Archival science, laboratories, and open science:

Contributions and challenges for document and data management

Archivología, laboratorios y ciencia abierta: contribuciones y retos para la gestión de documentos y datos / Arquivología, laboratórios e ciência aberta: contribuições e desafios para a gestão de documentos e dados

Paulo Elian dos Santos

PhD in History from the University of São Paulo (USP). Researcher in the Archives and Documentation Department of the Casa de Oswaldo Cruz/Oswaldo Cruz Foundation (COC/Fiocruz). Professor in the professional master's degree programs in Document and Archival Management at the Federal University of the State of Rio de Janeiro (Unirio) and Preservation and Management of Science and Healthcare Cultural Heritage at COC/Fiocruz, Brazil. paulo.elian@fiocruz.br

ABSTRACT

This article analyzes laboratories from the perspective of social studies of science, as the environment where scientific practices occur, and archival records are produced. It presents a panoramic view of archival thought committed to the identification, preservation, and methodological treatment of "science archives." Finally, it examines open science as a field that is favorable to the mobilization of archival science references, as well as the establishment of transdisciplinary relationships in the documental reality of science institutions.

Keywords: science archives; scientific laboratories; open science; research data management.

RESUMEN

El artículo analiza el laboratorio desde la perspectiva de los estudios sociales de la ciencia, como un entorno para las prácticas científicas y la producción de registros de archivo. Ofrece una visión general del pensamiento archivístico que se dedicó a la identificación, preservación y tratamiento metodológico de los "archivos científicos". Finalmente, aborda la ciencia abierta como un campo propicio para la movilización de los referentes de la archivología, así como el establecimiento de relaciones transdisciplinarias para acercarnos a la realidad documental de las instituciones científicas.

Palabras clave: archivos científicos; laboratorios científicos; ciencia abierta; gestión de datos de investigación.

RESUMO

O artigo analisa o laboratório na perspectiva dos estudos sociais da ciência, como ambiente das práticas científicas e da produção dos registros de arquivo. Traça um panorama do pensamento arquivístico que se dedicou à identificação, à preservação e ao tratamento metodológico dos "arquivos de ciência". Por fim, aborda a ciência aberta como um campo favorável à mobilização dos referenciais da arquivologia, assim como ao estabelecimento de relações transdisciplinares para tratarmos da realidade documental das instituições de ciência.

Palavras-chave: arquivos de ciência; laboratórios científicos; ciência aberta; gestão de dados de pesquisa.



Introduction

Science is a complex activity that has its own specific spaces, protocols, and practices. Viewed by scholars as part of social and cultural life, it absorbs, uses, and experiments with digital technologies, which are impactful vectors in the dynamics of production and circulation of data and scientific information.

Since the 1990s, authors from different currents of thought who have examined the history of archival thought in the 20th century have defended a vision of dialogue with the radical transformations and changes seen in the nature and use of documents, in the organizations that produce them, in administration systems, as well as in the cultural, legal, ethical, technological, and social aspects present in society (Cook, 2018).

This combination of new dynamics of creation, treatment, use and access to data and documents imposed on science and on archive management produces a material reality that favors initiatives that bring together theoretical, methodological, and practical knowledge for a transdisciplinary approach to documents and research data. In this regard, this article analyzes laboratories from the perspective of social studies of science as the institutional environment of scientific activities and practice, which is therefore the location where archival records are produced. To this end, the references used are the pioneering formulations of Bruno Latour (1997; 2000) presented systematically within a cartography of "science in action." Next, we present a panoramic view of the principal contributions of archivist thought regarding the work of identification, understanding, and methodological treatment of these records. Lastly, the article presents some revealing elements of the renewed interest from scientific institutions in the appropriation of concepts, methods, debates, and archival practices that provide a consistent framework for data management, dialogue with library science and information science, and the establishment of transdisciplinary relationships.

The laboratory: an article factory and place for archives

Scientific adventures follow a variety of paths. Their base is the laboratory, but this extends to "other places." Scientists teach, they invest in the training of young researchers, conduct field studies, create, and participate in scientific organizations, assume management duties, take positions in regard to public debate and work in national and international organizations. In this article, our reflections on scientific activity come from a privileged place for studying,

analyzing, and trying to understand scientific practice: the laboratory. A place for experimental study, laboratories associate scientific knowledge with practical objectives, and to this end, they bring together people, instruments and equipment, chemicals, and lab animals. Bruno Latour and Steve Woolgar (1997) describe a scene from the 1970s that reveals a routine that remains relevant to this day:

Technicians clean the dissection workbenches, weigh chemicals, and collect the pages of data from the computers; secretaries sit in front of typewriters and correct the handwritten notes; researchers trade comments about the day's activities and go to their workbenches; other employees are organizing lab animals, chemicals, and correspondence; blackboards are full of equations and notes; and computers output endless lists. (Latour; Woolgar, 1997, p. 11)

This description is a brief introduction to the intense and effervescent work of a "tribe" that spends most of its time codifying, marking, reading, and writing (Latour; Woolgar, 1997, p. 42). More than just associating scientific knowledge with practical objectives, the laboratories of modern science research institutes have been adapting to the techno-scientific changes of our current times. The appearance of new technologies, disciplines, and research topics, as well as the high cost of equipment and inputs, require more complex forms of research organization.

For example, biomedical laboratories are a radical example of the affirmation that this is the "place of scientific life," the place of students, of teaching activities. Since this is essentially the teaching of a modus operandi, a mode of scientific production that assumes a mode of perception, a set of principles about vision and division, the only way to acquire it is to see it in operation in practice or to observe how this scientific habitus "reacts" to practical operations (Bourdieu, 2007, p. 22).

According to Pierre Bourdieu, science historians and philosophers (above all, scientists themselves) have frequently noted that an important part of the knowledge in this profession is obtained through the use of entirely practical means of acquisition — the "pedagogy of silence." What leads to explicitation not only of the outlines transmitted, but also of the outlines used in transmission is, without a doubt, greater in a science in which content, knowledge, ways of thinking and of action are less explicit and less codified (Bourdieu, 2007, p. 22).

It is in the laboratory and through an in-depth examination of its daily activities that we can perceive the social construction of facts, the microprocesses

of scientific fact production. Through the chain that begins with ideas, we move through the system of "written notes," through the logical arguments, affirmations, and through the obtaining of evidence, we have the microprocesses of how scientific facts arise, the result of negotiation among the scientists who make use of multiple persuasive strategies for this purpose.²

The description of the day-to-day work of laboratories presented by Latour is our principal source of information for unveiling the relationship between the scientific work done on the bench and written notes and writings, material vestiges designed to convince others of the validity of the affirmations generated by them.

For Latour, writing is a component of a considerable part of the process of laboratory work, which, from the workbench to the office, is impregnated with text: "those who work at the benches write compulsively [...]. Every workbench has a large record book in which the members of that section meticulously note down what they have just done with a given code" (Latour; Woolgar, 1997, p. 41).

For researchers, the regular use of protocol notebooks in their workbench activities is a necessary procedure in what are known as the "experimentation sciences," even to write an article, the final result of their work, proof of a study, frequently concluded and successful, and an unparalleled communication tool and career management instrument (Welfelé, 2004, p. 68-69). Protocol notebooks and articles are like unique documents, witnesses and principal forms of information compilation that acquire this status of reference in laboratory archives.

In Latourian analysis, the "assembly line" of scientific activity begins with the biotests that use machinery to generate a full page of numbers that is used to feed data into the computer. It then goes to the printer, which prints out different lists, and reaches the still intermediary product of all this effort: a simple curve plotted on graph paper. A few days later, a new version of the curve will be ready to be published in a scientific article.

¹ This concept of written notes was taken from Jacques Derrida (1967) and designates the operation that takes place prior to the writing. Here it serves to summarize the notes, tasks, bullet points, histograms, record numbers, spectra, graphs, etc.

According to the arguments presented by Latour and Woolgar (1997), the construction of a fact does not simply end there. Pure and disinterested science gives way to rational science, in which scientists behave like "capitalist investors," cold, calculating and seeking to maximize their interests. The notion thus arises of the cycle of credibility, which represents the ongoing search for a gain that would allow reinvestment, and consequently new gains in credibility.

This process, which isn't always so harmonious and orderly, is nuanced through formal and informal communication channels.³ If well-structured written work represents formal communication, all the discussions and information exchanged in the laboratory address the different points covered in the published literature — the article. Going a bit further, informal exchanges always have a direct or indirect relation to the documents, devices, forms, technical notes, drafts of texts, etc. In his role as an ethnographer, Latour favors the dynamics of laboratories and highlights the function of written documents and annotations and note-taking devices as a way to avoid being imprisoned by the discourse of the actors — the scientists. This position indicates the critical role played by documents of a variety of types, as well as by these devices used to produce notes on a specific substance, which are in turn used to produce new articles (Latour; Woolgar, 1997, p. 59).

The intermediary stages are essential to understanding scientific activity. The final objective, which is the production of articles, scientific communications, and presentations, requires this "chain of writing operations," which goes from the first result scribbled down on a piece of paper to the "classification of the article published in the laboratory archives." The countless intermediary stages are related to what Latour calls literary production. Presentation outlines with projections, preliminary research notes, protocol notebooks with the results of tests and methods, and versions of articles make up the natural accumulation of laboratory archives.

The final publication does not show the different stages of construction, eliminating the traces of the path followed, thereby leaving the impression of order and rationality as found in "black boxes;" that is, stabilized scientific facts. Articles are the means for revealing "stable" and "natural" scientific facts, the result of a construction process with the peculiar characteristic that it can only be completed when it is capable of erasing any trace of itself.⁴ Laboratories, the place where scientists spend a good part of their lives, are locations where important aspects of the construction of scientific facts and theories takes place — in summary, places where science is practiced.

From the viewpoint of the sociology of science, we were able to identify some central elements of the nature of scientific activity, and what it produces

³ The importance that Latour and Woolgar (1997, p. 46) attribute to documents contrasts with a trend in the sociology of science to value the role played by informal communications in scientific activity.

 $^{4 \}quad That is the position that Latour and Woolgar (1997) seek to demonstrate based on some notions compared to the ethnographic material gathered.$

in terms of records and documental materials to achieve its principal function, which is to generate original knowledge about a given subject, making it a social asset. According to the curator Odile Welfelé (2004), archivists find it to be more difficult to use the word "archives" than to identify what scientific activity produces. This word, which is familiar to us, is totally outside that universe. For the author, what comes from observation, reflection, and experimentation goes through different forms and media before it reaches its final stage in some lines on paper — the article.

The task of professionals in the field of archival science is to understand what the different science document forms and records are, and what functions are performed by them, as a part of an organic whole that we call archives. The protocol notebooks and articles are part of this set, but their functions are shared with other documents and should (or can) be preserved to serve as proof and evidence, just as any other archives document.

These facets – the daily microcosm of investigation and institutionally organized science, with its dynamic relationship in the contact between peers in the scientific community and their interaction with society — will be present in these archives, in different proportions, according to the actions, facts, and events of the trajectory of the groups and individuals, always giving them meaning.

Science archives: contributions to open science

The process of institutionalization of the history of science, with variations in each country, began in the 1960s. One of the hallmarks of historiographic production on science took place in the 1970s, when scientific institutions, viewed as spaces where social and scientific interests meet, begin to be studied by historians, sociologists, and philosophers. Science, seen as a practice, begins to occupy the space that had previously been occupied by the history of scientific ideas. In the 1980s, there was an important change in perspective, as well as in the forms of organization and professionalization of the researchers in this field. Progress in this historiographic production only became possible with the recognition of the importance of the social history of science, which focused on the study of the social conditions of implementation of scientific activities.⁵

Within this movement, archival document sources became the focus of attention of researchers who pointed out the importance of understanding the

 $^{\,\,}$ Science historian Dominique Pestre (1996) wrote a reference article with a consistent historiographic balance.

relationships between the national and social contexts, and the meaning of science at each specific moment. Studies focused on institutional and/or personal paths gained a new dimension insofar as they established the essential links between science, politics, and society.

The search for document sources by science historians is only successful largely insofar as their research agenda converges with initiatives to preserve and value the collection of documents produced by institutions and persons with careers in the field of science.

A brief panoramic examination of preservation and research into scientific archives leads us to France. Whether due to the pioneering stance of its institutional initiatives or the originality and influence of its contributions since the end of the 1980s, worthy of note is the mission of the French National Archives at the Centre Nationale de la Recherche Scientifique (CNRS), which enabled the agents at the Ministry of Culture to organize the management and conservation of documents, prepare guidebooks, and classify and transfer historical archives. Beginning in 1993, the program Archives Issues des Sciences Contemporaines (ARISC), developed at CNRS, conducted studies focused on writing production modes within the scientific environment, visiting laboratories, examining documents, such as laboratory notebooks, and interviewing scientists, specifically in the fields of physics, geology, and chemistry (Welfelé, 2006).

Odile Welfelé, the Chief Curator of the National Archives of France, who was on a mission to CNRS at that time, noted the relationship between scientific practices and the creation and maintenance of various types of documents, with the challenges imposed on archivists by the appearance of "electronic networks." Supported by a perspective that engages in dialogue with science sociology studies, her articles "La souris et l'encrier: pratiques scientifiques et inscriptions documentaires" (1994); "L'éprovette archivée: réflexions sur les archives et les matériaux documentaires issus de la pratique scientifique contemporaine" (1996); "Organiser le désordre: usages du cahier de laboratoire em physique contemporaine" (1998-1999); and "Quels matériaux pour l'historien d'aprés-demain?

A more generalized interest in scientific articles arose at the end of the 1980s, as part of the preparations for several commemorations: the centennial of Institut Pasteur, the fiftieth aniversary of Centre National de la Research Scientifique (CNRS) and the bicentennial of the different institutions founded during the revolutionary period (Conservatório Nacional de Artes e Oficios, École Polytechnique, École Normale Supérieure), which enabled awareness of the importance of using and preserving archives to understand the history of educational systems and of scientific research in the country. See: Charmasson (2007).

⁷ This article was translated into Portuguese by Maria Celina Soares de Mello e Silva and was republished in the journal of the Sociedade Brasileira de História da Ciência, with alterations by the author (Welfelé, 2004).

Réflexions prospectives sur le devenir des archives scientifiques" (1999)⁸ represent important contributions for an examination of laboratory documents within the context of contemporary science.

The French understand science archives — a less ambiguous expression than "scientific archives" — to be all archival sources that allow the study of the general evolution of scientific investigation and teaching policies, the evolution of a specific scientific discipline, or even the contribution of a scientist to the development of knowledge (Charmasson, 2007). This definition separates the object into three principal categories, based on their origin: the archives of supervision at teaching and scientific research establishments; the files of these same establishments (institutes and universities), including laboratories and research units; and the personal files of scientists.⁹

Alongside Odile Welfelé, noteworthy are the works of Thérèse Charmasson, ¹⁰ Christiane Demeulenare-Douyére, Catherine Gaziello and Denise Ogilvie, in collective works and dossiers dedicated to this topic, published in journals in that country. In 1995, Charmasson, Demeulenare-Douyére, Gaziello and Ogilvie wrote Les archives personnelles des scientifiques: classement et conservation, a reference work that combines the characteristics of a guide and a manual, intended for scientists, archivists, librarians, and science history researchers. In Charmasson's (1999) view, the personal archives of scientists are of great interest and are similar to and become confused with laboratory archives, especially where scientists hold or held the functions of heads or group leaders.

In 1988, 1993 and 1996, the journées d'études organized by the Centre de Recherche en Histoire des Sciences et des Techniques resulted in articles being published in two special issues of La Gazette des Archives. The general topic of the archives des sciences is a broad cartography of the institutional fonds and personal archives of scientists in France, their states of preservation, their typologies, characteristics, and possible use by researchers. Nonetheless, two aspects

⁸ This article was translated into Portuguese in the book Caminho para as estrelas: reflexões em um museu [Path to the stars: reflections in a museum], organized by Ana Maria Ribeiro de Andrade (2007).

⁹ We can add a fourth category of archives, grouping together the archives of scientific associations and societies, such as science academies, which have been privileged places for the "militancy" of savants since the end of the 18^{th} century in Europe.

¹⁰ Chief Curator of Heritage (Archives), at the Centre de Recherche em Histoire des Sciences et des Tecniques, Cité des Sciences et de l'Industrie, Charmasson has a multitude of publications on sources for research on and the history of teaching in France. Among her works, we note: Archives et sources pour l'histoire de l'enseignement (2005); Histoire de l'enseignement, XIXe-XXe siècles. Guide du chercheu (2006); Les archives des scientifiques. XVIe-XXe siècle. Guide des fonds conservés en France (2008). Available at: http://www.cths.fr/hi/personne.php?id=2236. Accessed on: 16 Jun. 2021.

should be noted: identification of the original document types of scientific activities, especially those related to research conducted in laboratories; and the need to look at them differently to understand their full meaning.

At the end of the 1990s, in the journal Les Cahiers de l'École Nationale du Patrimoine, we came across a group of articles about the topic Sciences et archives contemporaines¹¹ [science and contemporary archives] which reaffirm questions relevant to the policies for protection of scientific heritage, the systemic treatment of institutional and personal archives, the transformations of science itself, and the impacts of the révolution informatique on work processes and document production.

This effort was not restricted to France. ¹² The "American line," expressed in the work of Helen Samuels, ¹³ offers us another perspective. Unlike the French line, it proposes a classification method that it calls "institutional functional analysis" and posits evaluation as a central problem of archival science, but also shows concern regarding science & technology documentation actions. Worthy of note is the manual Appraising the records of modern science and technology: a guide. This work presents the detailed results of an extensive project to identify and systematize science and technology activities and the documents generated by them (Haas et al., 1985).

The French authors recognize the valuable role played by the American studies. Odile Welfelé (1999a, p. 109-110), one of the principal names in archival production in this area, recognized that France lacked a work of a general nature describing the major stages of scientific work and of the documents, such as Appraising the records of modern science and technology: a guide.

The studies conducted by Helen Samuels (Haas; Samuels; Simmons, 1985; Samuels, 1995) are clearly centered on the process of technological research and development, the laboratory: choice of a problem, moving through the formulation of hypotheses, conducting of experiments, and on through patenting and

¹¹ Published in issue no. 3 (1999) of the journal, the texts were originally discussed at the round-table discussion coordinated by Odile Welfelé and organized by l'École Nationale du Patrimoine on November 20, 21 and 22, 1996.

¹² At the same time, it was in the 1990s that the first events appeared with the objective of promoting communication, the exchange of information, and collaboration among professionals at research and teaching institutions. In 1992, the International Council on Archives (ICA) created the Section on University and Research Institution Archives (SUV), open to professionals from universities, research institutions, science academies and scientific bodies.

¹³ Among her works, of special note is the article "Appraising the records of modern science and tecnology" (Samuels, 1995), published in the journal Janus, which is a synthesis of the guide (Haas et el., 1985), prepared in collaboration with Joan K. Haas and Barbara Trippel Simmons, and published in 1985.

communication. Samuels unveiled the principal functions of scientists and the institutions where they work. The post-war period left us "great science," characterized by interdisciplinary projects, research teams that favored the disappearance of personalization and the integration of science and technology. The author points out aspects that we believe merit emphasis. First of all, research is conducted in laboratories and may be carried out at institutions focused only on scientific and technological studies or at smaller units within multifunctional organizations, such as universities or large business corporations, at which research and development are only some of the many activities conducted. Moreover, scientists do not work in isolation; they depend on networks of peers/collaborators and administrators and communicate with them. This complex environment of internal and external associations requires that the archivist understand a universe of interconnected documents (Santos, 2012).

In Brazil, in the early 1990s, the topic of science archives was restricted to the holding of scientific events by universities or research centers. ¹⁴ In 2003, through an initiative of the Casa de Rui Barbosa Foundation and the Museum of Astronomy and Related Sciences (Mast), the First Scientific Archives Meeting ¹⁵ was held. This can be considered a landmark event, transformed into academic space for the exchange of experiences, covering not only university institutions, but also technological research and development institutes and centers of memory and documentation throughout the country. Within this context, reflections arose that, sustained by a critical view of the methodological procedures for treating science archives, searched for conceptual and methodological alternatives in archival science and in other disciplines.

Founded in the mid-1980s, the Casa de Oswaldo Cruz (COC), a part of the Oswaldo Cruz Foundation (Fiocruz), and the Museum of Astronomy and Related Sciences (Mast) represent the two most important, institutionalized experiences outside universities in regard to the surveying, gathering, founding, and technical treatment of science archives. Beginning in the 2000s, these institutions have been seminal in theoretical and conceptual reflections and in the preparation of methodological and technical instruments.

¹⁴ At the start of this decade, national university archives seminars were held at the State University of Campinas (Unicamp). There is a set of publications — that includes a commented bibliography in Portuguese (1985 to 1995), prepared by Prof. Mariza Bottino in 1996 — on the University Archives Section (SUV) page of the International Council on Archives (ICA). Available at: ica.org/es/suv-publicaciones. Accessed on: 22/Apr. 2021.

¹⁵ Held every two years, the seventh meeting took place in 2017, at which time they were interrupted. They all led to publications in annals or books, most of which are available in a digital version on the Mast site: http://www.mast.br/pt-br/publicacoes.html.

It was from these institutional spaces that the first works appeared, focused on an analysis of the origin of the documents and the characteristics of the archives resulting from scientific activity. Concerned with the study of classification or arrangement methods for the personal archives of scientists, and under the prism of archival science, Santos (2012) proposed a renewal of the organization criteria, capable of translating the functions and activities of scientists, establishing proximity with the scientific practices embodied in those archives. To this end, it seeks bases in the literature on archives, combined with the perspective of social studies of science, or of "science in action," developed by anthropologist Bruno Latour. The methodological procedures of classification and description proposed by Santos (2012) were soon adopted for the personal archives of the Casa de Oswaldo Cruz, and the perspective of adoption of a sociological focus on scientific activity became a reference for other institutions where documents are kept.

At the Mast Science History Archives, under the leadership of archivist Maria Celina Soares de Mello e Silva, the first research publications began to appear, focusing on personal archives, but also seeking to cover institutional archives. Aspects of document management in laboratories, document evaluation, document production and preservation of digital documents are examples of the topics she examined (Silva, 2006, 2007, 2009). In her dissertation Visitando laboratórios: o cientista e a preservação de documentos [Visiting laboratories; the scientist and document preservation (2007), she systematizes her research by analyzing the relationship between scientists and the documents produced in scientific and technological laboratories, while at the same time gathering elements for a science and technology archives preservation program. Her work, which covers more than one hundred laboratories, performs an in-depth examination of a set of questions related to the frontiers of the institutional and personal dimensions present in the production and accumulation of these documents and the view of scientists regarding the importance of scientific memory.

¹⁶ The subject of this study was the personal archives of the public health physician Rostan Soares, held by the Casa de Oswaldo Cruz/Fiocruz, and resulted in the master's thesis entitled Entre o laboratório, o campo e outros lugares: gênese documental e tratamento técnico em arquivos de cientistas, [Between the laboratory, the field, and other places: document origins and technical treatment in scientific archives] defended in February 2003 in the Graduate Program in Social History at the University of São Paulo. An integrated version of chapters 2 and 3 of this thesis was published in 2012, under the title Arquivos de cientistas: gênese documental e procedimentos de organização [Scientists' Archives: document origins and organization procedures] by the Association of Archivists of São Paulo (ARQ-SP) (Santos, 2012).

This broad cartography of the documental reality of scientific institutes has also led to other results. Maria Celina Silva co-authored Guia básico para preservação de arquivos de laboratório [Basic guide for preservation of laboratory archives] with Vera Lúcia Lopes Rego (Silva; Rego, 2010). This is a pioneering reference work, the result of a study conducted in the laboratories of the Ministry of Science and Technology between 2004 and 2008, and intended to provide guidance to institutions, managers, scientists, and archivists. With a similar objective, Santos, Pinto and Santos (2011) organized the Manual para gestão de documentos e arquivos de laboratórios das ciências biomédicas [Manual for laboratory document and archives management for the biomedical sciences], the result of research at the Oswaldo Cruz Institute, which offers, to a diverse public, procedures, orientations, and legal and normative references, as well as a video with scientists and technicians.

Learning more about work environments and scientific practices is at the center of studies on archives at science institutions. In the area of biomedical and health sciences, Fiocruz has a collection that dates back to the early part of the 20th century and includes dozens of personal archives of scientists, physicians, and sanitarians, many of whom have had careers linked to the institution. The concepts, visions, practices, and intentions of a group of scientists were mapped out by Verônica Martins de Brito (2002) in her thesis A preservação da memória científica da Fiocruz: o que pensam os cientistas [Preservation of the scientific memory of Fiocruz: what scientists think]. In this study, the author sought to achieve a panoramic view that would enable greater understanding or a closer view of the positions of these men and women on matters related to information, history, and "scientific memory."

What do scientists do? How does a scientific laboratory work? What concepts, methods, techniques, and archival practices should we mobilize to deal with the documents generated by scientific activity? These were some of the questions that guided the studies of Paulo Elian dos Santos, who focused on the Functional Genomic and Bioinformatics Laboratory of the Oswaldo Cruz Institute to write his 2008 dissertation entitled Arquivística no laboratório: história, teoria e métodos de uma disciplina¹⁷ [Archivists in the laboratory: history, theory, and methods of a discipline]. Published two years later, this dissertation is based on document research and interviews with laboratory scientists

¹⁷ Defended in October 2008, in the Graduate Degree Program in Social History of the College of Philosophy, Languages and Humanities of the University of São Paulo, under the guidance of Professor Heloísa Liberalli Bellotto.

and reconstitutes the scenario of activities represented in the archives and scientific practices, in order to reveal the processes of creation, use, selection, and storing of documents (Santos, 2010). For Ana Maria Camargo, 18 its principal merit lies in how the author articulates the essential elements of the sociology of science and archival theory.

The studies by Brito (2002) and Santos (2010), which took the Oswaldo Cruz Foundation as an empirical field, benefitted from having a significant portion of biomedical and public health scientists either agree or disagree with their activities, their relationship with the documents created, and their likely value for use in daily activities and in promoting memory.

Expansion of institutional experiences regarding the treatment of scientists' personal archives, together with advances in research and discussions on the matter, was the basis for the works gathered by Maria Celina S. de Mello e Silva and Paulo Elian dos Santos (2012) in the book Arquivos pessoais: história, preservação e memória da ciência [Personal archives: history, preservation, and scientific memory]. Presented during a themed symposium during the 12th National Seminar on the History of Science and Technology, in 2010, the texts examine the diversity of problems faced by institutions that house collections and cover different areas of science. One of the aspects emphasized in the work is the discussions on description and arrangement models and the preference for a functional classification capable of offering the necessary contextualization for the documents.

Without examining all of the reflections on models for classifying and describing personal archives, some of the studies in Brazil began to concentrate on the identification of document categories and types. Knowledge of typologies is crucial for several stages of the treatment of documents in archives, above all for the implications that the relationship between a document description and arrangement has for preservation and access to the archives. For Ana Maria Camargo (2015), identification of typologies is the basis of archival work. These translate the procedures through which the institution is able to achieve its objectives, reflecting routines and adaptations imposed by the passage of time. In this regard, the Glossário de espécies e tipos documentais em arquivos de laboratório [Glossary of document categories and types in laboratory archives], organized by Maria Celina S. de Mello e Silva (2014), has become a methodological reference for those who work with archives and documents produced and

¹⁸ A senior professor at the History Department of the University of São Paulo (USP), she was a member of the dissertation committee and wrote the book's preface.

accumulated; it also serves the function of establishing and sharing standards with scientists and technicians. The following articles by the same author are noteworthy: "Entre espécies e tipos, os documentos têm nome e sobrenome: as funções dos documentos" [Between categories and types, documents have first and last names: the functions of documents] (Silva, 2020), "Estudo de identificação tipológica de diploma em arquivo pessoal" [Study on diploma typology identification in personal archives] (Alves; Silva, 2017) and "Produção documental de cientistas e a história da ciência: estudo tipológico em arquivos pessoais" [Document production by scientists and the history of science: typological study in personal archives]; the latter was in collaboration with Márcia Cristina Duarte Trancoso (Silva; Trancoso, 2015). The studies by Maria Celina Silva on typologies focus mainly on personal archives, which, in their definition, present us with the challenge to define what we consider to be the "roles" of a person. As an alternative, the author deems to be more appropriate the use of the "activities" performed over the course of our lives, whether they be related to work and professions, or in regard to the different aspects of our personal histories, such as our family and social relationships, and individual interests of all kinds (Silva, 2015).

One of the objects of archival science in this area is the laboratory notebook. Used by scientists to record data, methods, procedures, and the stages of their experiments, it is a record frequently found in institutional and personal archives. This finding is explained by the different visions of the scientists themselves, who move between keeping them within the laboratory environment, as an institutional document, and its incorporation into the set of documents of a "personal" nature, generally gathered and protected in a residential environment (Silva, 2007; Santos, 2010).

The panoramic view of archival knowledge about scientific archives and documents presented here is the legitimate basis on which studies and methodologies focused on research data management should be based, as will be demonstrated below.

Open science, archive documents and data: challenges

The complex environment of scientific activity, with the creation of documents, data, and information, requires planning and management actions from institutions and scientists. It is within the complexity of laboratories, which produce digitalized documents and data, that we can visualize the problematic relationship between archives and data and the crisis of archival space as a place of authenticity.

Data has always existed and has been linked to archives. This data may be accumulated in documents, and the latter can be broken down into data, which are growing assets in any information system. What is known as "unstructured" data is data that significantly coincides with our perception of the document. A clear example of "unstructured" data is an email, a communication document, even though the data is organized in it. Laboratory notebooks are documents that contain structured and "unstructured" data. On the other hand, structured data is contained in databases, which are common in scientific laboratories. The new elements to be emphasized are the capacities of the technologies to deal with, analyze, process, and obtain value from large volumes of data.

It is within this environment that we envision an open field of studies to deepen our knowledge of historical and epistemological references about archival science and information science in Brazil, in their different trajectories, and to establish transdisciplinary relationships to understand reality. The different dynamics in the approach to document production require this vision, and we believe the movement of archival science researchers to undertake investigations is noteworthy.¹⁹

As an example, we can use the topic of open science, which entered the scientific world mobilizing not only scientists, but also managers, activists of open knowledge, social movements and a broad array of professionals that deal with the information, documents and data produced by scientific activity. This topic provides us with an opportunity to look for the theoretical, conceptual, and methodological interactions capable of being translated into the correct approach to the records produced, accumulated, and maintained under the care of science professionals and institutions — now guided towards an open perspective of knowledge.

Within the scope of the Fiocruz Política de gestão, compartilhamento e abertura de dados para a pesquisa: princípios e diretrizes²⁰ [Policy on management,

The presence of researchers and university professors linked to archival science in the journals in the "general area" of information science has been increasing. The Brazilian Journal of Information Science: Research Trends examined the relationship between these areas in its volume 11, no. 2, 2017, publishing ten articles by notable Brazilian researchers and professors, and one French author. In the introductory text, Natália Bolfarini Tognoli, the editor of this journal, speaks of a concept of interdisciplinarity supported by the idea of "reciprocity of conceptual, theoretical, and methodological contributions" among areas. She also notes that there are matters that are not fully accepted in the archival community, such as the idea that both disciplines move toward each other in the field of the object, based on movement within the scope of archives, from the document to information (Tognoli, 2017, p. 4).

²⁰ See: Fiocruz (2020).

sharing, and opening of data for research: principles and guidelines], approved in 2020 and the result of a process begun in 2017, the concept arose of a "methodology to identify types of research data," applied in some laboratories and research centers, as a pilot project.²¹

This initiative began from the perspective that research data is produced in institutional contexts and can be used to validate evidence and prove the results obtained; therefore, we note that the archival principles and qualities such as origin, organicity, reliability and authenticity can serve as a safe theoretical-methodological base for effective preservation, qualified availability and (re)use of this data.

Today, in regard to the matter of open research, we have many bibliographical publications in Brazil²² — with connections in Portugal and Spain — predominantly in the fields of library science and information science, which guides reflection and sets the definitions and terms on which a broader theoretical debate, and possibly an empirical research effort focused on the different areas of science, could be based.

There are many concepts, terms and expressions that could benefit from dialogue. The very definition of research data is still not universally agreed upon, and merits examination, as does the associated term "research data management." For some authors in the field of archival science, research data can have its own extensions, classifications, and formats, in addition to the traditional "document typologies and standards" already established. The use of terms that are very familiar to the theory, principles, methods, and practices of archival science seems evident. These information science and library science

²¹ Led by the Casa de Oswaldo Cruz, this initiative intermediates and contributes with actions within the scope of the institution's Archives Management System, and above all, offers the potential to implement procedures to identify aspects related to the authenticity, origin, organicity, and other aspects that support archival management in regard to the integrity and quality of the research data, focused on traceability, reproducibility, and reuse, as well as institutional memory.

²² Among the numerous books, scientific journal dossiers and articles, we wish to highlight the book Ciência aberta, questões abertas [Open science, open questions], organized by Sarita Albagli, Maria Lucia Maciel and Alexandre Hannud Abdo (2015), which contains a significant set of contributions by researchers from different areas and countries, including Brazil, working in and reflecting on the field of open science. Within the Ibero-American context, worthy of note is "Open science: the contribution of information science," in the minutes of the 8th EDICIC Iberian Meeting, coordinated by Maria Manuel Borges and Elias Sanz Casado, University of Coimbra, Center for Interdisciplinary Studies of the 20th Century – CEIS20, 2017. Available at: https://purl.org/sci/atas/edicic2017.

Jorge and Albagli (2018) use, as an example, the definition of scientific data or research data, which according to OECDO (2007) is: "records of facts used as primary sources in scientific investigation, which is generally accepted in the scientific community as necessary for validation of study results." In turn, "research data management" is a general term that covers the organization, structure, storage, and care of the information used or generated during a research project.

studies focus on laboratories, institutional spaces that are home to research, technological development, teaching, reference services, and science communication. Such activities generate a profuse quantity of archival documents, proof, records, and evidence.

Other similarities in the handling of the terms can also be seen. An exemplary case is the affirmation that "data is created to serve a specific purpose" (Jorge; Albagli, 2018, p. 417); an administrative purpose, for example, and become scientific research data when used in that manner. Clearly inspired by the ideas of the American archivist Theodore Schellenberg, administrative purposes remind us of primary value, the "quality inherent in the reasons for creating each document, typical of the initial phases of its life cycle" (Camargo; Bellotto, 1996, p. 78); and their use for other purposes, such as research, which leads us to their secondary value.

From the origin of archival documents and their usage values, we move on to the concept of life cycle or lifetime.²⁴ There are numerous examples. According to Janicy Pereira Rocha (2019), management practices and appropriate curatorship of research data serve to prevent the constant data losses mentioned in the literature. From the perspective of the Yellow Fever Project, undertaken by the Renê Rachou Institute (Fiocruz Minas), the modelling of "research data life cycles" and their "description" carry out the function of guiding and fostering good management and curation practices in laboratories (Rocha, 2019). To better understand this author, it should be noted that research data is recorded in documents on different types of media and different formats, which must be analyzed to determine their possible qualities of an archival nature. Research data is found in laboratory notebooks and databases, among many other document types, which indicates the need for work to describe its archival elements. At the same time, the life cycle of records²⁵ is a concept that was created in the 1940s in the United States, which had significant impacts on the later structuring of the U.S. record management system, and a strong influence in different countries, such as Brazil.

One of the archival documents that has received more attention from information science scholars is the laboratory notebook, especially in its open format. In an article that analyzes the benefits and disadvantages of using electronic

²⁴ To understand the origin and use of the concept of the life cycle of records, see: Jardim (2015).

²⁵ The expression "life cycle of records," according to the Dicionário de terminologia arquivística [Dictionary of archival terminology] refers to the "succession of phases that documents go through, from the moment they are created until their final destination" (Camargo; Bellotto, 1996, p. 15).

laboratory notebooks, Rocha, Sales and Sayão (2017) defend electronic notebooks as the format that would ensure the best management of data and its sharing. They defend the role of professional librarians as mediators in preservation of and access to this data and as curators of the information produced. Here, we identified the use of the terms preservation and access, which are widely used by different disciplines that deal with the treatment of collections, and the introduction of a new concept, the "curating" of information or research data, a term that is not common among these same disciplines (Jorge; Albagli, 2018). The definition of "digital curatorship" is the practice and study of the processes of selection, preservation, maintenance, collection and filing of digital data.²⁶

One of the principal approaches to notebooks is provided by Anne Clinio (2015). According to this author, the adoption of a digital laboratory notebook means more than just the migration of document information on paper to software, given the possibilities of interaction among the agents producing scientific knowledge that would create agility in the incorporation of data, dissemination of information and production of copies (including security copies and backups). The justification for adoption of open notebook science is to provide the researcher's raw data, recorded in the electronic notebook on the network, to any person who is interested in analyzing, interpreting, and using it (Clinio, 2015).

Notebooks are documents with the characteristics of an archive document and are not exclusive to the exact sciences and life sciences. Notwithstanding some differences, researchers in the humanities also use something similar, which are field notebooks. In anthropology, for example, field notebooks are the primary analysis tool, in which ethnographers write down not only their personal impressions, in the form of a personal diary, but principally onsite observations, such as the scenario and the actors involved in the investigation.

After a decade of studies on the laboratory archives of the Oswaldo Cruz Institute/Fiocruz, our diagnosis indicates a complex scenario in regard to the practices of document production, conservation, and management, including notebooks, which are given different names; in some cases, they refer to different activities, and are a permanent focus for work performed at the workbench. In certain laboratories, for example, they materialize in two stages, forming what the e-Arq Brasil requirement model defines as a "hybrid dossier," composed of digital and non-digital documents (Santos, 2020). At most laboratories, both non-digital (predominant) and digital notebooks can be found.

²⁶ See: Sayão; Sales (2012); Araújo (2018).

Concerned with research data produced by clinical trials within the context of open science, Pinto, Amaral and Santos (2019, p. 84) emphasize the need to implement research data management policies that ensure "criteria for the description, systematization, sharing, recovery, interoperability, preservation and reuse of data." Here there is a recurrence of the use of archival terms, such as the description, used for the treatment of documents in permanent archives.

In summary, we are not discussing distant disciplinary relationships about which we should not insist; rather, we must place them within a broader perspective, capable of favoring their theoretical, methodological, and practical knowledge, through academic reflection and projects focusing on dealing with real-world problems. Today, there is a reasonable consensus among the different fields of thought regarding the changeability of archival principles and concepts, which have been created at different times since the end of the 19th century, and which are therefore subject to reinterpretation by future generations. The trend towards the scientific maturation of archival science in Brazil may, in fact, contribute to dialogue with information science, provided that archivists pay attention to transformations in the broad organizational and information universe, without abandoning the essence of archival work.

Final considerations

In the past two decades, research in archival science on the topic of "science archives" has been increasing in Brazil, supported by the greater presence of document management programs at universities and the continued emphasis on memory institutions that care for these collections, especially the personal files of scientists. The possibility of expanding studies and methodological solutions applied to collections in the different areas of knowledge, particularly human and social sciences, seems to be essential for us to recognize the different ways in which their reference materials or evidence are produced and conserved, for example. Their respective traditions and rationales could help us understand the records we identify in these collections.

Alongside these trends, there is an agenda in the contemporary scientific world that is interconnected with this knowledge generated by studies in the archives and information fields. In this aspect, initiatives related to the management, sharing and opening of research data must not overlook this knowledge, since it represents a unique opportunity for innovation, generated through interaction, exchange, and reflection on the theoretical, conceptual,

and methodological components of archival science, of library science and of information science.

Laboratories, where science is produced, are hybrid environments, with a high degree of document complexity and diversity, which cannot be controlled by a specific field of knowledge or submitted to definitions that do not suit this complex, diverse reality. Laboratories are not subject to a single disciplinary viewpoint that reduces them to a type of organizational structure that has been surpassed by the scientific arrangements of the 21st century. We can turn to the idea of transdisciplinarity to deal with "transdisciplinary objects," in order to combine methodologies from different disciplines, in order to create new knowledge.

Many thanks to Naomi J. Sutcliffe de Moraes, who translated this article from Portuguese.

References

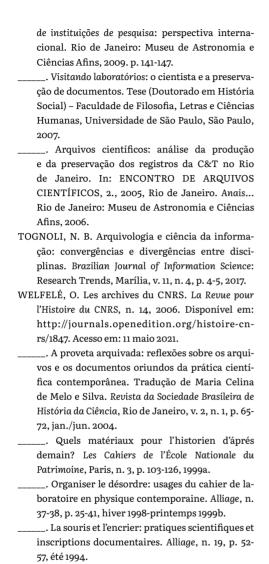
- ALBAGLI, S.; MACIEL, M. L.; ABDO, A. H. (org.). Ciência aberta, questões abertas. Brasília: Ibict; Rio de Janeiro: Unirio. 2015.
- ALVES, A. F.; SILVA, M. C. S. M. E. Estudo de identificação tipológica de diploma em arquivo pessoal. Revista do Arquivo, v. 4, p. 205-215, 2017.
- ANDRADE, A. M. R. (org.). Caminho para as estrelas: reflexões em um museu. Rio de Janeiro: Museu de Astronomia e Ciências Afins, 2007.
- ARAÚJO, C. A. Á. Um mapa da ciência da informação: história, subáreas e paradigmas. ConCI: Convergências em Ciência da Informação, Brasília, v. 1, n. 1, p. 47-72, 2018.
- BOURDIEU, Pierre. O poder simbólico. Tradução de Fernando Tomaz (português de Portugal). 10. ed. Rio de Janeiro: Bertrand Brasil, 2007.
- BRITO, Verônica Martins de. A preservação da memória científica da Fiocruz: a visão de quem faz ciência. Dissertação (Mestrado em Ciência da Informação) – Instituto Brasileiro de Informação em Ciência e Tecnologia/Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2002.
- CAMARGO, A. M. A. Sobre espécies e tipos documentais. In: _____ et al. Dar nome aos documentos: da teoria à prática. Apresentação de Danielle Ardaillon. São Paulo: Instituto Fernando Henrique Cardoso, 2015. p. 14-30.

- ; BELLOTTO, H. L. Dicionário de terminologia arquivística. São Paulo: Associação dos Arquivistas Brasileiros, Núcleo Regional de São Paulo, Secretaria de Estado de Cultura, 1996.
- CHARMASSON, T. Les archives des sciences en France: références bibliographiques utiles, fév. 2007. Disponível em: http://www.cn2sv.cnrs.fr/IMG/pdf/bibliographie-sur-les-archives-scientifiques.pdf. Acesso em: 20 maio 2021.
- Archives institutionelles et archives personelles. Les Cahiers de l'École Nationale du Patrimoine, Paris, n. 3, p. 13-23, 1999.
- _____ et al. Les archives personelles des scientifiques: classement et conservation. Paris: Archives Nationales, 1995.
- CLINIO, Anne. Por que open notebook science? Uma aproximação às ideias de Jean-Claude Bradley. In: ALBAGLI, Sarita; MACIEL, Maria Lucia; ABDO, Alexandre Hannud (org.). Ciência aberta, questões abertas. Brasília: Ibict; Rio de Janeiro: Unirio, 2015.
- COOK, Terry. O passado é prólogo. Uma história das ideias arquivísticas desde 1898 e a futura mudança de paradigma. In: HEYMANN, Luciana; NEDEL, Letícia (org.). Pensar os arquivos: uma antologia. Rio de Janeiro: FGV, 2018.
- DERRIDA, Jacques. La grammaticologie. Paris: Minuit, 1967.

- FUNDAÇÃO OSWALDO CRUZ (Fiocruz). Presidência.

 Vice-Presidência de Educação, Informação e
 Comunicação. Política de gestão, compartilhamento e abertura de dados para pesquisa: princípios e
 diretrizes. Rio de Janeiro: Fiocruz, 2020.
- HAAS, J. K.; SAMUELS, H. S.; SIMMONS, B. T. Appraising the records of modern science and technology: a guide. Massachusetts: Institute of Technology, 1985.
- JARDIM, J. M. Caminhos e perspectivas da gestão de documentos em cenários de transformações. Acervo, Rio de Janeiro, v. 28, n. 2, p. 19-50, jul./ dez. 2015.
- JORGE, V. A.; ALBAGLI, S. Compartilhamento de dados de pesquisa em saúde: iniciativas do National Institutes of Health (NIH). Reciis: Revista Eletrônica de Comunicação, Informação e Inovação em Saúde, Rio de Janeiro, v. 12, n. 4, p. 415-428, out./dez. 2018.
- LATOUR, B. Ciência em ação: como seguir cientistas e engenheiros sociedade afora. Tradução de Ivone C. Benedetti. Revisão de tradução de Jesus de Paula Assis. São Paulo: Unesp, 2000.
- ; WOOLGAR, S. A vida de laboratório: a produção dos fatos científicos. Rio de Janeiro: Relume Dumará, 1997.
- _____. Joliot: a história e a física misturadas. In: SERRES, M. (org.). Elementos para uma história das ciências. Lisboa: Terramar, 1989. p. 131-155.
- ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). Principles and guidelines for access to research data from public funding. Paris, 2007.
- PESTRE, Dominique. Por uma nova história social e cultural das ciências: novas definições, novos objetos, novas abordagens. Cadernos IG Unicamp, v. 6, n. 1, p. 3-56, 1996.
- PINTO, F. M. A. G.; AMARAL, J. C.; SANTOS, M. B. Curadoria de dados de pesquisa em repositórios de ensaios clínicos: uma revisão de escopo. Linc em Revista, Rio de Janeiro, v. 15, n. 2, p. 84-100, nov. 2019.
- ROCHA, J. A. P. Projeto febre amarela: ciclo de vida e tipologia de dados. Liinc em Revista, Rio de Janeiro, v. 15, n. 2, p. 161-176, nov. 2019.
- ROCHA, L. L.; SALES, L. F.; SAYÃO, L. F. Uso de cadernos de laboratório para as práticas de ciência aberta e preservação de dados de pesquisa. Ponto de Acesso, Salvador, v. 11, n. 3, p. 2-16, dez. 2017.
- SAMUELS, H. W. Appraising the records of

- modern science and technology. Janus: Revue Archivistique, Dordrecht, n. 2, p. 8-19, 1995.
- SANTOS, P. R. E. Documentos de archivo producidos por la actividad científica: un análisis de los cuadernos de laboratorio del Instituto Oswaldo Cruz. Comma: International Journal on Archives, v. 2.018, p. 127-138, 2020.
- . Arquivos de cientistas: gênese documental e procedimentos de organização. São Paulo: Associação de Arquivistas de São Paulo, 2012.
- ; PINTO, J. M. C.; SANTOS, C. B. dos. Manual para gestão de documentos e arquivos de laboratórios das ciências biomédicas. Rio de Janeiro: Casa de Oswaldo Cruz; Faperi, 2011. CD-ROM.
- _____. Arquivística no laboratório: história, teoria e métodos de uma disciplina. Rio de Janeiro: Teatral; Faperi, 2010.
- SAYÃO, L. F.; SALES, L. F. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. Informação & Sociedade: Estudos, João Pessoa, v. 22, n. 3, p. 179-191, set./dez. 2012.
- SILVA, M. C. S. M. Entre espécies e tipos, os documentos têm nome e sobrenome: as funções dos documentos. Anuário do Museu Imperial, v. 1, p. 77-90,
- . Os arquivos pessoais como fonte: reconhecendo os tipos documentais. In: GRATANO, Marcus (org.). Mast: 30 anos museologia e patrimônio. Rio de Janeiro: Museu de Astronomia e Ciências Afins, 2015. v. 1. p. 178-203.
- ; TRANCOSO, M. C. D. Produção documental de cientistas e a história da ciência: estudo tipológico em arquivos pessoais. História, Ciências, Saúde Manguinhos, v. 22, p. 849-861, 2015.
- _____. Glossário de espécies e tipos documentais em arquivos de laboratórios. Rio de Janeiro: Museu de Astronomia e Ciências Afins, 2014.
- ; SANTOS, P. R. E. (org.). Arquivos pessoais: história, preservação e memória da ciência. Rio de Janeiro: Associação dos Arquivistas Brasileiros; Faperj, 2012.
- ; REGO, V. L. A. Guia básico para preservação de arquivos de laboratórios. Rio de Janeiro: Museu de Astronomia e Ciências Afins, 2010.



Recebido em 31/1/2021 Aprovado em 1/9/2021