From Handwriting to Footprinting

Text and Heritage in the Age of Climate Crisis

Anne Baillot

https://www.openbookpublishers.com

Anne Baillot, *From Handwriting to Footprinting: Text and Heritage in the Age of Climate Crisis.* Cambridge, UK: Open Book Publishers, 2023, https://doi.org/10.11647/OBP.0355

ISBN Paperback: 978-1-80511-087-3
ISBN Hardback: 978-1-80511-088-0
ISBN Digital (PDF): 978-1-80511-089-7
DOI: 10.11647/OBP.0355

Cover image:
Yarn texture by Tim Mossholder (2023), https://unsplash.com/photos/AQjSAPNLjGI

Cover design: Jeevanjot Kaur Nagpal

https://www.openbookpublishers.com

ISBN Paperback: 978-1-80511-087-3
ISBN Hardback: 978-1-80511-088-0
ISBN Digital (PDF): 978-1-80511-089-7
DOI: 10.11647/OBP.0355

Cover image:
Yarn texture by Tim Mossholder (2023), https://unsplash.com/photos/AQjSAPNLjGI

Cover design: Jeevanjot Kaur Nagpal

Contents

Introduction **1**

1 Archiving text **5**

1.1 What archives do 7

1.1.1 Institutionalisation of archiving 8

1.1.2 Digital archiving 22

1.2 From writing to reading to archiving: becoming text . . . 32

1.2.1 Temporality and text as media: the power of genesis 33

1.2.2 Traces of the past in the present 44

2 Publishing, editing, and their digital transformation **55**

2.1 From text to book to œuvre 56

2.1.1 The deal between writer and publisher 58

2.1.2 Negotiating the œuvre 68

2.2 Text quality from scan to digital edition 78

2.2.1 Emancipation from the page? 79

2.2.2 Access made Open 95

3 What the climate crisis does to text **111**

3.1 The environmental cost of access to text 113

3.1.1 Assessing the environmental footprint of text . . . 114

3.1.2 Archiving text for tomorrow 125

3.2 Trying to make this book an environmental lightweight . 134

3.2.1 Writing, printing 135

3.2.2 Distributing, archiving — and the readers 142

Conclusion **151**

Index **154**

Bibliography **155**

Introduction

In his famous lecture entitled *Encyklopädie und Methodologie der philologischen Wissenschaften*, classical philologist August Boeckh frames achievements of the philological sciences as "die Erkenntnis des Erkannten", the knowledge of that which is known.[1] This definition of text-based scholarship is not as tautological as it may seem. It strives to embrace fundamental self-reflective efforts. Philologists — those who love texts — supply and interpret texts while they question the way in which they approach texts. In the same way, digital philology provides and interprets texts and reflects on the hermeneutical principles it follows. It also interrogates how digital media impacts textual transmission.

While text-based activities have remained essentially the same since Boeckh's time, including archiving, editing, commenting, publishing, and critiquing, the media shifts at play since the 20th century have opened new horizons and provided opportunities to engage productively with older cultural textual practices.

The advent of digital tools and resources has changed my own scholarly endeavours in a radical way. From a solitary and oftentimes repetitive activity, whose output would concern only a handful of the like-minded, it became a collaborative endeavour in which mechanical tasks were conducted by a machine, an undertaking in which the whole world could potentially take part. It loosened the boundaries of scholarship and made them more permeable to adjacent fields, in particular cultural heritage.

While I have learned a lot both from this development of textual studies through digital media, and from the manner in which it tightens the connections between philology and heritage institutions such as archives and libraries, this collaborative perspective is, to this day, still by no means predominant. Admittedly, it requires one to let go of established forms

1 Only one edition of the *Encyklopädie* manuscript currently exists. It was procured by Boeckh disciple Ernst Bratuschek in 1877 and presents all the available textual elements as one text, while the manuscript actually consists in over 20 layers corresponding to the updates that Boeckh made to his lecture script over time. Procuring a dynamic digital edition of this manuscript remains a desideratum. See Bratuschek, *Encyklopädie* [26], Horstmann, *Erkenntnis des Erkannten* [70], Baillot et al., *Neue Perspektiven* [15].

 https://doi.org/10.11647/OBP.0355.05

of authority, and with them, of power. Nonetheless, I am convinced that this approach is the only one that would make it possible to stay true to Boeckh's principle and, more generally, to root the social relevance of philology in the current context, which is one of digitisation, but also of the technical and cultural divide between the Global North and South, and of the world-wide climatic threat. Textual scholars and digital humanists will easily identify references to well-known theories in some of the following pages. Heritage professionals will also find attempts to popularise some of their advances for a non-professional audience. My writing is rooted in my personal experience as a scholar; it also reflects on a variety of theoretical and technological advances that have yet to benefit a wide user and reader community. It is my goal to address this community.

In this book, I discuss mechanisms that facilitate access to text based on my own scholarly experience of the last twenty-five years. I do so with a focus on early modern literary texts as an indicative part of cultural heritage. I present three core cultural and scholarly text-related practices — archiving, editing, and publishing processes — to frame them in a digital context, and ultimately shed light on the environmental impact of today's preservation and dissemination options for textual material. My goal is to show how theoretical and practical leverage can be gained from existing approaches in order to tackle the ecological and, with that, economical and social challenges we now face. I examine ways of preserving text and disseminating it to a wider audience from a historical, digital, and environmental perspective, and aim to foster reflections on today's archiving and publishing practices in terms of their sustainability in the context of the climate crisis.

Digital infrastructures have had a transformative influence on archiving and publishing ecosystems. Forms of media shifted, technical constraints appeared, costs rose, and reputation mechanisms took on new dimensions. I will argue, however, that modes of operation have not changed radically with the media transformations introduced during the 20th century. But, more recently, digital media have managed to move some philological lines, and the climate crisis calls for an even more enormous reassessment of practices of textuality. In the first two parts of the book, I move from general historical and theoretical remarks on archiving in Chapter 1 and publishing in Chapter 2 to their digitisation.

In the final part, Chapter 3, I explore the impact of the climate crisis on current archival and editorial practices.

Chapter 1 is dedicated to archiving strategies and their significance for our relationship to text and to the value of text. After a general presentation of archiving processes and the institutionalisation of archives in western societies, I go on to highlight the specificities of digital archiving. Trying to assess continuity lines and rupture points between traditional (analog) archiving and digital archiving, the argument focuses in this section on the materiality of both processes. Recording strategies are then presented as major leverages. The last section in Chapter 1 reflects on a more theoretical level about the contradictions inherent in text archiving as a balance between preservation and destruction of material.

Chapter 2 addresses publishing practices and strengthens the focus on literary texts. I present early modern publication processes as transformative events, and assess the impact of digitisation. The first section sheds light on the process of transforming a text into a book and ultimately part of an œuvre through edition and publication, examining authorship issues along the way. The relationships between writers and publishers play a key role here. I present a case study on two major German writers of the first half of the 19th century, Goethe and Tieck, in order to give a better sense of the pragmatics of authorship negotiation. The second section presents a panorama of current digitisation processes (scanning, automated text recognition, annotation, visualisation) and the way in which they multiply the forms of text representation available to readers. The closing section reflects on the implementation and social significance of Open Access both for archiving and for publishing.

Chapter 3 addresses the environmental footprint of digital archiving and digital publication. It suggests new approaches to the question of access to text in a context of digital sobriety, being frugal of one's use of diverse technologies. In the first section, I frame the overall environmental footprint of access to text in a digital context, underlining the difficulty of measuring it precisely and making informed decisions throughout the process of archiving or publishing. From a more speculative perspective, I then envision what sustainable access to text could look like in the future. The last section focuses more specifically on the production, dissemination, and storing of the book you are currently reading. It aims to identify the type of leverage that is necessary to offer a transformative, less resource-intensive approach to (digital) text.

I conceived these analyses as a journey through text. It begins in the following pages with old family papers found in a drawer and explores the different ways of preserving them and making them available for more readers. I argue that every reader is key in the transmission process: from the moment you are given access, you become one who can also give access. In this process, a new balance is established between the growth of digital transmission options, the almost magical immensity of content they provide, and their impact on the physical integrity of our planet. My argument is a material one: it is one of love for texts, especially literary texts, as one of the richest means of human expression — but one that pleads for a better embedding of textual activity in the complex materiality of our world.

1. Archiving text

Text presents itself to us in everyday life in a variety of forms, to the point that it has become key to more or less all social, political, economical, and cultural transactions in north-western cultures. Text is embedded in the immediacy of our environment. Yet not all textual material that surrounds us is a product of immediacy. I would like to begin here by looking at an approach that connects us to the past of textuality. What happens when you consider an old text?

Old papers have a fascinating quality that unfolds when you come into close contact with them. They can be kept in bundles or simply present themselves as loose sheets. Some might seem easy to decipher, while others are virtually unreadable. Sometimes they are composed of neatly numbered pages, sometimes of incoherent fragments. None of that makes much difference. What matters is that, most of the time, something happens to those whose path crosses old papers, whether the encounter happens on purpose or by chance.

The way in which old papers present themselves, their materiality, creates a distance. The paper differs from what one is used to flipping through in everyday life; it has aged. The shape of the letters or ciphers is also different, as are the ink and pen that are used. Nevertheless, it is a product of paper and writing and in that sense it is familiar. This mixture of familiarity and distance is captivating: it feels as though it would be possible to unlock the mysteries old papers contain — and at the same time it is clear that this can never be fully achieved.

Old papers transport you back in time while simultaneously anchoring you in the reality of your present time and space. Even if you manage to decipher their content, understanding fully what they are about will require much more effort than simply reconstituting the wording. After transcribing the letters and forming words from them, you need to understand where the papers came from, who the writers were, what their intention might have been, who or what they were alluding to, but also, why these specific papers have been preserved through time and finally landed into your hands.

 https://doi.org/10.11647/OBP.0355.01

Preservation is far from a general rule when it comes to inheriting the past. Old papers rather tend to disappear or disintegrate as time goes by. Paper technology is remarkable, but it resists water and fire rather poorly. Sometimes even air can induce material decay, especially if paper and air react chemically with aggressive ink types. And it is not only physical elements that accelerate decomposition: rodents too have a fair share of responsibility in the poor preservation of old papers.

And yet, when you find them, there they are, here and now. There is something compelling in the presence of old papers, as if they were a gift of time past. The specific piece of paper you are holding has resisted severe weather, wars, mice, neglect and mould, and it has done so for decades or centuries, as if it were all just to reach you, today. You might be afraid that you will damage it. The paper is probably becoming tattered, especially on the edges, and might be breaking up into smaller pieces where it was folded. But, for all its frailty, it has proved its extreme resistance simply by still being there.

While some old papers — family papers randomly found in a drawer, for instance — might not seem to obey a preservation logic destined to lead them to you intentionally, for a very long time there have been institutionalised forms of preservation, dedicated to a range of documents. Intentional preservation, with the goal of making items available at a later time, is in fact an accurate definition of what archives are and what they do.

Archives are, by and large, organised old papers. An archive consists of a coherent ensemble of old papers, as in "The Archive of the Paris Police". The word "archives" also designates institutions and/or the buildings hosting them, which take care of such organised old papers. The singular "archive" or its plural can be used to designate any one of them: the collection of papers, the institution, or the place where the institution is. The function of archives is to store, record, and present old papers from various times and places, and with various subjects. Whilst the content of an archive is not always restricted to handwritten papers, these generally make up the core of their stock.

The existence of early archives (and I will get to a historical overview of archival processes in section 1.1) suggests that archival intentions have existed in human cultures for quite some time. For centuries, there have been reasons for mankind to make an effort to preserve written traces of their political, economical, legal and cultural achievements. Since archiving had a purpose, and presumably followed a strategy, anyone handling old papers today can legitimately wonder why they had been archived and, more specifically, why one is, in this particular place and at this particular moment, holding these remnants of time past. Getting in close contact with old papers also involves acknowledging your responsibility to pursue the transmission momentum beyond your own personal experience — that special moment when you come to see, touch, unpack, unfold, present to the light and hold old papers.

Preservation does not generally happen from one day to the next. Archiving is a process. It is the process by which, on the one hand, contemporary material is turned into archival content, being stored, recorded, and made available for consultation in an institutionalised manner. But, on the other hand, it reflects the ongoing effort that is necessary to keep old papers remembered: transmission is never achieved once and for all, but needs to be actively pursued in order for the archive to continue to exist. In that sense, anyone consulting archived material is part of the archiving process. At the exact moment when you are consulting the archive, you become a recipient and a transmitter. Archiving is not exclusively the mission of institutions like archives: it engages each member of a society that is concerned with the past and its memory.

To begin this journey to and through text, this chapter will examine archiving as an old cultural technique — one that is strongly rooted in north-western societies. As a first step, I will present a historical approach to archiving techniques and go from there to the specificities of digital archiving, in order to highlight differences but, in particular, structural similarities between physical and digital archives. The second section is more theoretical: I propose a conceptualisation of approaches to archival material as a fleeting trace of things past. This chapter aims to show how past texts are embedded in our present lives, both materially and symbolically, and how we can draw from this to make our present become tomorrow's lively past.

1.1 What archives do

Archives and archiving relate to the material world: they concern the documents you find in a family home, the postcard collection you might purchase at a flea market, and the notary acts recorded throughout the history of your city, as well as medieval copies of illuminations. This meaning of "archive" is the primary reference of the word, as applied to archives as institutions, and archiving as a process related to this way of collecting and recording documents. The same words have a slightly different meaning in the digital context. I will first provide some historical background on the institutionalisation of archiving processes in the analog world, and then move on to the specific challenges of digital archiving.

1.1.1 Institutionalisation of archiving

The word archive itself comes from the Ancient Greek word αρχη, which generations of Hellenists have learned to identify as "the word that means either the beginning or the commandment". The point is that it actually means both, relating to something like a seminal order of things.[1]

It is quite clear how the word "archive" relates to the second meaning of the Greek word, that of commandment: archiving provides order. For each archive, there is a rule, or, more accurately, a set of rules, that is to be respected throughout the preparation of the material in order to achieve a satisfactory archiving process — one that will make it possible to preserve, stock, record, and make available its contents for a later consultation.

Recording strategies only work if the same logic is used to structure all the material that belongs together: this is the requirement that enables the organisation and listing of the documents that make up the archive in a manner that will make it possible to eventually find them again when one looks for them. Alphabetical order, for instance, is a simple way of arranging material. But you first have to decide what the alphabetical order applies to: the author's name (then what about unknown authors?), the document's title (provided there is one), or maybe simply the beginning of the document's text (easier to identify if the text is linear and running). And then you have to decide how to organise the material in the space

1 Derrida opens his small opus *Mal d'archive* with exactly these etymological considerations ("Ne commençons pas au commencement, ni même à l'archive. Mais au mot "archive" — et par l'archive d'un mot si familier. *Arkhè*, rappelons-nous, nomme à la fois le *commencement* et le *commandement*. Ce nom coordonne apparemment deux principes en un: le principe selon la nature ou l'histoire, *là où* les choses *commencent* — principe physique, historique ou ontologique —, mais aussi le principe selon la loi, *là où* des hommes et des dieux *commandent*.") [42], p. 11.

at your disposal: will you, for instance, fill first all the upper shelves on one wall, then the row below, or use separators such as the supports between shelves to switch level to proceed with the rest of the alphabet? Also, you might want to consider the fact that your collection could grow at a later point, which would require a shelving system that allows for enlargement.

Obviously, even for a very generic archiving task, there are many ways of organising documentary material, and what may seem intuitive to one person might not be self-explanatory at all to another. Organising material requires an explicit set of rules, decisions on a variety of sorting steps, and clear, accessible information about these decisions and rules.[2]

No wonder then that archives have contributed to the development of standardised recording methods across time. General archiving principles make it easier to consult different archives. Archival logic follows similar, general rules, regulated and reflected on by what has become a field of knowledge of its own: archival sciences as a branch of information sciences.[3]

Even if you have only ever had to deal with a small set of documents, you are certainly aware that a lack of systematic organisation inevitably leads to much time-wasting looking for information or documents you know you have but cannot remember where they are. Optimising organising and recording strategies is at the core of archiving processes, as archives, by definition, have to deal with large, and generally growing, amounts of material.

In Ancient Athens, the creation of the first institutional archive, the Metrôon, was initiated in order for citizens to be able to consult any law that was passed. This is the first known systematic, institutionalised preservation and accessibility device, including a dedicated building conceived to house one copy of each law that was ever passed.[4] Laws were

---

2 See Petra Gehring, "Archivprobleme", in *Handbuch Archiv* [58], pp. 17-18.

3 Anne Gilliland provides a rich overview of the history of archival sciences in Chapter 1 (entitled "Archival and Recordkeeping Traditions in the Multiverse and their Importance for Researching Situations and Situating Research") of *Archival Multiverse* [61], pp. 31-73; see especially p. 43 for a definition.

4 This does not mean that there were no other archival endeavours before the creation of the Athenian Metrôon. In the volume edited by Marie Brosius (*Ancient Archives and Archival Traditions*), the focus on the Athenian model of the 5th century B.C. is limited to one contribution, while many chapters examine other contexts. The contribution dealing with Greek archives by John K. Davies suggests that the mass of documents

transcribed on stone, not paper, and only full citizens were authorised to consult them. But it was already structured along the lines of what archives would be in the following centuries in European societies: a building dedicated to the systematic preservation of information pertaining to public matters, and a place where they could be consulted. In the case of the Metrôon, the information was legal in nature, and it could serve political purposes. Being able to refer precisely to pre-existing laws improved transparency in political and judicial matters. Instead of giving only an approximative account of a law, one could retrieve the actual wording.

The Metrôon building was conceived for the preservation of documents, which means that the format and storage conditions, as well as the fact that the stock would grow over time, were accounted for from the onset. Also, the preserved material did not consist of cultural artefacts, but documents, albeit carved on stone, relevant to the political well-being of the state, administrative material providing the history of the judicial branch of government.

These archives had a double function. First, they guaranteed the veracity of decisions taken in the past. People who needed to refer to these laws could do so confident of their accuracy, so that any suspicions about errors in the cited wording could be fact-checked. It was also a way to supply a precise remembrance of the known laws. People who might otherwise overlook them, not necessarily for manipulative reasons, but simply because their memory could not encompass all of the information inherited from previous generations (or even their own), would need some kind of memory extension that could provide access to this wide collection of all legal decisions made in the past.

Archives act as an ancillary service to human memory for the benefit of society at large. Archiving documentation connected to genocides like the Shoah, for instance, is strongly motivated by the wish for such an

archived in Athens was quite unique: "Argument has focused mainly on late fifth-century Athens. On the one hand, her headlong development, and her management of an Aegean empire, generated far more public documents than were ever cut in stone. [...] Nor was the practice purely Athenian. [...] Though no precise 'solution' to the problem is available, the general direction in Athens is clear. By the end of the fourth century BC at latest there was a reasonably well-organised public archive, located in the precinct of the Mother of the Gods, the Metrôon, wherein documents were lodged and could be found." (See [30], pp. 328–329). The main focus of Davies' analysis is the complex question of public access to the collection.

event never to happen again: archives provide details that people might end up forgetting, risking the denial of the event and its significance. World War Two is in many ways a turning point in public strategies of memory building in western societies, based on the theoretical work of the Frankfurt school of sociology, among others.[5]

More recent examples spring to mind, illustrating the relevance of the issue. During the invasion of Ukraine by Putin's Russia in the spring of 2022, alongside international manifestations of solidarity with Ukraine, there has been a noticeable effort worldwide to support the preservation of Ukrainian heritage.[6] The awareness that not only cultural, but also economical and political issues are at stake when cultural heritage is destroyed, has risen to the point that it is clear that its destruction would facilitate the negation of Ukraine's political existence as a nation. This sensibility is aligned with theoretical approaches that have developed since the 1970s, which shed light on issues that were previously chiefly the concern of archivists.

In his small book dedicated to archiving processes, *Mal d'archive*, French philosopher Jacques Derrida combines the notion that an archive would be an additional memory, or an extension to human memory, with the idea that it is the archiving process itself that generates the relevance of that which is archived. He presents archive as a memory supply ("prothèse ou technique hypomnésique") and as the generator of the event it records through the archival process ("archive archivante", "produit autant qu'elle enregistre"), but also the need for an infrastructure ("structure technique") backing both of these functions. Derrida postulates that archiving produces and records events: in his view, recording is that which produces the archive as such.[7]

5 See Horkheimer and Adorno, *Dialektik der Aufklärung* [69].

6 See, for instance, the SUCHO initiative for online content. Other initiatives tried to provide support for digitising what could be digitised, and protecting what could be protected. The symbolic and economic importance of cultural heritage preservation has been reported on, as far as this was possible, by journalists; see, for instance `https://www.nytimes.com/2022/04/30/world/europe/ukraine-scythia-gold-museum-russia.html?smid=tw-nytimes&smtyp=cur`.

7 See Derrida, *Mal d'archive* [42], p. 34: "Autre façon de dire que l'archive, comme impression, écriture, prothèse ou technique hypomnésique en général, ce n'est pas seulement le lieu de stockage et de conservation d'un contenu archivable *passé* qui existerait de toute façon, tel que, sans elle, on croit encore qu'il fut ou qu'il aura été. Non, la structure technique de l'archive *archivante* détermine aussi la structure du contenu *archivable* dans son surgissement même et dans son rapport à l'avenir. L'archivation produit autant qu'elle enregistre l'événement."

Derrida also provides a definition of recording ("consignation") that can be read as the affirmation of an identity that lies between recording and archiving.[8] Archival recording assembles elements according to one system, which thus generates unity. In these terms, the core principle of the archive is its system of recording.

Derrida's analysis reads like an abstract characterisation of archiving as a way of putting boxes and catalogue labels together and considering that they then constitute one unit. The specific space for assembling the archive is essential. As Derrida underlines: recording is a concrete, material process without which archives would not even exist.[9]

But it could also be read the other way round. If recording has an essential, transformative function in the archiving process, then maybe recording is sufficient to transform anything into an archive. Not so. You can find this out the hard way, as I did. When I was a young adult, I found some old family papers dating back to the late 19th century, and wanted to give them to the relevant local archives, which kept administrative papers recording regional history. What I brought were school papers and ledgers, but this was not in itself significant: I was sent back home with all of them altogether, because such material was only considered worth archiving if it was at least 150 years old. I have put the papers back in the cellar, where they still are to this day, waiting to be over 150 years old and deemed worth recording and professional archiving.

Obviously, if everything was to be considered valuable enough to be archived, the existing infrastructure dedicated to archiving would have to be much larger than it is. There would be whole buildings devoted to this: I imagine them to be an underground replica of actual cities, each one of them having a sort of "below the water part of the iceberg" underground mega-archive recording each city's story so far (a fantasy that probably unfolded in my imagination under the impression of Freud's depiction of Rome I mention at the end of section 1.2). In real life, and regardless of my personal dream of a mega-subterranean archive, archives have to

8 See Derrida, *Mal d'archive* [42], p. 14: "La consignation tend à coordonner un seul corpus, en un système ou une synchronie dans laquelle tous les éléments articulent l'unité d'une configuration idéale. [...] Le principe archontique de l'archive est aussi un principe de consignation."

9 The development of recordkeeping systematics has been central in the development of archival sciences since the 18th century, as accurately described by Gilliland in *Archival Multiverse* [61], especially pp. 38-39.

define priorities in order to achieve their missions. This can be stating that documents have to be at least 150 years old to be integrated to the archives, as was the case in my youthful experience; or deciding to collect only documents related to a specific topic, such as the legal focus they had in Athens; or being strictly dedicated to a particular political entity, such as a parliament, for instance.[10] This choice of a scope to be determined for each archival institution, which defines the extent of what this institution will preserve, record, and make available for consultation, is embraced in the concept of appraisal, championed by archivist Theodore Schellenberg in the United States in the 1950s.[11] From his perspective, the archivist's role is not simply to record documents, but also to decide which documents are to be kept and recorded.

Archives as institutions conceived to keep traces of the history of a state have developed their own purposes and practices alongside other cultural heritage institutions. While they sometimes converge in techniques and purposes, there are key differences between archives and libraries or museums. Archives preserve written documents that are, by definition, unique (or of which only few or, more likely, no other copies exist): in this, they are different from both museums and libraries. Museums are dedicated to artefacts and plastic works of art, while libraries provide single exemplars of books that were originally printed in several, oftentimes hundreds of copies, meaning that one would be likely to find the same book in another library. By definition, there is no automated reproduction of the content of an archive as opposed to a library, and the content consists of written material, in general handwritten, presented on paper as opposed to a museum. There are museums that display archival material as well as archives that preserve printed content. But there is a material difference that characterises the basic mission of each institution.

Another specific characteristic of archival structure concerns the manner in which unique documents are recorded. While alphabetical order and shelf structure, which I mentioned before, play a role in the organisation of boxes and their labels, modern archives have one central structur-

10 Similarly, the European Union provides an online archive of parliamentary minutes in the different official languages of the EU, see `https://historicalarchives.europarl.europa.eu/home.html`.

11 See Schellenberg, *The Appraisal of Modern Public Records*, [94]. See also the definition in the SAA *Dictionary of Archives Terminology*: `https://dictionary.archivists.org/entry/appraisal.html`.

ing principle — provenance.[12] The recording process of a document or a set of documents in archives is based on where they came from at the moment they were included in the archives' holdings. This principle is key to the structure of archives world-wide, not just the west, as it was exported in the wake of colonisation.

If all documents had always been archived on the basis of a mission-based, coordinated logic in, say, one political entity, this would be fairly straightforward. But the simple fact that political entities change over time already suggests that things are bound to be more complicated. We could imagine, for instance, that in the 19th century all documents to be preserved regarding the legal body of one nation were gathered in one place, and dispatched later on to regional archives for practical reasons: in that case, the inherent logic of the set of documents would remain untouched. But what really happened, especially in the late 19th and early 20th, is that a lucrative black market developed for manuscripts — often single documents or even pages, without their archival context — which led to the destruction of hitherto coherent collections. Additionally, both world wars saw forms of archival pillaging, or the displacements of documents away from combat zones, which led to the dismantling of collections and document loss.

Let me take an example here to get a better sense of the practical issues that the principle of provenance entails. I draw here on my interest in the collections of the *Staatsbibliothek zu Berlin*, which has a large manuscript department holding, among others, collections of handwritten documents pertaining to 19th-century intellectual and literary life, and in particular to Romanticism. There, I would happen to find manuscripts by the same author kept in different archival units. Indeed, manuscripts by a single author could be split between different boxes, or *Nachlässe*, while others were recorded as single manuscripts or *Autographen*. Following the principle of provenance, the documents were sorted depending on where the manuscript had come from when it was acquired and added to the stock.

For several decades, the acquisition policy of the institution has been to try and complete thematic archival stocks that were already substantial,

12 "Respect des fonds" in French, "Provenienz" in German. It was theorised as early as 1898 in the so-called *Dutch Manual* that served as a basis for archivists worldwide in the following decades. Gilliland provides an informative, albeit critical overview in [61], pp. 38-39.

and to document new additions as precisely as possible, especially their provenance. This impacts greatly on the collections connected to German Romanticism.

The "filling the gaps" strategy, especially in the context of manuscripts dealing with 19th-century intellectual life, owes a lot to manufacturer Ludwig Darmstaedter who donated his manuscript collection to the ancestor of the *Staatsbibliothek zu Berlin*, the Royal Library of Berlin. It also owes a lot to the fact that it was librarians — and not archivists — who determined the recording strategy for this collection and its later goal of "completion".

Darmstaedter lived at the end of the 19th century and was a great admirer of science and the history of science, a field in which he acquired many manuscripts. His private collection was structured according to scientific areas as they were defined at the end of the 19th century. What he eventually donated to the Royal Library was an extremely large collection of manuscripts of all kinds, documenting the many areas he considered relevant to the history of science in earlier centuries. Thus, although voluminous, the collection was not exhaustive, but rather focused on the then relevant fields of scholarship. But it was still large and encompassing, and its structure had a strongly systematic character.

This aspect was so pronounced that the documents that were acquired at a later point by the *Staatsbibliothek zu Berlin* were taken from their context of acquisition in order to integrate them into the Darmstaedter collection, with the goal of making it fulfil the mission the donor had originally assigned to the collection — much to the dismay of archivists. However, this structure has advantages, as Jutta Weber, former head archivist at the *Staatsbibliothek zu Berlin* argues, saying that dismantling some collections led to the construction of unequalled novel insights. Weber justifies *a posteriori* a strategy of destruction of an archival unit for the benefit, in particular, of exploiting correspondences and reconstructing historical networks. On the one hand the "context of collections of papers" is being destroyed, but, on the other hand, it facilitates the creation of a "cosmos", a whole universe.[13] This example encapsulates a virtual war between

13 See Weber, *Sternstunden eines Mäzens* [101], p. 49: "Natürlich wurden hier die Zusammenhänge von Nachlässen zerstört, aber was wurde gewonnen? Ein wissenschaftlicher Kosmos, der heute seinesgleichen sucht. Die Beziehungen von Wissenschaftlern zueinander wurden durch die gemeinsame Verwaltung ihrer Korrespondenz in einer Sammlung von Augen geführt."

two irreconcilable orders. The first made provenance the main principle of its organisation: it was to be dismantled in favour of the second one, the reconstruction of historical networks, which in this case gained the upper hand on the premise that it would provide greater advances in knowledge.

Today's archivists do not randomly dismantle acquisitions. On the contrary, they try to preserve provenance coherence. This means that, for instance, if the missing page of a manuscript is acquired separately at a later point, it will not be physically reintegrated at the "right" place in the order of pages because its provenance is different. It will most likely be presented as an autograph item in and of itself. The connection between the two archival units then needs to be explained for a reader to be able to put the pieces together again, and consult the manuscript as the whole it used to be.

A thorough documentation, laying out the content of all archival units, is required to make it possible for the reader to reunite elements that belong together — if not by way of provenance, then in terms of their context or topic — and to organise the consultation of these documents accordingly. This documentation takes the form of a catalogue or index, fittingly called a "finding aid" in archival contexts: they truly do help find the elements one is looking for. Considering collections as puzzles where time has separated pieces from one another, archives, while they may store the pieces in different boxes, provide finding aids, which indicate how to put them together to make the puzzle as complete as possible, if only for the duration of a single consultation.

Provenance logic has the intrinsic advantage of making shifts and movements between hosting institutions somewhat swifter. If we consider archival material that was displaced during World War Two, the political dimension of a change in the hosting institution and the conflict potential it involves, are not to be underestimated. Again, 19th-century manuscripts from the *Staatsbibliothek* offer a good illustration. During World War Two, part of the holdings dedicated to early 19th-century literary life was taken to the city of Merseburg in order to be kept safe from shelling: it was never recovered. Another part of the stock was taken to Krakow, and it is still there today. For decades now — at the very least since the end of the Cold War — there have been ongoing discussions between Germany and Poland concerning these manuscripts, Berlin still considering them theirs, as does Krakow.[14]

Part of the "Krakow deposit" transferred from the Berlin archives is the famous Varnhagen collection, a treasure for the study of German Romanticism. Just as with Darmstaedter's corpus, it had all started with an enthusiastic collector, in this case Karl August Varnhagen von Ense. Varnhagen first collected the papers of his late wife, Salonnière Rahel Levin, as well as her and his own correspondence with key actors of German Romanticism. He then extended his collection to further manuscripts documenting early 19th-century German intellectual life. Upon his death, his niece Ludmilla Assing inherited the collection, and eventually donated it to the Royal Library — the future *Staatsbibliothek*.[15]

The challenges of recording the Varnhagen collection extend to four areas: first, the physical place of deposit of the material artefacts; second, access to the catalogue that provides information on the content of the collection; third, the original geographical provenance of the stock; fourth, (political) relationships between two institutions, here the *Staatsbibliothek* on the one hand and the Jagiellonian University Library in Krakow on the other, or, more generally expressed, the countries who claim to be the legitimate hosts for the collection (the locations of the current hosting institution and the prior hosting institution). The preserved manuscripts are the bearers of all of these layers of ownership.

This type of challenges has a very concrete material impact on the manuscripts since each hosting institution has its own stamps and signs with which it marks — in the most literal way — the collections that are its own. Each has its way of numbering pages, too. The type of stamp or mark you can find on the manuscript (in some cases, on each page of a manuscript) also provides information on the period of time in which an institution was its owner. All in all, the integration of a handwritten document into a host institution makes the political, cultural, and ideological agenda of each collection visible on every document, sometimes on every page.

14 The presentation of the collection in the German general catalogue maintains a tactful ambiguity in that regard, see `https://kalliope-verbund.info/de/ead?ead.id=DE-611-BF-24146`.

15 Nikolaus Gatter inspects provenance issues regarding this collection in *Lebensbilder* [57] and in *Gift, geradezu Gift* [56].

The difficulties that arose in the wake of World War Two for the two Berlin collections I have mentioned show that provenance is not just a recording principle, but an authority issue, even if for different reasons. Discussions regarding the Varnhagen collection are still ongoing to this day: the case is not closed. Almost more dramatically even, the fate of the Darmstaedter collection epitomises German political history. It was transferred to Franconia during World War Two and successfully preserved, but when it was returned, the corresponding record cards, that is, the finding aid documenting the content of the collection, were sent to East Berlin, while the collection itself went to West Berlin.[16] The collection had to wait until German reunification to be merged again under one roof.

Recording, to speak in Derrida's terms, is not solely a question of organisation and structure ("consignation" as a "principe archontique"), but is essential in addressing the existence of archival material, its origin, and its integration into a specific archival context. While Derrida's stance is primarily theoretical, the examples I have given show that this theory is rooted in historical contexts: any archive is as much part of a historical setting as of the context of its current consultation. Any archive lives in at least two temporalities, and keeps recalling the multiple spaces and times in which it has been collected, kept, and eventually recorded and catalogued.

Ancient Athens had the advantage of being composed of little more than a city, which meant that taking decisions regarding the state and its way of working was materially not as challenging as it had become in early modern Europe. The Metrôon was one address, in the city centre, close to where all other central political institutions were located. In 18th-century Europe, political entities, such as states, had gained a relative political stability in terms of borders or the form of government, but they were much more complex, geographically, socially, and economically widespread structures than was the case in Athens in the 5th century B.C.. And so too was the political project behind their archiving strategies.

As my examples have already suggested, archiving strategies are strongly connected to political goals, and especially to the construction of national narratives. This involves not only administrative material,

16 See the article by Gabriele Spitzer in the exhibition catalogue *Sternstunden eines Mäzens* [101], here p. 28.

such as police, court, or parliamentary documents that are deliberately archived by a political authority, but also cultural artefacts. The shift from an administrative to an (also) cultural function was performed at different moments depending on the country. Even in a late-bloomer nation like Germany, perhaps even more so there, the choices that were made to conceive and implement national unity relied strongly on strategies in which particular writers were set up as symbols of national culture, not least through the creation of dedicated archival spaces.

One could hardly find a better example than Goethe to illustrate this point. Goethe lived to be 83 years old — long enough to look back on his own life and conceive not only several late, more or less final, editions of his works, but also an archive of his personal papers. In the last five years of his life, he pursued several projects in parallel (including the publication of his correspondence with Schiller, and completing *Faust II*), and he also set his secretaries onto the task of sorting and organising an archive of his life's achievements and work. He kept all these documents and partly reviewed them himself. Writing about it in his correspondence, he mentions estrangement ("I can't remember having ever given this topic any kind of attention"), the enormity of the task ("this is never ending", "it takes up so much space"), but eventually concludes he has to fulfil this mission for the sake of posterity, of the nation.[17]

I will not comment further on the obviously high regard Goethe had for himself and his work. What he is articulating here is interesting for other reasons. First, he is driven by the notion of the nation's interest. The nation is presented here as a *Kulturnation* by essence. What is relevant for the cultural elite is, in his view, relevant for the whole nation: politicising society at large means including cultural elements that, strictly speaking, actually concern only a fragment of the population (the well-read elite). Second, Goethe suggests that such a task is too much for a single person, even one with many diligent aids to assist and no pressing money issues like himself. The next logical step would be to say that "the nation", that is, political decision makers, should take the matter into their own hands.

Goethe was very well aware of the power of archiving. Controlling the way his work and life would be archived gave him a unique *post-mortem* hold over the reception of his work. This involved, obviously, sorting

17 See Baillot, *Moi, solitaire, tel Merlin* [11].

out which documents were to be kept and which were to be eliminated. But it also concerned the way in which they should be classified and presented. By presiding over the constitution of an archive of his life and work, Goethe determined the definition of his œuvre such as he intended others to see it,[18] and in the way he wanted to see the tradition established. After his death, the papers went over to his grandson, who made the great Duchess of Weimar sole heiress. She is the one who initiated the erection of a dedicated building, the current Goethe-und-Schiller-Archiv in Weimar. The Schiller papers, as well as those from other relevant writers, were then added to the collection. Although this archive with its dedicated building materialised only in the late 19th century, it was the first literary archive ever set up in Germany, strongly rooted in the emerging national feeling of the early 19th century — and at core drafted by none other than Goethe himself.[19]

Goethe had full authority over his papers as long as he was alive, just as he had full authority over his book production. This seems logical and legitimate: after all, it was his work. Who would be better entitled to make these kinds of choices over his production? Conversely, how can we be sure that a situation in which the author — in this case, an old man, and one of unequalled literary reputation — organised the archive would provide readers with a candid perspective on the work it presents? We cannot. There are plenty of reasons to believe that Goethe was keen on transmitting an ideal image of himself, an ideal he thought he owed to the nation, more than a reality.

Authority in the archiving process means projecting representations into the readers' expectations, and preparing to offer what one thinks it is that others would expect to find, trying to take into account the intentions they are likely to have. In other words, archivists, as representatives of an institution, aim to provide readers with what they think will interest them as regards the writer in question. In that, they stay true to archivist Hilary Jenkinson's position that evidence is key, and to that of Schellenberg, which states that the whole life cycle of documents has to be considered, meaning that they can have different functions and values at different

18 See section 2.1 for more on this aspect.
19 See *Goethe- und Schiller-Archiv* [50] on the history of this archive.

times. A handwritten letter, for instance, is first of interest to its recipient, but it may later turn out to be of interest to a wider audience.[20]

When the writer is the same person as the archivist, they have a unique overview of the evidence that is available, but at the same time, they might want readers to focus on some aspects of their work in particular, and be tempted to encourage a direction that suits them best. Whatever the collection, the person who has authority over the archiving process has to assess three fundamental issues: the source elements (indications given by the writer, for instance regarding relevance — in Jenkinson's words, the evidence); the target elements (expectations of the readers); and the institution's goals (specific missions of the institution hosting the concerned archive) — it is the latter two that define the scope of the document's secondary value according to Schellenberg's model. Statements of purposes for archives are key to understanding not only what they are keeping, but also why they are keeping it. There is not much other discursive space for archives to express the type of authoritative intervention that is at work in the processes of stocking, appraising, recording, and making available that they implement in shaping the character of their material.

The institutionalisation of archives over time has led to practices and to theoretical approaches in the archiving process that are strongly rooted in the materiality of the archive: what space it occupies and how; what role individuals and societies play in archiving as a process; and how archives are structured in relation to political orientations that impact social structures for decades.

The overview presented here provides a broad ouline of the way in which archives were institutionalised and how the definition of their missions evolved over time. In section 1.2, I will go into the archiving process of stocking, recording, and making available in more depth and from a more theoretical perspective. But, before delving into this topic, I will devote the next section to a general discussion of the elements of continuity between analog and digital archives.

20 See Cook, "What is Past is Prologue" [38], especially pp. 24-30 on the Jenkinson vs. Schellenberg debate.

1.1.2 Digital archiving

The emergence of IT-based technologies not only changed some of the technical infrastructure that hosting institutions can rely on: it also introduced a new meaning for the terms "archive" and "archiving" in the digital context. By "digital", I mean supported by a variety of devices able to perform computational tasks, such as computers, laptops, and tablets, connected to the internet, and in which content, including text, is represented by data.[21]

In the digital context, archiving means storing information that is encoded in such a manner that it can be read by a computer and retrieved at a later point. An archive, in the sense in which it is used in connection with IT technologies, provides a set of machine-readable information that makes it possible to retrace the evolution of a digital file. A digital file is the equivalent of what I called a document, or a manuscript, in section 1.1: it can be the letter you find in a drawer in your grandparents' home, or a random file you discover on a hard drive. In terms of its structure, a digital archive relies on the provision of different versions of the same reference document, making it possible to retrace the different steps of its completion up to the point when it is consulted in the present time. The process by which one keeps track of these different versions of the same document, called versioning, is key to digital archiving: it consists of storing information on a series of iterations of the file that reflect its evolution. Proper versioning also requires providing a hierarchy between the different versions (especially a temporal one: which version was there first, which one second, etc.) in order to facilitate navigation between these different versions. Each digital archive is a highly structured document, containing both the content of the file (data) and the information that makes it possible to retrace the evolution of the document's content (metadata).

Not that metadata is irrelevant for non-digital archives. All heritage institutions, including archives, museums, and libraries, have been working with metadata from the onset, and certainly already so in Ancient Athens or Egypt. Metadata are essential to any form of recording: what is being recorded by an archive is, primarily, metadata.

21 Data are organised in files, themselves structured in folders: in this area, the vocabulary is similar to that of the analog structure of document collections.

Metadata are “data about the data” — information about content, physical appearance, provenance, even in some cases the significance of a considered document or artefact. For the manuscript of a 19th-century letter, metadata will indicate that it is a letter; who the sender was; who the recipient was; when the letter was written (and, if known, when it was received); where it was sent from (and maybe where the recipient was staying when it was received); how many pages it has; on what paper it was written; whose hand has been at work on this manuscript (with what, ink or pencil, for example); which institution is currently in possession of the document (maybe also who has owned it before that); what kind of traces (stamps, foliation numbers, comments) were left on it in the course of archiving processes and by whom; whether the letter answers an earlier letter; whether there exists a response to this letter; whether there exist copies (maybe a printed version); and where those are to be found. These are the most basic metadata to be considered for a document like a manuscript of a letter. This information can be extended to a variety of more specific elements. And the answers provided to the questions asked — the metadata fields — can also be either very basic (just a name, a date, or a city), or elaborated upon in order to contribute additional context. Metadata, in and of itself, tells a story about a document, and this story can range from scant bits of information to a fully fleshed narrative. This is as true for digital files as well as for traditional archived material.[22]

Digital archiving relies strongly on metadata to identify different versions of the same document. Transposing the process of creating a digital archive for an analog document, taking again the example of the manuscript of a letter, the digital archive of a letter could for instance be structured along the following lines: first, include the initial text of the letter as it was drafted by the author when they sat down to write in the first place (version 1); then the version including the edits made by the author when they read what they had written and made some corrections to the letter before sending it (version 2); then the version deposited in a folder for preservation purposes, for instance, the box in which the recipient bundled all the letters from this specific correspondence partner, which is likely to include information on the position of the letter within

22 In *Managing Electronic Records*, Philip C. Bantin underlines the importance of metadata as a pivot between analog and digital in archival processes [20]; see especially p. 12.

the whole corpus, for example, its position in the chronological order of the correspondence (version 3); then the version generated when this correspondence folder was acquired by the archives and included in the structure of their holdings (version 4). In all of these versions, the content of the document (in this case, the wording of the text) varies very little. Also, most of the metadata describing the content of the document (sender, recipient, sending date, sending place, etc.) remains identical, although the information on the document, and the way it is part of a context, is susceptible to change, depending, for instance, on who acquires the document, how, and where it is preserved. This information can be recorded either by repeating all the information over again, or by identifying what varies from one iteration to the other and recording only this information. Here, just as with the analog context, there are choices to be made and informed decisions to be taken when shaping versioning mechanisms.

Versioning primarily involves the metadata related to the digital file: when the status of the file changes (when it is shifted from one folder to another, for instance), there is new information to be provided in terms of versioning. Digital methods have been created with the goal of recording only the elements that change from one version to another, and keeping the rest "as is" without having to record it anew. This allows one to identify differences between different versions swiftly. In order to gain yet more orientation, there exist techniques to structure the versions depending on the scale of the changes operated: master versions are the most important ones. This type of information provides readers with input on structure and hierarchy when they consult one specific version.

If we were to try to record every single version, every single change in document content or document status in every single digital document, this would lead to an information inflation of the same range as if we were to claim to archive "everything" in the analog world. There is not enough (virtual) room to keep track of every change that is made to every digital file, just as there is not enough physical room to keep every document ever written, including information on the context of its creation. But in terms of scale, things are rather different. Proper versioning, particularly in parsimonious computer languages, can be undertaken in such a way as not to take up much (virtual) space.

Here we can try a thought experiment. We can pretend to save an immense number of archived files on a variety of devices or cloud hosting, and assume that we would have enough server space to proceed with such an encompassing endeavour.

Even then, we would have trouble accessing information. Just as we might not be able to read some old handwritten scripts that have gone out of use, and need a Rosetta stone to help us translate them into languages we know, so computer-readable languages age and can become less and less readable over time. This is particularly true for proprietary binary formats developed by companies who make money from improving their products in such a way that they become more and more difficult to compute by standard machines, until one day, after yet another update, they become completely unreadable. Binary formats like those used in the office suite, for instance, are problematic because they embed raw text information and the formal features necessary to display it, making it impossible to access the raw text without the paraphemalia within which it is intertwined. This problem is well-identified, and there exist machine-readable languages that are of special interest because of their sustainability over time, as is, for instance, the Extensible Markup Language (XML) that has proven its stability over several decades. In its display of information, it separates semantic content from all that relates to formatting, but still makes the connection between them visible. More generally, open and free formats are conceived so as to be easier to adapt across technical evolutions, so that they remain readable.

Archiving strategies for digital files concern the infrastructure that supports them, but just as much as the computer language in which they are written. This directly concerns resources that need to be converted into a format that enables browser-based visualisation, such as HTML or formats that need to be convertible to HTML to be displayed. They rely on the development of technologies necessary for web visualisation and are susceptible to disappearance if one or the other of their features is not supported, for instance after a browser update.

Laments about online data ceasing to be accessible from one day to the next have become a topos of digitisation sceptics, often without any understanding of the clear distinction between the online display of the information and the source for this information, that is, its source code.

Let us consider the first type of online information: webpages that cannot be accessed after a while. The Internet Archive has proposed a new way of archiving webpages to address this problem. It implemented a tool called the Wayback Machine that makes it possible to go back in time to consult URLs in the way in which they were presented earlier.[23]. What is displayed by the Wayback Machine, though, is only the interface, that is, what is displayed on one specific URL at one specific point in time — like a screenshot from the past. It is not the original information (or source code) that is being archived, and it does not automatically reproduce its overall architecture including all the internal hyperlinks. The Wayback Machine does not archive the content of a database; it archives its presentation in a browser — it is not strictly speaking a digital archive since the source code is not what is archived, even if the HTML version is well preserved and can, to some extent, even be browsed. Although it has its limits, the Wayback Machine remains an endeavour of an unequalled scope. To fulfil its mission, it needs to keep browsing and harvesting the web for images of webpages whose content is likely to evolve over time — and that potentially means every single webpage, even if only because web design evolves and webpages follow this evolution. Of course, this is not done by hand, an automated process is at the core of this massive harvesting of web information. And it might well be that the precise snapshot of the precise URL you would like to consult has not been archived, and you are left with no result for your request. Nonetheless, the Wayback Machine remains a major asset when it comes to archiving internet-accessible content.

What it makes available for consultation, then, is only what is displayed, not the actual content of all of these webpages. This needs to be explained further, since it is essential in order to understand the structure of digitally available information in general. The key here is to understand the difference between, on the one hand, the information that machines display online, in general visually — their output, and, on the other, the information that they are provided with in order to display what they display — their input. Input and output provide different information and they do not rely on the same technologies to be accessible. Bridging the gap between them requires additional technologies to come into play

23 See `https://archive.org/web/`.

to process an input and transform it into an output. Yet other computer languages are used to process information that presents itself in a specific format or language, in order to make it readable, especially by humans.

One of the reasons why a language like XML (or LaTeX, which I am using to write this book) is so interesting is that, although they are used for writing source code (input), they remain readable both by humans and by machines. In these formats, text content coexists with information on how to display this content, all in a single file, although separated into two information levels (contrary to binary formats, which merge them, like a word document). It requires an additional intellectual effort to consider the different information layers presented in a file written in one of these languages, but it leaves room for interpretation when converting the information into a format intended for display. "What you see is what you get" text editors depend on display output. Moving away from that to start using languages, even as simple as Markdown[24] in which you can separate the content and how it is displayed, is certainly one of the most challenging aspects of digital information structure for non IT-specialists. It takes a while to understand that you can label a title line with a "title" tag, and decide later whether you want the title to be displayed centred, bold and red, or on the left, in small caps and black. You cannot see these output options in the basic text file as you do in a so-called WYSWYG text editor ("what you see is what you get", displaying from the onset the output layout), or, if this information is included, it is not intertwined with the text itself, but located in some other part of the digital document containing all the metadata pertaining to output display. Another advantage of these computer languages that describe the way the text presents itself is that they structure it. A digital file in XML, for instance, is organised as a tree with different branches containing information, all related to a common trunk. The hierarchies that structure the document follow the same tree-based logic, making it possible for a reader to easily gain orientation.

There is some complexity in digital archiving that goes beyond that of analog archiving. Taking the aforementioned aspects into account, what would an ideal digital archive look like? Ideally, a digital archive is presented in a stable language, saved on a sustainable server, and is so

24 See the Markdown Guide for more information: `https://www.markdownguide.org/`.

complete that someone wanting to consult any version of the file will be able to do so using only the information provided by the archive itself. It will contain at least the source code (a computer-readable file) and the information on how to transform this source code into human-readable information (output) such as, for example, an HTML file that can be displayed in a browser window. In that manner, even a reader that is not familiar with IT-based technologies or encoding languages can find all the information they need to access the information that is digitally archived.[25]

In that sense, "archive" has a somewhat different meaning when applied to the files stacked on a hard drive or to the old papers you found in a drawer in the family home. And yet, these are not two separate worlds, especially as digital formats are now also used to preserve information derived from the analog world (like those old papers from the drawer). Both techniques converge in the methods used for the digital archiving (and digital publishing) of analog sources. What is more, analog archiving and digital archiving have basic processes in common. Both rely on structured recording, and on updating information to make it understandable by readers. The ageing process is to some extent different: in the case of the analog archive, consultation can contribute largely to decay because it involves physical contact, while it does not modify the digital archive substantially. The digital archive loses its accessibility rather because of a lack of technical support or infrastructure for some formats. In both cases, however, one could generalise, saying that the cause for the lack of accessibility is not intrinsic to the archive itself, but to the fact that it is medium-dependent.

Digital archiving offers more food for thought than merely a polarised opposition to analog archiving: as we have just seen, there are more similarities than might be apparent at first glance. I would now like to turn to the notion of long-term archiving to consider how it can be implemented, taking into account not just the overall structure of digital archives, but also the readability of code. Long-term archiving is to some extent a contradiction in terms. What this expression actually means for the digital archive is "preserving data in their original format for about ten to twenty years". Ten years is certainly not a "long" time

25 For a more thorough presentation of the technical processes at work in digital archiving, see Ciaran B. Trace, *Beyond the Magic to the Mechanism* [100].

when applied to archives as we have known them historically, some of them being centuries old. But realistically speaking, in the context of IT technologies, ten years is a long time. We cannot be completely sure today which standards will apply then, which infrastructures will be available, which languages and formats will have been developed that are likely to make one specific type of files easier to read — or on the contrary impossible to execute. In the case of binary formats, it is almost certain that files written in a specific version of this format will not be readable with the version that is running ten years later since change is the basis of the business model they rely on. In other words, it is extremely risky to offer long-term archiving for longer than ten years, as you cannot be sure to be able to deliver on such a promise.[26] This has much to do with the format chosen for archiving, however, because the core text content, in a basic text format with no structure at all but just the sequence of signs, will, on the other hand, remain. Raw digital text is persistent.

Another difference, maybe on a more speculative level, concerns uniqueness. In the analog context, it is a basic difference between archiving and publishing that, while publishing by definition multiplies a text, archiving does not duplicate, but stores unique copies. The concept of archiving — preserving and recording a unique document for later consultation — is somewhat misleading in the case of digital information for which data has to be actualised, that is, overwritten, even perhaps transferred to another format, in order to still be consultable. Digital archiving calls for iterations or instantiations of archival material, while the uniqueness of the medium is essential to the definition of the analog archive. Taking again the example of the manuscript of a letter, there is only this one manuscript of this letter, and if we (or anyone else, even the authors themselves) copy it, the result will not be the manuscript of this letter, but a copy of it. It might preserve the wording better, for instance in case the paper is fragile and words become illegible for one reason or another, but it is only a copy. In the digital archive, the notion of an "original" as opposed to "copies" does not provide much orientation — every version is to some extent a copy, but also an original.

Readability of the digital archive by the computer does not mean that it can or should be generated without human intervention. On

26 These remarks apply to so-called hot data, that is destined to be opened and used. Cold data on the other hand (which is preserved, but not accessed and executed) is easy to store — but this does not encompass the scope of archiving at large.

the contrary, someone has to organise a digital archive just as someone has to place material in an archive box, store it on a shelf, and record it in a catalogue. In both cases, human effort is essential for defining metadata that will make retrieval of the archived information possible — in one case, boxes with numbers, in the other, metadata fields. But, in both cases, human intervention is key, even if some work steps can be automatised in one form or another. Again, here, languages at the interface between human-based reading and machine-based computing are a way of integrating these different dimensions. They represent a strong safeguard against a separation of computer processes and human approaches, a way to integrate large quantities of material and still have a sense of their quality. They are also a way of keeping archiving, at least to some extent, a human-based activity.

Let us consider a little longer this combination of human and computer activity in the archiving process specific to born-digital documents, created digitally at the outset, because they bring about yet other challenges. Transposing archiving processes of analog artefacts into the digital world can induce an intermediary work step that includes some human intervention, such as structuring the recording process. But the human activity in archiving born-digital productions is less straightforward. Ephemeral digital performances are of particular interest in that regard. They can be grasped statically (with a screenshot at one point in time), as by the Wayback Machine, but any dynamical presentation, such as rotating 3D model, or audio or video content is bound to elude this form of archiving. Today's authors are well aware of the type of power that can be exerted through the discontinuous availability of online material. Literature Nobel Prize laureate Elfriede Jelinek, for example, is known for randomly uploading literary content on her webpage — and erasing it randomly — making monitoring her webpage a tiresome activity for the research team dedicated to her work.[27] While Jelinek's course of action is in many ways similar to that of Goethe contemplating his personal archive, the seemingly random availability of information questions the author's role even more deeply. Subversive strategies like this one, inherited from 20th century forms of opposition to authority, can also be integrated in auctorial archival concepts that take specificities of the digital media into

27 Jelinek's webpage can be consulted at: `https://www.elfriedejelinek.com/`.

account. But whatever the digital staging of control and power on the author's side, if the text was online at some point, then it can be retrieved.

Born-digital archives bring other challenges too, especially for heritage institutions. Authors leaving behind hard drives instead of typewritten manuscripts, which in the course of the 20th century have taken over from the handwritten papers or notebooks I mentioned above, do not make archivists' lives easier. Hard drives do not take up as much space on the shelves as paper, but they too require specific temperature and humidity conditions to be well preserved. And they wear out. They also can very well stop working for no apparent reason and then all access is definitely lost. They need to undergo regular backup on other devices, not to mention actualisations in updated file formats, since popular text editors are proprietary and do not age well. They need to be updated to the next issue regularly, often at a certain cost.[28] In terms of curation, this can hardly be considered as an improvement compared to putting a stack of paper in a box on a shelf.

Questioning the quantity of information is all the more essential as digital archives are prone to never coming to an end, never being fully completed. There is no physical limitation to enriching a digital file, as opposed to a page that is, at one point or another, full. As a logical consequence, there is no need to consider that a digital resource is ever in its final and definite state. It can always be useful, compulsory even, to reopen and modify the file, version it and archive the new version.

One could certainly try to argue that the digital media does not require you to redraft the whole archiving process from scratch when dealing with material that was originally archived in an analog format. You may think that simply adapting this or that aspect (presenting a catalogue in a spreadsheet format for instance) suffices to achieve the shift into the digital.[29] When you consider what archiving means in general terms, the digitisation of archiving processes actually leads to major epistemological challenges, contradictions even. It involves changes in scale (the digital

28 Ageing concerns hard drives especially in the current context of global markets following a logic of planned obsolescence. As well as hardware, it also concerns software, i.e. the source code and the code that is necessary to execute it to provide access to file content. The Software Heritage project is dedicated to preserving software that is not in use any more. It curates source codes in order to provide long-term access to software content. See `https://www.softwareheritage.org/`.

29 See Glenn Dingwall, *Digital Preservation: From Possible to Practical* [44].

space being seemingly much more vast than the physical world, and yet, in another way, finite); the move to a duplication of the material (the variety of digital versions when a unique manuscript used to be the rule); the emergence of archival practices in a wide array of areas (archiving all kinds of material); and new approaches to text as a representational form (specifically in the archiving process itself, with the separation between source code and visual output).[30]

Every archive is a human-made construction of textuality at large. Digital archives are such a construction— perhaps even more so than any other — because of the balance they strive to maintain between contraction and expansion, preservation and accessibility. Broadly speaking, an archive is composed of different text layers, including data and metadata. We need to know how they are situated in their specific space and time: that is what gives readers a chance to position themselves in their relationship to these documents or files, be they old papers in a drawer or a digital file on a hard drive. We need to be able to state where we speak from when we speak about one specific piece of archive.

1.2 From writing to reading to archiving: becoming text

In the previous sections, I have considered archiving processes and their challenges in the analog context and in the digital context, showing how these processes are in some ways similar, but in others divergent. A key element on which to ground this comparison is the textuality that both processes rely on: the letters found in a drawer, the file on a hard drive, even to some extent the source code, are forms of textuality that have to be preserved to be made accessible.

The following section explores what text is made of more fully. As a first step, I look into the relationship between media character and temporality in text genesis. Then I put forward a step by step reconstruction of how text becomes a text, and through which process. Finally, I will consider the archival temporality specific to the preservation of textual traces.

These analyses aim at providing a theoretical framework for an approach of text that encompasses archival functions and other forms of

30 The principles I have presented in this section as an argumentative narrative can be found in a synthetic form under the title "Ending Principles for Digital Longevity" at `https://endings.uvic.ca/principles.html`.

public availability, and positions preservation challenges in this overall framework.

1.2.1 Temporality and text as media: the power of genesis

I have mentioned examples of authors who have attempted to monitor how their work would be archived: Goethe by gathering all kinds of papers at the end of his life and having them arranged by dedicated secretaries; Jelinek by using her webpage to create ephemeral archives of her writings. While these practices follow decisions taken by individuals, they also reflect the archiving habits of specific periods more generally, which are themselves strongly connected to the political context of each period. The contrast, even if only between the 19th and the 20th century, is striking. Manuscripts of the 19th century are often thoroughly doctored or even rewritten by complete strangers.[31] In the 20th century, it has become common practice to preserve, order and transcribe manuscripts by the letter, avoiding interventions. French critic and editor Jacques Petit considers authors who personally see to archiving their manuscripts as particularly affective for their readers.[32] He speaks of a dizziniess ("vertige") seizing whoever might try to take a stance on any part of the archival universe of an author without taking into consideration all that the author chose to preserve. The underlying assumption is that, if an author preserves all of their manuscripts, readers should go through all of them if they want to form an opinion on the author's work or even dare to make an interpretation of part of it. The commentator would have to take into account everything that had been written (and archived) by the author, paying attention to the wealth and variety of textual material available, and could only legitimately attempt an interpretation after acquiring a knowledge of the author's entire production. Petit suggests that the preservation strategy chosen by these authors is a way for them to pressure their readers (maybe more so well-read critics or literary

31 See for instance Sophia Zeil's and my analysis of interventions in the context of the *Tieck and Solger* corpus [19], or Anna Busch and Johannes Görbert's of Chamissos *Weltreise* [34].

32 Petit writes about early 20th century authors who archived almost all their work: "Et lorsqu'un écrivain comme Mauriac, Claudel, la plupart des contemporains, a conservé presque tous ses manuscrits, on se sent pris de vertige en constatant que tout ce qu'on en dit sans avoir *tout* lu, est *un peu* faux.". See "Pour une typologie des textes littéraires" in Hay, *Avant-texte* [66], p. 194.

scholars than "simple" readership though) and challenge their ability to embrace the whole of even a sole work. But does keeping all manuscripts, all drafts, all different states in the production of a text mean that every one of these work steps has to be taken into account at all times? Does it mean that it is necessary to know all of what is left in order to be able to address, consult, comment on a single text of an author? What do these diverse fragments mean in the overall puzzle of an œuvre, and what is their specific significance?

It is not technically possible to take into account every component of a textual ensemble every single time we read a text or every time we are confronted with a loose manuscript. What we interpret is generally one specific state of a text, one moment in its editorial history — the first edition of novel, for instance, or the primary draft of a manuscript. These are not identical with the second edition (the layout may have changed, typographical errors will have been corrected), or the following draft, which will probably have fewer edits. Hence the question: how can we address the fact that a text does not emerge suddenly like Athena, springing fully armed from Zeus' head, but has a history of becoming a text? How can we define it in such a way as to be able to situate our approach for each specific stage of the text, at each specific time in the process of its creation?

Depending on the consultation context, the position occupied by a specific version of a text can change. The materiality of the manuscript evolves — paper ageing, ink decomposing paper, pencil naturally erasing. The act of consulting a text, most of all a manuscript, has an inherently transitory quality: each consultation reveals a different iteration of the text, as the media it is inscribed in keeps evolving over time. Considering this changing situation, it seems necessary to address systematically what temporality, in relationship to its media, defines each moment in what we consider a text. Or does this inevitably lead to making a text eventually impossible to grasp? That is: does asking myself the question of what stage of the text I have in front of me make my reading easier and more legitimate; or, on the contrary, does it make reading it unnecessarily complex, even impossible to proceed with? What is it I hold when a manuscript passes through my hands? How can I know whether it is of major importance or virtually irrelevant?

The more one tries to grasp what constitutes a text as a part of a more general context, the more it slips through one's fingers and seems to lose its substance. Here, I will take the example of a literary œuvre for simplicity's sake, but will elaborate on the difference between text, book and œuvre in 2.1. The French school of the *critique génétique* constructed a differentiation between *avant-texte*, *texte* and *après-texte* (before-text, text and after-text).[33] Although this three-step structure provides helpful orientation, it lacks consequence in at least two points. First, it imposes a clear unidirectional linearity: the three steps follow one another and cannot be shuffled into a different order than that of before, during and after the composition of a fixed, core version of a text. And second, it suggests that the *avant-texte* could not exist without that which comes after it. The denomination *avant-texte*, *texte* and *après-texte* includes a pre-determined teleological bias as it revolves around the idea of one core text, chronologically situated in the middle. In these terms, one and only one version of the text gives meaning to the other versions. In practice, this ideal framework is not that simple to apply. There are often various copies or versions of a manuscript or a print, and one cannot always easily settle which one is authoritative, which is "the text".

Defining text is extremely easy and extremely complicated at the same time. French tradition of the late 20th century provides a few definitions that hardly resist media variability as we know it today, but still offer food for thought. Philosopher Paul Ricœur states: "Appelons texte tout discours fixé par l'écriture"[34], calling text any discourse fixed by writing — the focus is on the writing process, insofar as it is achieved. Derrida too focuses on writing, but he emphasises how it is rooted in origin rather than

33 In the eponymous seminal volume [66], the article by Jean Bellemin-Noël, who first proposes a definition of the "before-text" that struggles to emancipate from the notion of a "draft" that would simply be a prior, less achieved version (p. 162), but he then proceeds more convincingly: "Il est indispensable de substituer les métaphores spatiales aux images tirées d'un registre temporel, sous peine de réintroduire une téléologie, alors que la rédaction n'a pu être qu'une production toujours surplombée par l'incertitude et l'aléatoire. Le paradoxe à ne jamais oublier, le voici: ce qui a été écrit *avant* et qui n'avait a priori pas d'*après*, nous ne le connaissons qu'*après*, avec la tentation d'en faire un *avant* au sens de préalable, de cause, d'origine." (p. 163). Bellemin calls for moving from a spatial to a temporal register and sees it as an essential paradox that what was written "before" did not have an "after" when it was written, adding uncertainty and randomness to the process. In recent work by the *critique génétique*, even more complex concepts and vocabulary were developed, which can be consulted in the *Dictionnaire de critique génétique*: `http://www.item.ens.fr/dictionnaire/`.

34 Ricœur, *Du texte à l'action* [91], p. 137.

achievement: "'Qu'est-ce que l'écriture' veut dire 'où et quand commence l'écriture'"[35] — "what is writing" means "where and when does writing begin".

This remains general and cannot easily be transformed into a practical way of reading or interpreting texts. Anglo-American textual studies provide a more pragmatic grasp on socio-historical factors without giving up on a theoretical framework. American textual scholar David Greetham writes:

> [The text] is, on the one hand, a place of fixed, determinable, concrete signs, a material artefact, and yet, on the other, an ineffable location of immaterial concepts, not dependent at all on performance transmission. It is, on the one hand, a weighty authority with direct access to originary meaning and, on the other, a slowly accumulating, socially derived series of meanings, each at war with the other for prominence and acceptance. It is a place inhabited only by a sole, creative author who unwillingly releases control to social transmission, and it is also a place constructed wholly out of social negotiations over transmission and reception.[36]

Greetham argues along the lines of authority, control and concurrence. The metaphor of the battlefield is not the most reassuring one, but there truly are troop movements in the text: the author's, those of the reader(s), the pre-existing meaning of the words that are used, and the materiality of the text itself. Greetham describes them in a fight for attention, for authority within the text. There are different, sometimes opposing forces at work. Their plasticity is made all the more mouldable as they take account of the media options that are available. But these are not the sole driving forces at play. Greetham also emphasises the effort required from the author in order to enable transmission. Text happens between participants, and the "negotiations" take place both on a social and a material level.

The definitions I have cited all agree on the idea that text is in movement, both on the space of the page (the discourse "is being fixed", following Ricoeur; it "begins somewhere" following Derrida; or "is a

35 Derrida, *De la grammatologie* [41], p. 43.
36 Greetham, *Theories of the Text* [65], p. 63.

place", in Greetham's words) and in the temporality of it being written and read (Derrida also asks "when" it starts; Greetham suggests a "slow accumulation").

One way to focus these approaches would be to say that there is something like a geography and a history of a text, both anchored in its materiality and how it evolves in terms of media.

We can turn first to what I want to call its geography, or topography. It pertains to the way in which the signs are disposed on the page, and designates the topography of the chosen medium. Prosaically, it describes how darker and lighter bits — such as black ink on a white ground — alternate to shape letters and other glyphs on a page. It is not the letters that are written which produce meaning, but the way in which the different elements contrast with one another on a given space: it is the composition that matters. The topography of a text varies very little if at all when you compare different copies of the same edition; it varies a little more when you compare different editions of the same book; it varies remarkably when the two versions compared to one another are based on different media (for instance, a manuscript on the one hand and a printed copy on the other). Poorly executed print-on-demand is a good example of the fact that reproducing pages identically does not necessarily convey the same reading impression: bad print, imprecise reproduction, and lesser readability are more the rule than the exception in these practices. What this example also shows is that the geography or topography of a text is closely connected to the type of writing or printing equipment that is available at the time in which a text is either written or reproduced: ink type, paper type, printing equipment, and fonts do not vary randomly, they reflect the conditions in which the text is being produced.

For instance, Gothic print is strongly rooted in the development of print in the German-speaking area. Nowadays, these fonts are generally associated in the collective imagination with nationalist claims, as if they would programmatically incarnate the distinction from Latin print, and 'Latin' countries. In extreme cases, Gothic print can even be interpreted as an affirmation of the superiority of Germanic culture. Yet their cultural meaning used to be quite different. In the 18th-century German-speaking world, Latin script was considered a scholarly print, and more difficult to read for readers from the German-speaking areas. Gothic print was used for printing leisure reading and deemed more likely to foster pleasure in the reading process. This was the dominant point of view in the period when the publication market expanded dramatically, with the increase in leisure reading in the late 18th to early 19th century. A few German printers such as Unger, together with French printer Didot, engaged in creating a new set of printing letters (so-called types) that would combine characteristics of Latin and of Gothic print, but their attempt failed.[37] It would certainly have required at least a strong political will, if not an economical investment, to consolidate a technological effort that would eventually have facilitated the circulation of German prints on the European book market for many decades. Instead, by the middle of the 20th century, the cultural reference was not any more that of scholarly print vs. leisure print, but that of Latin vs. German print, with a strong political connotation. The materiality of text is itself political.

Choices in the elements that compose the print such as type or ink are related to what I have identified as the geography or topography of text. In addition to this dimension of space, each text has a dimension of time (what I would call its inherent history): text evolves over time. When textual geneticists state the distinction between *avant-texte*, *texte* and *après-texte*, they also suggest that one of the speculative difficulties is to consider this temporality in itself, to take into account the fact that the evolution of text is not necessarily organised along the lines of a straight time arrow, but is susceptible to leaps of time — in one direction or another —, affecting text and its materiality. The different steps in the dissemination of text are one possible matrix here (from the draft to the publication), but they should not be considered as an exclusive linear prescriptive power defining how text evolves, comes to being and is to be interpreted at all times.[38] Considering the evolution of text within time and yet not necessarily in a linear way is all the more difficult because we are used to reading in a linear order. It is rather speculative to stop believing that things evolve along a unidirectional time axis, but that is the idea proposed here: to try to reconstruct the inherent history of each text in a temporal framework that is not always moving in one direction from a beginning to an end in a direct line.

37 See Lehmstedt, *Ich bin nun vollends* [76].
38 See Vasak, "Analyse de système et textologie", in: Hay, *Avant-texte* [66], p. 199.

What I call the inherent history of a text commences with the act of writing in itself. Even this initial writing process can have a different significance depending on the author. For some writers, even the first draft presents some degree of achievement: they are *Kopfarbeiter*, people who draft mentally before starting to put anything down on paper. For other writers, it is the writing itself that initiates the process of drafting and redrafting until the meaning of the text emerges. The act of writing, as the first moment in the history of a text, informs the text both in terms of its content and its shape. A short poem will take little space on the page and be divided into stanzas. A prose narration will occupy the page space more completely. A letter will have a dateline, a salute, and the centred name of an addressee. While these elements can change at a later point (just as content can), they are drafted through this first writing act.

It is tempting to consider this stage of initial writing from a teleological perspective, that is, to interpret all the signs that are visible on the paper in such a way that they tend to become what the text will look like in its final stage. Reading into what one can see in a first draft, one could be tempted to interpret it along these lines of: here, the author started with this word, but then they preferred this other one because it addressed their intention more clearly, as the final version shows. Much can be projected in the interpretation process. Ricoeur objects to the risks of too strong a psychological interpretation.[39] The writing process itself, says Ricoeur, gives the text autonomy, disconnects it from its author's intention. The conclusion he draws from this observation is that psychological, as well as sociological creation conditions are insufficient to account for textuality. Textuality transcends its own creation and gives way to myriad readings, themselves anchored in different socio-cultural contexts. This argument is rooted in a reflection on the work of art in general. At its core, it calls for dissociating text from speculations on the author's intention from the onset of the writing process, and that is exactly what matters

39 Ricoeur, *Du texte à l'action* [91], p. 11: "l'écriture rend le texte autonome à l'égard de l'intention de l'auteur. Ce que le texte signifie ne coïncide plus avec ce que l'auteur a voulu dire. Signification verbale, c'est-à-dire textuelle, et signification mentale, c'est-à-dire psychologique, ont désormais des destins différents. [...] ce qui est vrai des conditions psychologiques l'est aussi des conditions sociologiques de la production du texte; il est essentiel à une œuvre littéraire, à une œuvre d'art en général, qu'elle transcende ses propres conditions psychosociologiques de production et qu'elle ouvre ainsi à une suite illimitée de lectures, elles-mêmes situées dans des contextes socio-culturels différents."

here: to consider a text as not strictly or exclusively correlated to its writer at the moment of writing.

There would be much to say about the initial moment in the inherent history of a text. The moment when writing begins is often interpreted in terms of the continuity and/or rupture elements that can be spotted on the media device —I will speak of a page for simplicity's sake. Does a variation in ink colour mean that the writer changed their mind? Does it mean they took a break at this point in the writing process — but if so, why? It is rather difficult, when considering the initial stage of writing, to actually dissociate text from author, materiality from individual. The writer is, after all, a human being, who sometimes gets disturbed by someone knocking at the door, just as much as they are a poet who needs to reconsider a verse for reasons of rhythmic balance. Who are we as readers to decide whether the author had a deep reason to pause or just had their cat knock over the inkwell? The question then is how we integrate the space outside the page, where things happen that we do not know much about, and what do we make of the signs we see on the page, or what we can see of them. One way of avoiding an over-interpretation of such text stages that would be too strongly led by authorial considerations is to primarily take the hand(s) that wrote into account, and not the person. For a draft, this approach facilitates interpretation.

One might object that not all writing processes formally require a hand and a paper manuscript. The hand considered here is not to be taken literally, but rather as a symbol for becoming written. At this point in the process, the text is not really yet a text, in the sense that it has not yet been read by someone other than the writer: the hand needs another person's eyes to realise the textual potential. This is, in fact, the next step in what I call the history of the text. The media history of a text does not fully take shape at the moment of writing, but in the next step, the moment of transmitting, of giving it to someone else to read. The transmission of a text can happen in a variety of contexts: in a private circle or to a public audience, in the form of manuscript or print. What these contexts have in common is the shift to a reader that is not the author — eyes that do not belong to the hand.

Writing for other eyes often means re-writing for the purpose of communication: what was a draft becomes a cleaner version, with fewer edits and more space dedicated to the final shape of the text. Some authors write their primary drafts using abbreviations: in the version for another reader, these will have to be expanded so that complete words are legible to those who are not familiar with the abbreviations. This first "clean" version of the text can be carried out by the same hand as the initial draft, but it can also be another hand's doing. Authors who can afford to dictate their initial draft to a secretary are less common than those who commission a third party to realise copies of their initial drafts. From a certain age, Goethe dictated almost everything he wrote — and yet one cannot say that he did not write.

The choice of the hand that will copy the draft has varied a lot over time. From the late 18th century and the development of editorial businesses as family endeavours, it was often a wife or a daughter who was in charge of copying, the activity that was considered not to entail any creative dimension. Wealthier writers could employ someone, a student for instance, or a professional copyist, to that end. It can be argued that this writing step, that of copying (in the literal sense of procuring a copy), is more creative than might be assumed: it is undoubtedly a material change that is likely to induce variations, shifts in the text. On the formal side, the disposition of the text on the page will most certainly change. In terms of content, it is prone to introduce errors (in the case of an erroneous reading of the draft, for instance, or a mistaken transcription by the new hand), or even intended corrections (if the copyist adds what they consider is a missing word in the text, or deletes one that seems superfluous). The *New Bibliography* differentiates, from this stage on, between *accidentals* and *substantials* in order to characterise the types of variations that can be found in a text. Some concern mistakes, typos, or other minimal alterations, while others involve more considerable shifts in the fabric of the text.[40]

Ideally, it should be uncomplicated to draw the line between modifications that qualify as *accidentals* and those that are *substantials*. But a great deal of textual phenomena operates in the grey zone in which the distinction between the status of each of these is not unequivocal, especially when it comes to stylistic alterations. Novels from the 19th century first published, often in form of episodes in journals, when their authors were young and penniless, could undergo a major rewriting process for

40 The importance of the New Bibliography is pointed out by Greetham in *Textual scholarship* [63].

later editions. This not so uncommon situation gives a sense of the type of challenges faced when interpreting these different versions of a text that has the exact same title. Greetham writes:

> [...] several major Victorian authors took advantage of reprints and new editions to rewrite the accidentals, often to undo the house-styling that had been foisted on their texts when they had been young struggling artists having to capitulate to publisher's styling demands. Dooley's argument is that the balance between intention and expectation has to be adjudicated for each text (and that includes each edition), for the equilibrium shifts as the relative power of author and publisher changes during an author's career.[41]

At this point of visibility or public dissemination of the text, the relationships between writer and publisher come into play. I will explore them more in detail in 2.1.1: encapsulating their significance is not as simple as it may seem. While the publisher is not the author, they can have a major impact not only on choices in layout, design, fonts, bindings, etc., but also on the wording. Looking at it, not from the point of view of the actors involved, but from the point of view of the text and its inherent history, what is at play in the intervention of a publisher is the relationship between primary intention and authorship. Who can be considered as the author, of, say, a printed version of a text in which the publication process has involved many variations in the wording, when compared to the manuscript that was sent by the writer in the first place?

These tensions are sometimes displayed very prominently in printed products, particularly books, as is the case where one finds *corrigenda*. They generally consist of a list of items that were printed incorrectly, such as word omissions or mistaken names that are likely to change the reading one can make of the text. They are presented together with the correction wanted by the author and the number of the page where the error is located. When there are *corrigenda*, they are often displayed beforehand, at the front of the book. The pre-eminence of this position, at the forefront of the publication, can be considered to reflect the authority the author has over the publisher: not only does the author re-establish the text as

41 Greetham, *Theories of the Text* [65], p. 194

it should be read, but they do so in the first place, sometimes before the title page itself. This is more than anecdotal. It shows that even in the printed version of the text, there is room for variations in disposition — for making text happen in one way rather than another.

Having distinguished a series of steps in which text is being modelled in such a way, first by a hand, then by alien eyes, and further a copying hand or a printing device, all contributing to making it become text, one last step is still missing. What happens when the text is being read, not simply by friends, family and publisher, but at large? This is the intersection between text genesis and text reception. Text genesis is defined as that which concerns the creation of text. Reception analyses how the text is being read and commented on by third parties. In general, the genesis of a text is interwoven with the reception of other texts: writers have read books before they themselves start writing. Any form of writing is always an intertext of some sort, as it is created in a context where other texts serve as a background that is more or less conscious, and more or less formalised as such, but that always interferes and contributes to co-create the new text in some way. The genesis of any text lies in the reception of other texts, and every text is always an intertext from the moment when it is read by someone. This intersection between writing and being written, reading and being read, generates a never-ending revolution in the circulation of texts, a movement that operates at the core of the history of any text.

This process affects all stages of a text and its publications, including editing, as Greetham concludes, and to some extent, as we will see, includes archiving as well:

> The business of editing, just like the business of literary criticism and the business of writing about textual and literary criticism, is relentlessly intertextual. For although there has always been the temptation for the editor to lay a deadening hand upon the text and fix it forever, such temptations will not withstand the forces of history or the enquiring intellect: these "fixed" texts will always have to be "done" again.[42]

Being aware that text is not fixed is central. One can usually spot movements or shifts in the writing process or in the transmission, that can be

42 Greetham, *Theories of the Text* [65], p. 461.

associated with writing hands at work. This helps to avoid authorial overload in the interpretation and prevents readers from projecting intentions into stages of text constitution. But in the end, one should consider that any interpretative gesture can always be questioned by the emergence of new texts, new intertexts.

The manuscript is an essential object around which the history of the text can be spun, and with it, its existence in the world. It is the bearer of information about the text that archives can record and preserve. Depending on the archived stage, it can reflect the editorial process that transforms a text into a publication to a greater or lesser extent. This complexity is reflected in archival systematics, but it also leaves room for interpretation on the archivist's side. What copy is the authorial one? What hand is that of the author? These are the questions archivists too have to answer, and that serve as a basis for our access to text.

In this section, I have examined text genesis, the process by which text becomes text, emphasising on the one hand the plasticity of text (always liable to change), on the other hand how strongly rooted it is in its media form. In the following section, I will examine this from the perspective of the archiving process proper.

1.2.2 Traces of the past in the present

We can now turn again to the old papers you found in the drawer of the family home, the postcard collection you purchased at the flea market, or even the hard drive a writer has deposited in literary archives, and examine them in the light of the text genesis process. What you find can be a draft; or it may be a copy; or it may be a printed version — it can be all these layers and versions, comprehensively bundled in a folder so that the reader will be able to follow the textual evolution. It may be a fragment of something one cannot even identify: a note? the draft of a letter? part of a diary? Has it been read before, or am I the first one to lay eyes on it? All these questions require you to remember that you are dealing with more than a text, but with what has become a document in the sense that it has gained recognition as a structured part of heritage.[43] This status as a document primarily concerns the epistemological value,

43 A group of French scholars has published an encompassing theory of documents under the collective name of Roger Pédauque, see "Document et modernités" [87]. Ricœur considers the relationship between document and trace within the framework of history

the knowledge it holds. In terms of its materiality, it consists of what are considered traces of the past that have reached a reader, traces that are anchored in both past and present. The edges of the page are certainly not the same as when the text was initially written, and the ink is not the exact same colour; what has reached the present is only partially identical to what it was in the past. What is more, considering texts as documents in the sense that they are traces of something past assumes that not everything that was produced in the past has reached us today. Archives do not actually contain all that was ever written, but solely a fragment of it, traces of time past. How does a document-based approach relate to an understanding of fragmentary traces?[44]

This is a challenge to anyone interested in the past. Trying to disentangle its complexity, we can start by asking ourselves: What are these traces of? What do they represent in terms of our knowledge of the past and of ourselves? Ricœur articulates this crux in *Time and Narrative*.[45] The paradox he points to there is that the traces we can see bear witness to an ageing process of which we can only see the result, and not the process itself — and even less so the original state of things. The trace, considered as the physical evidence on which historical research is based, lacks existentiality: it is neither what it has been by the time it was created and was not a document of something past yet nor is it a real thing of our time. It is constructed as a thing of the past while it is, in fact, a projection of the present on the past. It is also what Derrida points to when he writes: "la trace *n'est rien*, elle n'est pas un étant" — "The trace *is nothing*, it is not a being"[46]. The trace is a symbol of something that is

as a documentary science: "L'histoire en tant que recherche s'arrête au document comme chose donnée, même lorsqu'elle élève au rang de document des traces du passé qui n'étaient pas destinées à étayer un récit historique. L'invention documentaire est donc encore une question d'épistémologie. Ce qui ne l'est plus, c'est la question de savoir ce que signifie la visée par laquelle, en inventant des documents [. . .], l'histoire a conscience de se rapporter à des événements "réellement" arrivés. C'est dans cette conscience que le document devient trace, c'est-à-dire [. . .] à la fois un reste et un signe de ce qui fut et n'est plus.", see *Temps et récit III* [90], p. 13. In his view, history invents documents based on traces of the past, but should do so knowingly. The *SAA glossary* focuses its definition of document on its medium more than on its connection to the past: `https://dictionary.archivists.org/entry/document.html`.

44 See Baillot, *Reconstruire ce qui manque* [10].

45 Ricœur, *Temps et récit III* [90], pp. 217–218: "C'est bien là le nœud du paradoxe. D'une part, la trace est visible ici et maintenant, comme vestige, comme marque. D'autre part, il y a trace parce que auparavant un homme, un animal est passé par là; une chose a agi. [...] Où est alors le paradoxe? En ceci que le passage n'est plus, mais que la trace demeure [. . .]."

46 Derrida, *De la Grammatologie* [41], p. 110.

gone, a *symbolon* in the etymological sense of the word, a token made of one bit that is left behind where it originated in the past, and another one sent to the future world bearing the remembrance of the missing bit. This *symbolon* is only whole when both parts are physically united to form the original item. But the first part is lost for ever, swallowed by time, and all we have is the second one, always painfully reminding us that it is not whole, but a fragment. From this perspective, the awareness of what is missing, or at least of the fact that something is missing, makes it possible to actually structure knowledge.

Derrida brings the paradox to the limit of knowledgeability.[47] The trace does not only indicate the disappearance of what was at the origin, it signals that it was never itself at the origin — it was something else. It would, according to Derrida, require a completely different concept, that of an archi-trace, to actually embrace the trace that is at the origin. In fact, it is impossible to consider what the trace originally was, or to unite the notions of trace and of origin. And if the trace has no clear origin, then what is it a trace of? From this perspective, all our traces from the past are only ever but constructs of something that is not original, and whose relevance always needs to be reassessed.

And yet, for reasons sometimes unclear, something was left behind of time past following selection mechanisms that often lack transparency, and this trace remains the main artefact on which historical perspectives on text rely. Due to its loose temporality, to the uncertainty of its origin (where did it come from? what did it originally look like? what did it mean?), it remains to some extent difficult to grasp. We do not ever really know what it was the trace of. It is the trace of something that is not any more, and we cannot know whether what we think it emanated from has ever been, nor what it really was. All we know for certain, at a speculative level, is that the trace is not that which it is the trace of, and that it lacks its origin when it reaches us.

We do not know where the trace comes from, yet at the very same time, it is the well from which we draw. As if this was not embarrassing enough

47 In *De la Grammatologie*, he writes: "La trace n'est pas seulement la disparition de l'origine, elle veut dire ici [...] que l'origine n'a même pas disparu, qu'elle n'a jamais été constituée qu'en retour par une non-origine, la trace, qui devient ainsi origine de l'origine. Dès lors, pour arracher le concept de trace au schéma classique qui la ferait dériver d'une présence ou d'une non-trace originaire et qui en ferait une marque empirique, il faut bien parler de trace originaire ou d'archi-trace"; see Derrida, *De la Grammatologie* [41], p. 59.

from an epistemological point of view, the question of its preservation is also full of contradictions. Keeping everything is impossible. Choices and selections are being made. According to Ricoeur, this selection is natural, or at least congenial to the order of things. It suffices to do nothing for things to become ruins.[48] Not taking action is leaving it to time to do the work of letting things decay — and believing it is time itself that brings about destruction.

But time is not the only factor. Other mechanisms are at work, connected to recording and archiving processes, as was mentioned in section 1.1.1. Archiving means selecting, be it unwittingly, if physical decay or loss could not be avoided, or on purpose, in order to keep some documents rather than others — family papers only if they are 150 years old, novel manuscripts only if they are by famous authors. These are deliberate choices connected to the mission of purpose of heritage institutions, a principle officially embraced by the profession at least since Schellenberg's stance on appraisal.[49]

Whatever the mission or purposes, there is common ground: the fact is that neither institutions nor individuals are in a position to keep all traces from the past. Some artefacts, some pages, will inevitably have to disappear. There is not enough room to keep everything, and not enough time to record everything so that it can remain consultable.

But at the same time, keeping order in the traces of the past to archive them, recording them, means proceeding to actions that will both preserve and destroy content. There always lingers above archival documents the threat of destruction; this, too, is part of what defines them.

The threat comes from natural (uncontrollable, irrational) and human (deliberate, motivated, even rationalised) destruction, both at work during the archiving process. French archivist Arlette Farge describes the frailty of archival documents almost poetically, each a "piètre pièce

48 See Ricœur, *Temps et récit III* [90], p. 34: "En un sens, il n'y a là rien de mystérieux; il faut en effet faire quelque chose pour que les choses adviennent et progressent; il suffit de ne rien faire pour que les choses tombent en ruine, nous attribuons alors volontiers la destruction au temps lui-même."

49 In *The Appraisal of Modern Public Records*, Schellenberg states this in unambiguous terms: "A reduction in the quantity of such public records is essential to both the government and the scholar.[...] Scholars cannot find their way through the huge quantities of modern public records. The records must be reduced in quantity to make them useful for scholarly research." [94], p. 237.

[...] à conviction, toute en dentelles", a piece of lace.[50] Similarly, German philosopher Knut Ebeling characterises archives as a "Friedhof der Schrift",[51] a cemetery of writings, and goes one step further in the register of destruction, up to institutionalised rot and decay. Derrida too comments on the fundamentally destructive power of archiving: "L'archive travaille toujours et *a priori* contre elle-même"[52]: repetition, inscribed by recording and preserving in order to consult archived documents, is the process through which the archive works "always and *a priori* against itself". Each consultation of an archival document contributes to its decay. This means that it is not technically possible to consult the exact same document twice: between the two moments of consultation, the document will have evolved, and will not be exactly the same any more. (One could also say as much of the person who consults it too: one cannot encounter a manuscript twice in the exact same state of mind and body.) Consultation, the only action that can contribute to actualising the potential of preserved archival traces, accelerates their degradation. It induces an uneven time trial in which the intelligence generated by consultation has to keep up with, maybe even overtake, the degradation process accelerated by consultation, which itself contributes to loss of information.

This perspective also helps to shed to some light on the contribution of digitisation processes in the archiving process of physical artefacts. Scanning a document or an artefact that is degrading at a dangerous rate — at a rate deemed dangerous in the temporality of cultural appraisal and research processes — makes it possible to record the evolution of this degradation in two dimensions and with a pre-defined image precision. On the one hand, digitisation thus makes it possible to keep track of the history of archival traces. Thanks to it, we have the opportunity not to lose completely traces that are decaying before our eyes. The physical artefact becomes more valuable in the sense that an in-depth interpretation, taking into account its evolution, remains possible. The material life of the

50 Farge, *Le goût de l'archive* [48], p. 73: here with the words "meagre evidence" and "lace-like". Farge also uses the image of a puzzle whose pieces are scattered and can only be put together imperfectly in the form of obscure events: "Si l'archive sert effectivement d'observatoire social, ce n'est qu'à travers l'éparpillement des renseignements éclatés, le puzzle imparfaitement reconstitué d'événements obscurs." [48], p. 114.

51 Ebeling, *Archivologie* [45], p. 12.

52 Derrida, *Mal d'archive* [42], pp. 26-27.

document can, to some extent, be accounted for, visually at least. On the other hand, however, digitising a source document can generate the illusion that the image and what it is an image of are identical. The original document loses symbolic value, the scan gains a validity that may even be considered absolute: you could argue that the scan is better than the original since you can zoom in and see details the naked eye cannot grasp. But in this the temporality of the archiving process gets lost. Because the fact is that very few people engage in an elaborate consideration of the date and conditions of digitisation when they work with a scan. Technically, the scanning date and conditions should be added to the date of the source document to really define the temporality that is the one of its material textual traces when it is consulted as a scan on a computer monitor or on a smartphone display.

Not only the consultation of documents is a cause for decay: when paper is exposed to air, light, or heat, its degradation accelerates. And even if conservation operates with all thinkable and realisable precautions — optimising humidity and temperature, keeping sheets in protective boxes, finely separating them from one another, away from light — it too becomes part of the degradation process. Conservation does not stop time, it only slows down some of its effects. Acid inks will still corrode paper, for instance. This is part of the natural course of things — "time" in the words of Ricoeur — but also of the use of degradable cultural techniques and artefacts.

In the archiving process then, an archive deploys mechanisms of conservation that include a dimension of destruction. One of the consequences of this is that one cannot always dissociate conservation from destruction. Not all the traces we would need to reconstruct what we want to know of the past are available. We do not even have enough traces to know for sure what the blind spots are, to measure the extent of what has been lost over time. Derrida addresses this double issue in one question: "Comment peut-[on] prétendre faire la preuve d'une absence d'archive?"[53] — how can we pretend to prove an absence of archive? The emphasis is on "prétendre" more than on the fact of giving proof of what is missing: the problem is not so much that we do not have this lost material, but that we might be tempted to pretend to give an exact

53 Derrida, *Mal d'archive* [42], p. 103.

account, or at least assessment, of what we do not in fact have. Derrida's point concerns primarily the epistemological perspective: the best we can do would then be to admit the coexistence, next to the traces we perceive, receive, and know of, of an abyss of absence. Of this submerged part of the iceberg (or is it the ocean itself?), we can only ever know imprecise things. Its silhouette forever remains impossible to draw, and with it, the draft of a reality past that we can never get close to, be it materially or speculatively. From this perspective, the more one tries to get close to a documental trace, to embrace the whole of an archive, the more one surrounds oneself with chimeras, and the less one is able to know something about it.

This does not only concern the individual position of a person wanting to consult, work on, or work with traces of the past, and who will necessarily be confronted with what one could call a dialectic of preservation and destruction that is inherent to archival processes. As institutions, archives are equally confronted by this challenge, which, beyond the fate of one singular document, is also reflected in the archiving process in general. All three instances — archives as an institution, archiving as a process, and the single piece of archive — face the dilemma of survival: what remains is incomplete, imperfect, something whose essence is not guaranteed to survive and that can only exist through an act of destruction. Derrida insists on the notion of "survival", emphasising the somehow organic life of what could otherwise be considered as an old piece of paper. Contradicting archivist Terry Cook who argues that the "twin pillars of the archival profession, appraisal and arrangement/discussion"[54] should be front and centre in this debate, other philosophers desert all form of archival agency to embrace this dialectic by shifting agency to the social body.

Foucault situates it in discursivity, engaging culture at large.[55] What interests Foucault in one relevant passage of *Archeology of Knowledge*

54 See Cook, *What is Past is Prologue* [38], p. 20.

55 Foucault states the following in *Archéologie du savoir* [53], p. 177: "Par ce terme [= archive; A.B.], je n'entends pas la somme de tous les textes qu'une culture a gardés par-devers elle comme documents de son propre passé, ou comme témoignage de son identité maintenue; je n'entends pas non plus les institutions qui, dans une société donnée, permettent d'enregistrer et de conserver les discours dont on veut garder la mémoire et maintenir la libre disposition. C'est plutôt, c'est au contraire ce qui fait que tant de choses dites, par tant d'hommes depuis tant de millénaires, n'ont pas surgi selon les seules lois de la pensée, ou d'après le seul jeu des circonstances, qu'elles ne sont pas simplement la signalisation, au niveau des performances verbales, de ce

is neither the archives kept by a culture to document its identity, nor the institutions that record and preserve texts that are to be kept and made available. His focus is on how the discourses deployed in these contexts have emerged as verbal performances arising from an order of the mind and an order of things, as the result of relationships at play. In a way, Foucault and Derrida present similar views of archiving as an informing order: it is a selection process in which a material trace is not only preserved, and therefore destroyed in some of its dimensions, but also gains a place in the construction of a collective memory as a discourse, and an awareness of human limitation. This order of things is not established once and for all in the archive: archiving remains a process, and an archival piece keeps changing shape, shifting and evolving as time goes by. The complexity of the object and of the processes is in many ways an incitement to open up metaphorical spaces.

Barthes uses a positive image in *Writing Degree Zero* while delving into the processes of the making of text, namely that of the magic ink.[56] Any written trace reveals itself over time, with new layers appearing that display new and always denser connections to the past. As opposed to the notion of a permanent self-destruction, the magic ink suggests that the ongoing process makes way for new layers of meaning to appear over time. Rather than an impoverishment, the work of time is an enrichment: history as an accumulation of temporalities is inscribed in the new layers of meaning that appear, and not separated from them.

This idea of a fusion, or a coexistence of various, contradictory drives that all linger in the archival document — the trace — is famously illustrated by Freud in a passage of *Civilisation and its Discontents*, where he describes the impressions made by a walk through Roman ruins. The same metaphor used by Freud to describe the structure of human soul can serve just as well to describe the complexity of the archiving process.[57] At

---

qui a pu se dérouler dans l'ordre de l'esprit ou dans l'ordre des choses; mais qu'elles sont apparues grâce à tout un jeu de relations qui caractérisent en propre le niveau discursif."

56 See Barthes, *Le Degré Zéro de l'écriture* [22], p. 20: "Toute trace écrite se précipite comme un élément chimique d'abord transparent, innocent et neutre, dans lequel la simple durée fait peu à peu apparaître tout un passé en suspension, toute une cryptographie de plus en plus dense."

57 See Freud, *Das Unbehagen in der Kultur* [55], pp. 35-37: "Wir greifen etwa die Entwicklung der ewigen Stadt als Beispiel auf. Historiker belehren uns, das älteste Rom war die Roma quadrata, eine umzäunte Ansiedlung auf dem Palatin. Dann folgte die

first, Freud recapitulates the historical evolution of Rome following traditional historiography: Roma quadrata, Septimontium, Republic, Early Empire. He then notes that what he sees while wandering through the city are ruins, but of what has been rebuilt over the original buildings after fire and destruction. All remnants of Ancient Rome appear entangled in present, Renaissance and modern Rome, while many ancient remnants remain buried below the ground.

"Ruinen, aber nicht ihrer selbst" — ruins, but not of themselves, writes Freud: what is decaying here is not the original, but what modernity has made of old times ("nicht ihrer selbst, sondern ihrer Erneuerungen aus späteren Zeiten"). What was once buried is re-emerging, amidst the hustle and bustle of the present city ("in das Gewirre der Großstadt"). All the different layers of the city's history are intertwined, not to be dissociated from each other. The vocabulary used by Freud is not one that suggests an entirely organic growth of old and new together. It evokes moments of violence in both directions: violence imposed by the modern on the old ("Brände und Zerstörungen") as well as an unwanted emergence of the old in the new ("Einsprengungen" as an unexpected disturbance). In the lines that follow, Freud expresses a form of distress: the real Rome he is visiting is nothing at all like what he had read about it beforehand. It consists of an accumulation of disorderly traces in which it is impossible to distinguish what is antique from what is fake, what is old from what is new. Rome and its thousand intertwined layers call for some kind of order if you want to gain orientation. But Freud knows that the order he would need is not necessarily the same as what someone else would wish for. Any definition of order is bound by the requirements of one's own temporality. And it is, in fact, tempting to let nature have its way and let plants grow on old stones as they see fit, to let the natural decaying process take place undisturbed by the wish to impose rational

Phase des Septimontium, eine Vereinigung der Niederlassungen auf den einzelnen Hügeln, darauf die Stadt, die durch die Servianische Mauer begrenzt wurde, und noch später, nach all den Umwandlungen der republikanischen und der frühen Kaiserzeit die Stadt, die Kaiser Aurelianus durch seine Mauern umschloß. [...] Was jetzt diese Stellen einnimmt, sind Ruinen, aber nicht ihrer selbst, sondern ihrer Erneuerungen aus späteren Zeiten nach Bränden und Zerstörungen. Es bedarf kaum noch einer besonderen Erwähnung, daß alle diese Überreste des alten Roms als Einsprengungen in das Gewirre einer Großstadt aus den letzten Jahrhunderten seit der Renaissance erscheinen. Manches Alte ist gewiß noch im Boden der Stadt oder unter ihren modernen Bauwerken begraben."

recording and preservation strategies that are anchored in a time and space, and which are all the harder to define as one stands amidst it.

Applying this perspective to text as a trace of times past can be enlightening in many ways, revealing the tension between cultural constructs and natural evolutions of materiality in general. Gaining clarity on origin, transmission and perspective plays as much of a role in the archiving process as ensuring the right humidity, light and temperature: one needs both the one and the other, as well as awareness of the fact that neither of them will provide conditions that will make it possible to keep any documentary trace unchanged forever, nor to keep all of them. Change it must, and change it will.

Archive is a moving order of traces, caught in an ongoing process of self-definition. Distinguishing what happens to the single documentary trace and to archives as an institution does not make much difference: what matters in both cases is the archiving process. This process is a dynamic dialectics of preservation and destruction, destruction and preservation, that affects any trace of the past we might be tempted to interpret, including texts. Freud also points to the difference between the transmission discourse and the image that the city of Rome gives itself. The question is not about assigning blame to either the architects who covered ruins in concrete or the historians who describe what might be an imaginary city of the past: it is about accepting that we will never know what Ancient Rome looked like, and still make sense of Rome and of what historians and novelists write about it. Rome will always be Rome, and always be another Rome.

In this chapter, I have considered access to text from the point of view of archives. Although they have been in use for a long time, archival processes still challenge our vision of the past. In the archival profession, the increase in document mass in the 20th century, then the digitisation of records and sources in the late 20th century, have led to a shift in practices and the development of theories that encompass a more complex media and take into account the wider socio-cultural relevance of archiving. Philosophers, on the other hand, have come to challenge the notion of origin at the core of archival evidence. But both perspectives recognise the key role of archives as documents, as collections and as institutions in constructing discursivities throughout our history.

Archives provide structured access to text, and they have been doing so for a long time, alongside political evolutions. Taking them for granted is a cultural privilege, but one that should not prevent us from trying to understand underlying techniques (increasingly complex in the digital context) and epistemological challenges (such as the status of traces from the past), while making use of this privilege.