori delle performance a seguito dell'impatto pandemico e come la conseguente assenza dei dipendenti sul posto di lavoro possa aver influenzato i modelli organizzativi all'interno di specifiche realtà amministrative dove il rapporto umano è spesso pilastro per il corretto svolgimento delle attività. Il passaggio alla gestione dei procedimenti a distanza e svolta, conseguentemente, totalmente in forma digitale ha inevitabilmente portato conseguenze sui risultati attesi.

### PAROLE CHIAVE

Covid-19; Organizzazione lavoro; Smart working; Produzione documentale.

\* Si ringraziano per la collaborazione nel recupero dei dati: Pierluigi Feliciati dell'Università di Macerata, Andrea Passarino, Salvatore Ricchio, Antonio Filice e Anna Maria Bria, Università della Calabria, Maurizio Vitale, CNR e Paolo Ferragina, Università di Pisa. Gli autori, pur avendo lavorato collegialmente alla redazione del testo, hanno specificatamente redatto: Roberto Guarasci, paragrafo 1, Francesca Parisi, paragrafo 2, Roberto Elmo, paragrafo 3.

## 1. Lo scenario

Quando nei primi mesi del 2020 il mondo occidentale fu costretto a misurarsi con una pandemia, evento ormai dimenticato e rimosso anche dalla narrazione collettiva, visto che l'ultima grande pandemia – quella della Spagnola degli anni 1918-20 – era stata a sua volta oggetto di un processo di rimozione collettiva tanto che di essa si è ricominciato a parlare solo per fare degli impropri paragoni con l'esplosione dell'epidemia da Sars-Cov2, quasi nessuno si rese conto della portata di ciò che stava accadendo. Non solo da un punto di vista dell'emergenza sanitaria ma anche della modifica dei modelli di organizzazione del lavoro e, conseguentemente, di produzione dei documenti che di quelle attività sono la necessaria concretizzazione. Se è vero che i provvedimenti di contenimento adottati non si discostavano molto da quelli utilizzati un secolo fa ciò che ha fatto la differenza è stata la possibilità di virtualizzare processi ed attività realizzando un esperimento su scala planetaria impossibile da costruire senza quell'evento pandemico. Non abbiamo purtroppo studi quantitativi sulla produzione documentale durante la Spagnola sia per la difficoltà oggettiva, all'epoca, di misurare tali indicatori sia per la concomitante censura dell'ultimo scorcio della Prima Guerra mondiale. Aumentò certamente la quantità di articoli e rapporti scientifici sull'eziologia della malattia e sulle possibili cure (Ianni 2020, 1-13) così come è stato anche per il Covid – 19 (Mohadab et al. 2020; Teixeira da Silva et al. 2020). Non abbiamo dati sulla documentazione amministrativa sia per i citati eventi bellici che avevano già fortemente depauperato la macchina dello Stato ma anche perché gli uffici pubblici – all'epoca non virtualizzabili – dovettero gioco forza continuare a funzionare per garantire i servizi essenziali della collettività.

La pandemia più recente ha invece consentito il funzionamento da remoto di una larghissima parte della macchina amministrativa pubblica e privata senza soluzione di continuità. I pareri sull'efficacia e sulla opportunità del lavoro agile specie nella pubblica amministrazione sono a tutt'oggi contrastanti e discordi e non è certo questa la sede per valutazioni di merito anche perché ogni commento privo del sostegno di dati consolidati e verificati è poco più di una narrazione romanzata. La brevità del lasso di tempo che ci separa dai fatti ed una serie di ulteriori circostanze a contorno fanno sì che i dati sui vari eventi del periodo interessato dall'emergenza epidemiologica siano ancora non consolidati e non pienamente attendibili anche in virtù di metriche di calcolo non standardizzate e non universalmente accettate. Per tutti la ormai nota vicenda dei morti per Covid o con il Covid basata sulla individuazione della causa prevalente del decesso che in qualunque evento pandemico è sempre difficilissima da individuare visto che si tratta quasi sempre di patogeni poco noti o sconosciuti specie nelle fasi di esordio della malattia.

Quel che è certo – almeno in Italia – è che partendo da una situazione certamente non ottimale e non priva di forti elementi di criticità la macchina amministrativa pubblica è riuscita comunque a garantire il funzionamento dei servizi e delle attività necessarie. Ciò – con diversi gradi di operatività – è stato fatto non solo per strutture più grandi e complesse con la presenza di organici gruppi di supporto ma molto spesso anche nelle realtà più piccole dove spesso l'iniziativa dei singoli ha supplito alle carenze strutturali del sistema.

In questo contesto il sistema universitario nella sua generalità partiva certamente da una situazione privilegiata rispetto a gran parte della restante pubblica amministrazione italiana sia per la dotazione di attrezzature e competenze che possedeva sia per le attività di sperimentazione di erogazione a distanza di attività di vario genere condotte negli anni. Le attività erano però quasi tutte

ricadenti nell'ambito della didattica e dell'erogazione di percorsi formativi. I modelli organizzativi del lavoro amministrativo erano nella quasi totalità basati esclusivamente sull'attività in presenza con la sola eccezione degli atenei telematici.

Da un punto di vista normativo il lavoro da remoto nella Pubblica Amministrazione all'atto dell'insorgenza pandemica era espressamente consentito dall'articolo 4 della legge 191 del 1998 che forniva delle linee guida su uso della postazione, modalità di connessione e di autenticazione ai sistemi, comunicazioni tra uffici e utilizzo della firma digitale. Da ricordare anche la circolare Inps n. 52 del 27 febbraio 2015, con le istruzioni operative per l'attuazione dell'Accordo Nazionale sul progetto di telelavoro domiciliare. Era telelavoro più che lavoro agile con obbligo per il dipendente di lavorare da casa e trasferimento al domicilio del dipendente degli stessi obblighi previsti per il posto di lavoro. La legge 22 maggio 2017 n. 81 agli articoli 18 e seguenti introduce in maniera esplicita il concetto di lavoro agile specificando che lo stesso è attuato mediante accordo tra le parti e senza i vincoli di orario e di luogo che invece erano presenti nel telelavoro. Il primo marzo 2020 – in piena emergenza pandemica - viene emanato il DPCM “Ulteriori disposizioni attuative del decreto-legge 23 febbraio 2020, n. 6, recante misure urgenti in materia di contenimento e gestione dell'emergenza epidemiologica da COVID-19” che alla lettera g dell'art. 1 disponeva la chiusura di tutti i pubblici uffici ad eccezione dei servizi essenziali e al successivo art. 4 estendeva “a ogni rapporto di lavoro subordinato” la possibilità di applicare la modalità di lavoro agile “disciplinata dagli articoli da 18 a 23 della legge 22 maggio 2017 n. 81” anche in assenza di accordi individuali. Le pubbliche amministrazioni risposero in maniera diversificata a seconda della tipologia dei servizi erogate, della dotazione infrastrutturale e della capacità dei dirigenti di adeguarsi al cambiamento.

Per verificare se e quanto la necessitata modifica delle modalità di svolgimento del lavoro amministrativo avessero influito su quantità e qualità della produzione documentale abbiamo ipotizzato di cominciare a misurare le quantità di documenti prodotte o ricevute da alcune pubbliche amministrazioni nel periodo marzo-maggio 2020 raffrontandole con le quantità dello stesso arco cronologico dell'anno precedente. A valle di questa prima valutazione quantitativa, valutata la sostanziale riduzione delle quantità anche in relazione alla sospensione di alcune delle attività correlate alla produzione di specifiche tipologie documentali abbiamo provato a valutare, misurando i tempi di conclusione dei procedimenti, se la diminuzione registrata era da imputarsi esclusivamente al dato oggetti della diminuzione dei procedimenti amministrativi attivi ovvero se la pubblica amministrazione avesse, in misura variabile, colto l'occasione per una ridefinizione dei processi e quindi ridotto la documentazione circolante con conseguente positivo impatto su tempi di conclusione dei singoli procedimenti<sup>1</sup> (Molino et al. 2020).

---

<sup>1</sup> Una valutazione statistica sull'andamento del remote working nelle aziende a fronte della crisi legata al coronavirus è fornita da Molino (Molino et al. 2020) che prende a campione 878 partecipanti per uno studio sul livello di produttività/efficienza dei dipendenti in smart working, più precisamente sul technostress all'interno del contesto italiano. Nello specifico, le misure applicate per valutare il livello di sinergia tra produttività del team di lavoro e technostress che gli autori utilizzano sono riferite all'analisi dei dati descrittiva con deviazione media e standard, le correlazioni di Pearson e i coefficienti alfa di Cronbach.

## 2. L'Indagine

L'indagine mirava a definire la variazione delle quantità di documenti prodotti o ricevuti dalle amministrazioni pubbliche partecipanti alla rilevazione verificando, successivamente, se tali variazioni fossero o meno risultate influenti sui tempi di conclusione dei procedimenti.

Hanno aderito all'iniziativa l'Università della Calabria (Unical), l'Università di Macerata (Unimc), l'Università di Pisa (Unipi) e il Consiglio Nazionale delle Ricerche (CNR). I dati relativi al protocollo sono stati elaborati per tutte le strutture, mentre i dati relativi ai procedimenti sono stati analizzati solo per l'Università della Calabria e l'Università di Macerata. Queste ultime hanno circoscritto la rilevazione ad alcuni dipartimenti e alla Direzione risorse umane dell'amministrazione centrale.

I dati provenienti dal protocollo sono stati elaborati al fine di poter essere comparati in maniera aggregata, mentre i dati relativi alle performance di procedimenti più specifici sono stati confrontati calcolando la variazione (differenza) dei tempi e delle quantità riferiti ai periodi marzo - maggio 2019 e 2020.

In particolare, i dati relativi ai documenti in entrata e in uscita gestiti dalle strutture sono stati aggregati in modo da ottenere un indicatore complessivo per il 2019 e uno per il 2020. Per l'Università della Calabria e l'Università di Macerata sono stati presi in considerazione i documenti in entrata e in uscita gestiti dai dipartimenti che hanno reso disponibili anche i dati sulle performance. Tale scelta è stata effettuata per avere una omogeneità con le analisi successive. Una volta aggregate le quantità sono state poi calcolate le variazioni percentuali tra i due periodi evidenziando come quasi tutte le strutture hanno registrato una diminuzione dei documenti gestiti (Tabella 1/ Figura 1).

	Docu- menti in entrata 2019	Docu- menti in entrata 2020	Variazione Documenti in Entrata	Documen- ti in Uscita 2019	Documen- ti in Uscita 2020	Variazione Docu- menti in Uscita	Protocollo Totale 2019	Protocollo Totale 2020	Variazione Protocollo Totale
Unical	9367	7371	-21,31%	3308	1988	-39,90%	15087	11343	-24,82%
Unimc	5552	3704	-33,29%	7411	4424	-40,30%	14447	9259	-35,91%
CNR	19969	14711	-26,33%	29423	20719	2,55%	93368	69344	-34,64%
UniPi	32235	20678	-35,85%	36563	17230	-52,88%	81132	47259	-41,75%

Tab. 1. Variazione percentuale documenti protocollati

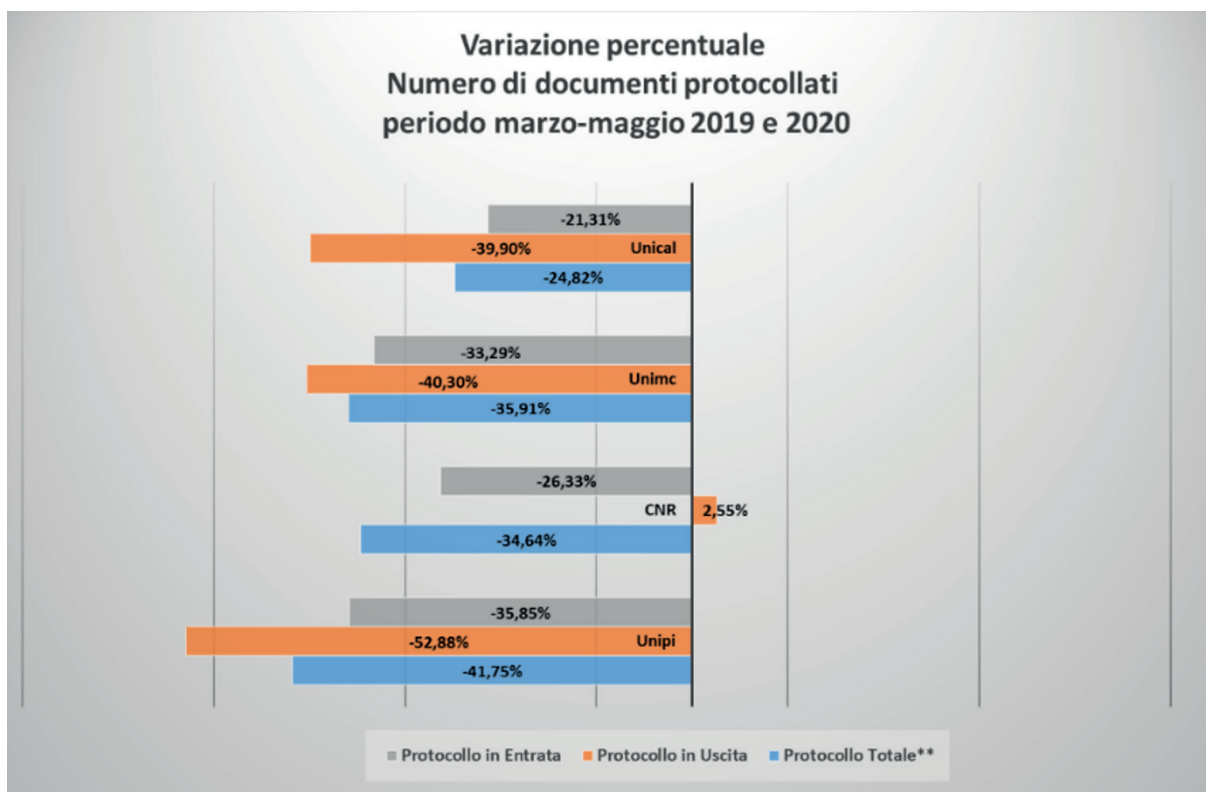


Fig. 1. Variazione percentuale documenti protocollati

I dati relativi alle performance dei procedimenti gestiti dall'amministrazione centrale sono stati analizzati confrontando le variazioni relative ai periodi marzo-maggio 2019 e 2020 sia in termini di quantità che di tempi. Il confronto è stato effettuato solo per l'amministrazione centrale dell'Università della Calabria e dell'Università di Macerata che hanno fornito questa tipologia di dati. Le variazioni a segno negativo indicano una diminuzione della quantità dei processi gestiti tra il 2019 e il 2020, viceversa quelle a segno positivo stanno ad indicare un aumento (Tabella 2, Figura 2). Tra i procedimenti presi in considerazione per il confronto, le approvazioni degli atti per le procedure di reclutamento, le autorizzazioni per lo svolgimento degli incarichi esterni, gli interpellati interni, e l'assistenza fiscale/retributiva/previdenziale.

	Variazione % Quantità di procedimenti	
	Unical	Unimc
Approvazione Atti procedure di reclutamento	53,33%	-22,22%
Autorizzazione incarichi esterni	-45,45%	-12,50%
Interpellati interni	-47,37%	-
Assistenza fiscale/retributiva/previdenziale	-2,30%	-6,67%

Tab. 2. Variazione percentuale quantità di procedimenti

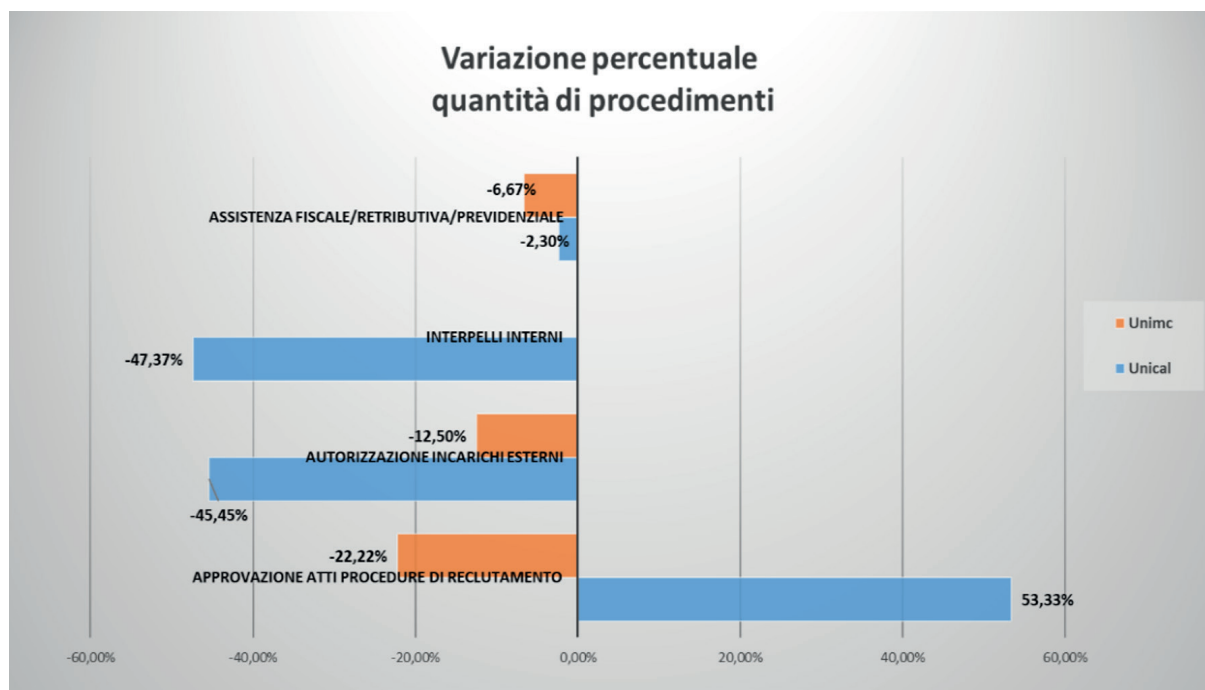


Fig. 2. Variazione percentuale quantità procedimenti

Per calcolare i tempi medi di conclusione del procedimento relativo all'approvazione degli atti delle procedure di reclutamento è stato considerato l'arco temporale dalla data di acquisizione dei verbali alla data di approvazione degli stessi; per le autorizzazioni allo svolgimento degli incarichi esterni sono stati considerati i giorni dalla data di richiesta dell'autorizzazione alla data di protocollo della nota che autorizza l'incarico; per gli interpellati interni sono stati considerati i tempi dalla data di protocollo della richiesta alla data di protocollo dell'avviso di interpellato; per i procedimenti relativi all'assistenza fiscale/retributiva/previdenziale è stato preso in considerazione l'intervallo compreso dalla data di richiesta di assistenza alla data di risposta della richiesta entrambe desumibili dalle mail (Tabella 3/Figura 3).

	Variazione % media dei giorni	
	Unical	Unimc
Approvazione Atti procedure di reclutamento	-49,72%	-41,19%
Autorizzazione incarichi esterni	-14,95%	4,43%
Interpellati interni	-32,94%	-
Assistenza fiscale/retributiva/previdenziale	-25,37%	0%

Tab. 3. Variazione percentuale media dei giorni

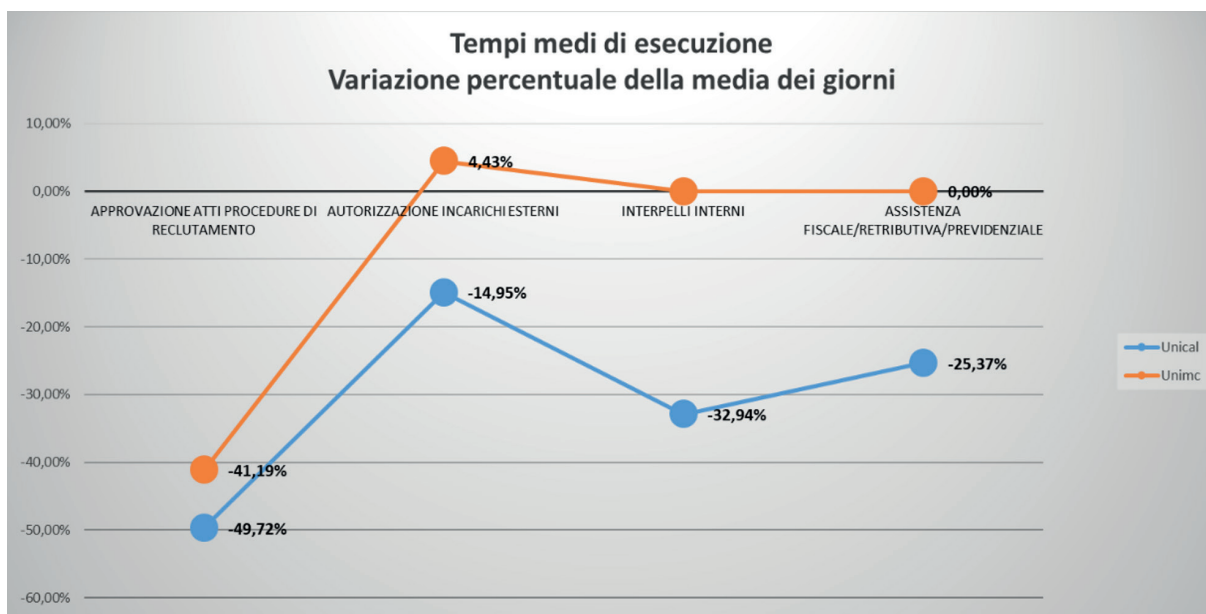


Fig. 3. Variazione percentuale della media dei giorni

Altro aspetto interessante per il confronto tra l'Università della Calabria e l'Università di Macerata ha riguardato l'analisi dei tempi medi impiegati per la conclusione dei procedimenti relativi agli ordini, alle fatture e alle missioni gestiti dai dipartimenti che hanno dato la loro disponibilità a fornire i dati. In particolare per l'Università della Calabria sono stati analizzati i tempi medi dei procedimenti gestiti dal Dipartimento di Matematica e Informatica (Demacs), dal Dipartimento di Culture Educazione e Società (Dices), dal Dipartimento di Farmacia e Scienze della Salute e della Nutrizione e dal Dipartimento di Economia, Statistica e Finanza (Desf) e ne sono state calcolate le variazioni percentuali nel periodo interessato (Tabella 4/Figura 4).

	Unical											
	Demacs Dipartimento di Matematica e Informatica			Dices Dipartimento di Culture, Educazione e Società			Dipartimento di Farmacia e Scienze della Salute e della Nutrizione			Desf Dipartimento di Economia, Statistica e Finanza		
	2019	2020	Var.	2019	2020	Var.	2019	2020	Var.	2019	2020	Var.
Ordini	6,3 gg	8,0 gg	1,75 gg	8,6 gg	3,8 gg	-4,8 gg	8,6 gg	3,8 gg	-4,8 gg	8,2 gg	16 gg	7,8 gg
Fatture	-11,33 gg	-2,26 gg	9,07 gg	-17,8 gg	-15,7 gg	2,06 gg	19 gg	-5,4 gg	-24,4 gg	18,9 gg	13,74 gg	-5,16 gg
Missioni	-	-	-	8,8 gg	23,4 gg	14,6 gg	19,6 gg	35,2 gg	15,6 gg	7,3 gg	9,7 gg	2,33 gg

Tab. 4. Variazione tempi medi dei procedimenti gestiti dai dipartimenti (Unical)

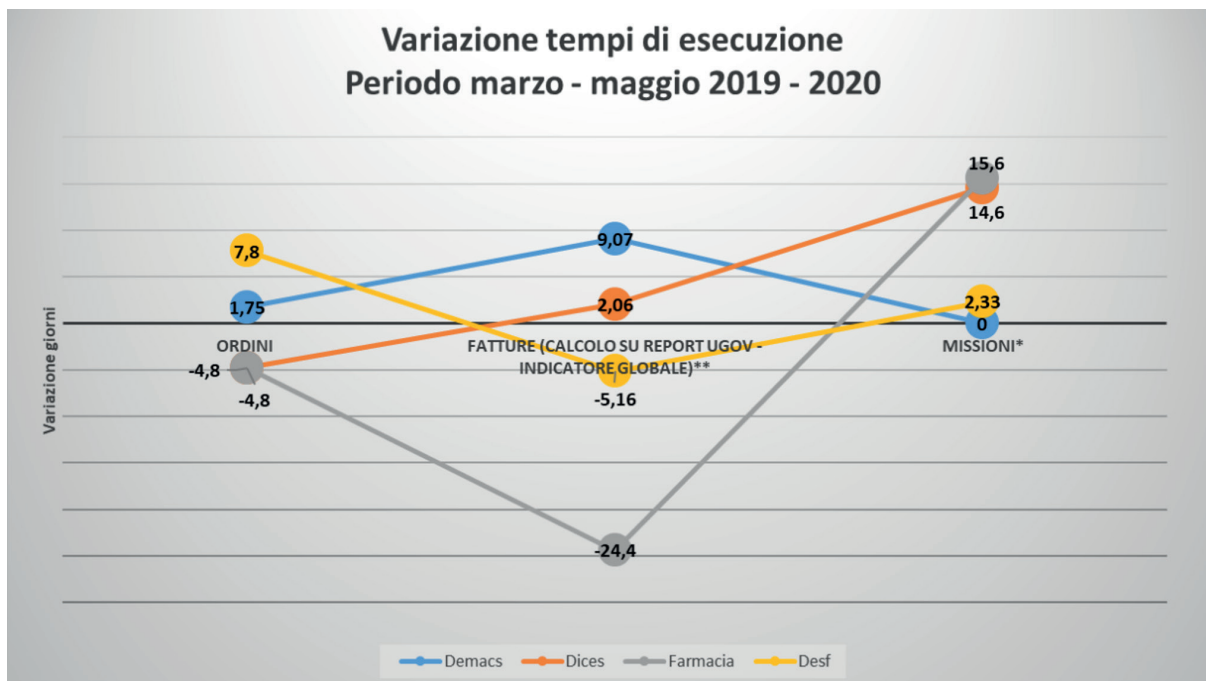


Fig. 4. Variazione tempi di esecuzione dei procedimenti gestiti dai dipartimenti (Unical)

Lo stesso calcolo è stato effettuato per l'Università di Macerata. In particolare (Figura 5), per il Dipartimento di Scienze della Formazione dei Beni Culturali e del Turismo (SFBCT), per il Dipartimento Economia e Diritto (DED) e per il Dipartimento Scienze Politiche Comunicazione e Relazioni Internazionali (SPOCRI). Per gli ordini si è tenuto conto del tempo medio dalla data di richiesta fino alla generazione dell'ordine, per le fatture il dato è riferito alla differenza tra la data di scadenza (30 giorni) e la data di pagamento, per le missioni è stato considerato il tempo medio dalla data di richiesta liquidazione fino alla generazione della missione su sistema gestionale U-GOV.

	Unimc								
	SFBCT Dipartimento di scienze della formazione, dei beni culturali e del turismo			DED Dipartimento Economia e Diritto			SPOCRI Dipartimento Scienze Politiche Comunicazione e relazioni internazionali		
	2019	2020	Var.	2019	2020	Var.	2019	2020	Var.
Ordini	8 gg	9 gg	1 gg	10,0 gg	10,0 gg	0 gg	10,0 gg	10,0 gg	0 gg
Fatture	18,42 gg	28,64 gg	10,22 gg	13 gg	28,44 gg	15,44 gg	14,76 gg	17,4 gg	2,64 gg
Missioni	25 gg	36 gg	11 gg	10,0 gg	10,0 gg	0 gg	10,0 gg	10,0 gg	0 gg

Tab. 5. Variazione tempi di esecuzione dei procedimenti gestiti dai dipartimenti (Unimc)



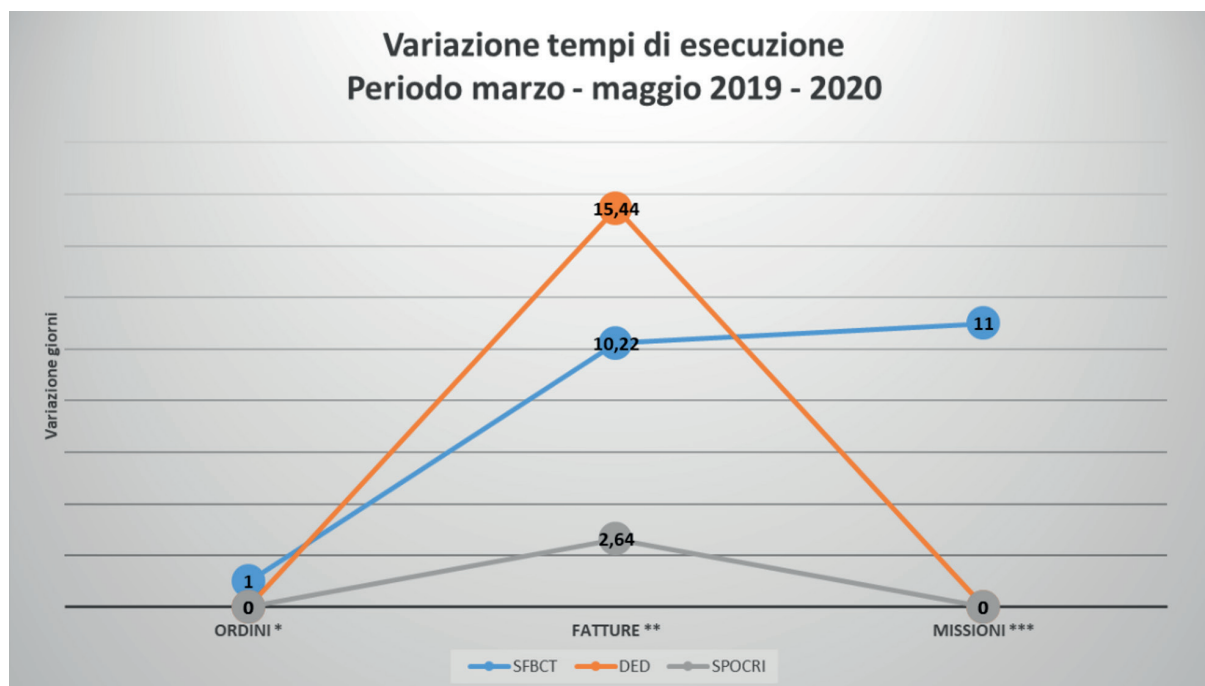


Fig. 5. Variazione tempi medi dei procedimenti gestiti dai dipartimenti (Unimc)

### 3. Lavoro agile e modelli organizzativi

Un sommario sguardo ai numeri che emergono dall'indagine conduce alla conclusione che, indipendentemente dalla variazione della quantità di lavoro registrata nei periodi oggetto di studio, le amministrazioni interessate - attraverso la rimodulazione delle modalità di esecuzione della prestazione lavorativa 'a distanza' dei propri dipendenti, impropriamente chiamata lavoro agile - hanno migliorato le proprie performance organizzative.

È altrettanto vero, tuttavia, che il tumultuoso e radicale cambio della modalità lavorativa è dipeso esclusivamente dall'emergenza sanitaria e non è stato accompagnato, contestualmente, dall'adozione di un modello organizzativo che potesse conciliare il binomio lavoro-esigenze di vita bensì quello lavoro-salute pubblica.

La differenza è sostanziale e merita un approfondimento.

Il sistema pubblico italiano e - nel caso che qui interessa - le Università e il CNR, hanno sperimentato forzatamente il modello del crisis management, inteso come il management delle situazioni eccezionali o che non rientrano nell'ordinario: tale modello può essere molto utile per capire quali elementi organizzativi, sperimentati in questa fase di emergenza, possono essere adottati in modo permanente dalle organizzazioni (Barnett et al. 2000).

L'obiettivo è quello di individuare i versanti sui quali operare per attivare un effettivo cambiamento che possa durare anche dopo la crisi.

Il punto di partenza è rappresentato dalla necessità di dover superare il concetto di una governance dell'emergenza in cui il cambiamento organizzativo (con tutti i limiti che il termine ha potuto significare nel pieno della pandemia) è stato imposto dall'esterno, anche in tempi rapidi,

e non motu proprio, dalla consapevolezza che il lavoro agile altro non è stato che una forma di delocalizzazione del lavoro (meglio, lavoro a domicilio) in quanto misura di contrasto e di contenimento del fenomeno epidemiologico piuttosto che espressione di una diversa e maturata cultura organizzativa.

Non basta, quindi, acquistare PC portatili, razionalizzare gli spazi, prescindere dalla presenza fisica del lavoratore, destrutturare il tempo e il luogo del lavoro per poter parlare di lavoro agile: occorre, molto più ‘semplicemente’, riscrivere un intero paradigma organizzativo del lavoro, per cogliere tutte le opportunità che questa metodologia di prestazione lavorativa offre e minimizzare i rischi ai quali, inesorabilmente, si va incontro.

L'emergenza sanitaria ha rappresentato, indubbiamente, un fattore di accelerazione del cambiamento e ha creato le condizioni affinché le organizzazioni votate ‘all’apprendimento’ possano ripensare i propri meccanismi di funzionamento attraverso il duplice problema di riflettere su come generare maggiore valore attraverso il lavoro agile rispetto alle tradizionali modalità di lavoro in presenza e su come trasferire e far percepire questo valore ai cittadini-utenti, partendo dal presupposto che alcuni lavori non possono (e quindi non devono) essere svolti a distanza (Valotti et al. 2000; Barbieri et al. 2020).

Quale potrebbe essere, pertanto, un percorso da seguire per approdare a un vero cambiamento organizzativo e culturale finalizzato all’adozione del lavoro agile e che accompagni anche un’evoluzione dei modelli organizzativi?

Nel 2019, la Commissione Europea, nella Comunicazione “The Workplace of the Future in the European Commission”, ha evidenziato che “future ways of working should enable the organisation, its teams and its staff to deliver results more effectively and efficiently”<sup>2</sup>: il principio organizzativo dal quale partire per attivare una virtuosa operazione di cambiamento potrebbe essere proprio questo.

In questo senso, le amministrazioni si troverebbero ad un bivio: incrementare il lavoro a distanza senza alcuna modifica al proprio modello organizzativo oppure adottare un nuovo modello organizzativo basato sul riconoscimento di flessibilità e di autonomia al lavoratore a fronte di una maggiore responsabilizzazione sui risultati.

La prima ipotesi è piena di insidie che si manifestano sotto forma di rischi organizzativi: anzitutto, un debole coordinamento manageriale nei confronti della linea operativa, una mancanza di motivazione e soddisfazione del lavoro, l’ennesimo impulso al mito del dipendente pubblico ‘fannullone’ che sarebbe esentato dalla timbratura del cartellino, il risultato inteso come unico parametro/indicatore di funzionalità del lavoro agile, una occasione di trasferimento della responsabilità manageriale al lavoratore e, per ultimo di questo esemplificativo elenco, quello che è stato definito come *time porosity* (Genin 2016; Fenoglio 2018) ovvero il rischio di una connessione ininterrotta in cui tempi di vita e di lavoro si sovrappongono e si confondono.

La seconda ipotesi potrebbe essere compendiata dal modello organizzativo più appropriato nella gestione del *crisis management* che è quello dell’assegnazione di obiettivi, della delega e della responsabilizzazione dei dipendenti attraverso l’incentivazione a sviluppare metodologie di soluzione di problematiche lavorative quotidiane e di capacità di pianificazione del lavoro. In tal

---

<sup>2</sup> <[https://ec.europa.eu/transparency/documents-register/detail?ref=C\(2019\)7450&lang=en](https://ec.europa.eu/transparency/documents-register/detail?ref=C(2019)7450&lang=en)>.

senso, il lavoro agile può rappresentare un'occasione proprio perché basato su una combinazione di flessibilità, autonomia e collaborazione (Cinque et al. 2020).

Il modello organizzativo appena accennato deve essere preparato, costruito, alimentato e gestito. E non può prescindere dalla presenza di tre fattori abilitanti:

- l'organizzazione e i processi (occorre una continua tensione all'innovazione degli assetti organizzativi e dei processi lavorativi);
- le tecnologie informatiche (incrementare gli investimenti tecnologici e infrastrutturali);
- le risorse umane (in particolare, garantire una progressiva 'intensità formativa' per allineare conoscenze e capacità organizzative alle nuove esigenze lavorative, per favorire un approccio culturale alla nuova modalità lavorativa e per consentire l'acquisizione e/o il consolidamento delle competenze richieste per lavorare in modalità 'agile').

L'esperienza ci ha insegnato che una organizzazione del lavoro snella e flessibile si adatta più rapidamente ai cambiamenti esterni ed è più facile da ridisegnare all'occorrenza: il decentramento di responsabilità, lo sviluppo delle relazioni orizzontali, il team work (anzi, i gruppi di lavoro 'ibridi' e trasversali, organizzati per processo e non per funzione), una direzione per obiettivi, una cultura organizzativa orientata alla condivisione dei valori, sono tutte peculiarità che contribuiscono a creare le precondizioni per l'adozione di una modalità 'agile' utile (per l'organizzazione) e consapevole (per i lavoratori).

Analogamente, il pensiero che esista il migliore modello organizzativo del lavoro è palesemente errato: ogni organizzazione deve comprendere quale modello può essere il più appropriato con il proprio modello in essere e con quello verso cui tendere nel futuro. E ciò è possibile solo se sono ben conosciuti il livello di complessità organizzativa e le peculiarità gestionali, compreso il grado di auto-motivazione dei lavoratori, intesa come coinvolgimento e adesione valoriale all'organizzazione.

Tra gli strumenti organizzativi finalizzati a promuovere il lavoro agile – anche in funzione di evitare il rischio di autoreferenzialità<sup>3</sup> - occorre evidenziare la misurazione e la valutazione della performance: non è questa la sede per proporre una riflessione sulla dimensione della misurazione del lavoro agile quale fattore abilitante dei nuovi sistemi di performance management delle amministrazioni ma è utile ribadire che la misurazione del lavoro agile dà concretezza a quello che, altrimenti, rischia di essere un auspicio frutto di una visione ottimistica ovvero l'idea che il lavoro svolto non in presenza possa garantire, solo perché attivato, di innovare i comportamenti delle persone, gli stili direzionali e, quindi, la produttività a livello individuale e organizzativo<sup>4</sup>.

L'adozione del modello organizzativo è funzione della tipologia di lavoro agile che si intende attivare: al di là delle tradizionali quattro dimensioni (dove si svolge l'attività lavorativa, chi svolge questa attività, quando si svolge la prestazione di lavoro e, per ultimo, i contenuti della stessa<sup>5</sup>) è fondamentale indagare la compatibilità effettiva del lavoro agile con le esigenze funzionali del

---

<sup>3</sup> Il lavoro agile non deve essere (solo) oggetto di misurazione quanto strumento di verifica per valutare se ha favorito il conseguimento degli obiettivi prefissati.

<sup>4</sup> Angeletti, Sauro. 2020. "Misurare il lavoro agile nelle amministrazioni pubbliche: profili, problemi e prospettive." *Rivista Italiana di Public Management* 3(1).

<sup>5</sup> In "Il lavoro agile: un'occasione da non perdere" si individuano le corrispondenti quattro dimensioni: 1) spatial flexibility, 2) numerical flexibility, 3) temporal flexibility, 4) functional flexibility (Valotti et al. 2000, 301-303).

servizio, con le aspettative di utilità di questa forma di organizzazione del lavoro nonché con l'accessibilità (sicura) da remoto delle piattaforme dell'amministrazione, la disponibilità di strumentazione nonché la conoscenza dell'utente esterno, in termini di possesso di un certo livello di alfabetismo digitale.

Lo step successivo è chiedersi la finalità del lavoro agile: misura di welfare aziendale per promuovere il benessere organizzativo dei propri dipendenti? oppure misura diretta ad incrementare la complessiva produttività e la performance dell'organizzazione? oppure ancora misura di sviluppo organizzativo inteso come strumento di lavoro utile a ripensare i processi e le modalità di erogazione dei servizi al fine di migliorare la produttività dell'ente nel lungo periodo? (Valotti et al. 2000). Ognuna di queste prospettive si porta dietro specifiche scelte in ambito di finalità, valori, logiche e modelli concettuali del lavoro agile.

A tal proposito, recentemente è stato detto di condividere l'esigenza di porsi nell'ottica del superamento della gestione emergenziale e di individuare quale via per la definizione del piano di sviluppo del lavoro agile nella pubblica amministrazione quella della contrattazione collettiva destinata ad occuparsi della definizione di una disciplina che garantisca condizioni di lavoro trasparenti, che favorisca la produttività e l'orientamento ai risultati, concili le esigenze delle lavoratrici e dei lavoratori con le esigenze organizzative delle pubbliche amministrazioni, consentendo, ad un tempo, il miglioramento dei servizi pubblici e dell'equilibrio fra vita professionale e vita privata, così riaffermando la duplice finalità dello strumento che si integra nelle pubbliche amministrazioni con la specifica attenzione ai servizi offerti alla utenza (Brunetta et al. 2021).

L'adesione a un modello organizzativo evoluto (quale può essere, ad esempio, la costruzione di una cultura organizzativa basata sull'innovazione e sulla propensione al cambiamento) potrebbe scontare un tipico vizio 'italico' da taluno (Ongaro e Valotti 2008) denominato "implementation gap": si tratta della significativa distanza tra quanto scritto nei documenti (di riforma e/o di programmazione) e quanto in realtà concretamente implementato.

Detto ciò, la necessità di un cambiamento organizzativo delle nostre amministrazioni attraverso l'adozione di un nuovo modello organizzativo non è più procrastinabile: il momento della 'resa dei conti' è arrivato e occorre operare isolando il condizionamento di 'tifosi' o 'detrattori' (entrambi interessati) del lavoro agile per abbandonare, definitivamente, modelli organizzativi inadeguati alle esigenze funzionali e con scarsa propensione alla qualità dei servizi offerti, ormai superati dai tempi e non più idonei a garantire, con reale efficacia, la sostenibilità di nuove e imprevedute modalità di lavoro.

## Riferimenti bibliografici

- Angeletti, Sauro. 2020. "Misurare il lavoro agile nelle amministrazioni pubbliche: profili, problemi e prospettive." *Rivista Italiana di Public Management* 3(1): 64-91.
- Barnett, Carole K., and Michael G. Pratt. 2000. "From threat-rigidity to flexibility: Toward a learning model of autogenic crisis on organizations." *Journal of Organizational Change Management* 13 (1): 74-88.
- Brunetta, Renato e Michele Tiraboschi. 2021. "Il lavoro agile o smart working nella pubblica amministrazione: prospettive e criticità." *Working paper* 10, ADAPT University Press: 4-17. <[https://moodle.adaptland.it/pluginfile.php/60594/mod\\_resource/content/1/wp\\_2021\\_10\\_brunetta\\_tiraboschi.pdf](https://moodle.adaptland.it/pluginfile.php/60594/mod_resource/content/1/wp_2021_10_brunetta_tiraboschi.pdf)>.
- Cinque, Pierpaolo, Patrizia Falzetti, e Giuseppina Le Rose. 2020. "Smart working ed emergenza sanitaria: una lettura prospettica a partire dall'esperienza dell'Invalsi del Covid-19." *Rivista Italiana di Public Management* 3 (2): 259-274.
- Genin, Émile. 2016. "Proposal for a theoretical framework of time for the analysis porosity, *International Journal of Comparative Labour Law and International Relations* 32 (3): 280-300.
- Fenoglio, Anna. 2018. "Il diritto alla disconnessione del lavoratore agile." In *Commentario breve allo statuto del lavoro autonomo e del lavoro agile*, a cura di Zilio Grandi, Gaetano e Marco Biasi, 547- 555. Italia: Wolters Kluwer Cedam.
- Ianni, Pierpaolo. 2020. "Cenni sulla Pandemia Spagnola: riflessioni su alcune fonti d'archivio parlamentari." *MemoriaWeb* 30: 1-13.
- Mohadab, Mohamed El, Belaid Bouikhalene, and Said Safi. 2020. "Bibliometric method for mapping the state of the art of scientific production in Covid-19." *Chaos Solitons Fractals* 139: 110052. <https://doi.org/10.1016/j.chaos.2020.110052>.
- Molino, Monica, Emanuela Ingusci, Fulvio Signore, Amelia Manuti, Maria L. Giancaspro, Vincenzo Russo, Margherita Zito, and Claudio G. Cortese. 2020. "Wellbeing Costs of Technology Use during Covid-19 Remote Working: An Investigation Using the Italian Translation of the Technostress Creators Scale" *Sustainability* 12, no. 15: 5911. <<https://doi.org/10.3390/su12155911>>.
- Ongaro, Edoardo, e Giovanni Valotti. 2008. "Public management reform in Italy: explaining the implementation gap." *International Journal of Public Sector Management* 21 (2): 174-204.
- Teixeira da Silva, Jamie A., Panagiotis Tsigaris, and Mohammadamin Erfanmanesh. 2020. "Publishing volumes in major databases related to Covid-19." In *Scientometrics* 126: 831-842. <<https://doi.org/10.1007/s11192-020-03675-3>>.
- Valotti, Giovanni, Marta Barbieri, Giorgio Giacomelli, Lorenza Micacchi, e Francesco Vidè. 2020. "Il lavoro agile: un'occasione da non perdere." *Rivista Italiana di Public Management* 3 (2): 299-319.

## A new archives definition

Gianni Penzo Doria<sup>(a)</sup>

a) University of Insubria, <https://orcid.org/0000-0002-7654-6638>

**Contact:** Gianni Penzo Doria, [gianni.penzodoria@gmail.com](mailto:gianni.penzodoria@gmail.com)  
**Received:** 31 January 2022; **Accepted:** 4 March 2022; **First Published:** 15 May 2022

### ABSTRACT

This paper aims to give a new definition of the word “Archives” as a result of an accurate survey of doctrinal orientations on this specific subject, analyzing - word by word – each particular lemma which makes up the new proposal. Even though there is no definition that is valid in every context, but archival theory seems to suffer largely from a heterogeneous vision, downstream of a rigorous scientific journey, but dense of conceptual and lexical unexpected contingencies. Each term - with meanings and contexts - does not constitute a simple signifier, but refers to a specific baggage of every disciplinary and cultural tradition. For these reasons, we will attempt to provide an overview of the definitions of archives, from the particular to the general.

### KEYWORDS

Archives; Archival Science; Definition; Archival Theory

## Una nuova definizione di archivio

### ABSTRACT

L'articolo propone una definizione della parola “archivio” a valle di un'indagine sugli orientamenti dottrinali in materia, analizzando – parola per parola – i lemmi in cui si compone la nuova proposta. Pur non esistendo una definizione valida in ogni contesto, il panorama dottrinale sembra soffrire in larga misura di una visione eterogenea, a valle di un cammino scientifico rigoroso, ma denso di imprevisti concettuali e lessicali. Ogni parola – con accezioni e contesti – non costituisce un significante semplice, ma rinvia a un bagaglio proprio di ogni tradizione disciplinare e culturale. Per queste ragioni, tenteremo di fornire una visione d'insieme, dal particolare al generale, delle definizioni di archivio.

### PAROLE CHIAVE

Archivio; Archivistica; Definizione; Teoria archivistica

## 1. Il nodo delle definizioni

L'archivio è il complesso dei documenti prodotti da un soggetto nel corso della propria attività. Prima di illustrare le ragioni che hanno condotto alla sintesi appena enunciata, è opportuno esaminare il nodo delle definizioni lungo il percorso evolutivo dell'archivistica e le ragioni sottese a questa scelta. Infatti, premesso che non esiste una definizione valida in ogni contesto, il panorama dottrinale sembra soffrire in larga misura di una visione eterogenea, a valle di un cammino scientifico rigoroso, ma denso di imprevisti concettuali e lessicali.

Partiremo dal particolare, cioè dalle singole parole di cui è composta la definizione, fino a giungere al quadro d'insieme come risultato provvisorio di una elaborazione convenzionale. Infatti, non sembrano esistere regole o linee guida condivise per poter redigere una definizione. Anzi, persino all'interno degli orientamenti più inclini a un formalismo rigido (ad es., la linguistica e, in particolare, quella d'orientamento generativista), il problema risulta poco sentito o scarsamente affrontato<sup>1</sup>. Inoltre, nel caso di definizioni semantiche e lessicali, i confini della trattazione sono oltremodo vasti, fino quasi a dissolversi. Preliminarmente, in questi casi risulta necessaria una panoramica d'orizzonte della lessicografia. La procedura è più simile a quella delle scienze giuridiche che a quella di discipline quantitative o a forte impianto matematico, quanto ad analisi con ispezione della dottrina.

## 2. Obiettivo e quadro di riferimento

Questo lavoro ha come obiettivo la ricerca di una definizione generale e di sintesi della parola archivio. È bene premettere che ogni tentativo in questo campo rappresenta una compartimentazione, una camicia di forza lessicale entro la quale ingabbiare un concetto.

Di per sé può risultare un'operazione rischiosa e transitare da una questione generale a una generalista, di solito poco efficace, anche perché ogni definizione assume un senso in un contesto determinato. Al tempo stesso, però, rappresenta un imprescindibile punto di partenza per dialogare con la comunità professionale di riferimento con metodo scientifico e condiviso. Si tratta, infatti, di stringere in una sinecdoche estesa a più lemmi un concetto che l'analisi di molte parole finirebbe per fuorviare, anche involontariamente<sup>2</sup>.

Tuttavia, nella relazione con altre discipline, prima di intraprendere un progetto di largo respiro, bisogna intendersi con le tecniche con cui ci esprimeremo. Ogni parola – con accezioni, significati e contesti – non costituisce un significante semplice, ma rinvia a un bagaglio proprio di ogni tradizione disciplinare e culturale. Per queste ragioni, tenderemo di fornire una visione d'insieme<sup>3</sup>.

In questa sede, ci riferiremo all'archivio in senso proprio, come magistralmente descritto da Filippo Valenti a valle di un ragionamento esteso<sup>4</sup>.

---

<sup>1</sup> Per la grammatica generativa e la grammatica trasformazionale, cfr. (Graffi 2008), anche per i riferimenti a Noam Chomsky e, sul potere delle parole, ai due concetti di *esecuzione* e di *competenza*.

<sup>2</sup> Per una disamina della terminologia tecnica – rivista dall'autrice – si veda il denso saggio di (Carucci 2014: 141).

<sup>3</sup> Un approccio ricco di spunti teorici per il nostro lavoro si ricava dalle teorie di (von Humboldt 2004), per il quale la lingua deve “fare un uso infinito di mezzi finiti”, anche nei rapporti tra filosofia del diritto e interpretazione giuridica. Sul punto (Palumbo 2016: par. 5.1).

<sup>4</sup> (Valenti 2000: 151-152), che riprende (Valenti 1981: 30). L'espressione, tuttavia, non è originale e risale almeno a (Bach-

Di contro, non ci occuperemo dell'archivio come istituzione, né come luogo, né delle accezioni ulteriormente desumibili dalla ormai considerevole dottrina, anche se la vecchia definizione di Ulpiano vive nei trattati eruditi di archivistica. Tra i molti, Trivelli alla fine del XIX secolo sostenne ancora come il termine archivio indicasse “il luogo ove gli atti si conservano” (Trivelli 1898: 1)<sup>5</sup>.

### 3. Come scrivere una definizione di archivio

La definizione dovrebbe essere elaborata seguendo almeno alcuni criteri di carattere generale:

- *sintesi*: troppe parole offuscano e rendono ambiguo un concetto; infatti, a scorrere le definizioni degli ultimi quattro secoli, con frequenza ci si imbatte in asindetici e ciò rende ancora più problematica la comprensione della proposizione principale e, in molti casi, di quella secondaria;
- *neutralità*: le definizioni che si estrinsecano nelle tecnologie, nei supporti, nelle modalità e, non ultimo, nelle caratteristiche indotte, dedotte o dipendenti, finiscono per appesantire ed eludere il cuore del problema definitorio, fino a circoscrivere la natura del concetto; in questi casi, i dettagli applicativi sono definibili nei corollari esplicativi;
- *anaggettivazione*: l'uso di aggettivi finisce incontrovertibilmente con perimetrare, in forma limitativa e orientativa, il contenuto lessicale;
- *ontologia*: molte definizioni perdono di vista la sintesi dell'essenza a favore dell'enucleazione delle caratteristiche; quest'ultime, invece, dovrebbero essere trattate nelle esemplificazioni, anche per evitare tautologie;
- *ana-esemplificazione*: qualsiasi esemplificazione, anche la più semplice, delimita e circonda in modo coercitivo un concetto; gli esempi devono essere rinviati ai corollari o alle esplicitazioni dei concetti principali e strutturali.

Per questo, in un discorso fluente, scritto e parlato, le esemplificazioni aiutano – come questa incidentale – a comprendere il significato con maggiori dettagli e a esplicitare i concetti. Di contro, all'interno di una definizione finiscono inevitabilmente per sfuocarla, fino a renderla limitata a un ambito settoriale di una disciplina<sup>6</sup>.

---

mann 1801: 2): “Archiv im eigentlichen Sinn heist die briefliche Schatzkammer eines Fürsten, worinnen die allerwichtigsten, nothwendigsten un kostebaren Urkunden (Originale, Dokumente, Diplome) und Akten verwahrt werden, welche das fürstliche Haus selbst, dessen Würde, Hoheit, Interesse, Vorrechte, Land und Leute, die innerliche und äusserliche Staatsverfassung in allen ihren Theilen betreffen” (l'archivio in senso proprio è la camera del tesoro epistolare di un principe, in cui sono conservati i più importanti, i più necessari e i più preziosi documenti (originali, documenti, diplomi) e atti che concernono la stessa casa reale, le sue dignità, sovranità, interessi, prerogative, territorio e uomini, la costituzioni interna ed esterna dello Stato in tutte le sue parti).

<sup>5</sup> Notizie puntuali in (Mottola 1998: 377), con spunti tratti da (Nicolini 1971).

<sup>6</sup> Analogo discorso di linguaggio tecnico, pur in contesto differente, può essere riferito alla nomenclatura dei gradi divisionali da scegliere nella redazione di un titolo. Anche in questo caso, una voce come *archivio e gestione documentale* oppure come *accesso civico e accesso documentale* non trovano agio in quanto definiscono in forma ultronea e, pertanto, inefficace, la sostanza. La gestione documentale, infatti, è ricompresa nella parola *archivio*, così come le tipologie di accesso possono essere efficacemente sintetizzate sotto la parola *accesso*: parole generali (non generiche) e inclusive.



## 4. La scelta delle parole: da *puzzle* ad *anagramma*

Le parole di una definizione sono come le tessere di un *puzzle*. Servono tutte, nel rispettivo ordine e nella loro posizione, ma non una di più, né una di meno. È un po' come ritrovarsi di fronte alla struttura anagrammatica di un archivio in disordine o in disordine apparente. Si inizia a ricostruire il senso della sedimentazione, a intuirne l'ordinamento e la struttura, per poi iniziare il riordino, con mille variabili aperte.

Inoltre, le parole utilizzate devono servire tutte in quanto indispensabili al complesso definitorio. Quelle in più, invece, possono essere espunte. Ogni specificazione può essere agevolmente colta dall'esegesi o da ulteriori corollari.

Quindi, una definizione dovrebbe essere proposta senza pause, senza parentesi, senza trattini incidentali, nel fluire di un concetto che reca l'essenzialità e non trova interruzioni, senza offuscare l'ontologia di un concetto.

## 5. Perché una definizione in regime di certezza provvisoria?

Una definizione accorcia la distanza concettuale tra le cose e il loro nome, sempre nella logica convenzionale della scrittura e del metalinguaggio come mediazione: descrivo un archivio senza usare l'archivio, ma la lingua italiana. Essa, dunque, rappresenta sempre una certezza provvisoria e, per fortuna, intrinsecamente instabile. Essa è, infatti, legata al contesto contingente sociale, linguistico e professionale di riferimento. Più di chiunque altro, Leopoldo Cassese aveva ben evidenziato i rischi e i limiti di questo lavoro: "Un'esatta definizione dell'istituto che è oggetto della nostra disciplina non è concepibile se non come il temporaneo punto di approdo di tutto un corso di esperienza culturale, e quindi come significazione diretta di una elaborazione concettuale collettiva" (Cassese 1959: 46). Anzi, potremmo dire che le definizioni di archivio rappresentano gli elementi più fragili della dottrina archivistica, che fin dalle origini non ha mai smesso di cercare un punto di sintesi. Probabilmente, esse hanno dignità maggiore nel luogo dell'effimero teoretico. Ad esempio, non sappiamo ancora quali strade evolutive prenderà il documento archivistico, tra dati e informazioni. Tuttavia, allo stato attuale, l'archivio può ancora solidamente essere definito come un complesso di documenti (e non di dati, né di informazioni) unitamente alle loro relazioni<sup>7</sup>.

## 6. Le definizioni più recenti di archivio e l'evoluzione dottrinale

Etimologia a parte, moltissimi autori si sono cimentati nella definizione del concetto di archivio. È possibile evidenziare tra tutti una sorta di minimo comune denominatore, un filo conduttore bipartito: da un lato, la doglianza circa la ciclicità di nuove definizioni in un susseguirsi di precisazioni e di aggiustamenti continui; dall'altro, un irrefrenabile impulso di proporre una rassegna, a volte anche copiosa, con ulteriori puntualizzazioni. Anche in questo caso, seguendo una tradizione internazionale, proveremo a districarci in questa materia solo apparentemente semplice.

---

<sup>7</sup> Ad esempio, la definizione di documento elettronico (non amministrativo, né archivistico) è stata riscritta in maniera "ecumenica" dal Regolamento UE 910/2014, *Electronic IDentification Authentication and Signature – eIDAS*, art. 3 "“electronic document” means any content stored in electronic form, in particular text or sound, visual or audiovisual recording”).

Preliminarmente, possiamo osservare che le definizioni coniate nella storia della disciplina archivistica evidenziano a chiare lettere un progressivo mutamento del modo di concepire l'archivio. In poche parole, la dottrina è tenacemente alla ricerca di un apparato definitorio stabile e inclusivo, non senza incoerenze.

Eugenio Casanova, enucleando due “rari” e un “rarissimi” a inizio paragrafo e prima di iniziarne una rassegna, censurava gli improvvisatori di definizioni di archivio, con toni sardonici e taglienti: “Rari sono, in Italia e altrove, coloro i quali sappiano che cosa sia un archivio; rarissimi, coloro i quali discernano a che veramente serva”.

E ancora:

«Rari sono, nella letteratura relativa, coloro i quali si siano astenuti dalla definizione dell'archivio. I rimanenti si sono espressi in proposito in tal modo che non può del tutto soddisfarci [...] Ultimi, infine, vengono coloro i quali si esprimono in modo vago, incerto, sì da redigere una definizione adattabile non solamente agli archivi, ma a qualsiasi consimile istituto, come a collezioni di autografi, a musei, ec. ec.” (Casanova 1928: VI, 12)<sup>8</sup>.

Casanova mostrava verso il continuo rincorrersi di definizioni segnali chiari di insofferenza, ormai quasi un secolo fa. Poi, però, esponeva la propria: “L'archivio è la raccolta ordinata degli atti di un ente o individuo, costituitasi durante lo svolgimento della sua attività e conservata per il conseguimento degli scopi politici, giuridici e culturali di quell'ente o individuo” (Casanova 1928: 19). Per ampiezza di argomentazioni e per esaustività di trattazione (precettistica, logico-filosofica e giuridica), restano tuttora insuperati – e, spesso, in piena disputa teorica – gli interventi di Elio Lodolini e di Filippo Valenti.

Il primo, in oltre settant'anni di saggi, più volte si è occupato degli aspetti legati alla concettualizzazione dell'archivio, accompagnati spesso da rassegne accurate (Lodolini 1980)<sup>9</sup>.

Invece Valenti, prima di cimentarsi nel proporre una definizione di archivio, ammoniva: “Le definizioni sono in genere sempre discutibili, astratte, inevitabilmente riduttive e quindi parzialmente false; ma possono riuscire utili come base per condurre un certo tipo di discorso”. Poi aggiungeva in tono deciso: “Quelle di ‘archivio’ che si possono leggere nella letteratura relativa sono fin troppe” (Valenti 2000: 150)<sup>10</sup>.

Eppure, nemmeno Valenti riusciva a sfuggire alla tentazione, tanto da enucleare una serie in una ventina di pagine, che occupava tutta la prima parte del suo celebre saggio di archivistica teorica<sup>11</sup>.

---

<sup>8</sup> In quest'ottica, può essere riletta, in particolare, la rassegna in esordio (Casanova 1928: 3-24).

<sup>9</sup> A onor del vero, Lodolini aveva già iniziato una rassegna più di vent'anni prima, anche con nuove proposte di traduzioni dal tedesco e dall'olandese: (Lodolini 1958: 321-323, nota 2). Ne parla anche (Cassese 1959: 28, nota 27).

<sup>10</sup> . Uno dei manuali dell'epoca era quello, peraltro interessante, di (Mazzoleni 1972), anche con una pre-edizione dattiloscritta in circolazione interna per le Scuole di archivistica, paleografia e diplomatica annesse agli Archivi di Stato, 1969/1971.

<sup>11</sup> L'argomento gli era caro, tanto che, analizzando la parola archivio prima scrisse – passando per Sandri e Moscati – “l'approfondimento teoretico del concetto di “archivio”, pur avendo avuto punte assai vivaci, si è esaurito ben presto nella sua stessa fondamentale vanità, ripiegandosi su sé medesimo e ben poco aggiungendo a quanto già acquisito e fin troppo dogmaticamente consacrato”, (Valenti 1975: 184-185). Poi, per mediare e concludere un percorso pluriennale, ribadì che “pochissimi termini, a questo livello di impegno teoretico e in quest'ambito di interessi, sono stati fatti oggetto con altrettanta frequenza e insistenza di esercitazioni definitorie e di tentativi di coglierne, come si suol dire, il concetto; di rispondere cioè alla domanda: che cos'è essenzialmente un archivio?”, (Valenti 1981: 9). Pur in un altro contesto, tratta la questione anche (Carucci 1978, 56-57; 1985: 409-436).

## 6.1 Le definizioni non tecniche

La storia dell'archivistica ci ha consegnato anche una serie di definizioni non tecniche – o, meglio, meno tecniche – fino quasi a sfiorare la forma affettiva. La più bella definizione non tecnica di archivio, qui in senso lato, con richiami alla funzione sociale svolta a tutela della trasparenza, proviene dalla *vision* del NARA: “... dal quale dipende la nostra democrazia...”<sup>12</sup>.

Un'altra definizione, più orientata al campo della filosofia del diritto e della sociologia, era stata coniata da Antonino Lombardo: “L'archivio può definirsi come la registrazione più autentica della vita organizzata di un popolo, di un ente o di una persona” (Lombardo 1958: 346).

Forse un po' eccentrica, ma aperta a spunti ulteriori, risulta quella proposta da Baldacci: “Archivi come sistemi generali della formazione e della trasformazione degli enunciati” (Baldacci 2016: 20).

Anche Londei, riprendendo alcuni celebri scritti di Claudio Pavone, espone la propria, ma orientata agli aspetti strumentali e di supporto, priva di ontologia (cosa *fa* e non cosa *è*): “diremo semplicemente che l'archivio è la struttura per mezzo della quale un soggetto organizza stabilmente la propria memoria documentaria” (Londei 2003: 53).

Tra le più recenti, segnalo quella di Valacchi: “Un complesso vitale di documenti prodotti da attività umane e abbracciati gli uni agli altri. Un archivio è la forza chiara del diritto, l'urlo della trasparenza, il baluardo contro l'oppressione del pensiero. E poi è sogno, immaginazione, retaggio vitale di mondi circostanti spalancati sullo spazio e sul tempo. Memoria, senso di sé, identità. La storia ascolta i sussurri e le grida degli archivi per inventare il passato con cui il presente nutre il futuro. Archivio è una parola d'amore. Per la vita” (Valacchi 2017: 20)<sup>13</sup>.

L'elenco potrebbe continuare, ma non sarebbe più strumentale al nostro discorso.

## 7. La logica incrementale verso una definizione generale

Da secoli si cerca la quadra della definizione di archivio e non ci sono dubbi che la questione durerà ancora. L'approccio al problema segue una logica incrementale ben nota. Anche i più recenti manuali di archivistica non disdegnano una rassegna pressoché esaustiva<sup>14</sup>.

Una collettanea recente ne riporta addirittura quattro nei vari saggi ospitati, a corroborare la lontananza della nostra disciplina dalle scienze dure:

- “L'insieme dei documenti redatti e ricevuti da una persona fisica o giuridica nel corso delle sue attività come loro strumento o residuo, e conservati per proprio riferimento da quella stessa persona o da un suo successore legittimo” (Duranti 2014: 21);

---

<sup>12</sup> National Archives and Record Administration, *2001 Annual Report*: “As McCullough so eloquently stated, the National Archives and Records Administration is a public trust on which our democracy depends. We allow people to see for themselves the workings of our unique government. We make transparent the story of the American people, our collective successes and triumphs as well as our blemishes and failures”; cfr. (Phillips 2004, 1). Rinvio a (Penzo Doria 2009: 27-28, nota 13), anche per la bibliografia ivi citata.

<sup>13</sup> Evidente il filo conduttore con (Valacchi 2015: 59-74).

<sup>14</sup> Sulla *Definizione di archivio*, (Bertini 2008: 11-15); e ancora (Bonfiglio-Dosio 2010: 13-19); (Romiti 2011: 123-158), in particolare l'esteso cap. XII dedicato al nodo delle definizioni, quest'ultimo contenente una proposta dell'autore che esamineremo *infra*; da ultimo (Franzese 2018: 44-53).

- “L’archivio è il complesso dei documenti prodotti o acquisiti da un soggetto nell’esercizio delle sue attività” (Grossi 2014: 35);
- “Ora si parla di *soggetto produttore*, ovvero dell’ente che ha prodotto l’archivio inteso, ovviamente, come insieme di documentazione spedita e ricevuta e, dunque, dell’ente titolare dell’archivio” (Carucci 2014: 167);
- “Archivio come ‘memoria formale dell’attività’” (Bologna 2014: 212)<sup>15</sup>.

Le definizioni di altri autori, a causa di una non adeguata riflessione sul lessico o di una ristretta visione dell’oggetto definito, non risultano condivisibili. Così, infatti, è da rigettare quella proposta da chi definisce l’archivio come “costituito da un complesso di messaggi (documenti), legati fra loro reciprocamente da un vincolo originario, spontaneo e determinato, prodotti e acquisiti da un’amministrazione individuale o collettiva nell’esercizio pratico delle sue funzioni e per il raggiungimento delle proprie finalità” (Lanza, e Golinelli 2006: 18)<sup>16</sup>.

Di tutt’altra impostazione risulta la visione di Leopoldo Sandri, il quale rivolgeva una critica dura a coloro che si attardavano in disquisizioni teoriche sulla definizione di archivio: “La vecchia tecnica del sezionare un capello raggiunge qui le forme dell’arte. Ma come quasi sempre accade in queste raffinatissime fatiche, la loro inconsistenza diviene anche subito palese. Ed infatti, alle tre espressioni non corrispondono tre diverse sostanze, per cui l’uso indifferenziato di quelle possa generare confusione o provocare il pittoresco e irrazionale passaggio da un oggetto all’altro, ma sempre una unica cosa e cioè l’archivio” (Sandri 1958<sup>o</sup>: 121)<sup>17</sup>.

Pertanto, ci concentreremo esclusivamente sulle definizioni in cui, in maniera diretta ed esplicita, si declina la produzione di documenti e delle rispettive tipologie (ricevuti, spediti, acquisiti, etc.). Questa parte lessicale, infatti, risulta strumentale al nostro ragionamento e alla proposta conclusiva per cui esamineremo esclusivamente le definizioni di archivio in lingua italiana, riconoscendo la linea di cautele poste in ogni tentativo definitorio<sup>18</sup>.

Infine, la definizione più sintetica ed efficace di archivio sembrava essere: “La memoria di un soggetto”. Come un lampo, quasi ungarettiana. Successivamente, le osservazioni puntuali di Valenti prima e di Vitali poi, hanno ridimensionato questa proposizione, anche per l’ontologia labile e instabile che il concetto di memoria reca con sé (Valenti 1981: 9)<sup>19</sup>.

---

<sup>15</sup> Non sfugge nemmeno (Paoloni 2014: 430-435), il quale ne propone una rassegna.

<sup>16</sup> Scrittura a parte (*sue e proprie* riferite a un medesimo soggetto nella stessa frase portano ad ambiguità), lessico approssimativo (*spontaneo* in luogo di naturale, *messaggi* al posto di documenti, che trovano agio secondario nella parentesi), la questione è la concezione limitativa dell’archivio: il termine *amministrazione* si riferisce in modo peculiare agli enti pubblici e, in questo modo, sembrerebbe escludere gli archivi di persona fisica, di famiglia, di impresa (tranne il *general management*), di associazione, etc.).

<sup>17</sup> Avrebbe censurato, pertanto, anche questo lavoro. Ci solleva il fatto, speriamo, che anche il grande studioso ha finito con il presentare la propria.

<sup>18</sup> Per il mondo nordamericano rinvio al recentissimo *Dictionary of Archives Terminology* nella nuova e altisonante (*fabulous*) versione rilasciata on-line il 29 aprile 2020: <https://www2.archivists.org/news/2020/new-dictionary-of-archives-terminology-launched>. Un compendio plurilingue si trova in (Duranti 2014: par. 1.2), dedicato interamente al tema *La definizione di archivio*.

<sup>19</sup> Si veda anche (Vitali 2001: 182). Il tema è caro all’autore, che lo riprende anche nel paragrafo *Le molteplici declinazioni della memoria* (Vitali 2007: 103-111).

## 8. Una parola per ciascun concetto

Esaminiamo ora le parole con cui potremmo comporre la definizione di archivio come nella ricostruzione di un *puzzle* o di un anagramma (a seconda della semplicità della struttura). Per ora, mettiamole di fronte a noi, scegliendo una sequenza appropriata al nostro scopo.

### 8.1. Raccolta, insieme o complesso?

Giorgio Cencetti fu certamente un innovatore quando scrisse: “Chiameremo archivio il complesso degli atti spediti e ricevuti da un ente o individuo per il conseguimento dei propri fini o per l’esercizio delle proprie funzioni” (Cencetti 1937: 10).

Questa definizione rappresenta uno spartiacque ideale rispetto a chi utilizzava – finanche Sebastiani e Casanova – la parola *raccolta* per definire l’archivio. Si tratta di un termine espunto, con qualche eccezione minore, dalla manualistica professionale italiana, proprio in ragione dell’antitesi del termine con il concetto stesso di archivio<sup>20</sup>.

Per le ISAD(G) esiste il termine fondo, non archivio. La sostanza non cambia, visto che in quel contesto il fondo è l’archivio in senso proprio: “L’insieme organico dei documenti archivistici, senza distinzione di tipologia o di supporto, formati e/o accumulati e usati da una determinata persona, famiglia o ente nello svolgimento della propria attività personale o istituzionale”<sup>21</sup>.

Non ultimo, giova ricordare che la definizione delle ISAD(G) può essere considerata una rielaborazione di quella già enunciata da Panella: “Chiamiamo archivio l’insieme dei documenti scritti di qualsivoglia specie, ordinatamente raccolti nella loro continuità temporale e fattuale, che un’autorità pubblica, un ente laico o ecclesiastico, una privata azienda, associazione o famiglia hanno prodotto ed accumulato, per fini esclusivamente pratici, durante il loro svolgimento storico, nei rapporti giuridici, sociali e politici” (Panella 1955: 65 e Cassese 1959: 29).

Veniamo al cuore del problema. Il termine *insieme* può essere sostantivo e avverbio: più significati per un significante. Qui faremo riferimento al primo, utilizzato nella lingua italiana per indicare in maniera indistinta le relazioni casuali, instabili o temporanee.

Infatti, in geometria, uno dei sinonimi di *insieme* è raggruppamento, il quale si realizza in forma volontaria (artificiale), ma anche accidentale e variabile. Siamo, pertanto, ancora nel campo delle antitesi al concetto di archivio. L’espressione “l’insieme dei miei interessi”, inoltre, allude a legami contingenti, effimeri e potenzialmente non duraturi nel tempo. In matematica *l’insieme* è una raccolta (termine ancora in antitesi di archivio) di elementi di qualsiasi tipo (geografico, numerico, morfologico o logico), individuati mediante una caratteristica o un criterio che li accomuna e, al

---

<sup>20</sup> Utilizzato ancora da (Sebastiani 1904: 10): “L’Archivio è una raccolta ordinata di documenti a scopo di amministrazione nel senso più lato, esistente o esistita, e che perciò possono emanare sia da una magistratura, sia da un ufficio pubblico o privato”. Ne tratta anche (Lodolini 1965: 107, nota 3). Significativo, infine, il ricordo di (Lombardo 1955: 301, nota 1).

<sup>21</sup> Interessante seguirne a ritroso lo sviluppo concettuale, International Council on Archives, *Statement of Principles Regarding Archival Description. First Version Revised*, pubblicato in Italia in “Archivi e Computer”, II, 1992 (2), 98-105; International Council on Archives, *ISAD(G): General International Standard for Archival Description*, Adopted by ad hoc Commission on Descriptive Standards, Stockholm, Sweden, 21-23 January 1993, Ottawa, Secrétariat de la Commission ad hoc sur les normes de description, 1994. Il testo inglese fu pubblicato nella “Rassegna degli Archivi di Stato”, LIV, 1994 (1), 133-153, con traduzione italiana a cura di Stefano Vitali, in “Rassegna degli Archivi di Stato”, LV, 1995 (2-3), 392-413.

contempo, li delimita. Ancora: una proposta politica raccoglie un insieme di voti, mai un complesso di preferenze. Infine, in scienza dell'informazione, tutti i dati rappresentano un complesso; quelli scelti rappresentano un *set* di dati, cioè un insieme di dati ordinati in maniera occasionale e volontaria come, ad esempio, i dati estratti mediante una *query*.

Il termine *complesso*, di contro, si riferisce in maniera diretta alla lucida intuizione antesignana di *universitas rerum*. Nessuno e nemmeno qualcuno tra gli archivisti più vicini all'ortodossia, infatti, si avventurerebbe a dire "ieri sono andato a cena con un *complesso* di persone". In una battuta, gli archivisti si occupano di relazioni stabili tra i documenti e, pertanto, di *complessi di documenti*, non già di *insiemi di documenti*.

Da ultimo, il termine *complesso* è utilizzato da Giuseppe Bonelli e da Giovanni Vittani attraverso un doppio passaggio di traduzione, dall'originale della Bibbia degli olandesi del 1898 alla traduzione tedesca di Hans Kaiser del 1905: "Archivio è l'intero complesso degli scritti, disegni e stampe, ricevuti o redatti in qualità ufficiale da qualunque autorità o amministrazione, o da qualsiasi impiegato di queste, purché tali documenti, conformemente alla loro funzione, debbano rimanere presso la stessa autorità o amministrazione o presso i suoi impiegati [...] L'archivio è un tutto organico" (Mueller, Feith, e Ruin 1908)<sup>22</sup>.

Pertanto, per definire un archivio in senso proprio utilizzeremo il termine *complesso*.

## 8.2. Il vincolo è insito nel complesso?

Strettamente connesso alla natura di complesso è il vincolo archivistico. Anche in questo caso, esso scaturisce da un'intuizione di Cencetti, il quale sistematizza il pensiero relativo alla organicità dell'archivio già messo a fuoco, come abbiamo visto, dalla Bibbia degli olandesi.

Ulteriormente, nella definizione di archivio è necessario enucleare anche il vincolo oppure possiamo ritenerlo ricompreso nella parola complesso, in ragione del quale potremmo lasciarlo a corollari successivi ed esplicativi della natura di archivio?

Come abbiamo visto *supra*, la differenza sostanziale tra complesso e insieme consiste proprio nella presenza nel primo di un vincolo naturale, necessario e determinato, grazie al quale l'*universitas rerum* può dirsi instante. Possiamo pertanto ritenere che, nell'ambito della definizione di archivio, le precisazioni sul vincolo possano trovare agio adeguato nei corollari esplicativi.

## 8.3. Scritture, atti o documenti?

Il termine *scrittura* ha una peculiarità spiccatamente giuridica, in maniera puntuale in ambito civilistico e processual-civilistico. Quante parole nel Codice civile è possibile riferire alla natura di documento? I sinonimi sono fortemente influenzati dalla tecnologia conosciuta nel 1942: carta, incartamento, scrittura e titolo.

---

<sup>22</sup> Partendo dalla definizione degli olandesi, Elio Lodolini giunse a proporre una traduzione differente, ma sostituendo *complesso* con la parola *insieme*: "Un archivio è l'insieme dei documenti - manoscritti, disegnati o stampati - ricevuti o redatti *ex officio* da una autorità o da un impiegato di essa, ecc." (Lodolini 1958: 322). Tuttavia, al termine dell'elaborazione teorica, il grande archivistica italiano non si discosterà molto dall'approdo elaborato in (Lodolini 1970: 355). Molto interessante anche l'ultimo intervento, soprattutto in virtù dei commenti di Eugenio Casanova (Lodolini 2015).

Ancora, il documento come prova è definito anche atto pubblico (artt. 2699-2701), scrittura privata (2702-2711), scritture contabili (2709-2711), riproduzioni meccaniche (2712), taglie o tacche di contrassegno (2713), copie (2714-2719) e atti di ricognizione o di rinnovazione (2720). Passando al Codice di procedura civile, si aggiunge il piego raccomandato (149). Nel processo penale, infine, si usa distinguere con la parola *atti* i documenti processuali, mentre con la parola *documenti* quelli extraprocessuali (Zacchè 2012: 7-14, 75-121)<sup>23</sup>.

Le tipologie appena menzionate, contrariamente alle convinzioni di Romiti, rappresentano – in punto di fatto e di diritto – documenti e non scritture. Ma l'autore chiarisce il proprio pensiero, affermando che “Alcuni archivisti sostengono che usando *documenti* potremmo comprendere altri elementi che non di rado fanno parte dell'archivio e che non rientrano nelle categorie delle scritture. Personalmente non condivido tale posizione poiché ritengo che quanto si trova incluso in un archivio, senza avere le caratteristiche scritte, non possa essere considerato elemento strutturale, bensì possa essere definito con la qualifica di *allegato*” (Romiti 2011: 158).

Tralasciando il verbo *includere*, la visione è da un lato estensiva del concetto di allegato, dall'altro risulta fuorviante per la natura giuridica delle deduzioni. In questa circostanza, risponde indirettamente Carnelutti, secondo il quale non bisogna confondere l'azione dello scrivere con lo scritto, anche per i margini di ambiguità della parola scrittura: “Poiché la scrittura è un mezzo permanente di manifestazione del pensiero, chi scrive esercita due funzioni diverse: *manifesta il pensiero e forma una cosa*, una *res nova*, che è la *carta scritta*. Noi chiamiamo “scrittura” il risultato di queste due funzioni diverse: così “lo scrivere” come “lo scritto”” (Carnelutti 1929: 510, nota 3)<sup>24</sup>.

Infatti, mentre un documento è sempre la rappresentazione in forma scritta di un qualsivoglia contenuto – giuridicamente e diplomatisticamente configurabile nella natura di *res signata* e di *res nova* – non tutte le scritture sono documenti in senso archivistico (Carucci 1987 e Duranti 1997)<sup>25</sup>. Ad esempio, le incisioni rupestri della Val Camonica e di Pompei sono certamente scritture, ma non possono essere definite come documenti d'archivio, così come qualsiasi lavoro al pirografo, un'insegna o un cartello autostradale. La portata del concetto di scrittura, dunque, è molto più estesa di quella di documento e, di fatto, decisamente in contrasto con la natura dei documenti archivistici. Per quanto, invece, concerne le amministrazioni pubbliche, è lo stesso legislatore che riconosce l'estrinsecazione dell'atto pubblico in documento e non già in scrittura, come puntualmente enucleato nella disposizione più diplomatistica del Codice civile (art. 2699), in cui si legano la certezza e l'affidabilità dell'agire pubblico alla necessità della forma scritta come documento e non come scrittura<sup>26</sup>.

---

<sup>23</sup> Si vedano in particolare i dettagliati interventi sulla *Nozione giuridica di documento* e su *Documenti processuali ed extraprocessuali*, nonché tutto il lungo capitolo sulla *Rappresentazione documentale*. Del resto, la questione sulle procedure di verifica delle prove era ben nota ad archivisti e diplomatisti, tra cui (Costamagna 1995). Infine, ricchissimo di spunti di diplomatica contemporanea è (Tondo, Casu, e Ruotolo 2004), cui si affianca temporalmente l'importante contributo di (La Torre 2004).

<sup>24</sup> Si veda anche (Carnelutti 1968). Sul valore tecnico della disciplina, anche Vincenzo Franco il quale, a margine, si occupa anche delle definizioni di archivio (Franco 1994, 503). In una visione interdisciplinare tra archivistica, diplomatica e diritto, si muove il lavoro, interessante ma non sempre coerente alle discipline, di (Crescenzi 2005: 15-51).

<sup>25</sup> Le due studiosi hanno impresso alla diplomatica una nuova vita, con osservazioni sempre originali o applicate in maniera innovativa e inaspettata.

<sup>26</sup> Codice civile, art. 2699: “L'atto pubblico è il documento redatto, con le richieste formalità, da un notaio o da altro pubblico ufficiale autorizzato ad attribuirgli pubblica fede nel luogo dove l'atto è formato”. Si tratta, dunque, della prova per eccellenza (pubblica fede) di un atto pubblico rappresentata in un documento amministrativo.

In definitiva, il termine *scrittura* ha necessità di specificazioni puntuali; per questo, possiamo preferire il termine *documento* nella definizione di archivio.

#### 8.4. La produzione e la selezione di documenti

Produzione, ricezione, spedizione o acquisizione di documenti servono a esplicitare o a rafforzare la definizione di archivio? Sono aspetti differenti delle tipologie di redazione o di trasmissione oppure esiste un termine che, nel linguaggio settoriale archivistico, può riunirli tutti?

Fino a tutta la prima metà del Novecento, nell'archivistica italiana non si era ancora consolidato il concetto di soggetto produttore<sup>27</sup>.

Inoltre, raramente nella definizione di archivio venivano ricordati i documenti interni: non solo quelli scambiati tra uffici del medesimo ente, ma anche quelli redatti per autodocumentazione (verbali, memorie, provvedimenti, etc.), che non possono essere ricompresi nella tipologia di quelli spediti e ricevuti.

Come sappiamo, la definizione di Cencetti influenzerà tutta la letteratura successiva per lungo tempo. Fu così per Sandri, secondo il quale “L'Archivio è “il complesso” delle scritture per cui mezzo si è esplicitata l'attività pratica di un istituto o di una persona, reciprocamente legate da un vincolo determinato dalla natura e competenza dell'ente o persona cui quelle scritture si riferiscono” (Sandri 1958b: 1002).

Fu così anche per Lodolini, secondo il quale “Archivio è il complesso dei documenti formati presso una persona fisica e giuridica (od un gruppo di uffici od organi di quest'ultima) nel corso dell'esplicazione della sua attività e pertanto legati da un vincolo necessario, i quali, una volta perduto l'interesse per lo svolgimento dell'attività medesima, sono stati selezionati per la conservazione permanente quali beni culturali” (Lodolini 1990: 124)<sup>28</sup>.

In questa definizione – e, in parte, in quelle di Cencetti e di Sandri – possiamo rilevare elementi accessori e non essenziali. Un archivio si può sedimentare in periodi determinati della propria vita anche in maniera indipendente dalla volontà del soggetto produttore e dall'esercizio di funzioni determinate. Si pensi, ad esempio, ai fermi amministrativi per gli archivi d'impresa, in cui da un lato non si ravviserebbero di certo *i propri fini* e in cui *le proprie funzioni* risultano giuridicamente inibite. Analogo discorso può riferirsi agli archivi depositati, anche in forma coattiva, ma non più ritirati e i quali, anziché restare distinti, finiscono per confluire nell'archivio proprio (si pensi a quanto accade quotidianamente nell'Ufficio dei messi notificatori di un ente locale).

---

<sup>27</sup> Sul tema del soggetto produttore si veda (Carucci 2014: 167-169).

<sup>28</sup> Tale definizione rappresenta, invero, un'evoluzione rispetto a quella di qualche decennio prima, contenente anche l'*excusatio* di essere stata strappata quasi a fatica (“Volendo dare una definizione, si potrebbe dire che”). Inoltre, risulta suddivisa addirittura in tre parti: “Il complesso delle scritture prodotte da una persona fisica o giuridica nello svolgimento della propria attività diventa *archivio* nel momento giuridico in cui le scritture stesse si configurano come *universitas rerum*”, cioè “nel momento in cui si estrinseca in atto fra i singoli documenti il vincolo necessario esistente fra di essi in potenza sin dall'origine”, cioè, ancora “nel momento in cui il complesso delle scritture assume funzione archivistica e diviene suscettibile di utilizzazione archivistica”, (Lodolini 1958: 321). Merita una rilettura anche il saggio di (Briguglio 1958: 304-305), con due definizioni di taglio più marcato di filosofia del diritto, sulla linea di pensiero di Carnelutti: “Intendiamo per *archivio come persona storica* una unità organica di volizioni e di azioni esprimenti un sistema di fini non soltanto attuali, ma anche potenziali” e, poco oltre, “Il concetto di *archivio come persona storica* si presenta dunque come la sintesi di finalità attuali e di finalità potenziali”.



Ancora: ci sono casi di corrispondenza ricevuta per errore o per delazione meschina, sia cartacea sia via e-mail, anche in modalità “nascosta” (bcc/ccn) che – inevitabilmente – persistono nell’archivio proprio in forma tradizionale o digitale senza alcuna funzione esercitata in capo al soggetto produttore.

Veniamo alla selezione. Si tratta di una delle condizioni che Lodolini pone affinché l’archivio possa essere tale *tout court* e senza aggettivazioni (corrente, di deposito o storico). L’esperienza, invece, insegna che molti archivi, anche coattivamente, sono versati senza selezione e così persistono nel tempo. Potremmo mai negare a queste *universitates* la qualificazione di archivio solo per uno scarto mancato? Per quanto importante, infatti, la selezione è un elemento fondamentale dei lavori archivistici, ma ontologicamente incidentale rispetto all’essenza dell’archivio.

In poche parole, l’attività pratica spesso non coincide con l’astrazione teorica e in molti casi – ed è questa l’avventura affascinante del lavoro in archivio – la varietà di casistica irride a una precettistica rigida, ancorché, come nel caso del grande Lodolini, basata sulla tradizione tedesca (e, in parte, anglosassone) e argomentata in maniera rigorosa.

Più di recente, ancora Romiti – a valle di un puntuale *excursus* – ha proposto un’ulteriore definizione: “L’archivio è un complesso di scritture che, legate da un vincolo naturale, sono prodotte da entità pubbliche o private nell’espletamento della loro attività, per il raggiungimento di finalità contingenti o per la conservazione della propria memoria” (Romiti 2011: 157)<sup>29</sup>.

La derivazione da Cencetti e da Sandri è dichiarata dall’autore medesimo poco più oltre, con alcune spiegazioni sulle parole *complesso* e, soprattutto su *scritture* ed *espletamento*, sulle quali è possibile essere d’accordo solo parzialmente. Esaminiamole insieme.

### 8.5. Persona, ente, entità o soggetto?

Il termine *persona* ha quasi sempre bisogno di un aggettivo (fisica o giuridica) e non si attaglia ad alcune tipologie di archivio. Nello scrivere *ente*, invece, si è troppo legati a un ambito pubblico. Inoltre, Valenti dedica all’*entità archivio* alcune pagine fondamentali, ma lo fa nell’ottica di fornire una nomenclatura adeguata a un oggetto d’indagine in divenire e non per una vocazione metafisica (Valenti 1981: 10-14).

Con l’introduzione dell’endiadi di *soggetto produttore*, l’archivistica italiana ha finalmente coniato un’espressione efficace e funzionale per tutte le tipologie d’archivio, siano esse pubbliche, private o ecclesiastiche. In questo modo, richiamando da un lato la *produzione* dei documenti, dall’altro un *soggetto*, la definizione porterebbe certamente a una chiarezza generale ed espositiva.

### 8.6. Nell’esercizio, nell’espletamento o nel corso?

I documenti sono opera dell’uomo. Pertanto, possiamo definirli – sotto un profilo strettamente giuridico – opere artificiali<sup>30</sup>.

---

<sup>29</sup> Sulle diverse tipologie di documento d’archivio, in senso lato, recentemente si è espresso in modo convincente (Brunetti 2018).

<sup>30</sup> Il primo fu (Carnelutti 1929), poi ripreso in (Carnelutti 1992).

L'espressione *nell'esercizio* sembra inadeguata, al pari di *espletamento*, perché il soggetto produttore riceve documenti anche senza esercitare alcunché e, a volte, anche senza aderire ai propri fini. Si rinvia, in particolare, alle lettere anonime di delazione personale in ambiti estranei all'attività istituzionale, ma anche a documenti depositati dalla polizia giudiziaria e mai ritirati, che finiscono per confluire – come abbiamo visto *supra* e in modo erroneo – nell'archivio proprio.

Nel fermo amministrativo, come abbiamo visto, non si esercita alcuna attività, pur continuando a ricevere (e quindi a “produrre”, in senso esteso, documenti). In questo caso, la produzione dei documenti è percepita come azione passiva (ma potenzialmente attiva) ed è pienamente inclusiva di tutte le fattispecie.

Scrivere *nel corso*, pertanto, appare più neutrale e indipendente da qualsiasi circostanza, in atto o in potenza. Per questo è da preferire.

### 8.7. Funzione, competenza o attività?

Funzione, competenza e attività non sono termini equivalenti e, soprattutto, non possiamo considerarli sinonimi. La funzione è generale e astratta, mentre la competenza è la funzione o la parte di una funzione esercitata nel tempo e nello spazio da un'unità organizzativa, da un ufficio o da una persona fisica.

Anche nel dettato costituzionale è specificato con chiarezza cristallina il concetto di competenza legato all'organizzazione degli uffici: “Nell'ordinamento degli uffici sono determinate le sfere di competenza, le attribuzioni e le responsabilità proprie dei funzionari” (Cost., art. 97). Enucleando le parole funzione e competenza, pertanto, si escludono almeno gli archivi di persona, di famiglia e di impresa.

Indipendentemente dall'essere pubblico o privato, ogni soggetto pone in essere una attività. Il termine attività, infatti, è polivalente e si attaglia tanto ai soggetti pubblici quanto ai soggetti privati. Nel caso delle amministrazioni pubbliche la legge 241/1990 aveva novellato espressamente in materia di attività amministrativa<sup>31</sup>.

Negli archivi d'impresa, inoltre, si parla di attività commerciale (e, a volte, di attività intendendo in senso lato una licenza o un esercizio commerciale) e non già di funzione commerciale. Infine, ogni soggetto privato esplica un'attività intellettuale, fisica, motoria, sociale, etc., così come ciascun partito politico svolge un'attività politica. Per queste ragioni possiamo scegliere il termine attività<sup>32</sup>.

## 9. La definizione e qualche conclusione, sempre provvisoria

A questo punto risulta possibile sussumere una definizione da quanto fin qui esaminato, generale, sintetica e neutrale rispetto alle età, agli scopi, agli utilizzi e agli effetti, enunciata *in limine*: l'archivio è il complesso dei documenti prodotti da un soggetto nel corso della propria attività.

In questa definizione abbiamo tolto alcune parole della tradizione, evitando di aggiungerne altre

---

<sup>31</sup> Legge 7 agosto 1990, n. 241, *Nuove norme in materia di procedimento amministrativo e di diritto di accesso ai documenti amministrativi*.

<sup>32</sup> Rinvio anche a International Council on Archives, *ISDF: International Standard for Describing Functions*, Desdra, 2007.

e scelto accuratamente quelle che – allo stato dell’arte della dottrina giuridica, linguistica, archivistica e diplomatistica – potrebbero servire al nostro scopo. In particolare:

- *complesso*: in quanto dotato di vincolo naturale, necessario e determinato, evitando le parole *insieme* e *raccolta* o analoghe riferibili a collezioni;
- *documenti*: perché negli archivi non si sedimentano atti, ma la loro rappresentazione sotto forma di documenti, né scritture, perché una scrittura è senz’altro un documento in senso lato, ma può non avere la caratteristica di documento archivistico;
- *prodotti*: dal momento che un soggetto redige, riceve, spedisce, acquisisce, ma anche rinviene documenti; ed è per questa ragione che si definisce soggetto produttore; pertanto, l’agire archivistico di un soggetto è definibile come produzione di documenti, nel senso più ampio del termine e senza ulteriori specificazioni sulla modalità di ingresso nella memoria archivistica (redatti, ricevuti, spediti, acquisiti, etc.);
- *soggetto*: non ente, il che limiterebbe la natura alle persone giuridiche, non entità, perché entreremmo nel campo della metafisica, ma soggetto, parola universale e neutrale applicabile a tutte le tipologie di archivio: pubblici, privati, ecclesiastici, di impresa, di partito politico, di famiglia, di persona, etc.;
- *nel corso*: rappresenta l’aspetto temporale, il fluire valutativo della sedimentazione dell’archivio, anche se non naturale (com’è, invece, il vincolo), in quanto avvenuta in maniera artificiale, cioè come opera dall’agire umano, interno o esterno al soggetto produttore;
- *propria*: e non sua, perché riferita al soggetto in maniera inequivoca;
- *attività*: perché termine universalmente applicabile, anche in questo caso, a tutte le tipologie di archivio.

In questa maniera non ci si attarda, all’interno della definizione, a illustrare il vincolo archivistico, le caratteristiche tipologiche, le nomenclature secondarie, gli scopi e gli utilizzi dell’archivio anche indiretti, rimanendo neutrali rispetto alle tre età.

Occorre tenere a mente che ogni definizione che si avvicini a una cosa in continua evoluzione, anzi poliedrica ma non polimorfa, come l’archivio, è necessariamente astratta, tanto da porre sempre dubbi nuovi, rischi e limiti. Proprio per questo, anche i teorici meno inclini alle definizioni finiscono con proporre.

In conclusione, le definizioni rappresentano un viaggio, non una destinazione. E i punti da cui si salpa e in cui si approda sono ogni volta diversi e non di rado sorprendenti. Anche se i luoghi sono gli stessi, basta cambiare prospettiva, vestirsi di duttilità mentale e cogliere qualche aspetto mai notato prima. Esiste, infine, la certezza inossidabile del metodo scientifico, secondo il quale il progresso si basa sulla distruzione delle teorie esposte fino a questo momento, in un convincimento temporaneo e contingente, soggetto ai cambiamenti dettati dalla conoscenza, che renderà caduca anche la definizione appena esposta.

## Riferimenti bibliografici

- Archivistica. Teoria, metodi, pratiche*. 2014. A cura di Linda Giuva e Maria Guercio. Roma: Carocci.
- Bachmann, Georg August, 1801. *Über Archive, deren Natur und Eigenschaften, Einrichtung und Benützung*. Amberg und Sulzbach: sine nomine.
- Baldacci, Cristina. 2016. *Archivi impossibili. Un'ossessione dell'arte contemporanea*. Monza: Johan & Levi.
- Bertini, Maria Barbara. 2008. *Che cos'è un archivio*. Roma: Carocci.
- Bologna, Marco. 2014. *La sedimentazione storica della documentazione archivistica*, in *Archivistica. Teoria, metodi, pratiche*. Roma: Carocci: 211-235.
- Bonfiglio-Dosio, Giorgetta. 2010. *Primi passi nel mondo degli archivi*. Padova: Cleup.
- Briguglio, Letterio. 1958. "Sul concetto di archivio". *Rassegna degli Archivi di Stato*, XVIII (3): 287-307.
- Brunetti, Dimitri. 2018. "L'archivio multitematico: definizione, descrizione e identità". *La Gazette des archives*, 249 (1): 39-49.
- Carnelutti, Francesco. 1929. "Studi sulla sottoscrizione". *Rivista del diritto commerciale*, (9-10): 511-543.
- Carnelutti Francesco. 1968. "Documento (teoria moderna)". In *Novissimo Digesto Italiano*, VI, Torino, UTET, s.v.
- Carnelutti Francesco. 1992. *La prova civile: parte generale. Il concetto giuridico della prova*. Milano: Giuffrè.
- Carucci, Paola. 1978. "Versamenti e ordinamento degli archivi degli organi centrali dello Stato". *Rassegna degli Archivi di Stato*, XXXVIII/1-2-3: 52-76, ripubblicato in *Antologia di scritti archivistici*. 1985. A cura di Romualdo Giuffrida. Roma, Ministero per i beni culturali e ambientali: 409-436.
- Carucci, Paola. 1987. *Il documento contemporaneo. Diplomatica e criteri di edizione*. Firenze: NIS.
- Carucci, Paola. 2014. *L'ordinamento*, in *Archivistica. Teoria, metodi, pratiche*. Roma: Carocci: 137-177.
- Cassese, Leopoldo. 1959. *Introduzione allo studio dell'archivistica, Prelezione al corso di archivistica speciale presso l'Università di Roma nell'anno accademico 1958-1959*, Salerno, Reggiani, ora in Cassese, Leopoldo. 1980. *Teorica e metodologia. Scritti editi e inediti di paleografia, diplomatica, archivistica e biblioteconomia*. A cura di Attilio Mauro Caproni. Salerno: Pietro Laveglia editore.
- Cencetti, Giorgio. 1937. "Sull'archivio come università rerum". *Archivi* (IV): 7-13, ripubblicato in Cencetti, Giorgio. 1970. *Scritti archivistici*. Roma: Il Centro di ricerca editore: 47-55.
- Costamagna, Giorgio. 1995. *Il notaio a Genova tra prestigio e potere*. Milano: Giuffrè.
- Crescenzi, Victor. 2005. *La rappresentazione dell'evento giuridico. Origini e struttura della funzione documentaria*. Roma: Carocci.

*Dictionary of Archives Terminology*. 2020. <https://www2.archivists.org/news/2020/new-dictionary-of-archives-terminology-launched>

Duranti, Luciana. 1997. *I documenti archivistici, La gestione dell'archivio da parte dell'ente produttore*. Roma: Ministero per i beni culturali e ambientali

Duranti, Luciana. 2014. *Il documento archivistico*. In *Archivistica. Teoria, metodi, pratiche*. Roma: Carocci: 19-33.

Franco, Vincenzo. 1994. *Teoria archivistica del documento*. "Rivista trimestrale di diritto pubblico", (2): 445-505.

Franzese, Paolo. 2018. *Manuale di archivistica italiana*. Perugia: Morlacchi editore.

Giuva, Linda, Stefano Vitali, e Isabella Zanni Rosiello. 2007. *Il potere degli archivi: usi del passato e difesa dei diritti nella società contemporanea*. Roma: Bruno Mondadori.

Graffi, Giorgio. 2008. *Che cos'è la grammatica generativa*. Roma: Carocci.

Grossi, Monica. 2014. *L'archivio in formazione*. In *Archivistica. Teoria, metodi, pratiche*. Roma: Carocci: 35-52.

Humboldt von, Wilhem, 2004. *La diversità delle lingue*. Roma-Bari: Laterza.

International Council on Archive. 1992. *Statement of Principles Regarding Archival Description. First Version Revised*, pubblicato in Italia in "Archivi e Computer", II (2): 98-105.

International Council on Archive. 1994. *ISAD(G): General International Standard for Archival Description, Adopted by ad hoc Commission on Descriptive Standards*, Stockholm, Sweden, 21-23 January 1993, Ottawa, Secrétariat de la Commission ad hoc sur les normes de description.

International Council on Archive. 2007. *ISDF: International Standard for Describing Functions*, disponibile anche nella traduzione italiana curata da Salvatore Vassallo, Pavia, 2009.

Lanza, Emanuela, e Golinelli, Paolo. 2006. *Elementi essenziali di archivistica teorica e pratica*. Bologna: Pàtron.

La Torre, Maria Elena. 2004. *Contributo alla teoria giuridica del documento*. Milano: Giuffrè.

Lodolini, Elio. 1958. "Identificazione dell'archivio". *Rassegna degli Archivi di Stato*, XVIII (3): 308-323

Lodolini, Elio. 1965. "Inventari di tre archivi comunali (Civitanova Marche, Corridonia, Morrovalle)". *Rassegna degli Archivi di Stato*, XXV (1): 107-122.

Lodolini, Elio. 1970. "Questioni di base dell'archivistica". *Rassegna degli Archivi di Stato*, XXX (2): 325-364.

Lodolini, Elio. 1980. "'Archivio', un concetto controverso nella dottrina e nelle leggi". *Rassegna degli Archivi di Stato*, XL (1): 9-25

Lodolini, Elio. 1990. *Archivistica. Principi e problemi*. Milano: Franco Angeli

Lodolini, Elio. 2015. "Postille di Eugenio Casanova al manuale degli archivisti olandesi". *Archivi*, X (1), 153-160.

- Lombardo, Antonino. 1955. "Il problema dello scarto degli atti di archivio". *Rassegna degli Archivi di Stato*, XV (3): 300-316.
- Lombardo, Antonino. 1958. *Archivi di Stato*. In *Enciclopedia forense*. I, Milano, Giuffrè, s.v.
- Londei, Luigi. 2003. *Elementi di archivistica*. Roma: Jouvence.
- Mazzoleni, Jole. 1972. *Manuale di archivistica*. Napoli: Libreria scientifica editrice.
- Mottola, Francesco. 1998. "Un "Manuale" di archivistica, paleografia e diplomatica di un notaio abruzzese di fine Ottocento". *Rassegna degli Archivi di Stato*, LVIII (2-3): 372-393
- Müller, Samuel, Johann Feith, e Robert Fruin. 1908. *Ordinamento e inventario degli archivi*. A cura di Giuseppe Bonelli e Giovanni Vittani. Torino: UTET.
- Nicolini, Fausto. 1971. *Un'epistola di Gabriele D'Annunzio intorno a un manuale di archivistica*, in Nicolini, Fausto, *Scritti di archivistica e di ricerca storica raccolti da Benedetto Nicolini*. Roma: Ministero dell'interno.
- Palumbo, Ciro. 2016. *Stato, diritto e linguaggio giuridico. Studi con Von Humboldt*. Roma: Nuova cultura.
- Panella, Antonio. 1955. *Scritti archivistici*. Roma: Ministero dell'Interno.
- Paoloni, Giovanni. 2014. *Il documento e le sue istituzioni. Archivi, biblioteche, musei*. In *Archivistica. Teoria, metodi, pratiche*, Roma: Carocci: 429-452.
- Penzo Doria, Gianni. 2009. "L'archivio come "bene della vita"", *Scrinia. Rivista di archivistica, paleografia, diplomatica e scienze storiche*, V (1-3): 21-37.
- Phillips, Robert. 2004. *American Governmental Archives*. In Fisher, Steven. *Archival Information. How to find it, how to use it*. Westport: Greenwood Press.
- Romiti, Antonio. 2011. *Archivistica generale. Primi elementi*. Torre del Lago Puccini: Civita editoriale
- Sandri, Leopoldo. 1958a. "La storia degli archivi". *Rassegna degli Archivi di Stato*, XVIII (1): 109-134.
- Sandri, Leopoldo. 1958b. "Archivi di Stato". In *Enciclopedia del diritto*. Milano: Giuffrè, II, s.v.
- Sebastiani, Ezio. 1904. *Genesis, concetto e natura giuridica degli Archivi di Stato in Italia*. Torino: Fratelli Bocca editori.
- Tondo, Salvatore, Casu, Giovanni, e Ruotolo, Antonio. 2004. *Il documento*. Napoli: Edizioni scientifiche italiane.
- Trivelli, Pasquale. 1898. *Disciplina degli archivi. Diplomi e carte antiche con una epistola di Gabriele D'Annunzio e con VIII tavole, di cui l'ultima segnata da Francesco Paolo Michetti*. Lanciano: Rocco Carabba.
- Valacchi, Federico. 2015. *Diventare archivisti*. Milano: Editrice bibliografica.
- Valacchi, Federico. 2017. *Archivio. Concetti e parole*. Milano: Editrice bibliografica.

Valenti, Filippo. 1975. “Parliamo ancora di archivistica”. *Rassegna degli Archivi di Stato*, XXXV (1-2-3): 161-197.

Valenti, Filippo. 1981. “Riflessioni sulla natura e struttura degli archivi”. *Rassegna degli Archivi di Stato*, XLI (1-2-3): 9-37.

Valenti, Filippo. 2000. *Nozioni di base per un'archivistica come euristica delle fonti documentarie*, in *Scritti e lezioni di archivistica, diplomatica e storia istituzionale*. A cura di Daniela Grana, Roma, Ministero per i beni e le attività culturali – Ufficio centrale per i beni archivistici: 150-169.

Vitali, Stefano. 2001. “Di angeli, di paperi e di conigli, ovvero dello strano mestiere dell'archivista”. *Archivi per la storia*, XIV (1-2): 179-186.

Zacchè, Francesco. 2012. *La prova documentale*. Milano: Giuffrè.

## Artists' archives

Dimitri Brunetti<sup>(a)</sup>

a) University of Udine, <https://orcid.org/0000-0003-2336-8118>

**Contact:** Dimitri Brunetti, [dimitri.brunetti@uniud.it](mailto:dimitri.brunetti@uniud.it)

**Received:** 11 February 2022; **Accepted:** 15 March 2022; **First Published:** 15 May 2022

### ABSTRACT

In the last few years, artists' archives have gained a central role in the debate on the nature, treatment and valorization of 20th century and contemporary archives. They are fascinating and precious aggregations, characterized by a plurality of documentary typologies and by specific requirements for their description. Starting from the definition of personal archives, the paper proposes a definition of artists' archives and then attempts to indicate their main characteristics, contents and elements of complexity. Finally, the paper attempts to establish a dialogue between archivists and artists, collectors, gallery owners, critics and curators in order to find common elements and differences in their actions.

### KEYWORDS

Artist's archive; Multi-typological archive; Personal archive; Description; Complexity.

## L'archivio d'artista

### ABSTRACT

Gli archivi d'artista hanno assunto negli ultimi anni un ruolo centrale nel dibattito sulla natura, il trattamento e la comunicazione degli archivi del Novecento e della contemporaneità. Si tratta di aggregazioni affascinanti e preziose, caratterizzate da una pluralità di tipologie documentali e da esigenze specifiche per la descrizione. Partendo dalla definizione di archivio di persona, il contributo ne propone una d'archivio d'artista e tenta poi di indicarne le caratteristiche principali, i contenuti e gli elementi di complessità. Infine, l'articolo prova a instaurare un dialogo fra archivisti e artisti, collezionisti, galleristi, critici e curatori per trovare nella loro azione elementi comuni e di differenza.

### PAROLE CHIAVE

Archivio d'artista; Archivio multitematico; Archivio di persona; Descrizione; Complessità.



## Introduzione

Oggi è sempre più frequente l'abitudine di costruire mostre che contengano porzioni più o meno ampie di documentazione del lavoro degli artisti, sia perché l'archivio in arte è diventato una presenza ingombrante che non può essere evitata, sia perché strumento per tentare di comprendere e riscrivere il passato (Scotini e Giacomelli, 2021). Così, nel mondo dell'arte l'archivio è ormai un elemento abituale della riflessione artistica e una testimonianza stabile del quotidiano e della creatività.

Da tempo i fondi di persona sono al centro di un diffuso interesse che ha allargato il tradizionale orizzonte degli operatori culturali favorendo lo sviluppo di una riflessione originale per definirne natura e caratteristiche, così da descriverli, organizzarli e comunicarli nel miglior modo possibile. Tale interesse è stato sollecitato tanto dalla stessa presenza negli archivi, nelle biblioteche e nei musei di fondi documentari dall'indiscutibile fascino, quanto da una storiografia che vede in questi materiali il grimaldello per guardare oltre la superficie e comprendere il Novecento e la contemporaneità attraverso gli occhi dei protagonisti. Questa attenzione agli archivi di uomini e donne si è subito rivolta ad una moltitudine di personaggi della cultura, delle scienze e della creatività e ha trovato terreno fertile nel mondo dell'arte. Da una parte gli archivisti, i bibliotecari, i museologi e gli operatori museali stanno sviluppando un efficace dibattito sul trattamento delle storie di persona, dall'altra artisti, critici, curatori ecc. hanno identificato l'archivio anche come oggetto di riflessione artistica, così che talvolta il confine fra l'opera e la documentazione diventa poroso (Scotini e Giacomelli, 2021: 50).

Nell'arte contemporanea, in particolare, l'archivio è stato adottato non solo come oggetto d'indagine, ma anche come medium per realizzare nuove visioni e ricostruzioni del mondo (Baldacci 2016: 9). Si assiste ad un lavoro in favore degli archivi in costante incremento che i luoghi di produzione e di conservazione, pubblici e privati, alimentano realizzando progetti di schedatura e di digitalizzazione, mostre ed esposizioni, l'allestimento di luoghi di ricerca e favorendo il dibattito. Ad esempio, la Biennale di Venezia è dotata di un formidabile Archivio Storico delle Arti Contemporanee e sul finire del 2021 ha annunciato di voler ampliare le proprie attività con il lancio del Centro Internazionale della Ricerca sulle Arti Contemporanee, ma sono ormai tantissime le iniziative di artisti, di musei e di galleristi, critici e collezionisti che in Italia e nel mondo tengono i riflettori accesi sulla memoria documentale, analogica e digitale, di un'epoca complessa come quella in cui siamo immersi e danno vita a forme di associazionismo prima impensabili.

A fronte di un dibattito d'arte che parla di anarchivi, antiarchivi e contrarchivi, di archivi disobbedienti e di archivi inarchiviabili, gli archivisti sentono la necessità di confrontarsi con una realtà fluida e multiforme così da allargare lo sguardo della disciplina, ma anche di ricondurre il dibattito archivistico entro logiche imprescindibili. Insieme, poi, artisti, archivisti, musei e istituti di conservazione lavorano da tempo per la salvaguardia di aggregati documentari fragili che purtroppo nella maggior parte dei casi vanno perduti o dispersi, o di cui si conservano solo gli elementi utili ad autenticare le opere.

## Gli archivi di persona e gli archivi d'artista

L'attenzione verso gli archivi privati è piuttosto recente e l'interesse per quelli di persona, che per lungo tempo sono stati trascurati, va collocato non prima degli anni Novanta. Da non molto, poi, lo sguardo si è rivolto anche agli archivi d'artista.

Nei primi anni dello Stato unitario l'interesse è tutto rivolto agli archivi pubblici e anche nei decenni successivi solo alcuni richiamano l'attenzione agli archivi privati. Fra questi nel 1897 Clemente Lupi che propone di imporre ai proprietari di "archivi privati" l'obbligo di ordinarli e di permetterne la consultazione (Lodolini 1997: 48) e nel 1937 Giorgio Cencetti che invita lo Stato a farsi carico anche degli archivi privati riconoscendone l'interesse sociale (Cencetti 1937). La nozione di archivio privato stenta ad essere accolta nella dottrina e i fondi documentari di famiglie, persone fisiche e persino di enti privati per lungo tempo non vengono qualificati come veri e propri archivi (Lodolini 1997: 34-35). Lo sguardo si allarga nei decenni successivi, ma solo sul finire degli anni Ottanta il Ministero per i beni culturali e ambientali avvia i lavori che porteranno nel 1991 alla pubblicazione del primo volume della Guida degli archivi di famiglie e di persone e, nel mese di settembre dello stesso anno si svolge a Capri il convegno internazionale di studi intitolato "Il futuro della memoria" che rappresenta il primo importante appuntamento dedicato agli archivi di persona e di famiglia e apre un nuovo settore di interesse negli studi archivistici, sebbene nessuno dei contributi presentati si orienta al mondo dell'arte (*Il futuro della memoria* 1997). Finalmente l'attenzione di molti si rivolge agli archivi di famiglia e di persona, determinando poco a poco una separazione degli studi fra gli archivi nobiliari (Casella e Navarrini 2000) e domestici da una parte e quelli propriamente di persona dall'altra, fra quelli antichi e quelli del Novecento e della contemporaneità (Barrera 2006; Bonfiglio-Dosio 2011: 364-366).

Oggi le comunità professionali degli archivi, delle biblioteche e dei musei stanno sviluppando un efficace dibattito sul trattamento delle storie di persona e gli archivi di persona si sono completamente affrancati dall'appellativo a lungo assegnatogli di "archivi diversi" assumendo un ruolo di primo piano nella ricerca storiografica anche grazie al fatto che sovente offrono una visione inedita e particolare per la comprensione e la ricostruzione della contemporaneità. Si sono moltiplicati gli interventi di salvaguardia, valorizzazione e condivisione, riservando attenzione non solo agli archivi di famiglia e nobiliari, dove talvolta si trovavano incastonati fondi di personaggi illustri, e a soggetti molto noti di ambito istituzionale, ma anche a quelli di un gran numero di personalità così da favorire la conservazione delle testimonianze più significative di una società profondamente mutata rispetto al passato (Santoro 2001: 58; Del Vivo 2012: 15-16).

I fondi personali sono complessi organici di materiali editi o inediti prodotti o raccolti da persone significative del mondo della cultura, delle professioni e delle arti, prevalentemente dalla seconda metà del XIX secolo in poi. Fra questi si possono identificare gli archivi di persona, gli archivi culturali e le biblioteche d'autore, ma anche collezioni di varia natura presenti nei fondi. L'elemento aggregatore rimane l'individuo e dunque il *corpus* preso in esame deve essere documento e testimone degli interessi, delle attività e delle relazioni della persona nel contesto storico e culturale in cui ha operato (*Linee guida* 2019).

## Gli archivi d'artista

Nell'ambito degli archivi di persona, gli archivi d'artista meritano un'attenzione speciale in quanto ne formano un segmento di peculiare interesse sia per la loro particolare natura e composizione, sia perché sempre più oggetto di cure da parte di storici, addetti ai lavori e appassionati. Senza addentrarci nelle finalità delle ricerche va comunque rilevata la straordinaria attrattiva di questi fondi che offrono contestualmente documenti personali, di lavoro e di contesto arricchiti da materiali non riconducibili alla sola dimensione documentaria tradizionale (Damiani 2020: 224-225). Si potrebbe inizialmente identificare l'archivio d'artista come un complesso di documenti di natura eterogenea prodotti da una persona significativa del mondo dell'arte nel corso della sua attività per finalità personali, professionali o artistiche. Si tratta, però, di una definizione generica e inevitabilmente imprecisa perché non considera ancora alcune variabili e si renderebbe necessario introdurre almeno degli elementi di precisazione riferiti alla tipologia d'arte, al segmento temporale e al soggetto titolare dell'archivio. Nel comporre una definizione di archivio d'artista vanno infatti considerate le molteplici qualità del mondo dell'arte e occorre essere consapevoli della sua vastità. Ciò che comunemente consideriamo "arte" abbraccia arti visive e performative di natura differente, ovvero la pittura, la scultura, l'architettura, la letteratura e la poesia, la musica, la danza, il teatro, il cinema e la fotografia. A questo primo elenco alcuni aggiungono il fumetto e i videogiochi, e ancora la radio e la televisione, la pubblicità, la realtà virtuale e il design. Approfondendo l'analisi si devono considerare anche le cosiddette arti applicate, fra cui, ad esempio, tutto quello che riguarda il design, il tessile, l'oreficeria e l'illustrazione. In tutti questi casi, poi, ci sono da tenere in conto i contesti ambientali, sociali ed economici, oltre al periodo con cui ci si confronta. Riguardo alla collocazione temporale, se è pur vero che rivolgendosi alle espressioni artistiche delle epoche più lontane poco o nulla è rimasto dell'artista e del processo creativo, rimangono comunque eccezionali testimonianze delle opere di grandi artisti, di pittori e scultori, di architetti, così come di importanti manifatture di ceramiche, arazzi e altre produzioni. Avvicinandosi poi alla contemporaneità i fondi documentari si moltiplicano e si intensificano nella quantità, nella varietà e nella complessità che esprimono. Rispetto, invece, alle figure coinvolte in questi archivi vanno evidenziate le differenze fra l'artista, il committente, il gallerista, il commerciante, il collezionista, il critico e il curatore. Sono appunto diversi gli interessi e gli obiettivi e si modificano i rapporti fra esigenza creativa, gestione economica, passione e coinvolgimento.

Considerando così tante variabili pare difficile immaginare una definizione generalista di archivio d'artista, ma astraendone i concetti e trovando elementi comuni si potrebbero forse introdurre nell'enunciato iniziale alcuni elementi che accennano alla varietà e alla complessità di cui si è detto. Riprendendo quindi la precedente definizione (un complesso di documenti di natura eterogenea prodotti da una persona significativa del mondo dell'arte nel corso della sua attività per finalità personali, professionali o artistiche) la si può articolare ulteriormente identificando l'archivio d'arte come un complesso di documenti omogenei o di natura differente, prodotti da un artista o da un soggetto significativo del mondo delle arti nel corso della sua attività per finalità personali, professionali o artistiche, conservate in quanto testimonianza della propria attività e della propria esistenza. Se il termine "d'artista" richiama un'immagine limitata all'opera di un singolo, la scelta di adottare la voce "archivio d'arte" si propone di essere inclusiva dell'opera e del lavoro di molti, di abbracciare vasti periodi e di ampliare l'orizzonte verso cui il nostro sguardo può dirigersi.

## Ma si tratta di archivio?

Nel sottolineare l'eterogeneità dei materiali presenti in un archivio d'arte insieme alla molteplicità dei soggetti e dei percorsi che portano alla formazione e strutturazione di questi materiali, si rende opportuna una riflessione circa l'attribuzione della qualifica di "archivio" ad ogni aggregato documentale. Talvolta, infatti, risulta più opportuno identificare i materiali di un archivio d'arte con una collezione o un centro di documentazione.

Nella maggior parte dei casi l'archivio di un artista è un vero archivio, anche se occorre osservarlo con la mente aperta, con lo sguardo libero da preconcetti dogmatici e nella consapevolezza sia delle particolarità degli archivi contemporanei, caratterizzati dalla pluralità e dall'eterogeneità dei documenti, sia del fatto che l'archivistica è una disciplina viva e dinamica, in perenne evoluzione, capace di confrontarsi con la realtà mutevole dei contesti e di adattarsi ad essi pur senza rinnegare i propri principi fondanti. Infatti, anche negli archivi d'artista si identificano senz'altro la spontaneità del momento della produzione e acquisizione dei documenti come risultato dello svolgimento delle funzioni caratteristiche dell'azione della persona, la naturalezza delle modalità di aggregazione, il rapporto diretto e inscindibile con il soggetto produttore e la sua attività, il vincolo chiaramente riconoscibile fra le parti, la finalità prettamente pratica e, infine, l'insita capacità di testimoniare i fatti: tutte condizioni queste che richiamano la natura più propria di un archivio. Sono archivi che si formano e crescono spontaneamente intorno all'artista, secondo il naturale scorrere del tempo e delle sue faccende. Si formano in modo naturale, senza che la persona ci abbia pensato troppo e di sicuro senza che abbia preso la decisione consapevole di creare un proprio archivio, un archivio fatto da elementi tutti collegati fra loro in maniera logica e necessaria, così come gli innumerevoli contesti personali, lavorativi, sociali ecc. che fra loro si completano e dialogano vicendevolmente restituendo una storia e un'informazione strutturata.

Non si può negare che osservando questi fondi il concetto di archivio si dilata e che le singole testimonianze che li compongono potrebbero talvolta essere viste come pezzi a sé stanti, consapevolmente acquisiti, benché elementi di un'attività complessa che si svolge su vari piani.

In relazione alle modalità di produzione, alla natura dei documenti e agli obiettivi di utilizzo in alcuni casi si potrebbe qualificare l'archivio d'artista con attributi particolari quali *speciale* o *arricchito*. Quasi sempre, infatti, un archivio di persona è un archivio "speciale" – anche in assonanza al concetto biblioteconomico che caratterizza alcune raccolte come riferite ad ambiti e soggetti specifici – considerando che si tratta di fondi fortemente caratterizzati, per qualità e quantità, dalla tipologia dei materiali, dai temi trattati o dalla collocazione temporale. Talvolta, invece, un archivio d'artista è un archivio "arricchito" poiché al nucleo naturalmente formatosi vengono aggiunti materiali affini e pertinenti che accrescono il complesso originale determinando quasi la formazione di un centro di documentazione. Si tratta di una particolare forma d'archivio abbastanza frequente, anche se non sempre identificata, dove, accanto ad un archivio normalmente formatosi, il soggetto che ne determina le modalità di crescita o conservazione raccoglie documentazione e materiali eterogenei da aggiungere a quelli già posseduti o custoditi, così da definirne meglio i temi o arricchirne il contenuto. È il caso, ad esempio, di quando per sviluppare un'idea o acquisire competenze si aggiungono a nuclei esistenti altre risorse che ne potenziano le qualità in funzione della realizzazione di attività proprie dell'artista (Brunetti 2018: 43).

Qui si è sul crinale, dove archivio, centro di documentazione e collezione si sfiorano, ma decidiamo

di proseguire ancora un poco a considerarlo un archivio e così l'archivio arricchito può identificarsi con un complesso di fondi, o *multifondo*, ovvero all'accostamento di archivi l'un l'altro autonomi benché affini nei temi e materiali. Fra questi fondi alcuni potrebbero anche essere archivi *riprodotti*, formati dalla copia integrale o parziale di altri archivi, altri invece veri e propri archivi a sè stanti di altri soggetti produttori (parenti, colleghi, altri artisti o soggetti del mondo dell'arte, mercanti e galleristi) acquisiti in vari modi e conservati e utilizzati. Nel prosieguo della riflessione si potrebbero anche intravedere gli archivi *inventati* o *neoarchivi*, cioè archivi che assemblano su base tematica materiali di diversa natura e provenienza estraendoli dai propri contesti di origine per ricrearne dei nuovi (Valacchi 2002; Vitali 2004: 116-119; Vitali 2006). Seguendo il filo dei pensieri non ci si può non riferire al concetto di archivio *improprio*, dove il vincolo naturale è assente o non sembra lo si possa ricondurre a tutte le parti dell'archivio. Però, pur constatando che con queste premesse sovente i fondi personali potrebbero essere archivi "impropri", e considerando non solo il vincolo, ma anche le modalità di formazione dell'archivio stesso, si giunge ad un reale ampliamento del concetto d'archivio (Navarrini 2005: 28). Su questa linea si potrebbero individuare esempi, coerenti con l'estrema vitalità e variabilità di forma degli archivi d'arte, così da formare un catalogo complessivo che inizialmente includerebbe, ad esempio, l'Archivio Zeri a Bologna, l'Archivio Paolini a Torino, l'Archivio Pallucchini a Udine, l'Archivio Crispolti a Roma, la Casa museo di Pellizza da Volpedo. Ecco, quindi, che l'archivio d'artista può essere inteso come la somma di due elementi: un archivio propriamente detto (di lavoro e familiare) affiancato da un aggregato documentario che raccoglie in sè sia fondi archivistici non direttamente collegabili al produttore principale, sia altri materiali identificabili come raccolte o collezioni. Una sorta di centro di documentazione personale in cui archivi, collezioni e materiale sciolto vengono trattati in modo unitario e hanno caratteristiche che li accomunano testimoniando un particolare ambito d'interesse.

Infine, richiamando ancora una volta la molteplicità di soggetti che, insieme all'artista, possono essere produttori o conservatori di documenti (committente, gallerista, commerciante, collezionista, critico e curatore), si può applicare anche nei loro confronti le medesime riflessioni fin qui condotte pur prevedendo un aumento delle collezioni formate sulla base di una specifica attenzione verso un artista o una forma d'arte e quindi con elementi senza vincoli originali fra loro.

In sintesi, riprendendo le definizioni d'archivio d'artista e d'archivio d'arte già sviluppate, e allargandone l'ambito un'ultima volta così da accogliere le diverse declinazioni di complesso documentale, si potrebbe definire l'archivio d'artista come "un'aggregazione di matrice archivistica (Damiani e Guercio 2019: 430) di documenti di natura eterogenea prodotti da un artista nel corso della sua attività per finalità personali, professionali o artistiche", e l'archivio d'arte come "un'aggregazione di matrice archivistica di documenti di natura eterogenea prodotti da un artista o da un soggetto significativo del mondo delle arti nel corso della sua attività per finalità personali, professionali o artistiche conservate in quanto testimonianza della propria memoria e della propria esistenza".

## La complessità degli archivi d'artista

L'archivio d'artista o d'arte rappresenta una realtà documentale complessa, caratterizzata da multi-formi componenti, variegati contesti, molteplici funzioni e interessi. Per affrontare un'aggregazione archivistica di questo genere non è sufficiente definirne la natura teorica, ma è anche necessario

identificare gli elementi che lo compongono. Così si è tentato di schematizzarne i nuclei documentari principali, tenendo presente dei diversi attori che agiscono nel mondo dell'arte. Un ipotetico archivio d'artista può contenere documenti di lavoro e corrispondenza, documentazione collegata alle opere e un archivio del prodotto, materiali di documentazione, un fondo di natura personale.

## **Le componenti di un archivio d'artista**

L'archivio di lavoro di un artista raccoglie la documentazione riguardante i progetti realizzati, le idee in cantiere, i rapporti di lavoro, la gestione ordinaria. Si tratta di materiale eterogeneo nella forma e omogeneo nella funzione che si riferisce all'attività del produttore e comprende progetti in bozza o strutturati, proposte, contratti, contabilità, documenti per l'acquisto dei materiali di consumo e corrispondenza. Una serie di fascicoli riferiti a un'idea, una lavorazione, un prodotto, un'opera che quindi possono includere disegni preparatori, schizzi e bozzetti accanto a prove di catalogo, layout di impaginazione e di pubblicità o di allestimento, corrispondenza, rassegne stampa. Tutto quello che in un modo o nell'altro è stato prodotto tanto in occasione del processo creativo, quanto nella gestione del rapporto con la committenza o l'espositore o per la messa sul mercato. Accanto ai fascicoli di progetto ce ne possono essere anche altri con materiali non direttamente riferibili all'idea in fase di sviluppo o realizzata, oppure al manufatto, ma collegati a ulteriori elementi che concorrono ugualmente al progetto creativo o di costruzione come, ad esempio, attinenti un argomento, un autore, un personaggio, luoghi reali o di fantasia, un materiale, un cliente, un editore, un gallerista, un critico. In questa porzione d'archivio, che costituisce la sezione più consistente e densa dell'intero complesso, si trova la maggior parte dei documenti che hanno seguito la realizzazione di un'opera.

Ancora riferibile alla documentazione di lavoro, l'archivio probabilmente contiene un nucleo di corrispondenza, ulteriore rispetto a quello riferito alla gestione formale degli incarichi e dei progetti di lavoro, che testimonia il dialogo con i colleghi e gli amici che si interessano della medesima attività, con i maestri e gli allievi, con un vasto insieme di persone che condividono gli interessi e l'attività dell'artista. Se presenti, questi scritti aiutano a ricostruire un dialogo corale che favorisce la comprensione del contesto entro cui si inserisce la produzione, la rete di relazioni, la genesi delle idee e delle opere.

L'archivio delle opere rappresenta senza alcun dubbio la porzione più interessante di un complesso archivistico riferito al lavoro di un artista. Qui si trovano i materiali che sono il risultato dei progetti creativi. Si tratta dell'insieme delle opere in forma non definitiva prodotte da un soggetto nello svolgimento della sua attività intellettuale e creativa. Sono tavole, disegni, modelli, prototipi, girati video, prove d'artista, istruzioni e note per la realizzazione di un prodotto, ossia tutto quello che conclude il processo di elaborazione, ma che non è ancora l'opera realizzata. Grazie a questi materiali di lavoro, rari a trovarsi, si possono ripercorrere le fasi di progettazione e realizzazione di un bene. Accanto a questo primo nucleo ci potrà essere una raccolta di prodotti finali, di opere artistiche di molteplice natura nella forma definitiva in cui sono state proposte al pubblico. Si tratta di opere conservate dall'artista o non commercializzate, di esemplari di opere multiple, di campioni, di cataloghi, immagini che testimoniano una realizzazione, un insieme necessariamente incompleto e disomogeneo, ma di straordinario valore che in altri contesti documentari viene

identificato come “archivio del prodotto”, ovvero come l’insieme dei beni prodotti e messi in commercio o comunque distribuiti (Bilotto 2002).

Chiunque nel corso del tempo raccoglie materiali disparati, ricordi di viaggio e di esperienze fatte, regali e gadget, piccole collezioni, oggetti e fotografie, stampe di lettere ed e-mail, libri, riviste, locandine e manifesti e chissà cos’altro che serve a ricordare ciò che si è fatto, a circondarsi di cose che piacciono e, consapevolmente o meno, come elemento di ispirazione e modello nel momento dell’ideazione di una nuova opera o di un allestimento. In questo grande aggregato documentario si trovano anche opere o prodotti vari di altre persone vicine al soggetto produttore principale che lo stesso ha ricevuto, scambiato, raccolto e collezionato anche come elemento di conoscenza del settore entro cui lavora e crea. Infine, è solitamente presente una biblioteca specialistica, pur se di dimensioni che variano da caso a caso, che contiene monografie, opuscoli e periodici (Petrucciani, 2020). Per questo terzo gruppo di materiali alcune delle categorie proprie dell’archivio vengono a mancare e, pur in presenza di elementi e qualità che legano le varie parti della raccolta, ci si trova di fronte a documenti consapevolmente ricercati, acquisiti e conservati e quindi, in questo caso, più che di archivio ci si può riferire ad una sorta di centro di documentazione. Però, pur ammettendo tale declinazione disciplinare, ci si confronta con un insieme documentario fortemente integrato all’archivio propriamente detto, a materiali che svolgono una funzione primaria per lo svolgimento dell’attività propria del soggetto produttore e a documenti che se presi singolarmente sono certamente corpi estranei, ma che nel loro insieme sono indispensabili a definire la funzione e la fisionomia dell’archivio.

Infine, si troverà un archivio personale e di famiglia, di solito presente in ogni casa, con lettere, diari, documenti personali e privati riguardanti la gestione della vita quotidiana, le proprietà, il percorso scolastico, le amicizie, i parenti, i ricordi familiari (Zagra 2013).

Nel caso ci si riferisca non all’artista, ma ad un altro soggetto comunque coinvolto della produzione, identificazione, esposizione o commercializzazione di un’opera, si avranno, a seconda dei casi, nuclei documentari più corposi di corrispondenza e contabilità, di cataloghi e materiale di documentazione, di fotografie e stampe. L’ordine o il disordine, così come le presenze e le mancanze oppure le quantità, di cui si dirà più avanti, saranno determinati dal livello di attenzione dei soggetti produttori e aggregatori verso i materiali, in quanto utili per il lavoro o la memoria, e determinanti per definire l’efficacia del complesso documentario nell’essere strumento d’uso corrente o custode dei percorsi creativi e delle opere.

## **Elementi di complessità**

Questa prima elencazione di aggregati documentali riferiti all’archivio d’artista, così come la riflessione sulla sua stessa natura, ne tratteggiano un profilo polimorfo, articolato, fluido e certamente complesso. In effetti, sembra che la complessità sia la cifra distintiva di questi sistemi archivistici, che peraltro contrassegna una gran parte degli archivi della contemporaneità soprattutto di ambito personale e non istituzionale. Tale elemento caratterizzante si rende poi più evidente quando il soggetto produttore principale svolge un’attività creativa, non esattamente codificata ma con forti elementi di variabilità legati agli aspetti artistici, dell’inventiva personale, del genio. Se poi si considerano anche le forme documentarie, la pluralità di soggetti e attori coinvolti e la trasformazione

digitale, allora si può certamente affermare che gli archivi personali del mondo dell'arte sono una delle forme archivistiche più complesse e, nel contempo, più interessanti e affascinanti.

Questi archivi sono il risultato di una sedimentazione quotidiana che risponde a criteri del tutto individuali, lasciati alla volontà o non volontà della persona (Navarrini 2005: 53-55). Si tratta di una documentazione poco strutturata, raccolta e accumulata con criteri individuali riferiti agli interessi, alle attività e alle relazioni personali che inevitabilmente definiscono il modo di aggregarsi dei materiali, determinando anche quella distinzione del vincolo archivistico, proprio e improprio, di cui si è accennato. Sulla forma della raccolta influiscono anche, e molto, le decisioni consapevoli o inconsapevoli dei soggetti conservatori che sovente rimodellano questi archivi ponendo maggiore (o esclusiva) attenzione alle opere e all'archivio del prodotto piuttosto che all'intero *corpus* documentario e alla qualità di testimonianza.

Un altro elemento di complessità è dato dalla quantità dei materiali accumulati. In questo caso non è, naturalmente, una questione di metri lineari o di numerosità di fascicoli e documenti, ma di quantità e qualità dei materiali in relazione all'attività dell'artista. Si tratta di valutare la dicotomia fra pieni e vuoti, fra presenze e assenze e di valutare questi elementi attribuendo loro un significato nel tentativo di conoscere quanto più possibile la persona, l'artista, i contesti. In alcuni casi questi archivi sono colmi di una massa davvero considerevole di materiali accumulati nel corso degli anni, conseguenza di quel "mal d'archivio" (Derrida 1996) che spinge alcuni a raccogliere il maggior numero possibile di documenti, testimonianze e oggetti. Altre volte, al contrario, c'è reale scarsità perché prevale la volontà di distruggere quanto più possibile nel convincimento che non si tratti di cose importanti, che le opere "parlino" loro stesse e nulla d'altro sia necessario, o che sia meglio non conservare nulla o quasi nulla della propria vita. Quindi la tenuta dei documenti dipende molto dalla persona che ha prodotto l'archivio, dagli eredi (Cardinale 2019) o da chi ce l'ha in custodia e in altri casi si afferma il desiderio, se non la determinazione, di proporre una visione soggettiva di sé e del proprio operato, di trasmettere una memoria precostituita (Allegrezza e Gorgolini 2016: 15) modellando assenze e presenze, aggregazioni e riferimenti. In questo caso i documenti possono essere tanti o pochi, ma certamente non saranno sinceri. Bisogna mettere in conto che siano avvenute operazioni più o meno intenzionali di autorappresentazione (Giorgi 2019), di selezione, montaggio, attribuzione di senso, ispirate a esigenze funzionali, a preoccupazioni ideologiche o ad altri fattori, facendo perdere all'archivio una certa sua spontaneità originaria (*Archivi familiari e personali* 2012). Infine, può invece anche darsi il caso che la perdita sia stata causata da situazioni accidentali, da scarti involontari oppure anche ad opera di persone terze per scelta o per incuria.

Gli archivi d'artista, così come molti altri del Novecento e della contemporaneità, sono caratterizzati dalla presenza contestuale di documenti di forma differente. Archivi *multitipologici* dove accanto a documenti scritti di natura tradizionale, si trovano fotografie, disegni, giornali, oggetti di piccole e grandi dimensioni, opere di vario genere, strumenti di lavoro e altri materiali eterogenei formati e sedimentati nel corso della normale attività da un soggetto identificato, che testimoniano e definiscono un artista, un percorso artistico, un'opera, una vicenda creativa e umana. Tessere colorate e disuguali di un mosaico complesso che se pazientemente ricomposte danno forma a immagini meravigliose. Rappresentazioni multiformi di un "secolo breve" e di una contemporaneità che non differenzia più fra le composte raffigurazione di sé e della memoria, e che utilizza indifferentemente lo strumento più appropriato per memorizzare i propri documenti. Un archivio "allargato" o multitipologico è un archivio tradizionale composto da documenti di varia natura e



forma che ci sfida sul piano della descrizione della sua fluidità per comunicarne la bellezza (Brunetti 2016; 2018).

Descrivere in modo appropriato gli archivi e le collezioni d'artista rappresenta uno degli elementi di maggiore interesse e difficoltà nel lavoro degli operatori, naturalmente insieme agli aspetti legati alla conservazione e al loro utilizzo nella ricerca. Descrivere la complessità multitipologica e ibrida di questi archivi, soffermandosi sia sugli elementi che sui contesti, rappresentandone le particolarità e le relazioni che intercorrono fra le unità e fra queste e le entità di produzione, gestione e conservazione coinvolte costituisce una vera sfida e richiede in uguale misura professionalità e sensibilità (Vitali 2014; Valacchi 2016). La precisa comprensione dell'archivio è un fattore determinante per la protezione e la sicurezza dell'integrità dei materiali, ma la corretta rappresentazione dei beni rende soprattutto possibile l'accesso, la condivisione, l'utilizzo, l'approfondimento culturale e la comunicazione (Del Vivo 2012; Allegrezza 2019; Ciaccheri, Cimoli e Moolhuijsen 2020). In questo contesto, pur senza entrare nel merito delle operazioni di schedatura e di riordino, così come in quelle di produzione degli strumenti di ricerca, occorre certamente porre attenzione ad alcuni elementi per il trattamento di questi fondi di natura plurale. Richiamata la necessità di conoscere a fondo le vicende personali e artistiche del soggetto produttore, così come il contesto entro cui svolge la sua attività e circolano le sue opere, la principale difficoltà è quella di identificare il modo migliore con cui descrivere in maniera dinamica gli elementi dell'archivio. Piuttosto che applicare le stesse categorie descrittive a testimonianze differenti fra loro, è opportuno fare ricorso a tracciati scheda appropriati per ciascuna tipologia documentale, attingendo anche a modelli catalografici di tradizione non archivistica così da costruire un insieme di idonee rappresentazioni (Corbo 2017). Naturalmente una tale operazione va realizzata con molta attenzione perché gli standard d'ambito non sempre condividono i presupposti disciplinari e l'accezione terminologica, o garantiscono relazioni orizzontali e gerarchiche stabili e autoesplicative. La sfida quindi è quella di elaborare descrizioni analitiche e pertinenti, di sviluppare un'architettura archivistica coerente ad albero o a rete (Brunetti 2020) che renda la complessità del contesto e del contenuto nel modo più semplice possibile. In tale processo oltre agli standard, alle regole e alle raccomandazioni nazionali e internazionali si potrà fare ricorso ad elementi qualificati di raggruppamento ulteriori oltre a quelli tradizionali del complesso, del produttore e del conservatore, quali l'opera, la persona, il luogo, il soggetto giuridico, l'evento, l'anno o le cose notevoli, nonché alle strategie semantiche, alle ontologie e ai RIC (Di Marcantonio e Valacchi 2018), che permettono adeguati approfondimenti multidimensionali e la registrazione di relazioni non uniche.

Nel mondo dell'arte, poi, a volte accade che si intervenga sull'archivio quando l'artista è ancora disponibile. In quei casi fortunati, sovente generati dall'intelligenza di personaggi lungimiranti, le operazioni di descrizione vanno condotte facendo il più possibile ricorso ai ricordi, al vissuto e all'esperienza del produttore, alla sua capacità di tessere relazioni fra i documenti e di arricchire i contesti.

## **Gli archivisti e il mondo dell'arte**

Il tema che si è cercato di affrontare è evidentemente molto ampio. Per ora si è scelto di accennare alla maggior parte delle questioni disciplinari e terminologiche pur approfondendone solo alcune, mentre altri temi, sebbene di grande importanza, sono stati temporaneamente accantonati e tro-

veranno spazio in ulteriori lavori: fra questi le prassi conservative adottate dagli artisti e l'impatto che la trasformazione digitale ha avuto nella produzione, circolazione e fruizione dell'arte e, di conseguenza, sugli archivi degli artisti contemporanei.

Parlare di archivi d'artista vuol dire tentare di confrontarsi con molte variabili dove la complessità è la cifra distintiva. Una complessità che richiede competenze, sensibilità e il dialogo fra tutti gli attori coinvolti nel processo di formazione, tenuta, utilizzo e conservazione. Infatti, un conto è parlare di un artista autonomo che nel suo studio "produce" il proprio lavoro, un conto riferirsi a un professionista che entra nei processi industriali e che punta a elaborare un prodotto riproducibile. Un conto è affrontare la documentazione formata da un artista e un altro è confrontarsi con un museo, una galleria, un collezionista, un curatore o un critico. Un conto è guardare alle raccolte antiche o lontane nel tempo, altra cosa è accostarsi ad aggregati documentari dal secondo dopoguerra ad oggi. Un conto è relazionarsi ad un artista vivente, un altro trattare con gli eredi o con chi si occupa della conservazione.

Gli archivisti considerano l'archivio d'artista come un sistema di relazioni e non come un semplice contenitore di informazioni, però la loro percezione dell'archivio non sempre sembra coincidere con quella dei soggetti del mondo dell'arte. Così «L'archivio, come tale, è una collezione di materiali eterogenei» (Scotini e Giacomelli, 2021: 52), ma anche «L'archivio è il luogo della collezione e della conservazione di ciò che è passato e che non deve andare perduto» (AitArt, n.d.), e poi «L'archivio consiste quindi in un database di opere, in cui sono presenti immagini, descrizioni tecniche e informazioni su esposizioni e bibliografia» (Art and the Cities 2017). Quindi nel contesto artistico sembra prevalere l'idea di raccolta o elenco di materiali disparati da organizzare e conservare così che non vadano perdute le tracce della vita e della personalità di un artista per testimoniare l'opera e la figura (AitArt, n.d.).

In effetti, nella generalità dei casi l'archivio d'artista viene identificato come la raccolta dei documenti legati alle opere, se non proprio come l'insieme delle opere stesse (l'archivio del prodotto di cui si è detto) finalizzato alla tutela del diritto d'autore e alla certificazione, ovvero ad accertarne l'autenticità (operazione che può essere realizzata dal solo autore) o all'attribuzione della qualifica di "archiviato" (se, per tutelare dalla falsificazione, il soggetto che detiene i documenti ha già l'opera in esame fra quelle da lui riconosciute o accetta di ammetterla nel catalogo). Infatti, «L'archivio d'artista è un prezioso alleato per la tutela e la promozione del patrimonio artistico ed economico della vita di un artista, a favore del futuro della sua produzione, in aiuto dei galleristi e a supporto dei collezionisti. (...) Gli archivi diventano per galleristi e collezionisti garanti dell'autenticità delle opere dell'artista e punto di riferimento nel mercato dell'arte in qualità di certificatori dell'originalità dei lavori in circolazione». (Art Rights, 2019).

Per svolgere al meglio la loro funzione di ente certificatore, oltre che di custode di un percorso artistico, gli archivi d'artista vivente, piuttosto che i musei o i soggetti privati di conservazione e promozione conducono un'intensa attività di catalogazione di documenti e opere, realizzando anche campagne di digitalizzazione, come quella imponente del Guggenheim Museum di New York, che valorizzano le opere, ma non sempre presidiano relazioni e contesti.

L'azione di questi soggetti del mondo dell'arte, pur talvolta divergendo dalla condotta che l'archivista vorrebbe fosse tenuta, assolve a funzioni essenziali di memoria d'identità e di buona conservazione. E proprio in merito alla salvaguardia degli archivi d'artista va anche rilevato che fino a non molto tempo fa gli artisti non avevano, in genere, alcuna percezione dell'utilità di conservare i

documenti della loro opera e che oggi le cose sono un po' cambiate solo in funzione della necessità che molti hanno di disporre di un proprio portfolio. Inoltre manca per lo più negli artisti la percezione della fragilità del digitale che invece si sostituisce con una grande fiducia nella tecnologia. In conclusione, il dibattito sulla natura, le modalità di trattamento e l'uso degli archivi d'artista rimane aperto e oggi rappresenta uno dei percorsi di riflessione più interessanti in seno alla comunità professionale degli archivisti, ma anche all'interno del mondo dell'arte dove soggetti diversi, musei e associazioni hanno intensificato le iniziative in favore di una rinnovata attenzione verso i materiali e di una nuova coscienza archivistica. Ne sono la prova, ad esempio, i corsi di formazione di archivisti d'arte contemporanea del Castello di Rivoli Museo d'Arte Contemporanea o dell'Associazione Italiana Archivi d'Artista, così come i frequenti incontri che coinvolgono professionalità differenti in funzione di un dialogo comune.

## Riferimenti bibliografici

AitArt (Associazione Italiana Archivi d'Artista). nd. *Home page*. Ultimo accesso il 29 gen. 2022. [www.aitart.it](http://www.aitart.it).

Allegrezza, Stefano e Luca Gorgolini, a c. di. 2016. *Gli archivi di persona nell'era digitale. Il caso dell'archivio di Massimo Vannucci*. Bologna: Il Mulino.

Allegrezza, Stefano. 2019. "Gli archivi di persona tra consultabilità, privacy e diritto all'oblio". In *Rodolfo Pallucchini: storie, archivi, prospettive critiche*, a cura di Claudio Lorenzini, 417-427. Udine: Forum.

*Archivi familiari e personali Qualche consiglio per difenderli meglio*. 2012. Torino: Soprintendenza archivistica per il Piemonte e la Valle d'Aosta. [www.sato-archivi.it/Sito/index.php/strumenti/archivi-familiari.html](http://www.sato-archivi.it/Sito/index.php/strumenti/archivi-familiari.html).

Art and the Cities. 2017. <https://artandthecities.com/2017/07/06/cosa-un-archivio-dartista/>.

Art Rights. 2019. *L'archivio d'artista*. Ultimo accesso il 29 gen. 2022. [www.arrights.me/l-archivio-d-artista/](http://www.arrights.me/l-archivio-d-artista/).

Baldacci, Cristina. 2016. *Archivi impossibili. Un'ossessione dell'arte contemporanea*. Cremona: Johan & Levi.

Barrera, Giulia. 2006. "Gli archivi di persone", in *Storia d'Italia nel secolo ventesimo. Strumenti e fonti*, a cura di Claudio Pavone, III: 617-657. Roma, Direzione generale per gli archivi.

Bilotto, Antonella. 2002. "L'archeologia del documento d'impresa. L'«archivio del prodotto»", *Rassegna degli Archivi di Stato*, LXII (1-3): 293-306.

Bonfiglio-Dosio, Giorgetta. 2011. "Gli archivi privati". In *Archivistica speciale*, a cura di Id., 363-366. Padova: Cleup.

Brunetti, Dimitri. 2016. "La lente archivistica, per rendere convergenti percorsi catalografici paralleli. Appunti sulla multidisciplinarietà della descrizione", *Archivi*, XI (1): 101-114.

Brunetti, Dimitri. 2018. "L'archivio multitematico: definizione, descrizione e identità", *La Gazette des archives*, (249/1): 39-49.

Brunetti, Dimitri. 2020. "L'albero e l'edera. La descrizione da ISAD a RIC-CM". In *Controluce. Spigolature d'archivio*, a cura di Laura Giambastiani e Annantonia Martorano, 129-142. Torre del Lago Puccini: Civita Editoriale.

Cardinale, Eleonora. 2019. "Il ruolo degli eredi nella trasmissione dell'archivio", *JLIS.it*, vol. 10 (3): 71-82.

Casella, Laura e Roberto Navarrini, a c. di. 2000. *Archivi nobiliari e domestici. Conservazione, metodologie di riordino e prospettive di ricerca storica*. Udine: Forum.

Cencetti, Giorgio. 1937. "Sull'archivio come Universitas rerum", *Archivi*, IV: 7-13.

Ciaccheri, Maria Chiara, Anna Chiara Cimoli e Nicole Moolhuijsen, a c. di. 2020. *Senza titolo. Le metafore della didascalìa*. Busto Arsizio: Nomos edizioni.

Corbo, Alessandra. 2017. “Dalla somiglianza alla sinergia. Il patrimonio culturale tra descrizione e rappresentazione digitale”, *Il mondo degli archivi*, 7 dic. [www.ilmondodegliarchivi.org/rubriche/in-italia/570-dalla-somiglianza-alla-sinergia-il-patrimonio-culturale-tra-descrizione-e-rappresentazione-digitale](http://www.ilmondodegliarchivi.org/rubriche/in-italia/570-dalla-somiglianza-alla-sinergia-il-patrimonio-culturale-tra-descrizione-e-rappresentazione-digitale).

Damiani, Concetta e Maria Guercio. 2019. “Il ruolo degli archivi d’arte nel mondo contemporaneo”. In *Rodolfo Pallucchini*, a cura di Claudio Lorenzini, 429-441. Udine: Forum.

Damiani, Concetta. 2020. “Gli archivi di persona e gli archivi d’artista”. In Calabrese, Gaetano. *Gli archivi storici non statali: ordinamento e conservazione*, 221-227. Viagrande: Algra editore.

Del Vivo, Caterina. 2012. “Accostarsi ad un archivio di persona: ordinamento e condizionamento”. In *Archivi di persona del Novecento. Guida alla sopravvivenza di autori, documenti e addetti ai lavori*, a cura di Francesca Gheretti e Loretta Paro, 15-38. Treviso: Fondazione Benetton studi e ricerche e Fondazione Giuseppe Mazzotti per civiltà veneta con Antiga edizioni.

Derrida, Jacques. 1996. *Mal d’archivio*. Napoli: Filema.

Di Marcantonio, Giorgia e Federico Valacchi, a c. di. 2018. *Descrivere gli archivi al tempo di RICCM*. Macerata: Eum.

Giorgi, Andrea. 2019. “Se peindre pour ne pas se perdre. Considerazioni marginali in merito all’archivio come autorappresentazione”, *JLIS.it*, vol. 10 (3): 59-70.

*Il futuro della memoria. Atti del convegno internazionale di studi sugli archivi di famiglie e di persone. Capri, 9-13 settembre 1991*. 1997. Roma: Ufficio centrale per i beni archivistici.

*Linee guida sul trattamento dei fondi personali*. 2019 (vers. 15.1, 31 mar.). Commissione nazionale biblioteche speciali, archivi e biblioteche d’autore dell’AIB. [www.aib.it/struttura/commissioni-e-gruppi/gbaut/strumenti-di-lavoro/linee-guida-sul-trattamento-dei-fondi-personali/](http://www.aib.it/struttura/commissioni-e-gruppi/gbaut/strumenti-di-lavoro/linee-guida-sul-trattamento-dei-fondi-personali/).

Lodolini, Elio. 1997. “Archivi privati, archivi personali, archivi familiari, ieri e oggi”. In *Il futuro della memoria*, 23-69. Roma: Ufficio centrale per i beni archivistici.

Navarrini, Antonio. 2005. *Gli archivi privati*. Torre del Lago Puccini: Civita Editoriale.

Petruciani, Alberto. 2020. “Fondi e collezioni personali: alcune questioni”. In *Storie d’autore, storie di persone. Fondi speciali tra conservazione e valorizzazione*, a cura di Francesca Gheretti, Annantonia Martorano e Elisabetta Zonca, 31-36. Roma: AIB.

Santoro, Michele. 2001. “Archivi privati: esperienze a confronto”. *Biblioteche oggi*, 19 (8 ott.): 56-66.

Scotini, Marco e Marco Enrico Giacomelli. 2021. “Tutti pazzi per gli archivi. Ennesima moda o c’è di più?”. *Artribune*, XI (63 nov.-dic.): 48-55.

Valacchi, Federico. 2002. “Il web per gli archivi e gli archivi nel web”, *Archivi e Computer*, XII (3): 7-16.

Valacchi, Federico. 2016. “Pezzi di cose nel mondo. Il processo di integrazione delle descrizioni archivistiche nei sistemi interculturali”, *JLIS.it*, vol. 7 (2): 331-367.

Vitali, Stefano. 2004. *Passato digitale*. Milano: Bruno Mondadori.

Vitali, Stefano. 2006. “Come si ‘diventa digitali’ negli archivi”, *Bibliotime*, IX (1). [www.aib.it/aib/sezioni/emr/bibttime/num-ix-1/vitali.htm](http://www.aib.it/aib/sezioni/emr/bibttime/num-ix-1/vitali.htm).

Vitali, Stefano. 2014. “La descrizione degli archivi nell’epoca degli standard e dei sistemi informatici”. In *Archivistica. Teorie, metodi, pratiche*, a cura di Linda Giuva e Maria Guercio, 179-210. Roma: Carocci.

Zagra, Giuliana, a c. di. 2013. *Conservare il Novecento: lettere, diari, memorie. Convegno Ferrara, salone internazionale dell’arte del restauro e della conservazione dei beni culturali e ambientali 30 marzo 2012. Atti*. Roma: AIB.

## Certification and preservation of artworks in digital environments

Concetta Damiani<sup>(a)</sup>

a) University of Campania “Luigi Vanvitelli”, <https://orcid.org/0000-0002-3301-0075>

**Contact:** Concetta Damiani, [concetta.damiani@unicampania.it](mailto:concetta.damiani@unicampania.it)

**Received:** 4 February 2022; **Accepted:** 15 March 2022; **First Published:** 15 May 2022

### ABSTRACT

Reflecting about the documentary heritage of artists requires the consideration of archival, legal, historical-artistic and museological factors. If the analogue productions of artifacts and the correlated documentation have been studied, transferring the creative perspective into a digital environment asks to face a reality in turmoil: the ways of expressing and representing contemporary art, in balance between digital and crypto art, require new visions in the production and management of works. In this redefinition of roles and relationships, which almost exclusively concern the documentary structure and the technological and digital components, document systems assume a primary role.

### KEYWORDS

Archives; Artist; Blockchain; Certification; Preservation; NFT

## La certificazione e la conservazione delle opere d'arte in ambiente digitale

### ABSTRACT

Una riflessione sui patrimoni documentali degli artisti impone la considerazione di fattori archivistici, giuridici, storico-artistici e museali. Se le produzioni analogiche di manufatti e il relativo corredo documentale sono stati oggetto di studio e trattamento, spostando la prospettiva creativa in ambiente digitale troviamo una realtà in fermento: le modalità di espressione e rappresentazione dell'arte contemporanea, in bilico tra arte digitale e crypto arte, richiedono nuove visioni nelle attività di produzione e gestione delle opere. In questa ridefinizione di ruoli e rapporti, che attengono quasi esclusivamente alla struttura documentaria e alle componenti tecnologiche e digitali, i sistemi documentali assumono un ruolo primario.

### PAROLE CHIAVE

Archivi; Artista; Blockchain; Certificazione; Conservazione; NFT

## Gli archivi d'artista

Gli archivi d'artista rispecchiano realtà riconducibili a diversi ambiti espressivi e rimandano quindi, oltre che a soggetti attivi nel campo delle arti figurative, della musica, della poesia e della letteratura, anche a personalità dello spettacolo: dal teatro al cinema e alla televisione<sup>1</sup>.

Il trattamento di tali eterogenei patrimoni documentali, che spesso includono elaborati, manufatti e prototipi della produzione artistica, impone la considerazione di fattori archivistici, giuridici, storico-artistici e museali. L'organizzazione e la gestione degli archivi, la cura per gli aspetti relativi alla creazione, alla circolazione e alla conservazione delle opere, la tutela del diritto d'autore e le attività di certificazione, le scelte di promozione, le relazioni documentali tra artisti, eredi, curatori, collezionisti e studiosi richiedono infatti confronto e raccordo.

Se tutto ciò vale in ambito analogico, la riflessione si amplifica in ambito digitale: gli artisti che esprimono la propria creatività servendosi delle tecnologie informatiche includono i profili più svariati; pertanto le varietà tipologiche si moltiplicano e con esse le modalità produttive. I principi di fruizione, certificazione e conservazione di tali opere di arte contemporanea devono perciò essere ripensati in base a nuovi criteri, che attengono quasi esclusivamente alla struttura documentaria e alle componenti tecnologiche e digitali.

L'archivio rappresenta una risorsa per l'artista, sia perché ne documenta il percorso consentendo anche rivisitazioni e recuperi di temi e tecniche, sia perché si pone come elemento indispensabile per la tutela della produzione artistica<sup>2</sup>. Proprio in relazione a questo secondo aspetto, i documenti che riguardano l'auto-rappresentazione e la produzione/riproduzione delle opere hanno acquistato una forte valenza e sono diventati indispensabili alla produzione del cosiddetto *archivio delle opere*. Si tratta di una sorta di catalogo ragionato dei manufatti artistici, corredato da apparati cronologici, da note storiche e da riferimenti alla provenienza, alla bibliografia e alla storia espositiva, strumento indispensabile per la tutela dell'autorialità delle opere e per il rilascio dei certificati di autenticità.

Il catalogo può essere realizzato in forma di volume a stampa o di sito web e rappresenta la fonte ufficiale per i dati anagrafici, la rassegna stampa e la repertoriatura delle opere. Se i cataloghi a stampa vantano una consolidata tradizione, anche la produzione di cataloghi in rete è in crescita: tra i primi ad apparire si segnalano quelli di Daniel Buren (Buren 2012) e Gerhard Richter (Richter 2020), concepiti e realizzati per la fruizione on-line in libero accesso.

Tra le istituzioni italiane la Fondazione Giulio e Anna Paolini, costituita per la salvaguardia e la diffusione della conoscenza delle attività di Giulio, ha in corso la catalogazione ragionata delle opere, realizzata con impostazione inequivocabilmente archivistica (Fondazione Paolini 2021).

Al di là delle iniziative individuali, l'attenzione per la memoria documentaria delle opere d'arte e la condivisione di buone pratiche gestionali trova un aggregatore nell'AitArt, Associazione italiana archivi d'artista. L'organismo è nato nel 2014 a Milano, con l'intento di sensibilizzare curatori e

---

<sup>1</sup> Il recente convegno organizzato a Firenze dall'Associazione nazionale archivistica italiana "Le Muse in archivio: itinerari nelle carte d'arte e d'artista" ha avuto, tra gli altri, il merito di offrire una panoramica su tali diverse tipologie d'archivio, evidenziandone caratteri, specificità e affinità di trattamento (ANAI, 2021).

<sup>2</sup> Tra i numerosi archivi d'artista oggetto di intervento di riordinamento e inventariazione si segnala l'archivio Galileo Chini (Martorano 2014; 2018, 23-42), per il quale è stata praticata anche un'attività di sperimentazione con l'applicazione delle linee guida RDA alla rappresentazione della descrizione del fondo (Bettazzi 2021).



artisti alle tematiche di tutela e trasmissione della produzione e della memoria degli artisti (Damiani e Guercio 2019, 434).

L'effettiva funzione di certificazione e di attribuzione certa di un'opera trova le sue ragioni e la sua attendibilità principalmente nel fondo documentario propriamente detto: i curatori d'archivio svolgono un ruolo sempre più marcato di *gatekeepers*, tra esigenze di mercato e prassi cautelative (Donati 2016; Visconti 2020). Il quadro normativo italiano non specifica però quali siano i soggetti cui spetta il diritto di autenticare un'opera d'arte e di rilasciare le relative certificazioni<sup>3</sup>.

Si colloca in questo contesto l'adozione da parte del Ministero della cultura di PACTA - Protocolli per l'autenticità, la cura e la tutela dell'arte contemporanea (MIBACT 2017), un modello di certificato realizzato dai componenti di un tavolo tecnico di lavoro per l'acquisizione di opere d'arte contemporanea. Il modello è stato definito per garantire la corretta conservazione e permanenza nel tempo dell'opera d'arte e, anche a garanzia di supporto all'art. 64 del Codice dei beni culturali, viene proposto quale documento integrante del contratto di acquisto dell'opera. I PACTA sono stati creati per i musei statali e per musei pubblici con statuti diversi, ma se ne auspica un'adozione diffusa, anche da parte di altri attori (Zorloni 2018).

Intanto si va facendo strada, come vedremo, un sistema di certificazione digitale del tutto nuovo, legato alla registrazione su blockchain.

## La produzione non tradizionale: dall'arte digitale alla crypto arte

Nell'arte digitale rientrano tutte le opere che comportano l'impiego della tecnologia e dei sistemi informatici come parte essenziale del processo creativo e della successiva presentazione e fruizione.

Siamo di fronte a una forma di arte contemporanea caratterizzata da un processo di elaborazione o da un'interfaccia digitale che darà come risultato un'opera che potrà essere prodotta in serie (Concas 2020, 138) e che ha trovato una delle prime esplicitazioni nell'operato dell'accademico tedesco Georg Nees negli anni '60 del Novecento.

In senso generale si è poi assistito a un moltiplicarsi di categorie che hanno investito i diversi campi delle arti con una differente combinazione di elementi tecnologici

“Media Art, Digital Art, Interactive Art, Electronic Art, Software Art, Computer Art, Intermedia Art, Time-Based Media Art, Bio Art, Robotic Art, ma anche per il versante cinematografico, Database Cinema, Digital Cinema, Cinema quantico, Software Cinema, Web Cinema e così via, sono tutti termini, spesso complementari, spesso equivalenti, che nascondono un tentativo di coprire un universo sempre più complesso e di difficile definizione” (Catrìcalà 2016, 28).

---

<sup>3</sup> Il Codice dei beni culturali stabilisce espressamente l'obbligo da parte del venditore di consegnare all'acquirente la documentazione di autenticità o almeno di attribuzione e provenienza (D.Lgs. n. 42/2004, art. 64), rimandando poi alla legge sul diritto d'autore (Legge 22 aprile 1941 n. 633 e successive modifiche) che garantisce agli artisti il godimento sia dei diritti morali sia dei diritti patrimoniali connessi all'opera realizzata. In base alla legge l'artista, pur cedendo l'opera, continua a vantare sulla stessa i diritti morali e, salvo patto contrario, anche i relativi diritti economici connessi all'uso commerciale (dal diritto di riproduzione al diritto di seguito). Il diritto ad autenticare l'opera, riconosciuto al solo autore, se e fino a quando è in vita, viene poi fatto proprio da eredi, fondazioni e archivi, periti.

Nei fatti le opere d'arte digitali sollecitano una riflessione sulla necessità di accorte strategie di gestione: tali produzioni, infatti, senza un'accurata programmazione d'interventi rischiano l'impossibilità di fruizione e conservazione in tempi straordinariamente brevi.

Dal punto di vista istituzionale il panorama internazionale presenta iniziative di eccellenza quali quella dello ZKM (*Zentrum für Kunst und Medien*), fondato nel 1989 e interamente dedicato all'arte multimediale e interattiva o quella di Rhizome, che dal 2006 ha avviato un programma di conservazione per il proprio archivio di opere digitali (Cattricà e Quaranta 2020, 9-10).

L'Italia sconta invece un forte ritardo programmatico e organizzativo, anche se è notizia recente quella dell'istituzione a Milano di un museo dell'arte digitale, Mad, per il quale è stata individuata sede nei suggestivi locali del dismesso Albergo diurno di Porta Venezia, prevista l'apertura per il 2026 e scelta la direttrice, nella figura della storica dell'arte e curatrice Ilaria Ester Bonacossa. Taglio, struttura e contenuti restano da definire, nell'ambito di "produzione e presentazione di contenuti digitali" (MIC 2021).

Dal punto di vista di artisti e curatori, la migrazione su supporti digitali sta interessando buona parte dei lavori time-based, collezionati su supporti in misura sempre maggiore obsoleti; si ravvisa tuttavia scarsa coscienza delle problematiche di conservazione, mentre la natura digitale delle collezioni richiederebbe interventi e soluzioni mirate.

La portata e la natura della collezione, con la specifica di formati e durata, le prospettive di crescita, la valutazione dei casi in cui l'opera d'arte è accompagnata da uno specifico hardware, che sua volta richiede un trattamento dedicato, rappresentano gli elementi da cui partire per operare scelte virtuose.

Si registra inoltre un aumentato interesse alla "documentazione dell'opera d'arte", nella doppia accezione dell'azione del documentare l'opera e del tenere traccia delle risultanze del documentare:

"un legittimo processo di conservazione dovrebbe essere disciplinato nel rispetto dell'intenzionalità dell'artista, coinvolto in prima persona nella documentazione dell'opera, nella descrizione delle informazioni relative ai materiali, alle tecniche e all'individuazione di quegli aspetti essenziali e coerenti all'originario significato estetico e storico dell'opera, senza i quali la ripresentazione futura rischia di essere un atto discrezionale del curatore, o del conservatore. La documentazione dell'opera d'arte si pone quindi come momento fondante del processo conservativo, anche in considerazione di una condivisione delle informazioni raccolte" (Barreca 2020, 62-63).

Si rafforza quindi il valore documentario del percorso artistico, ma non si accenna minimamente alla gestione documentale complessiva; bisognerebbe invece interrogarsi su procedure e modalità di conservazione integrata dell'opera e della documentazione che ne accompagna creazione e sviluppi.

Va segnalato che nel momento in cui la mancanza di fissità, tipica della Media Art, rilancia i temi di autenticità e originalità di un'opera contemporanea, gli artisti ricorrono alla soluzione della certificazione di autenticità che rappresenta "il bene scambiabile del mio lavoro" (Lozano-Hemmer 2015), quasi più dell'opera stessa. E come è evidente le attività di certificazione hanno necessità di un fondo archivistico di riferimento, stabile e stabilizzato.

Nel caso della Net art, ma l'impostazione è replicabile, la documentazione è vista sempre più come lo strumento per una strategia di conservazione adeguata. Le installazioni digitali sono accompa-

gnate da descrizioni testuali (su contenuti, storia e tecnica), istruzioni di re-installazione, interviste agli artisti, fotografie e documentazione video. Si prova inoltre a documentare il rapporto con il pubblico, provvedendo a raccogliere e conservare opinioni e impressioni, nonché testimonianze dell'interazione tra l'opera e il pubblico, se prevista.

Per dirla con Gaby Wijers “Quando la documentazione risulta fondamentale e solo le opere con la miglior documentazione hanno possibilità di sopravvivere, si può parlare di «archiviazione darwiniana»” (Wijers 2020, 125).

Naturalmente, in prospettiva archivistica, bisogna impegnarsi perché tutte le opere abbiano la documentazione idonea a qualificarle e certificarle correttamente e a contribuire a una corretta gestione documentale con il beneficio, tra gli altri, di una sempre maggiore trasparenza del mercato dell'arte. Michele Foti nel 2018 ha presentato una rassegna delle strategie di conservazione delle opere d'arte digitale, accompagnata da un'ampia casistica. Le buone pratiche devono tener presente che l'arte digitale, in qualsiasi delle declinazioni a cui si presta, è legata a componenti software e hardware; è penalizzata da fragilità e obsolescenza; è un oggetto digitale complesso; è multipla, variabile e in continuo divenire; tende a dissolversi in un network di persone, programmi e dispositivi distribuiti potenzialmente ovunque, grazie all'utilizzazione delle reti. Per una corretta conservazione vanno pertanto identificate le proprietà significative dell'opera d'arte; soltanto in un secondo momento ha senso individuare la strategia, o la combinazione di strategie, più consona a preservare l'oggetto digitale (Foti 2018, 176).

Una riflessione meriterebbe anche la presenza sempre più decisa dell'Intelligenza Artificiale nel mondo dell'arte, sia per quanto riguarda la sua applicazione nella gestione delle strutture museali, sia per quanto concerne la sperimentazione nella produzione artistica: la creazione di opere d'arte create usando algoritmi di IA aprono alla valutazione di una nuova dimensione estetica e alla considerazione, per la macchina, di un ruolo sempre più netto in termini di partneraggio nella creazione dell'opera (Pedrazzi 2021).

## “Crypto art is now”

Le opere di crypto arte sono frutto di una precisa intenzione: l'artista decide, sin dall'inizio del processo creativo, di realizzare un'opera destinata ad esistere e a essere promossa, esposta e tracciata sulla rete blockchain.

Il processo si sviluppa con l'uso di una particolare tecnologia, rappresentata dagli NFT (non fungible token). Gli NFT sono token crittografici, creati su blockchain e caratterizzati dalla proprietà di contenere informazioni uniche non alterabili e non intercambiabili. Tali token - dotati di infungibilità, prerogativa che gli permette di possedere una specifica individualità e di essere pertanto insostituibili - consentono di identificare e certificare in modo univoco un oggetto digitale, ma anche un diritto inerente ad un bene fisico di cui garantiscono la titolarità e il diritto di proprietà. Un'immagine, un disegno, un video, un'animazione, una foto, un brano musicale possono perciò essere rappresentati da un token “unico e insostituibile”.

Banco di prova degli NFT, sicuramente non esclusivo ma ampiamente sperimentato, è rappresentato dal circuito dell'arte nei diversi termini di produzione di opere d'arte in digitale e di certificazione delle medesime e/o di opere su supporto tradizionale.

Per le opere d'arte in digitale è sempre più diffuso l'uso dei token ERC-721<sup>4</sup>, sviluppati nel 2018 da William Entriken, Dieter Shirley, Jacob Evans e Natassia Sachs. Tali elementi hanno caratteristiche identitarie specifiche che li rendono nettamente distinguibili uno dall'altro: ciascun token, ha un suo DNA (una stringa alfanumerica 256 bit che lo identifica) ed una serie di attributi. I token ERC-721, dunque, sono utilissimi per identificare oggetti unici, non escludendo la possibilità che facciano parte di insiemi omogenei o ampie classi.

Una delle caratteristiche dei token ERC-721 è la possibilità di contenere metadati; questo oltre a consentire di individuare il token o gli oggetti fisici o digitali che rappresenta, introduce la possibilità di identificare un'immagine che rappresenta l'oggetto stesso cui il token è legato. Nella sostanza, in una scala valori di affidabilità ERC 721 si qualifica molto meglio di altri token ma prevede costi più elevati: maggior numero di dati significa anche maggior costo di elaborazione, maggior sicurezza e maggiore tracciabilità (Annunziata e Conso 2021, 29-31).

Tra le criticità appurate, non vanno trascurati il dispendio di energia e i furti d'arte.

Nel primo caso recenti ricerche hanno dimostrato che il ricorso alla blockchain e agli NFT richiede un consumo di energia molto elevato. Lo scambio di un file di crypto art può richiedere la stessa quantità di elettricità necessaria per alimentare lo studio di un artista per due anni; la creazione di un'opera sfrutterebbe inoltre una quantità di energia pari al fabbisogno elettrico dell'intera Argentina.

Il processo di generazione di blockchain e di NFT si fonda su complessi calcoli che richiedono ai computer prestazioni molto elevate: decine di processori lavorano contemporaneamente per generare i files digitali servendosi di strumenti di raffreddamento molto efficienti e dispendiosi in termini di energia, indispensabili per evitare il surriscaldamento delle macchine. Sembra pertanto urgente e indispensabile sperimentare soluzioni volte a ridurre l'impatto ambientale (Trione 2021).

Sul tema degli illeciti, tra le frodi in area blockchain, si segnala quella subita da Todd Kramer. Al gallerista newyorkese sono state sottratte opere per un valore di oltre due milioni di dollari, in seguito ad un attacco hacker attuato con un uso improprio delle credenziali di accesso ai documenti digitali legati alla collezione di NFT.

La tecnologia blockchain e gli NFT sono caratterizzati da decentralizzazione; l'assenza di soggetti regolatori rende le transazioni libere dai costi delle attività di mediazione ma, di contro, non prevede interventi in caso di illeciti, il che può creare seri danni al soggetto derubato. In questo caso però c'è stato un intervento da parte della galleria digitale OpenSea, che ha congelato la possibilità di utilizzare sulla sua piattaforma gli NFT sottratti, rendendoli di fatto un bene rubato e ha consentito a Kramer di recuperare molti dei suoi token (Exibart 2022).

Il problema resta aperto, ma la buona notizia è che non si è trattato di una falla del sistema, bensì di un comportamento ingenuo del gallerista, vittima di una truffa di phishing.

---

<sup>4</sup> Per il 2021 La Power 100, classifica stilata da Art Review e attesa febbrilmente nel mondo dell'arte contemporanea, ha posizionato al primo posto gli ERC-721.

## La blockchain

Il discorso sugli NFT rende indispensabile una riflessione sulla blockchain e sulle sue caratteristiche.

Per la produzione, ma anche e soprattutto per la catalogazione e la certificazione delle opere d'arte si guarda con grande interesse alla blockchain, che è un registro digitale di transazioni. Si tratta di un database decentralizzato, in cui le informazioni vengono registrate sequenzialmente, attraverso un'organizzazione a blocchi.

È proprio il blocco a rappresentare l'elemento di base della blockchain: i dati relativi alle transazioni sono registrati all'interno di un contenitore permanente e immutabile che funge da aggregatore di più transazioni. L'insieme dei blocchi, collegati in sequenza cronologica, compone la catena. Ciascun blocco possiede un ID univoco (hash), creato sottoponendo a un algoritmo crittografico l'ID del blocco che lo ha preceduto e i dati memorizzati nel blocco corrente. Questo garantisce l'integrità di tutti i dati memorizzati sulla blockchain. Se i dati di un blocco venissero alterati, infatti, l'algoritmo di hash restituirebbe risultati differenti. La catena propone caratteristiche di tracciabilità, verificabilità, sicurezza, salvaguardia della privacy.

Una blockchain utilizza una rete di nodi decentralizzata: i dati sono diffusi e replicati in una rete *peer-to-peer* e ciascun nodo della rete salva e aggiorna in maniera autonoma una copia del registro. Non vi è alcun amministratore che abbia l'autorità per sovrintendere alla buona tenuta del sistema e validare i blocchi di transazioni; è altrettanto vero però che è la blockchain stessa ad agire da intermediaria, garantendo l'osservanza di rigide regole. Come abbiamo descritto, i comportamenti degli utilizzatori possono però creare criticità.

Giovanni Michetti ha analizzato la blockchain, in rapporto ai sistemi documentali, verificandone anche l'attendibilità in funzione dei parametri dello standard ISO 15489-1 (Michetti 2020, 103). Riprendendone l'assimilazione alla tipologia documentaria di registro, ha evidenziato come il modello di blockchain sia stato ideato per realizzare un sistema di valuta virtuale, cioè per registrare transazioni di natura economica. Nonostante le sue origini, la blockchain non è però vincolata all'ambito finanziario: l'aspetto rilevante e più generale è semmai la capacità di registrare una transazione, con particolare riferimento alla provenienza e alla destinazione della proprietà del bene transatto, qualunque sia la sua natura.

Le transazioni possono riguardare quindi qualunque bene tangibile (mobile o immobile), ma anche beni intangibili, come ad esempio i diritti relativi alla proprietà intellettuale. Nei fatti l'introduzione della blockchain è sempre più ricorrente in diversi settori, da quello dei beni di lusso, a quello dei servizi, a quello delle attività produttive, al mondo delle biblioteche (Morriello 2019). La vera novità della blockchain è nella visione che le sottende e nel tentativo di offrire una soluzione tecnologica avanzata ad un problema che in prima istanza è di natura socio-culturale: la gestione della fiducia nella società digitale (Michetti 2020, 103).

Per una serie di ragioni legate anche a quelli che sono apparentemente i suoi punti di forza, la blockchain non potrà mai diventare ambiente esclusivo di gestione dei sistemi documentali ma potrà, certamente, parteciparvi.

## Opere d'arte in NFT: alcune esemplificazioni

Benché presenti nel mondo digitale almeno da un paio d'anni addietro, gli NFT iniziano a suscitare attenzione e interesse dal 2017, anno in cui viene proposto dagli sviluppatori di software canadesi Matt Hall e John Watkinson dello studio Larva Labs, il progetto sperimentale Crypto Punks. Si tratta di 10.000 figurine in 24x24 *pixel art* in 8 bit che rappresentano degli NFT a tutti gli effetti: sono generate da algoritmi; ciascuna è imm modificabile, in esemplare unico e dopo un'iniziale distribuzione gratuita sono state commercializzate. Il sito dedicato è periodicamente aggiornato con i valori di scambio degli NFT delle icone punk, che hanno raggiunto transazioni nella criptovaluta ETH corrispondenti al valore di 7,58 milioni di dollari.

L'evento, primo di una serie, ha suscitato scalpore e ha aperto una strada battuta poi, nel 2021, anche dalle più celebri case d'asta. Christie's l'11 Marzo 2021, ha venduto l'opera digitale *Everydays: The first 5000 Days*, del graphic designer Mike Winkelman (alias Beeple), per la somma di circa 69 milioni di dollari, seguita a ruota da Sotheby's che ha organizzato dapprima l'asta *The Fungible Collection* (che ha raccolto circa 17 milioni di dollari), cui ha fatto seguito a una seconda asta, *Natively digital: A Curated NFT Sale*, tenuta dal 3 al 10 giugno 2021.

Nel campo dell'arte digitale agli NFT viene attribuito il merito di rendere un contenuto non replicabile, ponendo quindi l'oggetto digitale al riparo dalle criticità collegate alla riproduzione massiva e alla distribuzione non autorizzata sui canali web.

La casistica, nella stragrande maggioranza dei casi, fa rilevare che avere nel proprio portafoglio virtuale un NFT per l'acquirente non significa possedere l'opera digitale, che resta nella proprietà dell'autore o di terzi: la proprietà infatti è sull'NFT, che comunque si traduce in un diritto sull'opera stessa. Iniziano però a nascere e consolidarsi piattaforme in cui la compravendita di un'opera digitale e del correlato NFT comprendono sia la certificazione di autenticità sia la piena proprietà dell'opera (Annunziata-Conso 2021, 35).

## NFT e opere d'arte tradizionali: riproduzioni digitali e certificazione congiunta

Un interessante aspetto è senza dubbio rappresentato dall'applicazione degli NFT al mondo dell'arte tradizionale, con la "tokenizzazione" di opere fisiche.

Nel maggio 2021, le Gallerie degli Uffizi hanno affidato la digitalizzazione in scala 1:1 del Tondo Doni di Michelangelo con il sistema di cifratura e crittografia DAW® (Digital Art Work, brevettato in esclusiva dall'azienda Cinello), ottenendo la prima serigrafia digitale in edizione limitata e certificata da un NFT che attesta, sulla blockchain, la proprietà della riproduzione. La vendita del primo esemplare ha fruttato al museo un incasso di 70.000 euro e ha rappresentato l'esempio di una nuova modalità di reperimento fondi: l'accordo tra le parti prevedeva infatti che l'azienda liquidasse agli Uffizi una somma pari al 50% del ricavo netto del prezzo di vendita (Pirrelli 2021). L'intenzione è quella di creare attraverso gli NFT delle versioni digitali di capolavori tradizionali; una sorta di mercato secondario ma, al contempo, capace di accrescere le entrate delle istituzioni museali in misura significativa.

Siamo di fronte a "multipli digitali" delle opere d'arte, creati con il consenso dell'istituzione che possiede l'originale; qualcuno arriva a sostenere che un DAW possa essere considerato un "originale" digitale, sia perché l'utilizzazione di una tecnologia di digitalizzazione così particolare

potrebbe condurre a un'elaborazione dell'opera che superi la dimensione esclusivamente meccanica, sia perché la tecnologia stessa rende l'opera non riproducibile e ne garantisce l'unicità (Minio 2020).

## **NFT e certificazione digitale di opere su supporto tradizionale**

Non va trascurato infine che l'NFT è considerato come potenziale elemento di facilitazione nelle transazioni di opere artistiche fisiche, di cui dovrebbe essere in grado di certificare immutabilmente originalità e proprietà.

In questa direzione, Verisart è stata la prima compagnia al mondo ad offrire certificazioni d'arte su blockchain. Di recente è balzata agli onori della cronaca la piattaforma InArtNFT, ideata con l'obiettivo di creare l'«Internet del valore» per un mercato dell'arte sempre più tracciabile, sicuro e trasparente grazie alla certificazione in blockchain decentralizzata, verificabile e non falsificabile. Il catalogo propone dipinti e sculture di pregio acquistabili in criptovaluta (Binance Coin). Insieme al valore dei beni la piattaforma permette di trasferire all'acquirente anche un NFT, collegato alle singole opere e collezioni, come certificato digitale di autenticità e come documento di legittimazione che accompagna la consegna dell'opera e dei documenti originali.

Nello specifico la piattaforma ha proposto in vendita, nel dicembre scorso, il dipinto “Giuseppe e la moglie di Putifarre” realizzato dal Guercino nel 1631 e ritenuto disperso per quattro secoli. L'opera è oggi ricomparsa nelle disponibilità di un collezionista privato che ha scelto InArtNFT come intermediario digitale per la vendita; il dipinto è valutato 5 milioni di euro, da corrispondere in criptovaluta (Olivari 2021). È legittimo chiedersi se la proposta di vendita sia reale o se non si tratti di una strategia di marketing, volta a creare attenzione intorno agli NFT e al ruolo che potrebbero assumere nella certificazione e nelle operazioni di vendita di opere sinora proposte nei soli circuiti tradizionali del mercato.

Occorre considerare, quindi, cosa può essere associato a un NFT: il token può riferirsi a un bene nativo digitale ma, come indicato, sembra poter essere efficacemente associato anche ad un bene fisico digitalizzato. In questo caso, si rileva la necessità di individuare soluzioni idonee a collegare in modo univoco l'opera fisica al token digitale, che dovrebbe certificare l'autenticità.

Se si applica la tokenizzazione di beni fisici sembra possibile identificare due beni, consistenti in un'esemplare fisico e in un'esemplare digitale, entrambi originali, convenzionalmente trasferiti a mezzo di un'unica cessione e – ancora convenzionalmente – destinati a un'esistenza congiunta, benché sul piano pratico potrebbero essere oggetto di circolazione parallela ed autonoma.

Particolarmente interessante, sotto questo profilo, la scelta di quegli artisti che, anche per evitare il corto circuito che potrebbe derivare dalla circolazione parallela dell'esemplare fisico e di quello digitale, una volta praticata la digitalizzazione, provvedono ad eliminare l'esemplare fisico.

Sembra opportuno citare a questo proposito il recente progetto *The currency* di Damien Hirst, che consiste in una collezione di 10.000 opere d'arte fisiche custodite nel Regno Unito, cui corrispondono 10.000 NFT. Tali opere sono state cedute in NFT, ma l'acquirente deve scegliere se intende possedere l'NFT o l'opera fisica. Nel caso in cui la scelta ricada sul token, l'opera fisica viene bruciata; viceversa, ove la scelta ricada sull'opera fisica, sarà il token a subire l'operazione di *burning*, consentendo di mantenere in esistenza il solo esemplare cartaceo dell'opera.

Il progetto, in pochi mesi dal suo lancio, ha generato un valore di oltre 25 milioni di dollari grazie alla successiva circolazione dei token in questione (Simeone 2021).

## Ancora sulla funzione di certificazione degli NFT

Gli artisti, il mercato dell'arte e il collezionismo necessitano di strumenti di attribuzione certa e di certificazione di autenticità, pertanto potrebbero trovare una soluzione soddisfacente nella registrazione univoca su blockchain.

Come abbiamo visto, un NFT può rappresentare lo strumento di certificazione di unicità di un *digital asset* e ciò produce un certificato digitale, non duplicabile, quale attestazione della proprietà dell'NFT stesso. Il token permette quindi di stabilire la provenienza di un determinato oggetto digitale, consentendo l'identificazione del creatore e del possessore – o dei possessori che si sono succeduti – e l'esplicitazione dei diritti che incorpora, con la specifica di chi può esercitarli legittimamente. Ciò è reso possibile dalla presenza di un registro sulla blockchain, che codifica la provenienza, e di uno smart contract, che definisce le condizioni del trasferimento dell'NFT. Più precisamente, l'NFT non contiene l'oggetto digitale di proprietà cui si riferisce, ma soltanto la sua descrizione, i dati di proprietà, l'hash e/o puntatore a dove la proprietà dell'oggetto digitale è digitalmente memorizzata (nodo IPFS o normale web server).

Restano però ancora da definire molti aspetti che riguardano, in particolare, l'esercizio alla protezione dei diritti, che negli NFT sono incorporati.

Solo per citarne alcuni, si pensi alla discussione sulla attendibilità della certificazione: sulla base dei primi esempi in circolazione, è ancora dubbio che il certificato possa assicurare una generale definizione di che cosa voglia dire “essere il proprietario” di un oggetto digitale con NFT, nonché i quali diritti rimangano in capo, invece, al suo creatore; ancor più dubbio il regime di quegli NFT che prevedono un meccanismo di *royalty* a favore di sviluppatori, artisti, creatori ad ogni rivendita. In termini di inquadramento giuridico, in attesa di una regolamentazione dedicata e considerate la funzione di certificazione digitale e di attestazione della titolarità dei diritti in essi incorporati, alcuni propendono per assimilarli nella fattispecie, tipica dell'ordinamento italiano, ai titoli di legittimazione (Annunziata-Conso 2021, 39-40).

Dal punto di vista normativo il cosiddetto “Decreto Semplificazioni” (D.l. 135/2018, art 8 ter), ha introdotto la definizione delle tecnologie basate su registri distribuiti (blockchain) e degli smart contract. Il decreto prevede anche che la memorizzazione di un documento informatico attraverso l'uso di tecnologie basate su registri distribuiti produca gli effetti giuridici della validazione temporale elettronica di cui all'articolo 41 del regolamento UE n. 910/2014 del Parlamento Europeo e del Consiglio, del 23 luglio 2014.

L'Agenzia per l'Italia digitale non ha però ancora individuato gli standard tecnici a cui dovrebbero parametrarsi le tecnologie basate su registri distribuiti, ai fini della produzione degli effetti probatori.

Il non fungibile token è un investimento che non si sottrae alle logiche di mercato e che rappresenta una modalità tecnologica vantaggiosa, poiché permette agli utenti di trasformare, di fatto, qualsiasi asset materiale o dematerializzato, in una forma proprietaria che può essere compravenduta velocemente e senza necessità di intermediazione.



Gli NFT hanno già conquistato interesse nel mondo della proprietà intellettuale e sembrano poter offrire, in prospettiva, opportunità interessanti per una gestione efficiente, trasparente e decentralizzata dei relativi diritti. Possono altresì migliorare il grado di certezza negli scambi commerciali grazie alla capacità di attestare in modo univoco l'autenticità di un bene digitale e tenere traccia della sua circolazione, garantendo inoltre la giusta remunerazione negli autori per il valore via via acquisito dalle loro opere.

Tuttavia, come ogni innovazione che esprime valore, il mercato di riferimento ha bisogno di consapevolezza e, soprattutto, di regole certe. Come è già accaduto per altre crypto attività anche per gli NFT vi è la concreta evenienza che la mancanza di un quadro regolamentare specifico consenta, ancora oggi di sfuggire all'applicazione di ogni norma sovraordinata e a quei controlli che, al pari di altri settori, hanno garantito una relativa solidità anche al mondo fisico dell'arte.

Le opere in crypto arte, quindi, sono elaborati artistici in alta definizione, realizzati con sofisticati strumenti digitali, concepiti per stare solo dentro la Rete. Tali opere possono essere esposte attraverso schermi; in termini di fruibilità va segnalato il lancio, da parte di una nota multinazionale, di una piattaforma che consentirà la navigazione tra i principali siti che offrono NFT, la possibilità di incrementare la propria collezione e di visualizzare le opere anche come sfondo dei televisori, quando non utilizzati (Ansa 2022). Le produzioni, infine, possono essere commercializzate ma non duplicate, perché criptate grazie alla tecnologia blockchain e all'applicazione NFT; la prima le certifica, ne legittima e ne garantisce l'autenticità e l'unicità, la seconda consente la creazione di file collezionabili, dei quali è possibile individuare proprietà, valore, tracciabilità.

## Conclusioni

Le tematiche di produzione, gestione e conservazione degli archivi d'artista e delle opere d'arte sono al centro di cambiamenti e meritano un'attenta analisi.

I principi di fruizione, certificazione e conservazione preventiva di molte delle opere di arte contemporanea create in digitale devono infatti essere ripensati in base a nuovi criteri, che attengono quasi esclusivamente alla struttura documentaria e alle componenti tecnologiche e digitali.

Se per le produzioni in digitale in modalità "classica" sono necessarie proposte e interventi, più che mai vanno monitorate le produzioni della crypto arte. Al fianco poi di questo ambito, che interessa in termini documentali il trattamento delle opere d'arte in una maniera più incisiva e determinante rispetto a quanto accade con la cosiddetta arte fisica, vanno considerati anche i parametri di produzione e gestione del relativo contesto documentario.

Alla produzione artistica criptata può affiancarsi la relativa produzione documentale, comprensiva di certificazione; ma, come abbiamo visto, la crypto certificazione può riguardare anche le opere prodotte in ambiente analogico. Il prossimo passo potrebbe essere quello della realizzazione di cataloghi ragionati in ambiente blockchain.

Molti tra gli addetti ai lavori valutano che, per la crypto arte e il suo nascente collezionismo, gli NFT possono rappresentare la soluzione ideale per garantire proprietà, autenticità, provenienza e "rarietà" del bene/diritto incorporato nel token, anche in termini di quantità di copie immesse sul mercato.

Un NFT, l'abbiamo precisato, è un insieme di metadati che include un *hash*, ossia un codice identificativo univoco del bene digitale cui si riferisce, che può essere scambiato per mezzo di *smart contract*. Tali metadati possono altresì includere un link (spesso un URL) al bene digitale cui si riferiscono, disponibile online su altri supporti o spazi di archiviazione. Si tratta di “oggetti digitali” di vario genere, tra i quali immagini e foto, video, musica, tweet, estratti, codice sorgente, testi, *domain name*, opere d'arte digitali, *collectible*, ma anche versioni digitali di beni fisici.

Cambia dunque molto nel modo di intendere e rapportarsi al “documento” in accezione di oggetto della produzione artistica, di documento certificatorio e, non ultimo, di oggetto di eredità digitale in relazione al valore di beni digitali a contenuto patrimoniale che gli NFT assumono (d'Arminio Monforte 2022).

Il discorso, anche nell'ambito artistico inteso in senso più ampio, si va aprendo a nuove sperimentazioni: di recente Alessandro Baricco ha creato un NFT, lanciato con il titolo *Novecento. The Source Code* e realizzato con la registrazione della propria interpretazione vocale del testo

“alla fine mi chiedevo cosa avevo in mano con quel codice sorgente di *Novecento*: e la risposta è stata: un tipico NFT. Leggero, elementare, ruvido, feticistico, immateriale” (Baricco 2022).

Tanto l'introduzione della blockchain, come già indicato da autorevoli studiosi, quanto le diverse attività in ambito NFT implicano un coinvolgimento degli archivisti nella riflessione scientifica e metodologica. In questa ridefinizione di ruoli e rapporti, i sistemi documentali hanno l'opportunità di assumere un ruolo di primario soggetto di riferimento e sarebbe un grave errore non entrare in partita e non provare a partecipare a processi decisionali che interessano e gravano, come abbiamo visto, sull'ambito documentale.

## Riferimenti bibliografici

ANAI (Associazione nazionale archivistica italiana). 2021. “Le Muse in archivio: itinerari nelle carte d’arte e d’artista”. [http://media.regesta.com/dm\\_0/ANAI/anaiCMS//ANAI/000/1741/ANAI.000.1741.0002.pdf](http://media.regesta.com/dm_0/ANAI/anaiCMS//ANAI/000/1741/ANAI.000.1741.0002.pdf).

Annunziata, Filippo e Andrea Conso, a c. di. 2021. *NFT: L’arte e il suo doppio. Non fungible token: l’importanza delle regole, oltre i confini dell’arte*. Milano: Montalbano.

Ansa. 2022. “Samsung, Nft come opere d’arte da visualizzare su televisori. Creazioni digitali diventano quadri da sfogliare sugli Smart Tv”. [https://www.ansa.it/sito/notizie/tecnologia/hitech/2022/01/03/samsung-nft-come-opere-darte-da-visualizzare-su-televisori\\_dd9afa2e-5daf-4899-87b7-e43d801081d0.html](https://www.ansa.it/sito/notizie/tecnologia/hitech/2022/01/03/samsung-nft-come-opere-darte-da-visualizzare-su-televisori_dd9afa2e-5daf-4899-87b7-e43d801081d0.html).

Baricco, Alessandro. 2022. “Il mio Novecento all’asta come NFT”. *la Repubblica*. 15 gennaio: 36-37.

Bettazzi, Sofia. 2021. “The application of RDA to archives: criticalities and advantages of using the universal metadata standard. The case of Galileo Chini Archive”. *JLIS.it* 12 (3): 119-128. doi: 10.4403/jlis.it-12746.

Buren, Daniel. 2012. *Catalogue raisonné 1967-1972*. <http://www.catalogueraisonne.danielburen.com/artworks/index/page:5/period:All>.

Catricalà, Valentino e Domenico Quaranta, a c. di. 2020. *Sopravvivenza programmata: Etiche e pratiche di conservazione, dall’arte cinetica alla Net Art*. Roma: Kappabit.

Damiani, Concetta e Mariella Guercio. 2019. *Il ruolo degli archivi d’arte nel mondo contemporaneo*. In *Rodolfo Pallucchini: storie, archivi, prospettive critiche*, a cura di Claudio Lorenzini. Udine: Forum editrice universitaria udinese, 429-441.

d’Arminio Monforte, Alessandro. 2022. “NFT e collezioni private. Come funziona l’eredità digitale”. *Artribune*. <https://www.artribune.com/progettazione/new-media/2022/01/nft-collezioni-private-eredita-digitale/>.

Donati, Alessandra. 2016. *Rilevanza giuridica dell’archivio d’artista*. In *Impresa cultura. Creatività, partecipazione, competitività. XII rapporto annuale Federculture*. Roma.

Exibart. 2022. “4 Bored Apes e 2 milioni di dollari in NFT rubati dal portafoglio di un gallerista”. 5 gennaio. <https://www.exibart.com/mercato/4-bored-apes-e-2-milioni-di-dollari-in-nft-rubati-dal-portafoglio-di-un-gallerista/>.

Fondazione Paolini. 2021. *Catalogo ragionato*. <https://www.fondazionepaolini.it/ita/catalogo-ragionato>.

Foti, Michele. 2018. “Una panoramica delle strategie di conservazione di opere d’arte digitale”. *Umanistica digitale*, 3. doi: 10.6092/issn.2532-8816/8209.

Lozano-Hemmer, Rafael. 2015. *Best practices for conservation of Media Art from an artist’s perspective*, <https://github.com/antimodular/Best-practices-for-conservation-of-media-art>.

Martorano, Annantonia. 2014. *L’Archivio di Galileo Chini: introduzione, inventario*. Lucca: Istituto Storico Lucchese.

Martorano, Annantonia. 2018. *Un uomo e un artista nello specchio delle sue carte*, in Laura Giambastiani (a cura di), *Percorsi negli archivi*. Lucca: Maria Pacini Fazzi.

Michetti, Giovanni. 2020. *Introduzione alla blockchain: Una guida per archivisti*. Napoli: Editoriale Scientifica.

Minio, Federica. 2020. “La digitalizzazione delle opere d’arte tra beni culturali e diritto d’autore”. *Artribune*. <https://www.artribune.com/professioni-e-professionisti/diritto/2020/01/digitalizzazione-beni-culturali-diritto-autore/>.

Ministero dei beni e delle attività culturali e del turismo. 2017. *Circolare 7/2017 DG-AAP: Certificato PACTA – Protocolli per l’autenticità, la cura e la tutela dell’arte contemporanea e relative linee guida per l’utilizzo*. <https://www.beniculturali.it/comunicato/circolare-7-2017-dg-aap-certificato-pacta-protocolli-per-l-autenticita-la-cura-e-la-tutela-dell-arte-contemporanea-e-relative-linee-guida-per-l-utilizzo>.

Ministero della Cultura. 2021. *Museo nazionale dell’arte digitale*. <https://www.beniculturali.it/ente/museo-nazionale-dell-arte-digitale>.

Morriello, Rossana. 2019. “Blockchain, intelligenza artificiale e internet delle cose in biblioteca”. *Aib studi*, 59 (1-2): 45-68. doi: 10.2424/aibstudi-11927.

Olivari, Arianna. 2021. “Odissea nel Metaverso: anche un dipinto di Guercino diventa NFT”. *Exibart*. <https://www.exibart.com/mercato/odissea-nel-metaverso-anche-un-dipinto-di-guercino-di-venta-nft/>.

Pedrazzi, Rebecca. 2021. *Futuri possibili: Scenari d’arte e intelligenza artificiale*. Milano: Jaca Book.

Pirrelli, Marilena. 2021. “Gli Uffizi sdoganano il Tondo Doni in versione NFT”. *Il Sole 24 Ore*. 18 maggio. <https://www.ilssole24ore.com/art/gli-uffizi-sdoganano-tondo-doni-versione-nft-AEuiMFK>.

Richter, Gerhard 2020. <https://www.gerhard-richter.com/it/>.

Simeone, Mario Francesco. 2021, “Il primo progetto di arte in NFT di Damien Hirst vale 25 milioni di dollari, per ora”. *Exibart*. <https://www.exibart.com/arte-contemporanea/il-primo-progetto-di-arte-nft-di-damien-hirst-vale-25-milioni-di-dollari-per-ora/>.

Trione, Vincenzo 2021. “Tutt’altro che sostenibile, la Crypto Art consuma energia quanto l’Argentina”. *Il corriere della sera*. 24 maggio. [https://www.corriere.it/pianeta2020/21\\_maggio\\_24/tutt-altro-che-sostenibile-crypto-art-consuma-energia-quanto-l-argentina-ecce91e-bc74-11eb-ab-b7-46b8b952d96c.shtml](https://www.corriere.it/pianeta2020/21_maggio_24/tutt-altro-che-sostenibile-crypto-art-consuma-energia-quanto-l-argentina-ecce91e-bc74-11eb-ab-b7-46b8b952d96c.shtml).

Visconti, Arianna. 2020. “Il mercato dell’arte: Contraffazione di opere d’arte e posizione del curatore d’archivio”. *Aedon*, 1, <http://www.aedon.mulino.it/archivio/2020/1/visconti.htm>.

Wijers, Gaby. 2020. “Etica e prassi della conservazione della Media Art dal punto di vista di un artista”. In *Sopravvivenza programmata: Etiche e pratiche di conservazione, dall’arte cinetica alla Net Art*, a cura di Valentino Catricalà e Domenico Quaranta, 91-136. Roma: Kappabit.

Zorloni, Alessia. 2018. “L’archiviazione autentica con i PACTA”. *Il Giornale delle fondazioni*. <http://www.ilgiornaledellefondazioni.com/content/l%E2%80%99archiviazione-autentica-con-i-pacta>.