



Ca' Foscari  
University  
of Venice

Master's Degree Programme  
in Digital and Public  
Humanities  
ex D.M. 270/2004

Final Thesis

**Planning and Experimenting a  
Digitization Project for  
Small Cultural Institutes**

The Case of the Ateneo Veneto  
Early-Printed Medicine Books

**Supervisor**  
Ch. Prof. Dorit Raines

**Graduand**  
Kristian Gandin  
Matriculation Number  
873821

**Academic Year**  
2022 / 2023



# DEDICATION

This thesis is dedicated to my dearest grandparents Luisa, Maria and Domenico, and to the memory of my grandpa Guido. You are the most precious heritage to preserve.

To my parents Mirca and Roberto, who have accompanied me with love and effort in this journey. Your strength and perseverance are my source of inspiration.

To Ketty and Sergio. Best of luck for your future projects.

To all my friends, who have shared their words of advice and waited for the end of this work.

And to my Almighty God. May I always remember that You love me more than I can.



# ACKNOWLEDGEMENTS

This thesis and the project behind it would not have been possible without the exceptional support of my supervisor, Prof. Dorit Raines. Her enthusiasm, knowledge, resourcefulness and exacting attention to detail have guided my work far beyond what I could have imagined. Thank you for directing me in moments of indecision and for continuous teachings and pleasant conversations far beyond the scope of your duties.

I am extremely grateful to the Ateneo Veneto for the unforgettable and rewarding experience I could enjoy. In this regard, I express my warmest thanks to the President Dr. Antonella Magaraggia and the Vice-president Esq. Alvise Bragadin, who trustfully allowed me to work on and get inspired by the first digitization project of the institution. I want also to express my gratitude to Daria, Valerio, Marina, Silva, and Elena, who contributed in various ways to the success of the project. It has been an honour to be a temporary member of your staff.

I extend my grateful thanks to the staff of the digital platform Phaidra in the person of Gianluca Drago and Linda Cappellato, who provided me with technical support and valuable advice during the digitization project.

I also would like to thank all the people who helped and supported me in the project experience or the writing of this thesis. Thanks, Dad, for your technical advice, design resourcefulness and practical work. Thank you, Giorgio, for the valuable exchange of ideas about the project and Elisabetta for your tips on the thesis layout. Thank you so much, Samuel and Sebastiano, for sharing part of this journey with me and thanks a lot, Sara, for the many study days spent together in the library.

Finally, I would like to express my profound gratitude to my University, which, through the professors, faces met, experiences, and opportunities allowed me in these five years to become what I am now.



# ABSTRACT

The funds of the Next Generation EU Plan and the National Recovery and Resilience Plan allocated to promote the digital transformation of the cultural sector in Italy present a challenge to small-sized organizations. These tend to embrace a resourceful approach to compensate for their inherent weaknesses, especially the absence of state-of-the-art equipment, yet this adaptive approach may potentially lead to a diminution in the quality of their outputs. In searching for funding opportunities, they must surmount a labyrinth of structural impediments of economic, logistical, and HR dimensions. The thesis delves into the realm of cultural digitization planning, with a specific emphasis on small cultural institutes, using as an example a project promoted by the Ateneo Veneto in Venice, Italy. It proposes a meticulous examination of the requisite steps in digital project planning and underscores the pivotal decisions that these institutes must undertake. The dissertation comprises four chapters. The first one scrutinizes the features and objectives underpinning the digitization phenomenon. The second chapter deals with the best practices that pertain to small institutions embarking on digitization projects, elucidating the minimum standards prescribed by national guidelines. The third and fourth chapters are dedicated to the intricate processes entailed in the digitization and dissemination through a long-term repository of cultural artefacts, in particular early-printed books.

Keywords: cultural digitization, project planning, small cultural institutes, early-printed books, Phaidra.





# TABLE OF CONTENTS

INTRODUCTION.....	11
CHAPTER 1: INTO THE REALM OF DIGITIZATION .....	15
1.1    DIGITIZATION, DIGITALIZATION, DIGITAL TRANSFORMATION. THE NEED FOR A CLARIFICATION .....	17
1.2    THE NATIONAL PLAN FOR THE DIGITALISATION OF CULTURAL HERITAGE. VISION AND STRATEGY .....	24
1.3    WHAT OPPORTUNITIES FOR SMALL CULTURAL INSTITUTES?.....	31
CHAPTER 2: PLANNING AND BEST PRACTICES FOR A DIGITIZATION PROJECT .....	39
2.1    OBJECTIVES .....	40
2.2    SELECTION OF DOCUMENTS.....	44
2.3    INSTITUTIONS AND PROJECT TEAM.....	51
2.4    PLANNING THE PROJECT.....	53
2.4.1    The Choice of the Long-Term Archival Platform.....	54
2.4.2    Budgeting and the Choice of In-House Digitization .....	57
2.4.3    The Choice of Hardware and its Optimization .....	60
2.4.3.1    Planetary Scanner .....	61
2.4.3.2    Computer and Monitor.....	66
2.4.3.3    Prototyping of the Digitization Workstation .....	67
2.4.3.4    Other Equipment.....	72
2.4.4    Data Management.....	73
2.5    GENERAL WORKFLOW AND TIMELINE OF THE PROJECT .....	76
CHAPTER 3: DIGITIZING RARE AND EARLY-PRINTED BOOKS.....	79
3.1    PRELIMINARY ACTIVITIES .....	80
3.2    CALIBRATION OF THE ACQUISITION SYSTEM .....	83
3.3    SCANNING.....	88
3.4    QUALITY CONTROL.....	95
3.5    POST-CAPTURE IMAGE PROCESSING AND COLOUR CORRECTION .....	97
3.6    PRODUCTION OF DERIVATIVE FILES.....	101
3.7    A PARTICULAR DIGITALIZATION CASE: ACQUAPENDENTE'S <i>DE FORMATO FOETU</i> .....	101
3.7    SOME CONSIDERATIONS.....	105
CHAPTER 4: REACHING OUT TO THE PUBLIC. THE PHAIDRA REPOSITORY .....	107
4.1    THE REPOSITORY.....	108
4.2    METADATA.....	116
4.3    THE PUBLICATION PROCESS.....	123
4.4    SOME CONSIDERATIONS.....	128
CONCLUSION.....	131

REFERENCES .....135

LIST OF TABLES AND FIGURES .....145

# INTRODUCTION

*“How far that little candle throws his beams!  
So shines a good deed in a naughty world.”*

William Shakespeare, *The Merchant of Venice*, 1594-1598, Act 5, Sc.1

In 2022, a small cultural institute located in Venice, Italy, called Ateneo Veneto, took advantage of the funding call *Cultura digitale* provided by the Fondazione di Venezia and planned its first digitization project. One year later, after careful evaluations and repeated tests, more than a hundred early-printed books have been successfully digitized and archived in the open-access platform Phaidra developed by the University of Padua. Inspired by the significant expertise gained during the project, this thesis discusses cultural digitization and investigates the type of digital strategies that small cultural institutes can plan.

The author had the privilege of witnessing the planning, development and conclusion of the “Digitization Project of a Hundred Rare and Valuable Medicine Books”. He was the technical manager and operator of the project run by the Ateneo Veneto under the scientific responsibility of Dorit Raines, Academic Librarian and professor at Ca’ Foscari University of Venice. The project was an intensive, multi-step work that required meticulous planning. On several occasions, national and international guidelines needed to be adjusted in order to underscore pivotal decisions and formulate an effective digital strategy, even if the structural limits impeded the institute. While planning and experimenting, the author thought that lots of the problems and outcomes encountered during this digitization project could be compared to those that other small cultural institutions have to face.

In this context, planning a digital strategy that can fulfil the requirements expressed in the *National Plan for the Digitalisation of Cultural Heritage* (NPD) could indeed represent a real challenge.<sup>1</sup> Restricted staff – which can hardly contemplate a variety of professionals –, economic reasons – which will preclude the purchase of state-of-the-art equipment –, and logistical aspects – the working space may not be consistent with standards<sup>2</sup> – lead sometimes to the formation of a labyrinth of inherent

---

<sup>1</sup> Ministero della cultura, Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library, ed. (2022), p. 7.

<sup>2</sup> Federal Agencies Digitization Guidelines Initiative (FADGI) (2023), p. 23.

weaknesses. These must not prevent small institutes from seizing funding opportunities and planning digitization projects, especially now that digital transformation is lately at the heart of national policies. At the same time, the phenomenon of small cultural institutes constitutes a relevant issue since they are the hard core of the Italian cultural scenario. The number of these cultural organizations could be around 45.1% of Italian museums<sup>3</sup> and 63.8% of libraries<sup>4</sup>, which are the ones that have a staff composed of less than five operators. This is the reason why their case is relevant when discussing digitization in Italy.

This thesis is structured as follows. Following this introduction, Chapter 1 establishes the theoretical and conceptual frameworks of the dissertation. After a necessary terminological disambiguation and definition of digitization practices, the chapter delves into the national digital strategy expressed in the NRRP and pinpoints the consequences of digital transformation for the cultural sector. The vision and the strategy of the NPD will serve as a premise for developing a digital strategy for small cultural institutes, whose specificities and growth opportunities are taken into consideration in the last part of the chapter.

Chapter 2 deals with digital project planning. Starting from the definition of the project's objectives and the selection of the originals, passing through the team composition, the chapter gets to the crux of the matter and explicates the requisite decisions underpinning a digitization project. The choice of a long-term digital platform, budgeting, as well as the purchase and optimization of the equipment, are contextualized inside the possibilities of small cultural institutes and the requirements and standards established by international and national guidelines. All these steps are explained through the lens of the project planned by the Ateneo Veneto.

Chapter 3 scrutinizes the methodologies and good practices entailed in the digitization phase. It examines the procedures for the creation of digital objects, such as the originals' handling, the calibrating of the workstation, the scanning and the post-capture image processing. While illustrating the process, the chapter clarifies the expedients and solutions adopted during the Ateneo Veneto's project to compensate for the absence of state-of-the-art instrumentation.

Chapter 4 focuses on the final strategic passage required by a digitization project, namely the online publication of digital objects. It explains the underlying reasons for the choice of Phaidra, the permanent digital platform managed by the University of Padua and developed in collaboration

---

<sup>3</sup> Elaboration by the author on data available on Istituto Nazionale di Statistica (ISTAT), *Statistiche culturali – Anno 2021* (2022) (“Tavole Musei”, Tav. 1.12).

<sup>4</sup> Elaboration by the author on data available on Istituto Nazionale di Statistica (ISTAT), *Statistiche culturali – Anno 2021* (2022) (“Tavole Biblioteche”, Tav. 4.14).

with the University of Vienna. In this regard, a description of the platform's functionalities clarifies the principles of permanent hosting, interoperability, and cultural dissemination. Furthermore, the chapter presents the fundamental passage of metadata management and its publication together with the digital objects inside the repository.

A final assessment is undertaken in the conclusion and reveals that, in the attempt to balance persistent limits, resourceful approaches are at the heart of digitization and good planning can represent a trump card not only for the project but also for the renewal of small cultural institutes.



# CHAPTER 1: INTO THE REALM OF DIGITIZATION

The contemporary world is marked by the persistent use of digital technology. Current lifestyle habits are impregnated by the massive use of interconnected electronic devices that are increasingly equipped with artificial intelligence (AI). The computational power has stopped being inherent in uncomfortable, voluminous machines for many years and has conquered new, invasive forms. Laptops, tablets, smartphones, consoles, cameras, smart TVs, connected cars, and digital clocks, just to name a few, have appeared alongside traditional computers, which have also undergone recurrent changes. The march of technological progress has culminated in the ubiquity of smartphones, now a pivotal tool for modern communication, nestled in the pockets of nearly everyone. A varied series of services and facilities that were unthinkable until a few decades ago are now brought into the hands of users, streamlining mechanisms to travel, read news, listen to podcasts, access streaming entertainment platforms, learn, even buy groceries, order takeaway food, manage the home automation, make financial transactions. Altogether, humanity is undergoing a digital transformation that is comparable - if not superior because of the impact it is having on everyone's life - to the race to industrialization that characterized humanity more than two centuries ago.<sup>5</sup> The digital world is influencing more and more the current perception of reality. Internet, with web platforms and digital technologies, affects human beings, introducing for example “new communication scenarios, sharing and exchange.”<sup>6</sup> It involves new habits, creates new needs, and stresses new necessities. Among these necessities it can also be recognized the practice of reproducing the cultural heritage into the digital world, trying to take advantage of the benefits that the IT systems can offer.

In the meantime, great attention was paid to digital transformation during the COVID-19 pandemic, when the world had to stop and fight the health emergency. Italy was the first European country to be affected in early 2020, becoming the first country worldwide to introduce a national lockdown. During repeated periods, citizens were confined to their homes and were forbidden from any unjustified movement, while all sport activities and events and even the closure of museums, places of culture and sports centres were suspended. For more than two years measures

---

<sup>5</sup> FTA Online News (2020). The development of computer sciences and IT technologies has marked the birth of the Fourth Industrial Revolution. Its starting date has not yet been established, but the effects are more than obvious. Internet, cloud computing and the no longer so much futuristic interaction of cyber-physical systems are gradually investing all areas of contemporary man's life, starting from the industrial sectors. It is the age of digital transformation.

<sup>6</sup> Ministero della cultura, Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library, ed. (2022), p. 3. Author's translation.

were taken to contain the spread of the virus, with procedures of isolation, social distancing, compulsory use of protective masks, hand sanitization and vaccination campaigns. The return to normality was then accompanied by the introduction of the “Green Pass”, the EU Digital COVID certificate that was required to participate in sports and cultural events, as well as to access indoor places such as bars, restaurants and gyms. The restrictions imposed by COVID-19 containment measures and the difficulties of communication led to an increasing demand for digital technologies. Work environments adopted more and more smart working and increased e-commerce, evaluating digital transformation strategies at the same time. Schools and universities were compelled to pass from face-to-face teaching to remote learning. Public institutions that by their nature are called to dialogue with the public, especially cultural institutions, have activated social media accounts and promoted digitization projects. This clear attempt to continue institutional activities at a distance accelerated digital transformation.<sup>7</sup>

It is in this pandemic context that some possibilities and advantages of the digital world might have been recognized by public institutions. Not coincidentally, the recovery plan launched by the European Union in support of member countries, “NextGenerationEU”, has conceived the digital as one of its strategic objectives.<sup>8</sup> In the wake of this recommendation, the Italian Government has approved the National Recovery and Resilience Plan (NRRP) placing the digital transformation as the primary objective of the First Mission of the plan.<sup>9</sup> Digitization (or digitalization?), which was already a burning issue of discussion, has thus become one of the primary objectives on the agenda of the national strategy. What is particularly interesting is that the plan refers to digital transformation as a transversal necessity. Not least, great attention has been paid to cultural

---

<sup>7</sup> Gandin, K. (2021), p. 33.

<sup>8</sup> European Union, *NextGenerationEU* (n.d). Next Generation EU (NGEU) can be considered the EU Marshall Plan to support the EU member states in recovering from the COVID-19 pandemic. The investment is equal to 750 billion euros and was adopted on 14 December 2020. It will operate from 2021 to 2026. The entire European Commission initiative is structured around three pillars: support for Member States for investment and reform, boosting the EU economy by encouraging private investment, and learning from the pandemic. NGEU will also bring additional funds to other European programmes. One of the most interesting programmes for this dissertation is Horizon Europe, which is the 7-year EU scientific research initiative and successor of Horizon 2020. With a total budget of 95.5 billion euros, including 5.4 bn allocated by the NGEU, Horizon Europe is the world’s largest transnational research and innovation programme. The program’s recommendations for digital transformation will be taken into account in this chapter.

<sup>9</sup> European Commission, Italian Government, *National Recovery and Resilience Plan* (2021), p. 87. The National Recovery and Resilience Plan (NRRP), called *Italia Domani* (translated, “Italy Tomorrow”) traces the objectives, reforms and investments that Italy intends to achieve through the use of the plan Next Generation EU. Italy is the biggest beneficiary, by absolute value, of the NGEU funds. The total amount of resources is equal to 191.5 billion euros, which must be integrated with a Complementary Fund, endowed with additional resources equal to 30.6 billion euros. Six strategic missions are planned: “Digitalization, Innovation, Competitiveness, and Culture” (49 bn), “Green revolution and Ecological transition” (68.6 bn), “Infrastructure for sustainable mobility” (31.5 bn), “Education and Research” (31.9 bn), “Inclusion and Cohesion” (22.6 bn), and “Health” (18.5 bn). Amounts include both fundings allocated by the NRRP and the Complementary Fund.



digitization, which is the scope of interest of this dissertation. Adopting the prospect of a small cultural institute, it will be interesting to investigate the national purposes of digital transformation as expressed in the NRRP and the National Plan for the Digitalisation of Cultural Heritage announced by the Italian Ministry of Culture. It will be first necessary to make terminological disambiguation of processes which in documents and common language are usually confused or considered equivalent. All this will then serve to try to understand in this first chapter how cultural institutions, in particular small ones, fit into the realm of digitization, what opportunities are offered to them and what actions they can take.

## 1.1 Digitization, Digitalization, Digital Transformation. The Need for a Clarification

The National Plan for the Digitalisation of Cultural Heritage (NPD) represents the concretization of the NRRP's digital transformation strategy in the field of cultural heritage. The plan is part of the First Mission of the document, called "Digitalisation, Innovation, Competitiveness, Culture and Tourism", which aims to give a decisive impulse to the relaunch of the competitiveness and productivity of the country through digital technologies. The digitization strategy of the Ministry of Culture in favour of the support and revitalization of cultural heritage is in reality one of the numerous interventions planned for the period 2021-2026 within the first Mission, representing almost 1% (500 million euros) of the total investment of the mission.<sup>10</sup> It is part of the third Component, called "Tourism and Culture 4.0", where other initiatives focused on the digital revitalization of these economic sectors are designed to improve cultural operators' capacity building and create a digital tourism hub. Much more funds are allocated for the digitalization and bureaucratic simplification of the public administration (First Component of Mission 1) and for the digital transition of manufacturing sectors, as well as more fluid national connectivity for citizens, public institutions and enterprises (Second Component). Such different interventions find their common denominator in the term "digital transformation", which appears not to be a peculiar phenomenon only for the cultural sector. Digital transformation is a large-scale process that affects the entire economy of the country. This assumption is further confirmed by the expectations of improvement held by the European Union concerning the Digital Economy and Society Index

---

<sup>10</sup> Author's elaboration on data available on European Commission, Italian Government, *National Recovery and Resilience Plan* (2021), p. 108.

(DESI), a classification of EU Member States according to their level of digitalization.<sup>11</sup> The 2022 index ranks Italy 18<sup>th</sup>, which is below the European average. This fact highlights the strategic importance of the NRRP to stimulate Italian growth in several spheres, including the acquisition of digital skills, increase of connectivity, integration of digital technologies and digital public services. The race for digital transformation characterizes more sectors than what can be expected. Prior to focusing on the digital strategies adopted by the NPD, it seems more than necessary to define the phenomenon. One could wonder whether all these interventions that are expressed in the NRRP refer to the same concept of digital transformation or whether specificities persist in the different fields of application. Is digitalization then a synonym for digital transformation? What is the difference between terms that seem so similar, namely digitalization and digitization?

As already remarked by Nicola Barbuti<sup>12</sup>, researcher in Archival studies at the University of Bari and expert in Digital Humanities, national legislation struggles to adopt and present a clear terminology about the phenomenon without leading to ambiguity. The Italian language does not even help to clear up the confusion because of the lack of specific terms. The words “digitalizzazione” and “trasformazione digitale” are commonly used promiscuously to express at the same time the technical process meant to create digital objects and the use of digital technologies to introduce changes from an administrative and productive perspective, as well as to designate all those implications on the social and economic spheres. There is no unambiguous interpretation of terms even in the English language. However, a reference to the Anglo-Saxon world would facilitate this semantic investigation.

The literature frequently distinguishes between “digitization”, “digitalization”, and “digital transformation”. Digitization, which derives from the root “digit”, literally means to “convert into a sequence of digits” and describes the original process of converting information made of continuous, analogue data into discrete, digital values, so that they can be interpreted and processed by modern computer according to Boolean algebra.<sup>13</sup> More broadly, digitization can be considered

---

<sup>11</sup> European Commission, *Italy in the Digital Economy and Society Index (2022)*. Italy’s score is 49.3 while the EU average is equal to 52.3. The result is all the more intolerable considering that Italy is the third economy of the European Union. Both national and community production would benefit from Italian technological efficiency. Nevertheless, Italy is advancing at a very fast pace if it is taken into account the progress of its DESI score over the past five years (Italy ranked 25<sup>th</sup> in 2020).

<sup>12</sup> Nicola Barbuti is CEO of the university spin-off digitization project D.A.BI.MUS. S.r.l. and co-inventor of the National Patent of the ICRPad system for intelligent digital recognition and graphic matching of digital resources related to historical documents. Moreover, his book *La digitalizzazione dei beni documentali. Metodi, tecniche, buone prassi* (translated “The digitization of documentary items. Methods, techniques, good practices”), published in 2022, was taken as a reference for the formulation of this dissertation.

<sup>13</sup> Online Etymology Dictionary (2023), under “digital”. The term “digital”, connected to “digit”, comes from the Latin *digitus* (finger or toe) and literally means “pertaining to fingers” but assumed a numerical sense from the mid-15<sup>th</sup> century to indicate numbers below ten that were precisely counted on fingers. Subsequently, the term was adopted to

as the set of all the operations that take advantage of the possibilities of digital technologies to create and organize information in a computer-readable format. This format can be achieved either by converting existing data presented in analogue form – as the literal meaning expresses – or by producing it directly through digital creation processes. As a result, digitization can create two typologies of digital objects: digital resources derived from analogue entities, such as raster images (produced with scanners or professional cameras) or converted analogue audio (produced using a sampler), and resources called “born digital”, such as vector graphics and digital audio.<sup>14</sup> Reproduction using an analogue-to-digital converter is the most common digitization technique and allows the creation of facsimiles from images, sounds, and videos. The final result is a digital object that has the feature of originality. In the case of predominantly textual items, digitization would further use specific technologies, such as OCR<sup>15</sup>, to transcribe and create digital texts. Specific aspects of digitization are explained subsequently.

Much more complex and ambiguous is the concept of digitalization. In this case, literature has difficulty in finding a unanimous definition because the term is closely associated with the previous one. Nevertheless, it is possible to find some peculiar features. Digitalization may be considered as the subsequent step of digitization. From an etymological perspective, the term focuses on the word “digital”, as if the existence of digits can be taken for granted. For this reason, this process would not correlate computers with the external world through the production of digital information – as in the case of digitization. On the contrary, it would indeed use digital technologies in the opposite direction to influence and modify the external world. The American business magazine *Forbes*<sup>16</sup> and the glossary compiled by the American technological research and consulting firm Gartner, Inc.<sup>17</sup> agree to use the term “digitalization” concerning the innovation in industrial production processes fostered by modern technologies. Business operations can be redesigned and organizational processes can be made more efficient by the use of computers and IT technologies,

---

represent discrete values as opposed to continuous numbers and assumed the acceptance that is currently being used “of something relating to computers” only after c. 1945. At that time, computers started to be programmed using structured sequences of zeros and ones, namely the binary system, which forms the core of the Boolean algebra and enables data processing through algorithms.

<sup>14</sup> Barbuti, N. (2022) and Commissione Tecnologie Digitali di ICOM Italia, *Glossario Museologia digitale e nuove tecnologie* (2020), under “Artefatto digitale / computazionale”.

<sup>15</sup> IBM Blog, *What Is Optical Character Recognition (OCR)?* (5 January 2022). OCR or Optical Character Recognition is the mechanical or automatic recognition of handwritten, typed or printed text and its conversion into machine-readable text. OCR systems use a combination of hardware and software and can take advantage of artificial intelligence (AI) to implement more advanced methods of intelligent character recognition (ICR), like identifying languages or handwriting styles. The process of OCR is most commonly used to turn hard copy legal or historical documents into PDF documents so that users can edit, format and search the documents as if created with a word processor.

<sup>16</sup> Bloomberg, J. (2018).

<sup>17</sup> Gartner, *Gartner Glossary* (2023), under “Digitalization”.

providing “new revenue and value-producing opportunities.”<sup>18</sup> Automation is a major part of digitalization and contributes to changing the business model by shifting work roles and adopting new productive or administrative techniques. Moreover, the term is widely used in public-sector organizations. Consider for example a paperless (or paper-free) office. This case illustrates the correlation between digitalization and digitizing operations. A paperless office is a work environment in which the use of paper is eliminated or greatly reduced.<sup>19</sup> All the decisions connected to paper reduction, electronic automation, redesign of working roles, and more generally the innovation of production processes deal with digitalization. This is then realised thanks to the conversion of papers into digital form and the adoption of an electronic infrastructure to enable the production, transmission and storage of digital information (e.g., the use of PDFs, databases, emails, digital signatures, digital platforms, web services...). This series of operations concerns digitization.

When talking about digital transformation, the evidence suggests that it does not consist of either a series of operations or a single initiative. The forementioned sources tend to narrow the scope of application of this word to the organizational dimension, stating that it pertains to the invention or rethink of new digital business models, with particular attention to customer-driven strategies and the social impact they have. In reality, if the term is literally interpreted, it can be understood as a complex phenomenon caused by the application of digital technologies, affecting the whole society, and producing changes not only in the technological and managerial dimensions but also in the cultural, social and even creative aspects of human society. For this reason, digital transformation is broader than the preceding two processes of digitization and digitalization and originates from their combined action. It is the phenomenon that is causing the ongoing Digital Revolution or Fourth Industrial Revolution.<sup>20</sup> Returning to the previous example of the paperless office, digital transformation concerns IT optimization strategies and social implications for both employees and customers, who for example would need to possess different IT capabilities and competencies to access and fill an electronic form. Any impact on the individual life, including the technology demand and social interactions, but also disadvantages of using digital devices, is part of digital transformation.

Digitization, digitalization and digital transformation refer therefore to different processes. The scope of investigation of this dissertation is limited to the practices of digitization of cultural items, with particular attention to early-printed books. However, this is just one, minor declination of a

---

<sup>18</sup> *Ibid.*

<sup>19</sup> Wikipedia. The Free Encyclopedia (2023), under “Paperless office”.

<sup>20</sup> *Cfr.* [Footnote 5](#).

vast, polyhedral phenomenon that the NRRP attempts to intercept. The plan primarily aims to use digital technologies to promote the bureaucratic simplification of the public administration (First Mission, Component 1) and the digital transition and automation of the production sectors (First Mission, Component 2), which are both aspects of digitalization. The same approach is then adopted for the sectors of infrastructures (Second and Third Missions), schools (Fourth Mission) and health (Fifth and Sixth Mission), in order to increase productivity, streamline processes, and improve quality. All these governmental strategies will include digitization practices, following the model of the paperless or paper-free office. The digital transformation would also affect the tourism and cultural sectors (First Mission, Component 3), thanks to a series of initiatives aimed at creating digital strategies, revitalization plans, and the development of both databases and digital collections. It is important to note that the predominance of digitizing practices in the cultural sector does not exclude the implementation of other processes. Digital transformation has to embrace all the operational areas of cultural heritage (from logistics to collection management, from HR training to marketing and communication, from service design to management models, etc.).<sup>21</sup> For this reason, in the cultural sector, the digitization of cultural artefacts has to be accompanied by the digitalization of administrative, managerial and productive processes, with the aim to create new modalities of interaction with the public, rethink the operational models and improve the conservation modalities.

With regard to cultural digitization, some peculiarities may be highlighted in comparison to the practices adopted in other contexts. Firstly, cultural items are characterized by preciousness, rarity or uniqueness and for this reason they are regulated by state protection measures. The techniques of digital reproduction must be non-invasive in order to preserve the support. In other cases, there is no reproduction because the use of technologies leads artists or other creative operators to create born-digital artefacts. Secondly, the process must follow the principles of trustworthiness and authenticity, so that the public can access faithful and certified resources. Thirdly, cultural items are typically meant to be published in comparison to other type of objects. Digitization is in fact considered as part of cultural institutes' strategies formulated to accomplish their public mission. Organizations would also ensure the durability of the objects so that they can be accessed in the long term. Finally, it is important to provide a contextualization, namely a set of data and metadata that describes the physical object, its history, and its cultural value. In contrast to many practices that aim at digital reformatting – in other words, the creation of a surrogate of the original –,

---

<sup>21</sup> Ministero della cultura, Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library, ed. (2022), p. 81.

cultural digitization creates digital objects with the aim of an integration with the original, except for those cases where conservation limits the fruition. It would be therefore essential to provide the objects with all the information that is needed for a proper consultation and interpretation or that is not possible to digitize.

Another issue arises when discussing what digitization is and concerns the fact that by its nature it is not a static phenomenon. Digitization has been described as the process of converting the analogue into the digital by use of digital technologies, but what are digital technologies? It is evident that technological tools, methodologies and techniques, which all influence the final appearance of digital objects, evolve according to technical advancements. New applications in the field, such as the development of the Semantic Web, have recently increased the computational power of digital devices, implicating new operating standards.<sup>22</sup> In the future, discoveries and new technological inventions may offer new possibilities for human society and perhaps digital scent technology will allow to digitize perfumes or holography will succeed in generating sophisticated three-dimensional images. Digitization is a fluid concept in constant evolution. It follows technical changes, adopts their applications, accepts their improvements and abandons them if new higher-quality technologies are developed. This evolution is evident in the case of the cultural sector too. Digital objects are increasingly recreating in a realistic way the original appearance of physical artefacts. There are no more only bidimensional objects, such as images. Laser scanning and three-dimensional photogrammetry are creating accurate 3D models of objects and environments using photos taken from different angles. The use of Augmented Reality (AR), Virtual Reality (VR) or even Immersive Reality have established themselves during the pandemic and today constitute virtuous examples of cultural heritage's fruition from home or new experiential forms in presence.<sup>23</sup> In the case of digital art and valuable digital objects, technological advancements have then developed digital systems that use non-fungible tokens (NFTs) and blockchains to extend the criteria of uniqueness and scarcity even for purely digital objects, making them collectable.<sup>24</sup> Digitization practices differentiate over time according to the type of cultural item and the technological progress. In the attempt to intercept exigences of creation, fruition and preservation that are also changing over time, digitization can be considered a potentially never-ending process.

---

<sup>22</sup> Wikipedia. The Free Encyclopedia (2023), under "Semantic Web". The Semantic Web or Web 3.0 is an extension of the World Wide Web developed in 2006 in order to make Internet data machine-readable. Standards set by the World Wide Web Consortium (W3C), such as the Resource Description Framework (RDF), the Web Ontology Language (OWL), and Extensible Markup Language (XML), aim to transcend the limitations of the Hypertext Markup Language (HTML), which is usually used to format webpages and hyperlinks. By adopting RDF, OWL, and XML standards, webpages can properly describe specific categories, contents and entities.

<sup>23</sup> Cecotti, H. (2022), p. 84.

<sup>24</sup> Trček, D. (2022), p. 6.

Despite this mutability, it is still important to recognize some constant traits in digitization, at least in the cultural sector. Taking inspiration from the guide that Patrick Sahle, professor of Digital Humanities at the University of Wuppertal, formulated for digital scholarly editions, four principles can be adapted and highlighted.<sup>25</sup> The first one is completeness, which means that the digital object has to be a “full representation of the subject in question”. The final form of the object would depend on the purpose of the digitization campaign, which may be meant for the creation of facsimiles or a digital scholarly edition. The second principle is to use a critical approach. Trustworthiness and usefulness can be achieved only by adopting rules for the processing of the materials and following national and international recommendations. Academic quality is then required and would derive from the scrupulous application of the aforementioned rules. Quality and accuracy make the digital object reliable and useful for the public. Finally and most importantly, the digitization campaign must follow what Sahle defines as the “digital paradigm”. In other words, the creation of digital objects must make use of the possibilities offered by digital technologies. The use would be conditioned by economic factors and by the purpose of the project, but in any case, the application of the digital paradigm would not make the object “printable without a major loss of content and functionality”.<sup>26</sup>

Digitization can be finally seen as a peculiar cultural practice of the contemporary era. From a historical perspective, it will be considered as part of the long process of migration of human knowledge from one vehicle to another that has preserved and transmitted cultural heritage through the centuries.<sup>27</sup> This perspective would clearly show the burden that digitization has in relationship with the past and its strategic and preservative mission towards the future. It is not sufficient to consider digitization a mere means of cultural replication, though. The European Union has increasingly stated the exigency to create a fertile representation of modern society and its technological evolutionary processes.<sup>28</sup> In other words, the conversion of cultural items into digital form has to encourage the creation of a real Digital Cultural Heritage (DCH), made of “not only visual and structural information, but also stories and experiences (stored in language data),

---

<sup>25</sup> Sahle, P. (2016), p. 19. The professor conceives the four rules mentioned above in order to define digital scholarly editions. In doing it, he distinguishes between digitized and digital editions, asserting that the latter “cannot be given in print without significant loss of content and functionality” and is “guided by a digital paradigm in their theory, method and practice”. It is important to remember that the subject in question is very specific and refers to academic textual items produced with a critical encoding. For this reason, the mere process of scanning would not be adequate in this context. In others, instead, the creation of facsimiles is the first step towards the digitization of cultural objects.

<sup>26</sup> *Ibid.*

<sup>27</sup> Docs Italia, *Linee guida per la digitalizzazione del patrimonio culturale* (2022), p. 6.

<sup>28</sup> Barbuti, N. (2022).

together with their cultural and socio-historical context, as well as their evolution over time”.<sup>29</sup> This awareness would enable digital objects to become complex cultural resources that offer a contemporary, digital interpretation of the originals.

## 1.2 The *National Plan for the Digitalisation of Cultural Heritage*. Vision and Strategy

Along these lines, the Italian Ministry of Culture planned the birth of a national digital cultural heritage. This objective is expressed in the *National Plan for the Digitalisation of Cultural Heritage* (NPD), the document that creates the strategic context for the NRRP’s investment M1C3 1.1 *Digital strategies and platforms for cultural heritage*.<sup>30</sup> A specific institute, the “Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library”, was founded in 2020 to coordinate and promote programmes of digitization of the cultural heritage on behalf of the Ministry of Culture. The principal aim of the NPD is to bridge the existing “digital gap”, both in terms of infrastructure and systems and at the level of personnel skills. An overview of the Italian cultural heritage will clarify the reasons for the interventions.

Italy’s cultural heritage counts almost 29,000 places of culture.<sup>31</sup> There are more than 6,200 museums, monuments, and archaeological areas. Over 9,500 are archives, 12,000 are libraries and

---

<sup>29</sup> European Commission, *Curation of digital assets and advanced digitisation* (2020). The existence of a Digital Cultural Heritage (DCH) was defined by the EU Council in the Art. 2 of the conclusions of 21 May 2014 on “Cultural heritage as a strategic resource for a sustainable Europe”. On this occasion, the need for the preservation and enhancement was applied to all forms and aspects of cultural heritage – tangible, intangible and even digital (born digital and digitized). For more details, *cf.* European Union (2014).

<sup>30</sup> For the sake of clarity, this investment is part of the First Mission of the NRRP, “Digitalisation, innovation, competitiveness, culture and tourism”, Component 3, “Tourism and culture 4.0”, Measure 1, “Cultural Heritage for the next generation”. The investment (for a total investment of 500 million euros) consists of twelve autonomous projects, or sub-investments, complementary to each other and aimed at creating four different areas of services: enabling services, production services, services for document management and preservation, and access services. For more information, *cf.* <https://digitallibrary.cultura.gov.it/pnrr-cultura> and <https://pnrr.cultura.gov.it/misura-1-patrimonio-culturale-per-la-prossima-generazione/1-1-piattaforme-e-strategie-digitali-per-l'accesso-al-patrimonio-culturale/> [Accessed 17 September 2023].

<sup>31</sup> Italian places of culture are defined and regulated by Legislative Decree no. 42/2004, *videlicet* the *Code of Cultural Heritage and Landscape*. Art. 101 includes in the category of place of culture the typologies of: museum (“a permanent structure which acquires, catalogues, preserves, orders and exhibits cultural goods for educational and study purposes”); library (“a permanent structure which collects, catalogues and preserves an organized set of books, materials and information, in any case published or published on any medium, and ensures that they are consulted in order to promote reading and study”); archive (“a permanent structure that collects, inventories and preserves original documents of historical interest and ensures their consultation for study and research purposes”); archaeological area (“a site characterised by the presence of fossil remains or prehistoric or ancient artefacts or structures”); archaeological park (“a territorial area characterized by important archaeological evidence and the coexistence of historical, landscape or environmental values, equipped as an outdoor museum”); monumental complex (“a set of buildings built even in different periods, which with time have acquired, as a whole, an autonomous artistic, historical or ethno-anthropological relevance”). For more information, *cf.* [https://www.bossetiegatti.eu/info/norme/statali/2004\\_0042.htm](https://www.bossetiegatti.eu/info/norme/statali/2004_0042.htm) [Accessed 17 September 2023].



approximately 1,000 are institutes of the performing art sector.<sup>32</sup> Cultural heritage is considerable and it is also distributed over the whole national territory. Despite this, it has the weakness of being fragmentary and it predominantly consists of small and medium-sized institutes, as it is explained in next Subchapter.<sup>33</sup> This fact means that the majority of artworks and cultural items are kept inside not very large places, with obvious administrative, economic and managerial repercussions. It is from this complexity that the NPD plans its interventions. At the same time, it recognizes a backward situation in terms of digitization. Digitalization campaigns started systematically in the 1990s. The Ministry of Culture reports that:

“Overall, the IT systems managed by the Central Institutes of the Ministry, which today constitute the main point of reference in methodological and technological terms, contain over 37 million cataloguing descriptions associated with about 26 million images; this wealth of information has been consulted by over 100 million visitors in the last five years. It must also be added what is managed at a territorial level by regional IT systems or by local applications, and not merged into national databases; a heritage certainly considerable but fragmented, of which it is difficult to give not only an interpretation but even a reliable quantification.”<sup>34</sup>

In the last decades, much of the Ministry’s action has been focused on digital cataloguing, which is an indispensable prerequisite for the creation of digital objects. Nevertheless, it is no longer sufficient. If the current objective is to create a national digital cultural heritage, a lot of work still needs to be done to digitally reproduce at least the majority of cultural objects. The cultural sector demonstrates a poor or insufficient application of technologies in both the digitization and digitalization processes. Only 68% of museums state to have a computerized cataloguing system, 76% said they have no strategic digital plan, 21% have not achieved any digitization and only 23% of the institutes digitized more than 75% of the collection.<sup>35</sup> Percentages regarding other institutes of the cultural sector, such as libraries, and archives, are expected to be roughly on the same level. As a result, the state of the art of digital transformation of Italian culture is hence underdeveloped and inhomogeneous. Most of the time it advantages big cultural institutions, where digitization was launched many years ago and is near completion, in contrast to small-sized organizations, where it is not planned, yet to be started or at an early stage.

---

<sup>32</sup> Ministero della cultura, Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library, ed. (2022), p. 5.

<sup>33</sup> *Cfr.* [Subchapter 1.3 What Opportunities for Small Cultural Institutes?](#)

<sup>34</sup> Ministero della cultura, Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library, ed. (2022), p. 4. Author’s translation.

<sup>35</sup> Pirrelli, M. (2022).

Furthermore, the NPD identifies some weaknesses that arose during the recent COVID-19 pandemic. First of all, the problem of institutional closure and restricted accessibility of heritage affected many cultural organizations. By intersecting data provided by both the Ministry of Culture and ISTAT, 34% of Italian libraries and 31% of museums were closed during 2021, as a result of the pandemic or for internal reasons.<sup>36</sup> In addition, when considering the opened ones, three libraries out of ten and almost half of museums reduced their opening hours to the public.<sup>37</sup> The closing process brought with it a consequent failure to access the heritage preserved and precluded the possibility of enjoying it, studying it, and getting inspired by it. The problem got all the more urgent with those categories of visitors who are not just tourists. Students, researchers, and experts needed to get access to heritage for study and work. The critical situation highlighted the necessity of ensuring digital accessibility and it indeed boosted both processes of digitization and digital communication. In 2021, for example, only 5% of Italian museums investigated by the Observatory for Digital Innovation in Heritage and Culture of the Politecnico di Milano did not use a website and 81% of them started the process of digitization.<sup>38</sup> The COVID-19 pandemic, drastically reducing the possibility of maintaining physical relations with the cultural heritage and its users, emphasized the need to accelerate the processes of digital transformation. It also marked the exigence of overcoming non-shared and short-lived logics that reveal limited sustainability over time due to the rapid obsolescence of data, applications and infrastructures not conceived in terms of networks of interconnected subjects.<sup>39</sup> Thirdly, the pandemic marked the need for specific digital skills that are often lacking in the cultural sector.

Starting from these premises, the NPD clarifies that it is essential to adopt a national strategy to share cooperative lines of action and enable the interoperability and dialogue of hundreds of existing information systems. For this reason, the creation of a national digital cultural heritage will be possible only if the Ministry first enables the conditions to structure and grow a digital ecosystem. This is also a fundamental prerequisite to allow Italy to contribute to the Horizon Europe programme known as the “European Collaborative Cloud for Cultural Heritage” (ECCCH).<sup>40</sup>

---

<sup>36</sup> Author’s elaboration on data available on *Statistiche culturali – Anno 2021*. Istituto Nazionale di Statistica (ISTAT) (2022) and Ministero della cultura, Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library, ed. (2022), p. 4.

<sup>37</sup> Istituto Nazionale di Statistica (ISTAT), *L’accessibilità di musei e biblioteche*. (2022), p. 4.

<sup>38</sup> Gandin, K. (2021), pp. 33-34.

<sup>39</sup> Ministero della cultura, Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library, ed. (2022), p. 6.

<sup>40</sup> European Commission, Directorate-General for Research and Innovation, Brunet, P., De Luca, L., Hyvönen, E., et al. (2022). The ECCCH is part of the Horizon Europe Framework Programme. The European Commission will invest 110 million euros by 2025 to finance projects related to building a collaborative cloud infrastructure. This infrastructure

The vision for the digital transformation of cultural heritage presented in the NPD follows three trajectories. These trajectories combine values, objectives and opportunities of the digital transformation in the cultural sector. The first trajectory is called “From objects to relationships” and is founded on the value of “cultural landscapes”, which is a theoretical model focused on the relationships between cultural heritage, communities and physical/virtual contexts. The traditional notion of cultural items is therefore expanded to include the social context where they are placed. In this way, it would be possible to overcome the historical isolation of cultural objects inside the places of culture and intercept new typologies of the public. The objective of this trajectory is indeed to expand the form of access to cultural heritage. This can be achieved in two ways: extending the quantity of digital resources available online and improving the quality of access, as well as the ways of use and reuse. The conceptual passage from closed cultural systems to open ecosystems would offer new opportunities for participating, experimenting, interacting, disseminating and reusing a public cultural heritage.

The second trajectory, called “The digital as an environment”, has its core value in the concept of “digital cultural heritage”. This term, which was mentioned in the previous subchapter, includes both born-digital resources and digital objects derived from analogue entities through digitization processes. Raw data and digital reproductions are not elements of cultural value per se. They are part of the cultural heritage when they present the elaborate and organized form of digital objects that can produce relationships with other elements.<sup>41</sup> The digital cultural heritage can produce new horizons of meaning and be considered an expression of contemporary society’s culture. For this reason, the trajectory treats it in the field of conservation, research and enhancement in its own right, as ontologically independent from the original physical objects. The purpose of this action is to operate a comprehensive digital transformation that can produce these complex digital objects according to the EU suggestions.<sup>42</sup> This assumption will lead to new opportunities for cultural dissemination, based on services capable of offering individuals innovative knowledge processes understood as authentic experiences of cultural growth.

The third trajectory of the NPD’s vision is inspired by the principle of “A heritage without borders”. The birth of a digital cultural heritage facilitates interconnections between cultural items and the creation of new levels of meaning. The semantic capital of relationships would be possible

---

will provide technologies for digitizing artefacts, scientific research and data documentation. Furthermore, it will represent the digital ecosystem where the EU digital cultural heritage can be built.

<sup>41</sup> Ministero della cultura, Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library, ed. (2022), p. 77.

<sup>42</sup> *Cfr.* [pp. 23-24](#).

thanks to the creation of interdependent ecosystems, capable of breaking down information barriers existing between databases and directing the multitude of organizational systems towards a common purpose. Along this transformation, it would be important to move from the self-sufficiency institutional perspective to the logic of interdependence between institutions. The sense of belonging to a national, common cultural ecosystem would also allow the exchange of resources, practices and methodologies. New models of knowledge would be possible for a variety of organizations: not only for a “consolidated” segment, represented by the GLAM (Galleries, Libraries, Archives and Museums) institutions that constitute the cultural heritage, but also for an “operational” segment, consisting of scholars, and an “open” segment, that is a dynamic and changing group of users, students, associations, tourists and cultural and creative industries.

The strategy adopted in the NPD outlines the concrete actions to be realised at the national level as a precondition for enabling the development of a digital ecosystem of culture. The trajectories discussed beforehand suggested three strategic objectives: to prevent the dispersal of the digital cultural heritage (the principle of unification), to allow it to be freely used by everyone (accessibility), and to maintain it over time (stability).<sup>43</sup> It is essential to develop the existing cultural databases and digital collections in the direction of the PND Vision, placing the current fragmentation into the perspective of a unique, complex national cultural heritage. This measure will be effective only by ensuring the long-term use and accessibility of digital objects. At the same time, the strategy has to enlarge the opportunities for relationships and exchange with citizens, passing from the vision of the user to that of the person. It would be important to rethink the digital as human-centred not only in relation to the categories of the public but also in favour of cultural operators. For this reason, the strategy would adopt training actions aimed at increasing skills and developing the potential for co-creation. Finally, new collaborations with businesses and innovative start-ups would facilitate the growth of a complementary market for cultural services.

Three cornerstones support the strategy proposed by the NPD: enabling technologies, processes and people. Technologies are needed to create the digital infrastructure where innovative processes can take place. These processes are then overseen by people, in such a way that people can benefit from these processes. The first line of action aims therefore to create an enabling technological context. This action includes the creation of a software infrastructure for the national cultural heritage to manage, preserve and enhance data, as well as a platform for integrated access to cultural heritage, and other platforms and digital systems for data archiving, digital identity certification,

---

<sup>43</sup> Ministero della cultura, Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library, ed. (2022), p. 29.

and services for citizens, developers and businesses. The most important measure is the creation of a software infrastructure for the national cultural heritage (in Italian, the *Infrastruttura Software per il Patrimonio Culturale – ISPC*) and it will enable the forementioned national digital ecosystem. It is supported by sub-investment 1.1.4 for a total amount of 73 million euros and it is based on a cloud infrastructure that will be boosted by sub-investment 1.1.3 (25 million euros). The ISPC consists of a series of services that enhance current digital systems operating at national and territorial levels and offer functions for the archiving, management and enrichment of digital resources – based on models and schemes, rules and ontologies and artificial intelligence algorithms (AI).<sup>44</sup> Furthermore, it presents an extensive catalogue of APIs to process domain and cross-domain data. This digital ecosystem will be the first national infrastructure that keeps the digital cultural heritage of the country. Cultural institutions will be offered to join the platform with flexibility and autonomy, following two modalities: integration – when the institution provides all its data and digital objects, which are then stored and managed within the ISPC – or federation – if the institution wants to maintain digital resources on its digital platform and just allows the ISPC to acquire descriptive data and to create references on the platform. As a result, the ISPC will serve as a DAMS (Digital Asset Management System) and a platform of services.<sup>45</sup> All the information and digital resources stored in the ISPC will be then accessed through a platform for integrated access to cultural heritage, namely the Digital Library. Citizens, experts, researchers and a wide variety of users will have the possibility to customize their navigation inside this hub of the Italian digital cultural heritage, which will be constantly increased by digitization projects.<sup>46</sup> The platform will be created with sub-investment 1.1.10 for a total amount of 36 million euros. Furthermore, the NPD strategy aims to enrich this digital ecosystem with other platforms. A Digital Preservation Pole (sub-investment 1.1.8, 58 million euros) will collect and preserve digital archives produced by the central and peripheral structures of the Ministry and those declared of historical interest while adopting uniform strategies and tools. A national system of digital identity certification of cultural assets (sub-investment 1.1.2, 16 million euros) will be developed to provide both physical goods and born digital items with a digital identity, on the model of the Public Digital Identity System (SPID). This would be necessary to unequivocally identify a cultural item and connect it to administrative procedures and existing digital resources. Moreover, a Portal of procedures and services to the citizen (sub-investment 1.1.9, 10 million euros) will be created to design services and models for cultural, social and economic value. Another platform, called Platform of Digital

---

<sup>44</sup> Cerullo, L., Negri, A. (2023), p. 41.

<sup>45</sup> Cerullo, L., Negri, A. (2023), p. 42.

<sup>46</sup> Ministero della cultura, Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library, ed. (2022), p. 54.

Services for Developers and Businesses (sub-investment 1.1.12, 45 million euros), will enable the development and integration of complementary services among creative operators and enterprises. A final platform will be created to stimulate the reuse of digital resources and generate mechanisms of co-creation and crowdsourcing (sub-investment 1.1.11, 10 million euros).

Most of the investment funds are established to promote processes of digitization throughout the country (sub-investment 1.1.5, 200 million euros). The aim is to increase the digital cultural resources available on the platforms, considering the participation as well in the Horizon Europe programme of the European Collaborative Cloud for Cultural Heritage. Funds are allocated to finance four digitization projects of national interest. The first regards the digitization of manuscripts' microfilms kept in the National Centre for the Study of the Manuscript (CNSM) at the National Central Library in Rome (9.2 million euros). The second project concerns the digitization of land registries and post-unification newspapers collected in the Central National Libraries of Florence and Rome and of the National Libraries of Milan and Naples (27.9 million euros). A third project is about the photographic archives of the Soprintendenze Archeologia, Belle Arti e Paesaggio, i.e. the government departments responsible for monuments, the environment and historical buildings (16,9 million euros). A fourth, final national project regards the cultural items kept in the museum depositories of over 70 relevant institutions (18.6 million euros). Funds for a total amount of 70 million euros were then allocated to the individual Regions and Autonomous Provinces of Trento and Bolzano to digitize the public cultural heritage of the territory of competence.<sup>47</sup> Together with the investment, the strategy of the PND aims to share methodologies and tools to plan, execute, and monitor digitization projects. In this regard, the Ministry developed a series of guidelines, in particular those about the digitization of cultural heritage, which are a valuable contribution to the promotion of good practices.<sup>48</sup>

Considering the human capital, it seems essential to bridge the existing digital gap in personnel skills. Operators need to know how to create and manage information and content through information and communication technologies. Their competencies must include technological skills, methodological knowledge, management attitudes and mental processes. There would be no sustainability in creating a digital cultural ecosystem without the development of digital capabilities. This is the reason why sub-investment 1.1.6 (together with sub-investment 1.1.7 for the operative support, for a total amount of 25 million euros) plans a lifelong learning programme in favour of

---

<sup>47</sup> Ministero della cultura (2022).

<sup>48</sup> Cfr. Docs Italia, *Linee guida per la digitalizzazione del patrimonio culturale. Versione 1.0* (2022).

the internal staff of the ministry and all professionals working in the cultural sector. The training plan coordinated by the Fondazione Scuola dei beni e delle attività culturali.

Before the end of December 2025, the PND expects that a number of at least 40 thousand operators will be included in the training plan and that almost 75 million new digital resources will enrich the digital cultural heritage of the Digital Library. The creation of a digital infrastructure, based on a wide range of online services and platforms, constitutes the fundamental premise for an aware, participatory, shared, sustainable and inclusive future digital transformation that can go beyond the scope of the NRRP. Afterwards, the challenge will be to succeed in feeding this regenerative flow that has started in recent years in the context of NextGenerationEU. It will be essential to find financial resources and above all to take advantage of the opportunities offered by this digital ecosystem to attract the multiplicity of cultural institutions scattered throughout Italy, offering valuable growth opportunities. This fact would signify a real digital transformation for small cultural institutes.

### 1.3 What Opportunities for Small Cultural Institutes?

The context of the digital transformation and the strategy promoted by the National Plan for the Digitalization of Cultural Heritage presents a challenge to small-sized organizations. It seems extremely important to reflect on this topic not only because this dissertation focuses on the practices that these institutes can plan, using as an example a project promoted by the Ateneo Veneto in Venice, Italy. Other elements should be considered. Primarily, because Italian cultural heritage is fragmented in a multitude of small-sized organizations that embark on digitization projects in the here and now or more and more of these institutes do so in the future. Reflecting on the specificities of these institutes, it would be possible to understand how on the one hand they can benefit from the digital transformation and on the other hand the weaknesses and constraints to be taken into consideration.

First of all, small cultural organizations can be defined as places of culture, such as museums, libraries and archives, characterized by insufficient human and financial resources in respect to the cultural heritage objects they own. Even if they own a sizable collection, they sometimes do not enjoy a large number of visitors and, above all, can count only on limited staff. These institutions can be either public or private. However, they are not considered of national or regional importance. In this regard, the Ministry of Culture has not provided any comprehensive analysis of the phenomenon of small cultural institutes that clearly quantifies their number through an intersection of data concerning the number of visitors, the cultural assets held and the staff. One

can only provide some data on the individual categories of the survey in order to estimate the amount of small organizations. In 2021 official data provided by both the Ministry of Culture and ISTAT revealed that the majority of Italian museums (45.1%) had staff composed of less than five operators, whereas 32.1% of them had staff between six and fifteen operators and only 15.9% had more than fifteen operators.<sup>49</sup> At a national level, the majority of institutions with a limited number of internal staff are accompanied by a majority of institutions with a limited flow of visitors. However, no aggregated data is available to express the percentage of institutions affected by restricted values for both the categories. With regard to the flow of visitors, 46.5% of Italian museums had less than 1,000 visitors and more generally 82.1% of them had less than 10,000 visitors.<sup>50</sup> The flow of visitors is an important variable to consider since it expresses on the one hand the attractiveness of the institutions and on the other their possibility to finance themselves through ticketing if services are not provided free of charge. Each institution will evaluate the number of visitors to be admitted and the relative access policy according to the nature of the institution, economic needs, the institution's size, conservation needs, etc. Concerning Italian libraries, the same observation arises. 56% of libraries had a number of total physical accesses recorded inferior to 1,000 and, above all, more than half (63.8%) had less than five operators in their staff.<sup>51</sup> The picture that comes out is relevant. On a national level, cultural institutions that can be defined of medium and small size are indeed the majority. As for their territorial distribution, this typology of institutes appears in large urban centres and increases in number in the small urban centres where they constitute a peculiarity. An idea of this distribution may derive from the trend in the number of persons employed by the institution in relation to cities or densely populated areas, small towns and suburbs with an intermediate population density and rural or sparsely populated areas. This trend passes from 21 operators to six for museums and from seven to three for libraries.<sup>52</sup>

The predominance of small and medium-sized institutions requires particular attention when digital transformation is concerned. Some observations can be made taking as an example the case of

---

<sup>49</sup> [Cfr. Footnote 3.](#)

<sup>50</sup> Author's elaboration on data available on Istituto Nazionale di Statistica (ISTAT), *Statistiche culturali – Anno 2021* (2022) (“Tavole Musei”, Tav. 1.5).

<sup>51</sup> Author's elaboration on data available on Istituto Nazionale di Statistica (ISTAT), *Statistiche culturali – Anno 2021* (2022) (“Tavole Biblioteche”, Tav. 4.9 and Tav. 4.14). Concerning the number of total physical accesses, 4.7% of Italian libraries had no users in 2021, 41.3% had less than 500 users, 10.0% between 501 and 1,000 users, 20.2% between 1,001 and 5,000 users, 5.8% between 5,001 and 10,000, and 5.6% more than 10,000 users. 12.3% of the libraries did not answer the survey. With regard to the staff size, 63.8% of the libraries had less than five cultural operators, 19.1% between six and fifteen, only 3.3% more than fifteen and 1.3% no operators.

<sup>52</sup> Istituto Nazionale di Statistica (ISTAT), *Statistiche culturali – Anno 2021* (20 December 2022) (“Tavole Musei”, Tav. 1.12) and *Ibid.* (“Tavole Biblioteche”, Tav. 4.14).



some academies and their libraries located in Lombardy and Veneto.<sup>53</sup> Despite some specificities, it will be possible to formulate some features that are valid not only for these organizations but also for the wider category of small cultural institutes. Inspired by the model of Plato's *Akademia*, modern academies were founded in the course of the 16<sup>th</sup> and 17<sup>th</sup> centuries as a “social space where a group of people gathered to share knowledge, experience, and information without the need for drafting membership rules or statutes”.<sup>54</sup> Their development was influenced by the Neoplatonist revival that accompanied the Italian Renaissance and responded to the need for the promotion of an alternative model of culture in comparison to scholastic philosophy that continued to be adopted in the majority of the universities. Over time, many academies produced innovation and experiments in disciplines such as literature, music, science, arts, theatre, and philosophy by promoting an eclectic and encyclopedic approach to knowledge.<sup>55</sup> A circle of luminaries, mostly belonging to the local noble classes, was responsible for these places and oversaw the activities of culture dissemination. As in the past, academies are depositories and guarantors of knowledge and are made up of institutional members who are luminaries in specific disciplinary fields. Scholars meet inside those places to discuss topics related to their subject and share ideas with other colleagues. Even though academies might still be commonly considered old-timers and niche institutions in the diminishing sense, in reality, they are places that in time knew how to reinvent themselves and today they fulfil a social role through the spread of knowledge among the larger public.<sup>56</sup> Events, meetings, and conferences are occasions of debate and cultural growth. In this context, the presence of member scholars who provide authoritative and certified knowledge is a great strength. Furthermore, academies have represented places of conservation and re-enactment of the urban contexts in which they are inserted. Here the memory and the identity have settled over time in private libraries and cultural items of various types – which continue to be acquired through donations and bequests. The libraries of these institutions are usually composed of a historical collection consisting of a few thousand books and a majority of 18th-19th-century and contemporary publications. This heritage, which is all the most valuable because it embraces an encyclopedic knowledge that manifests a symbiotic connection between humanistic and scientific disciplines, was originally intended to be consulted by hundreds of

---

<sup>53</sup> Andreolli, A., ed. (2022). The institutes considered in the conference proceedings are: the Accademia Roveretana degli Agiati; the Ateneo di Brescia; the Ateneo di Salò; the Ateneo di Bergamo; the Accademia Nazionale Virgiliana in Mantua; the Accademia Galileiana in Padua; the Ateneo Veneto in Venice; the Accademia di Agricoltura Scienze e Lettere di Verona; Accademia Olimpica in Vicenza; and the Accademia dei Concordi in Rovigo.

<sup>54</sup> Testa, S. (2015), p. 3.

<sup>55</sup> Testa, S. (2015), p. 5.

<sup>56</sup> Stocchi, M.P. (2007). *Conservazione della memoria e promozione della cultura: il ruolo degli Atenei*. In Simionato, G. ed. (2007), p. 55.

members of the Academic Assembly.<sup>57</sup> The administrative evolution, fortunately, disclosed this precious heritage from internal members to the greater public. These institutes transformed into non-profit associations – the Ateneo Veneto was qualified as ONLUS in 1997 – and this fact implied not only that they assumed a public function but also that they used their rich heritage to achieve their cultural and social objectives.<sup>58</sup> As a result, their library collection is nowadays accessible to everyone through online local and national library networks' databases.

The necessity of fruition and the public mission clashes with some structural constraints that typically limit the activity of small cultural organizations. First of all, the small dimensions of these institutes lead to a direct consequence on the managerial capabilities. They are generally conducted by small staffs, which are busy on several fronts, and they do not or rarely make use of trainees who are a valuable resource for carrying out core or additional activities. In fact, the current legislation on curricular internships – set by the Law 92/2012, State-Regions Agreement of May 25, 2017, and the subsequent amendments made by the Finance Act 2022 – allows institutions to host a number of trainees proportional to their internal staff.<sup>59</sup> For this reason, small cultural institutes are able to host only one trainee, with the result that they cannot count on large external help. In other cases, the lack of staff with technical skills can slow down or stop key processes, such as the cataloguing of the library and documentary heritage.<sup>60</sup> The tight staff structure can be potentially deleterious if these institutions cannot count on the continuous exchange of skills that is normally present in larger institutions or on capable operators who possess transversal skills. Without proper competencies, it would be hardly possible to face new challenges, such as digital transformation, and the institute can reach an impasse. Moreover, restricted financial resources have led them to “live with public subsidies or with the help of some bank”.<sup>61</sup> In addition, costs for the maintenance of buildings recognized as exceptional cultural heritage and the preservation of the collections have further burdened finances. Their perspective still turned to the past, their preservative and conservative approach and the concept and the preconception of being a milieu for a few have often affected their attractiveness to public and private funding. In other cases, these institutions do not have ample and dedicated spaces for the preservation and fruition of the book

---

<sup>57</sup> Ateneo Veneto, *Statuto* (20 March 2023), art. 9.

<sup>58</sup> *Ibid.*, art. 30.

<sup>59</sup> *Cfr.* <http://archivio.statoregioni.it/DettaglioDoc6350.html> [Accessed 15 September 2023]. It depends on the regional regulations to fix the number of traineeships that can be activated simultaneously by a host enterprise in proportion to the production unit. Nevertheless, there are quotas - from which apprentices are excluded - equal to: a trainee for host subjects employing 0 to 5 employees; two trainees for host subjects from 6 to 20 employees; 10% for host subjects with more than 20 employees.

<sup>60</sup> Ferraglio, E. (2022). *La biblioteca dell'Ateneo di Brescia. Stato e prospettive*. In Andreolli, A., ed. (2022), p.93.

<sup>61</sup> Gottardi, M. (2007). *Conservazione della memoria e promozione della cultura: il ruolo degli Atenei*. In Simionato, G. ed. (2007), p. 61. Author's translation.

heritage and the simultaneous public activity towards citizenship. These spaces are often subject to protection restrictions due to their historical importance, which does not make their use easy.

The solution proposed by the same academies during two conferences<sup>62</sup> was to go beyond the singularity and individuality of the institutes and dare to reinvent themselves. It is not necessary to deny the identity, the history, or the vocation of the institutions, but to rethink the course of action to make their public function more effective. The first way to do this is to aggregate together in a network of similar institutions, where individual action can be strengthened by the forces of other entities. It would be essential to establish relationships with other associations in the same area and to create connections that can broaden the horizons of the institution. Above all, it will be essential to dialogue and interface with a wide range of the public. The cultural mandate of every cultural organization, small or large, is an ethical and civil mandate and requires using cultural assets for the greater good of the community. Academies, and more generally cultural institutions, can develop “knowledge protocols capable of combining the indispensable recognition of the past with the sum of the data relating to the present”.<sup>63</sup> They can disseminate qualified culture – that is not obvious today – and develop the public’s critical sense. They really can contribute to the change and stop the process of identity loss that the realities in which they are placed are undergoing.

In this context, digital transformation arises as a possibility of reinvention. In a world that is increasingly impregnated by the digital, one must accept the challenge and use the opportunity to influence the transformation. Ten-fifteen years ago an online catalogue and a website were the basic prerequisites for greater visibility of the institution and accessibility of the cultural heritage.<sup>64</sup> Nowadays, the development of technology and the use of social media offer more and more the possibility of increasing visibility and popularity. It allows even unknown people to become influencers. Small cultural institutes will not be asked to become influencers but rather to take advantage of digital communication to enlarge their public. The means to do that might be found in the creation of alternative contents on the Web. The Web is a new world that has been systematically reinventing itself for some decades and presents risks and challenges. Although it seems that everything is accessible, complete and above all trustworthy on the Web, in reality the current situation is far from it. Digital contents too often suffer from the problem of homologation,

---

<sup>62</sup> The first conference was held in Treviso on the 21<sup>st</sup> and 22<sup>nd</sup> October 2006. The second one was held in Rovereto on the 4<sup>th</sup> of October 2019. The conference proceedings are respectively Simionato, G. ed. (2007) and Andreolli, A., ed. (2022).

<sup>63</sup> Simionato, G. (2007). *Gli Atenei delle città e per le città: identità e testimonianza*. In Simionato, G. ed. (2007), p. 47. Author’s translation.

<sup>64</sup> Gottardi, M. (2007). *Conservazione della memoria e promozione della cultura: il ruolo degli Atenei*. In Simionato, G. ed. (2007), p. 61.

inaccuracy and unreliability. Cultural organizations must stimulate the development of the digital world as a more reliable, human place through their qualified culture dissemination. By focusing on quality, it will be possible for them – and for small cultural institutions, including the academies taken as an example – to gain the esteem of users and expand the circle of audience. The digital could therefore enable them to follow their public mission while reinventing themselves.

Digitization could be a way out of an impasse. The objectives that guide the digitization projects of small cultural institutes are discussed in more detail in Chapter 2. One of the most important advantages of digitization can be measured in terms of accessibility and this might have been understood by small cultural institutes only during the recent COVID-19 pandemic. Measures for the containment of disease closely affected those institutions – and forced the same Ateneo Veneto in Venice to interrupt its institutional activities. While all events were rapidly canalized into online participation, the portion of collection that was not digitized remained inaccessible. It stands to reason that digitization has to be considered a useful tool not only when cultural institutions are closed, but especially when they are regularly open in favour of their activity. This is the will and the wish expressed by the Ministry of Culture in the National Plan for the Digitalization of Cultural Heritage. Small cultural institutes will benefit from the creation of a national digital ecosystem. Choosing between the two modalities of integration or federation, they will participate in the software infrastructure for cultural heritage (ISPC), where they can become part of a cultural net and establish relationships with other institutions. Moreover, they will access a digital environment that is meant to be interoperable and permanent, with the certainty that their digital objects are preserved. But all this will only be possible as of 2026 if there are no hitches in the realization of the NPD.

What is to happen in the meantime? It will be strategic not to wait so many years and start the process of digitization and digitalization ahead of time according to the possibilities of the institution. In this perspective, finding potential sources of funding and formulating a digital strategy are essential. Concerning the economic aspect, which is the *sine qua non* of these future prospects, the Ministry of Culture has allocated 200 million euros to promote a systematic series of digitization campaigns throughout the country.<sup>65</sup> Almost 72.6 million euros are currently financing four digitization projects of national interest, whereas more than 70 million euros have been allocated to Regions to promote local digitization projects. No information is available on what the remaining 57.4 million euros will be used for.<sup>66</sup> It would be desirable to allocate this sum

---

<sup>65</sup> Cfr. [p. 30](#).

<sup>66</sup> The report on the status of implementation of 1.1 investment is available at: Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library, *Investimento 1.1* (2023).

in favour of a digital regeneration of small cultural institutions. Nevertheless, this would involve verifying that small institutions can develop projects in accordance with national standards, and have the proper equipment and adequate staff at their disposal. It would be appropriate but it might be not feasible at the government level. In formulating the public announcements for the digitization of cultural heritage, some Regions have indeed established some conditions for regulating the expression of interest. The Veneto Region, for example, decided that only cultural institutions with professionalized personnel and adequate space could apply for the NPD's regional funding.<sup>67</sup> It is a procedure that reasonably limits project risks but at the same time is questionable because it does not use the NRRP funds – the NextGenerationEU funds – as an opportunity for betting on small institutions and giving them the possibility to reinvent themselves. If they cannot obtain public funding, cultural organizations must use self-financing or find private funds. Private funds can be received from bank foundations – as in the case of the Ateneo Veneto –, associations, sponsors, or crowdfunding campaigns. The need to network is more than ever understandable for this economic aspect.

With regard to the digital strategy, small institutes should conceive digitization projects as occasions to exercise their public function and increase their public. This does not mean the abandon of the cultural collection. On the contrary, it is important to take advantage of a digital collection or a series of digital objects to disseminate culture, look for the actualization of the contents, and achieve real involvement with users. Digitization should not be undertaken merely as a fleeting trend, but rather as a purposeful practice with meaningful objectives. If executed with a strategy, it can encourage the creation of a valuable hybridization between the physical library and the digital cultural heritage, where the physical and the digital complement each other. Small cultural institutes can join digital platforms to promote a scientific, qualified dissemination of contents. Academies, which by their nature are called to spread culture to a wider audience, can reinvent themselves in the digital world by becoming certified content creators. They could use their cultural assets as a starting point for their dialogues with users and digital platforms as places of debate and discussion. The success of this operation will be measured not only online but also in a greater turnout of activities carried out in presence.

To seize opportunities that arise from digital transformation, a successful digitization project requires meticulous planning. During the project planning, objectives, financial resources, technical possibilities and staff capabilities are intersected. National standards and guidelines are taken into consideration as well. In the case of small cultural institutes, planning will be a mandatory step and

---

<sup>67</sup> Regione del Veneto, *Digitalizzazione del patrimonio culturale - Avviso pubblico PNRR (2023)*.

must balance the specificities and structural limits of these organizations with technical requirements in order to avoid bad practices. Such a challenging task is faced in Chapter 2.

## CHAPTER 2: PLANNING AND BEST PRACTICES FOR A DIGITIZATION PROJECT

Having discussed the state of the art of digitization in Italy today, this chapter will analyse the issue whether small cultural institutes can afford digitization projects and will further try to set a model for project planning for such institutes. In fact, small-sized institutions that will undertake digitalization projects in the future will surely be bound by economic and logistical limits and it is hard to see how they will be able to dispose of state-of-the-art equipment. Hence the need to indicate the modes of action that are compliant with the planning of digitization projects.

Project planning is the first step in any digitisation project. As the Minerva Working Group <sup>68</sup> had occasion to highlight as long ago as 2004, “time spent on planning the project will pay dividends in the easier management and execution of the project.”<sup>69</sup> The use of a strategic approach prevents many problems and overcomes many limits. It is certainly the first card that a small institution can play when dealing with a digitalization project. If everything goes according to plan (which inevitably include in the plan B options just in case something goes wrong), the little institute’s planning becomes its trump card.

In order to deal with digitization planning, reference is made to a specific project carried out by the Ateneo Veneto in Venice. The “Digitization Project of a Hundred Rare and Valuable Medicine Books” was developed as part of a funding call promoted by the Fondazione di Venezia to support and enhance the propagation of digital culture. The Ateneo Veneto responded to the call planning a strategic action in order to be selected and develop a long-term digitization practice at the same time. Its heritage, in particular its prestigious collection of early-printed books, was the driving force and the main objective of the project. As explained in more detail below, the Ateneo Veneto, whose full name is Ateneo Veneto di Scienze, Lettere ed Arti, has a library of among fifty thousand books and scientific journals, out of which three thousand are early-printed books. The institute

---

<sup>68</sup> Fondazione Istituto Internazionale di Storia Economica F. Datini - Minerva Project, *The Minerva Project* (07 January 2014). The MINERVA project – “MINisterial NETwoRk for Valorising Activities in Digitisation” – is a project financed by the European Commission in the ambit of the IST program (Technologies for the Information Society). It has given life to a network of European Ministries for culture, coordinated by the Italian Ministry, to coordinate and harmonise the digitisation of scientific and cultural heritage. The project started in March 2002 and ended in January 2005.

<sup>69</sup> Minerva Working Group 6, ed. (2004), p. 15.

could indeed count on its typographical rarities to propose itself for a digitization project and enhance its heritage.

The decisions and experiences gained at the institution will serve as an illustration for a more methodical examination of what a project planning entails. Many strategic points have to be discussed and implemented: the definition of the objectives, selection of materials, institutions involved, the choice of instrumentation and its optimization, the choice of repository, together with other aspects. All of these are discussed as follows.

## 2.1 Objectives

When you proceed to project a digitization campaign, the first and most important question to raise is “why”. The reasons for the project serve as a basis for the whole process and have a direct effect on the criteria used to select materials and technical methods.<sup>70</sup> The cultural institute always has a leading role in digitizing its collection and in determining its purposes. Nevertheless, in some cases, the mission and execution of the project may not be unilaterally set by the institution itself but be influenced by other organizations, depending on the underlying assumptions of the project. For example, an institute can obtain the necessary funds to develop the project through a funding call or through sponsorship. This is often the case with small cultural institutions. In these cases, the conditions for accessing the contribution set by the funding organization are decisive. That was the case of the project planned by the Ateneo Veneto. The definition of the project’s mission was conditioned by the requirements stated in the funding call “Cultura Digitale” promoted by the Fondazione di Venezia. These requirements were:

1. support the activities of the Museums, Libraries, and Archives of the Metropolitan City of Venice through sustainable and replicable initiatives, capable of making concrete contributions to the enhancement and dissemination of culture and knowledge;
2. activate and enhance distinctive innovative skills and planning on the territory;
3. offer young people opportunities for professional growth in the field of culture.<sup>71</sup>

The three objectives reflect the nature of the promoter. Founded in 1992, the Fondazione di Venezia is a banking foundation that since 1993 has been operating through the use of funding calls with the mission of investing its financial assets and making them profitable in favour of the Venetian territory. During these years, it has been particularly interested in sustaining the vibrant

---

<sup>70</sup> Minerva Working Group 6, ed. (2004), p. 16.

<sup>71</sup> Fondazione di Venezia, *Bando “Cultura Digitale”* (11 April 2022), p. 3.



culture of Venice, considering the heritage and the great number of cultural activities that characterize the city. This covers the first mission of the Fondazione di Venezia. As for the second and the third ones, the call “Cultura Digitale” shared the premises that guide the PNRR plan. As already mentioned in Chapter 1, the two keywords required by the Ministry of Culture for digitization are “planning” and “skills”. This is the reason why they are both included in the mission of the forementioned funding call of the Fondazione di Venezia. The first requirement, “planning”, regulated the modality for every digitization project proposal presented. To this, the second requirement, “skill”, was interpreted by the funding institute in the sense of enhancing young skills. In fact, it asked the funded institute to include in the project “young students or recent graduates from high schools or university and/or the retraining of internal staff through training and professional growth”.<sup>72</sup> In this way, there can be support for a process of upskilling and reskilling of cultural operators or a growth in youth employment in the GLAM (Galleries, Libraries, Archives and Museums) sector. The Ateneo Veneto aimed for the digital competencies possessed by the new generation, taking advantage of the presence of the master’s degree in “Digital and Public Humanities” at the Ca’ Foscari University of Venice and the interdisciplinary knowledge both in the humanities and in the digital field of its students.

Once the objectives of an eventual external funding are clear, the institution must ask itself which objectives will guide its digitisation project. There are usually four main goals that can be taken into consideration. Each institution then adapts them according to its specificities. Starting from the most important, they include:

1. Accessibility
2. Preservation of the originals
3. Enhancement
4. Transparency

A digitization project aims in the first place to take advantage of the high power of technologies and improve access to cultural heritage. Digital items archived in online platforms can be potentially consulted by everyone, anytime and anywhere. Physical distances that separate the institution from the users are reduced. Conservation issues that often prevent or limit the consultation of the items can be evaded. Artworks can be digitally admired and books can be read even if they are in such a state of preservation that physical accessibility is prohibited because repeated direct contact leads to a further deterioration of items.<sup>73</sup> Access is given to all those items that are out or for just internal

---

<sup>72</sup> Fondazione di Venezia, *Bando “Cultura Digitale”* (11 April 2022), p. 3.

<sup>73</sup> Docs Italia, *Linee guida per la digitalizzazione del patrimonio culturale* (2022), p. 5.

consultation or that cannot be seen in presence because deposited in storage room. Furthermore, fruition is guaranteed for a large amount of purposes and different typologies of user. They can be curious visitors that are exploring the online collection or students that are doing a school assignment or even researchers and experts that need to consult cultural items for their work. Digitization projects can help cultural institutes to fulfil their institutional aims towards the public and make their heritage accessible beyond the contingencies of the moment, as happened during the COVID-19 health emergency. Whilst the institution may be temporally or permanently closed, users of today and tomorrow can still benefit from digital reproductions of collections.

Concerning the specific case of this dissertation, the provision of accessibility was an extremely important objective for the Library of the Ateneo Veneto, whose collection – as already mentioned – is composed of fifty thousand books out of which three thousand are priceless, having been printed between the 15<sup>th</sup> and 18<sup>th</sup> centuries. The digitization of a portion of its early-printed books was part of strategic action to disseminate culture. The Ateneo Veneto organizes each year hundreds of cultural and scientific events open to the public free of charge which are moments of training and debates that come to life each Academic Year.<sup>74</sup> At the same time, the awareness of the enormous cultural value and fragility of these books had led the institute to limit their direct consultation. It is a common trade-off that characterizes many institutes' conservation policy toward the originals, especially when artworks are in such a state of conservation that repeated direct contact leads to a further deterioration of the items.<sup>75</sup> Thanks to the various activities of digitization, today preservation and fruition are both possible and the Ateneo used this solution to ensure “virtual” access. It is true that the direct consultation of artefacts, especially when it is rare and precious, is not comparable with the digital copy that the user can see on their monitor or smartphone. Nevertheless, in the case of books, this problem is much less cogent. The textual form represents the principal, cultural value of books and can be easily reproduced through digital images. The transmission of knowledge is preserved. As a result, the Ateneo Veneto's project was therefore an exigence required by both the nature of the institute and the preciousness of the books.

Preservation of the originals is the second aim of digitization projects. The creation of a digital copy of artefacts helps to preserve cultural objects through the study and diagnosis of materials. It allows the institute to map the item's state of preservation at the exact moment when the digital image was produced. This is particularly useful for restoration practices and in case of fast material

---

<sup>74</sup> Ateneo Veneto, *Ateneo Veneto - il più antico istituto culturale attivo per tutti a Venezia* (27 February 2023).

<sup>75</sup> Docs Italia, *Linee guida per la digitalizzazione del patrimonio culturale* (2022), p.5.

degradation processes, as well as for safety reasons and the occurrence of unfortunate accidents or thefts. As already mentioned in Chapter 1, digitization does not certainly create reproductions that are intended to be substitute items of the originals. Several physical features cannot be directly appreciated and sensory experiences, such as touch, smell, sight of the originals, can be expressed with difficulty though digital images. Nevertheless, if seen with a wider look, digitization can be regarded as a new vehicle to convey and preserve cultural information, in the same, revolutionary way as the invention of printing during the 15<sup>th</sup> century.

Book digitization is not only a matter of replication or reproduction of items. Today digital infrastructures provide remarkable advantages in terms of connections and correlations between items. The enhancement augments the knowability of books thanks to the digital environment where they are “kept”. This is the third objective of digitization. Considering the Ateneo Veneto’s project, the repository that was used in the project, Phaidra, was chosen for several reasons that are explained thereafter. Among these, the platform gives the possibility to create virtual collections, where books can be not only presented one by one to users and described in many sections, but they can also be connected in thematic batches among them or with other digital objects that were not digitized by the same institute. The connection between materials belonging to different contexts fosters interdisciplinary studies while ensuring the searchability and availability of all relevant information thanks to metadata and permanent links. Moreover, this process increases the institution’s visibility and stature and promotes collaboration between diverse cultural institutions.<sup>76</sup>

Finally, a fourth and last objective of digitization regards transparency. This concept concerns more the methodologies of the project – how to perform it – than its outputs – the digital objects themselves. The project team should aim to carry out digitization with practices that can be codified and “decrypted”. Data and documents have to be properly organized, operations have to be clear, and written reports are desirable. At the end of the project, the institution would enjoy the considerable benefit of a complete and documented work. It can ideally replicate the project, even if some components of the project team are no longer involved.

---

<sup>76</sup> International Federation of Library Associations and Institutions (IFLA) (2015), p. 7.

## 2.2 Selection of Documents

Generally speaking, the selection of documents is part of the strategic planning of a digitization project, and it is the result of its objectives and selection criteria. One of the prerequisites for selecting and digitising cultural items is to first have some knowledge regarding their contents and significance and then rely on databases and resources that describe them. Some information can be found in print or – more comfortably – online catalogues. It is useful to understand the typologies and characteristics of the items, as well as to estimate the amount of work to do.

With reference to the “Digitization Project of a Hundred Rare and Valuable Medicine Books”, the interest of the Ateneo Veneto was focused on its historical book collection and on a particular section of it dedicated to medicine. The presence of early-printed books concerning this discipline lies in the history of the institute and in the Società Veneta di Medicina from which it originated.<sup>77</sup> The institution was in fact founded on 12 January 1812 following a Napoleonic decree dated 25 December 1810 ordering the merging of the medical association Società Veneta di Medicina, the Accademia dei Filareti, and the Accademia Veneta Letteraria. The main purpose of its foundation, which continues to be the same for over two centuries, was social solidarity and cooperation in the development and dissemination of sciences, humanities, and arts, in all their manifestations. At that time, since 1808, the Società Veneta di Medicina was already settled inside a building of remarkable historical and artistic interest, the Scuola Grande di San Fantin or “Scuola dei Picaì”.<sup>78</sup> The Secretary of the Society, the Venetian physician Francesco Aglietti (Brescia, 1759 – Venice, 1836) understood the importance of having a scientific library inside the institute and asked the authorities the right to select appropriate books from the large number of books formerly owned by the libraries of suppressed convents. At that moment, these books were kept inside a temporary deposit in Santa Maria dell’Umiltà, eventually demolished. On 17 April 1809, 2,095 volumes were delivered to the Società Veneta di Medicina, selected by the Abbot Sebastiano Ongin Polacco. They had the following provenance: San Francesco della Vigna convent (955 books), San Salvador

---

<sup>77</sup> Raines, D., *catalogo della mostra a cura di* (2007), pp. 17-21.

<sup>78</sup> Ateneo Veneto, *Storia* (21 April 2023). The Scuola Grande di San Fantin was built to house the two brotherhoods of San Girolamo and Santa Maria della Consolazione that had merged sharing the same welfare purposes in the mid-fifteenth century. The main one was the spiritual support to those sentenced to death by hanging, from which the institution derived the popular name of “Scuola dei Picaì” (“of the hanged men”). Destroyed by fire in 1562, the building underwent a profound renovation, with the intervention of important architects, painters, and sculptors (Alessandro Vittoria, Tintoretto, Palma il Giovane, Veronese, Leonardo Corona). They shaped the building in the forms we can see today between the end of the sixteenth and the beginning of the seventeenth century. At the beginning of the nineteenth century, the Scuola di San Fantin ceased to exist with the fall of the Republic of Venice and the subsequent suppression of the Scuole Grandi. The building was precisely later assigned to the Società Veneta di Medicina on 18 January 1808.

convent (600 books), and Santo Stefano convent (540 books). Another delivery of 123 volumes took place on 12 December 1811. This nucleus came from the Somascan convent of Santa Maria della Salute (62 volumes), from the Capuchins of the Redentore (24 volumes), the Franciscans of Santa Maria Gloriosa dei Frari (24 volumes) and the Carmelites of the Scalzi (eight volumes), as well as the convent of Santa Maria della Fava (five volumes). These are the books that constitute the original collection of the Ateneo Veneto and they also regard the disciplines of natural philosophy and science in general, in addition to medicine. The selection of the materials for the digitization project was facilitated by the catalogue of the exhibition *Anatomia di una biblioteca. Cinquanta volumi di medicina dalla collezione storica dell'Ateneo Veneto*<sup>79</sup>, curated by the Academic Librarian prof. Dorit Raines. The exhibition was held at Palazzo Nervi-Scattolin in Venice – headquarters of the Cassa di Risparmio di Venezia which financed the exhibition – from April 21 to May 11, 2007. It exhibited around fifty of the most precious books of the historical collection of the Ateneo Veneto and had the merit of disclosing to the public the great value of the items kept inside the institution. It was the first occasion to study the collection and to evaluate the contribution of these books in defining the history of medicine.

This catalogue constituted the starting point for the project's theme and a basis for selecting the materials to be digitized. After an accurate analysis of the physical state of the books, the criteria principally used in the selection were: quantitative requirements; preciousness and rarity; feasibility. It should be noted that by responding to the call, the Ateneo Veneto already set as quantitative criteria the number of books to be digitized: a hundred (although at the end 104 books were digitized). This number included the most remarkable early-printed medicine books of the library, including those fifty of the 2007 exhibition and selecting fifty others that stood out for rarity and value. Priority was given to the most ancient books of the collection and, in particular, to books printed during the 15<sup>th</sup> and 16<sup>th</sup> centuries. A technical parameter – the size of the book – influenced the choice. Books have different typographical formats (folio, quarto, octavo, duodecimo, sextodecimo) and smaller formats were preferred because they greatly facilitated the digitisation process, as explained in Chapter 3. *Figure 1* shows some of the selected volumes.

The collection of digitized books consists of one incunable printed in 1485, the *Libellus de medicorum astrologia*<sup>80</sup> falsely attributed to Hippocrates (Kos, c. 460 BC – Larissa, c. 377 BC); 34 books from the 16<sup>th</sup> century; 45 books from the 17<sup>th</sup> century and 24 from the 18<sup>th</sup> century. As for the formats,

---

<sup>79</sup> Translated: “Anatomy of a library. Fifty medicine books from the historical collection of the Ateneo Veneto”. *Cfr.* Raines, D., catalogo della mostra a cura di (2007).

<sup>80</sup> de Beauval, F., Hippocrates, d'Abano, P., Ratdolt, E. (1485). *Opusculum repertorii pronosticon in mutationes aeris tam via astrologica quam metheorologica vti sapientes experientia comperientes voluerunt perquam vtilissime ordinatum...*. Venetijs : impressus est arte ac diligentia mira Erhardi Ratdolt de Augusta. Shelfmark: 27.A.33.

the majority (57) were quartos, followed by octavos (20), folios (12), duodecimos (10) and sextodecimos (five). The larger one and certainly among the most remarkable books of the entire collection was the masterpiece written by Girolamo Fabrici d'Acquapendente (Acquapendente, 1533 – Padua, 1619), *De formato foetu*<sup>81</sup>, printed in 1600 in Venice by the editor Francesco Bolzetta.



*Figure 1* Some of the Ateneo Veneto's books digitized during the project. ©Ateneo Veneto

The author is considered one of the pioneers of studies in Embryology and the book indeed is extraordinarily important because of 33 large engraved anatomical tables that analyse the formation of the foetus inside the uterus not only in human beings but also in other mammals such as sheep, cattle, horses, pigs, mice, together with other animals like fish and reptiles. The size of the book required a particular digitization process, which will be thoroughly explained.

The historical and cultural value of these books can be highlighted by seeing how many libraries belonging to the Italian Online Public Access Catalogue (OPAC)<sup>82</sup> own the same titles that were

---

<sup>81</sup> Acquapendente, G.F. (d'), Valegio, G., Pasquato, L., Bolzetta, F. (1600). *Hieronymi Fabricii ab Aquapendente De formato foetu*. Venetiis : per Franciscum bolzettam (Patauii : ex typographia Laurentij Pasquati, impress. almae Vniuersitatis iuristarum, 1601). Shelfmark: 8.E.34.

<sup>82</sup> OPAC is the collective catalogue of the National Library Service (SBN), which is the network of Italian libraries promoted by the Ministry of Culture in cooperation with Regions and Universities, with the coordination of the Central

digitized by the Ateneo Veneto. Throughout the country, four of the digitized books are kept only at the Library of the Ateneo Veneto; 21 are present only in other four libraries; 36 in other nine libraries and the remaining 42 in less than 37 libraries. It is not just a matter of numbers: these precious books are a valuable contribution to the study and research on the history of medicine. They “let us remember how research on the internal and external [medical] environment has been constant, in the history of medicine and yet how both the inside and the outside are thought in very different ways.”<sup>83</sup> The number of historical sections described in the 2007 catalogue presents the wealth of knowledge in the discipline of medicine throughout the ages. To the original seven sections, in which the catalogue of the 2007 exhibition was divided, a further, eighth one was added to gather together books from the 18<sup>th</sup> century.

The first section is dedicated to Hippocrates, who is considered the Father of Medicine in recognition of his early contributions to the field, such as the use of prognosis and clinical observation. His formulation of the humoral theory had an extraordinary fortune over the centuries and was definitively disproved only in the 1850s. His teachings were transmitted through his famous *Aphorismi*, which are presented for example in the book *Hippocratis Coi Aphorismorum sectiones* <sup>784</sup>.

The second section is dedicated to Arabic medicine, presented by Galen (Pergamon, 129 – Rome, c. 201) and to the revolution promoted in the Renaissance by Paracelsus (Einsiedeln, c.1493 – Salzburg, 1541). Over the centuries Galen, along with Hippocrates, enjoyed a reputation in the field of Medicine that was unprecedented and was especially known for his studies of anatomy. He described the four humours of Hippocrates’s theory as black bile, yellow bile, blood, and phlegm. Five of the 13 volumes kept at the Ateneo Veneto and part of the edition<sup>85</sup> of Tommaso Giunta printed in Venice in 1625 were digitized. These five volumes are bound with parchment together in one single, folio book and decorated with woodcut initials and headings. Arab Medicine, one of

---

Institute for the Union Catalogue of Italian libraries and for bibliographic information (ICCU). The number of participating libraries increases constantly and the network gathers universities, academies, and local, national, public, and private institutions. Libraries participating in SBN are grouped in local clusters that manage all the services of the linked libraries, through automated procedures. The Poles are in turn connected to the SBN Index system, the central node of the network managed by ICCU, which contains the collective catalogue of publications acquired by the SBN member libraries. The Ateneo Veneto joined the Local Venetian SBN Pole, Polo-VEA SBN, in 2008 and, therefore, is part of the National Library Service. For further information: <https://opac.sbn.it/en-US/opac-del-sevizio-bibliotecario-nazionale>.

<sup>83</sup> Semi, A.A. (2007). “Le radici dei medici”, in Raines, D. catalogo della mostra a cura di (2007). Author’s translation.

<sup>84</sup> Hippocrates, Leonicensis, N., Brasavola, A.M., Frambotto, P. (1638). *Hippocratis Coi Aphorismorum sectiones 7. Nicolao Leonicensi interprete. Accessit octava ex Ant. Musae Brasavoli commentarijs. Cum gemino indice. Item Prognosticorum Libri 3.* Patavii : impensis ac typis Pauli Frambotti. Shelfmark: 6.C.48.

<sup>85</sup> Galenus, C., Paolini, F., Giunta, T. (1625). *Galenus Opera ex nona Iuntarum editione. Quae, quid superioribus praestet, pagina versa ostendit. Ad amplissimum Venetorum medicorum Collegium.* Venetiis : apud Iuntas. The first five volumes are bound together in a unique book with shelfmark 8.D.44.

the most advanced medicines during the Middle Ages, influenced medieval medical practices of Western Europe and remained highly regarded in the humanistic era as well. The books of some of its great masters, Al-Razi (Rey, 864 – 930), Avicenna (Afshana, 980 – Hamadan, 1037) and Averroes (Córdoba, 1126 – Marrakesh, 1198), were digitized and made available to the public online. One among all that can be mentioned is Abū Bakr al-Rāzī's book *Opera parua Abubetri filii Zacharie...*<sup>86</sup>, printed in 1511 in today's Lyon. The features of the book are those of a post-incunabile, written in blackletter, with woodcut initials, a red and black frontispiece and a valuable woodcut illustration (Saint George's slaughter of the dragon). Two works of well-known intellectuals belong to the same section: *De vita libri tres*<sup>87</sup> of the Neoplatonic and close collaborator of Lorenzo the Magnificent Marsilio Ficino (Figline Valdarno, 1433 – Careggi, 1499), and *Fasciculus Paracelsicae medicinae veteris et non novae*<sup>88</sup> written by Paracelsus, the Swiss physician who was among the promoters of the Medical Revolution in Renaissance through early studies in Toxicology and Pharmacology.

The Medicine between the 15<sup>th</sup> and 16<sup>th</sup> centuries is the theme of the third section. Some works still insist on the systematic study of the great works of the past. Oddo degli Oddi (Padua, 1478 – Padua, 1558), for example, strongly supported the theories of Galen, as it can be seen in his *Expositio, in librum artis medicinalis Galeni*<sup>89</sup>. Georgius Agricola (Glauchau, 1494 – Chemnitz, 1555), instead, was completely of another opinion and offered a larger sight on the discipline, reaching out for other disciplines, such as mineralogy and natural sciences. The Ateneo Veneto digitized one of his works, *De mensuris & ponderibus*<sup>90</sup>, a comprehensive study on weights and measures in use in Roman and Greek times. During the 16<sup>th</sup> century, disciplines such as astronomy, geography,

---

<sup>86</sup> al-Razi, A.M.i.Z., da Cremona, G., de Portonariis, V., Giolito de Ferrari, G., Villiers, G. (1510). *Opera parua Abubetri filii Zacharie ... que in hoc paruo volumine continentur sunt. Liber ad Almansorem decem tractatus continens cum nonnullis additionibus interlinearibus Gerardi Cremonensis ... Tractatus de egritudinibus juncturarum. De morbis puerorum. Aphorismorum ejusdem libri sex. Paruum antidotarium ipsius. Tractatus de preseruatione ab egritudine lapidis. Liber introductorius paruus in medicinam. De sectionibus et cauteriis ac ventosis. Synonyma ejusdem. Liber diuisionum cum nouem capitibus in fine additis et ab alijs impressoribus semper obmissis: quibus operibus additus est Constantini Monachi Viaticus.* Venundantur Lugduni : in vico Mercuriali sub signo Angeli (Impressus Lugduni : expensis Vincencij de portonarijs de tridino montisferrati : per Gilbertum de villiers, 1510 die vero viii mensis nouembris). Shelfmark: 7.A.257.

<sup>87</sup> Ficino, M., Ricci, G., Al segno della Speranza (1548). *Marsilii Ficini ... De vita libri tres, recens iam à mendis situque uindicati, ac summa castigati diligentia. Quorum. Primus, de studiosorum sanitate tuenda, secundus, de vita producenda, tertius de vita caelitus comparanda. His accessit Epidemiarum antidotus, ex idiomate Thusco.* Venetijs : ad signum Spei. Shelfmark: 6.C.17.

<sup>88</sup> Paracelsus, Dorn, G., Feyerabend, S., Spies, I. (1581). *Fasciculus Paracelsicae medicinae veteris et non nouae, per flosculos chymicos et medicos, tamquam in compendiosum promptuarium collectus ... Gerardo Dorneo interprete.* Impressum Francoforti ad Moenum. Shelfmark: 13.C.11.

<sup>89</sup> degli Oddi, O., Zara, O., Presegni, C. (1607). *Oddi de Oddis ... Expositio, in librum artis medicinalis Galeni. Nunc primum in lucem edita, et castigata laboribus, & vigilijs Marci Oddi medici eiusdem filij, ...* . Brixiae : ex typographia Comini Praesenij, ad instantiam Orlandus Zara bibliopola patauini. Shelfmark: 8.B.30.

<sup>90</sup> Agricola, G., Wechel, C. (1533). *Georgii Agricolae medici Libri quinque de mensuris & ponderibus, in quibus pleraque à Budaeo & Portio parum animaduersa diligenter excutiuntur.* Parisiis : excudebat Christianus Wechelus in uico Iacobaeo, sub scuto Basiliensi. Shelfmark: 7.A.281.



and botany, were deepened in order to see their relations with medicine and their direct or indirect relationship with the human body. The polymath Girolamo Cardani (Pavia, 1501 – Rome, 1576) was particularly interested in studying mathematics, astrology and astronomy, and physics. In one of the three books kept by the Ateneo Veneto, *Hieronimi Cardani, In Cl. Ptolemaei De astrorum iudiciis...*<sup>91</sup>, he dealt with the *Quadripartitum* of Claudius Ptolemy and the influence of stars on human beings.

The fourth section is completely dedicated to the aforementioned work of Girolamo Fabrici d'Acquapendente, *De formato foetu*.

The fifth one investigates the empirical approach, made up of experiments and remedies carried out directly in the field, that characterized the Renaissance when the medical tradition started to wane. The so-called practical medicine conducted studies in the field of anatomy and surgery. One of the most evident practical studies was the work *Exercitatio anatomica de motu cordis et sanguinis in animalibus*<sup>92</sup>, a duodecimo written by William Harvey (Folkestone, 1578 – Roehampton, 1657). He was a disciple of Girolamo Fabrizio and the first known physician to describe completely, and in detail, the systemic circulation and properties of blood. He, therefore, demonstrated that the quantity of blood in human beings is limited, in contraposition to the Galenic groundless assumption. Also noteworthy is the work *Delle fontanelle*<sup>93</sup> by Domenico Galvani (d. 1649), who explains the theory and practice of soft spots as a remedy to relieve and cure cysts and ulcers and other pathological formations. Moreover, there was an initial interest in the treatment of infections and their prevention (today's Epidemiology), as in the case of *Politica medica per il governo conservativo del corpo humano*<sup>94</sup> written by Bartolomeo Pietrigrassa.

Books dealing with botany, natural remedies and medicinal plants are grouped in the sixth section. In the 16<sup>th</sup> century, the birth of the first botanical gardens and herbariums marked the beginning of the emancipation of studies concerning plants. *Dell'istoria de i semplici aromati, et altre cose che*

---

<sup>91</sup> Cardano, G., Dasypodius, K., Henricpetri (1578). *Hieronimi Cardani, In Cl. Ptolemaei De astrorum iudiciis, aut (vt vulgo appellant) Quadripartitae constructionis lib. 4, commentaria ab autore postremum castigata, & locupletata. His accesserunt eiusdem Cardani De septem erraticarum stellarum qualitatibus atque uiribus liber posthumus, antè non uisus. ... Conradi Dasypodii ... Scholia et resolutiones seu tabulae in lib. 4. Apotelesmaticos Cl. Ptolomaei: ... Denique breuis explicatio astronomici horologii Argentoratensis, ...*. Basileae : ex officina Henricpetrina. Shelfmark: 14.A.49.

<sup>92</sup> Harvey, W., de Back, J., Leers, A. (1660). *Guilielmi Harveji ... Exercitationes anatomicae, de motu cordis & sanguinis circulatione. Cum duplici indice capitum & rerum. Accessit Dissertatio de corde doct. Jacobi de Back ...*. Roterodami : ex officina Arnoldi Leers. Shelfmark: 14.C.56.

<sup>93</sup> Galvani, D., Crivellari, G. (1620). *Delle fontanelle trattato di Dominico Galvani filosofo, e medico della mag. Comunità di Este. Diviso in duo libri. L'vno pertinente alla teorica, ... l'altro alla pratica, ...*. (In Padoua : appresso Gasparo Criuellari. Ad Instantia dell'auttore, 1620). Shelfmark: 13.C.16.

<sup>94</sup> Pietrigrassa, B., Magri, G.A. (1649). *Politica medica per il governo conservativo del corpo humano. Divisa in due trattati. Nell'uno si discorre d'alcune cose proemiali, nell'altro dell'aria. Con la cui salutare disposizione si mantiene la sanita. Opera ... di Bartolomeo Pietrigrassa fisico collegiato, lettore ordinario di filosofia nella regia Vniuersità di Pauia sua patria, et Accademico Affidato. Dedicata all'eccell.mo Senato di Milano. In Pavia : per Gio. Andrea Magri stampatore della città. Shelfmark: 8.E.50.*

*vengono portate dall'Indie Orientali pertinenti all'uso della medicina*<sup>95</sup> is the result of the work of the Portuguese physician and herbalist Garcia de Orta (Castelo de Vide, c.1501 – Portuguese Goa, 1568). He worked in Goa and Bombay in Portuguese India and there he learned an experimental approach to the identification and the use of herbal medicines, different from the approach of traditional medicine. The first part of the work, divided into four books, discusses all the plants used for medical purposes and coming from the East Indies, whereas the second part, in two books, discusses those coming from the West Indies. It therefore represents a complete dissertation of drugs and natural goods that at that time were cultivated in exotic countries and then were transported to Europe.

The 17<sup>th</sup> medicine is the theme of the seventh section of the Medicine collection of the Ateneo Veneto. Some of the physicians of this time were influenced by the scientific revolution of Galileo Galilei (Pisa, 1564 – Arcetri, 1642) and described the functioning of the human body as that of a machine. As a result, the studies in anatomy and surgery tried to understand the parts that constitute the human body and the way they interact with each other. In this regard, of interest, for example, is the anatomical study provided by Domenico Marchetti (Padua, 1626 – 1688) in his *Nova observatio, et curatio chirurgica*<sup>96</sup>. There is an illustration that shows a thumb and its anatomical division.

Finally, a new section, the eighth one, was dedicated to medicine books of the 18<sup>th</sup> century. Among various digitized authors that can be mentioned to understand what medical practices at the time were conducted, the Dutch physician Herman Boerhaave (Voorhout, 1668 – Leiden, 1738) definitely offers the most complete work. His seven-book work *Praelectiones academicae*<sup>97</sup>, collects the teachings that this great master developed directly on the field. They discuss chyfication, blood circulation, body organs and systems, external and internal senses, sleep, pathology, semiotics, hygiene, and therapeutics, and make it clear how medical studies were structuring themselves in complex science. Not by chance, he is regarded as the founder of clinical teaching and of the modern academic hospital. In Leiden, where he worked and taught, he reintroduced the practice of training at the patient's bedside. Students watched from a gallery overlooking the wards, while Boerhaave discussed for about two hours on each case. Among these students, one can imagine

---

<sup>95</sup> da Orta, G., de L'Ecluse, C., Monardes, N., Briganti, A., Salis, G. (1616). *Dell'istoria de i semplici aromati, et altre cose che vengono portate dall'Indie Orientali pertinenti all'uso della medicina. Di don Garzia Dall'Horto medico portughese, con alcune breui annotationi di Carlo Clusio. Parte prima, diuisa in quattro libri. Et due altri libri parimente di quelle cose che si portano dalle Indie Occidentali; con un Trattato della nene, & del beuer fresco. Di Nicolò Monardes ... Hora tradotti ... da messer Annibale Briganti, ...* In Venetia. Shelfmark: 7.A.253.

<sup>96</sup> Marchetti, P., Martini, I., Pasquati, G.B., Cadorino, M. (1654). *Petri de Marchetis ... Nova observatio, et curatio chirurgica, in lucem edita a Iacobo Martini Germano, D. Patauui* : typis Io. Bap. Pasquati. Shelfmark: 13.C.40.

<sup>97</sup> Boerhaave, H., von Haller, A., Occhi, S. (1743-1745). *Hermanni Boerhaave ... Praelectiones academicae in proprias institutiones rei medicae. Edidit, et notas addidit Albertus Haller [...] Tomus primus [- sextus]*. Venetiis : apud Simonem Occhi sub signo Italiae. The seven books have shelfmark 8.B.11-17.

there was Gerard Van Swieten (Leiden, 1770 – Vienna, 1776), who became the personal physician of Empress Maria Theresa of Austria. The digitization of the four books of his *Commentaria in Hermanni Boerhaave Aphorismos...*<sup>98</sup> gives access to the comments and notes made by him in reference to the master Boerhaave. Other interesting books are the two volumes of *Corporis humani anatomia*<sup>99</sup>, written by Philippe Verheyen (Verrebroek, 1648 – Leuven, 1710), show the latest studies in anatomy and surgery at the beginning of the 18<sup>th</sup> century. The two books are accompanied by 45 anatomic illustrations and the engraved portrait of the author.

From an operative point of view, all the relevant information about the books was managed inside a spreadsheet, where the convenience of Excel formulas and cells easily allowed the process of data management and made it possible to combine and analyse data. This list of originals included shelfmarks, sections of the collection, data about the author and the work, and features of the books. They were used to create the nomenclature of files and folders. Nomenclature and data management will be discussed thereafter.

## 2.3 Institutions and Project Team

Human resources are a crucial point in planning digitization projects. The major projects carried out by national libraries and museums, or the most important private cultural institutions, usually have large corps of internal staff that are gathered in a project team. Each of them has specific skills and is responsible for specific tasks, such as: selection of materials (conservator and restorer); transport and relocation (logistician); cleaning and restoration of the material (restorer); production of inventory data and cataloguing (cataloguer); choice of hardware and its optimization (technical manager); digitization (technical operator); data management (IT specialist and computer assistant).<sup>100</sup> These experts are coordinated by the project manager, who supervises the project and is responsible for the achievement of the results, ensuring the effective use of resources assigned and assessing objectives and constraints (time, costs). The project manager reports to the institution that is carrying out the project and – in case of the public sector – in particular to the Sole Manager of the Procedure (R.U.P.), who is part of the administration and legally monitors compliance with contracts. Besides the project team, the digitization project directly involves the institution which

---

<sup>98</sup> van Swieten, G., Boerhaave, H., Pellicchia, A. (1755). *Gerardi L.B. Van-Swieten ... Commentaria in Hermanni Boerhaave Aphorismos de cognoscendis et curandis morbis Una cum Boerhaave ipsius vita ex ejus schedis collecta. Tomus primus [-?]*. Neapoli : apud Alexium Pellechia. The four books have shelfmark 8.B.8-10.

<sup>99</sup> Verheyen, P., Selsoni, F., Gessari, B. (1734). *Corporis humani anatomiae liber primus in quo tam veterum, quam recentiorum anatomicorum inventa, methodo nova, & intellectu facillima describuntur, ac tabulis aeneis representantur. Authore Philippo Verheyen ...*. Neapoli. The two books have shelfmark 13.B.24 and 13.B.25.

<sup>100</sup> Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library (2022), p. 8.

is the promoter of the project. In some cases, projects of great importance interest a net of institutions, contributing to the project in different ways. There is the main cultural institute that owns and keeps the collection and promotes its digitization, involving its internal staff; the research centres, which may be consulted for technical opinions or the supply of experts or equipment; universities, which can collaborate by offering human resources, particularly students, led by a professor, for study and research purposes; enterprises, which usually provide private hardware or sponsorship; other entities, foundations and associations, which may be funders or partners in the project.

With reference to small cultural institutes, it is clear that they cannot count on large-scale human resources nor on a high number of specialists. Many cultural organisations lack a sizable staff that “have a great deal of free time to carry out digitisation projects, over and above their usual duties.”<sup>101</sup> This type of institute may use project contracts to involve external personnel (as happened in the case of the Ateneo Veneto), selecting capable figures among university students who need to conduct their curricular internship coming from universities, or involving high schools and recruiting young people with school-work alternation. The principal advantage in choosing young people is that they are naturally predisposed to the use of technology. What is important in this case is to preserve the contribution of skills required by digitization. It is better to have fewer but capable people, who are proactive and equipped with transversal skills, rather than investing in some experts in the field, who instead are not in symbiosis with each other and lacking in digital skills. Specialization in the discipline is an important requirement for achieving high quality, but at the same time, the implementation of similar projects requires interdisciplinarity, a combination of humanistic and digital knowledge and skills (digital humanities), problem-solving, practicality, skills that go beyond a single profession. And most importantly, the success of each project, including digitisation, is also determined by a group of figures which is coordinated and result-oriented.

If, on the one hand, small cultural institutes cannot count on all those figures described above for a project, they have a great advantage: the sense of belonging to an institution and a drive to build a project that can emphasize the institution and its heritage. The important thing is that one internal figure assumes the role of project manager and works on the planning and the procedures. This figure can serve as a technical manager at the same time. It might be desirable that one or two members of the internal staff, dedicate some of their time to contribute or to assist the project team, depending on the roles they hold in the institution. The figure who is in charge of the conservation of the heritage can evaluate the items to be digitized, select them, and provide

---

<sup>101</sup> Minerva Working Group 6, ed. (2004), p. 8.

information about them, in view of the metadata process. In case of the public sector, the legal representative of the institute can be the R.U.P. As for the IT specialist, this figure may be absorbed by the technical executor, who must have suitable computer skills. This will lead to a situation where few people working entirely on the project can transversally fill in for experts who for economic reasons could not be involved.

Regarding the Ateneo Veneto, its internal staff consists of a general and administrative secretariat, a coordinator of library services and events manager, a librarian, a press officer, and a technician, who did not directly plan and execute the project, but who in various ways contributed to it. They worked on the institute's proposal to the funding call of the Fondazione di Venezia, they managed the expenses and other technical aspects, provided information regarding the collection and the available instrumentation, carried out archival research, updated the cataloguing cards and promoted the initiative. The technicians working for the Library System of the University of Padua, the university that manages the designated repository, Phaidra, provided advice and solutions during the project. The project team was essentially composed of two people: Prof. Dorit Raines and the author. This makes it clear how few hybrid figures are essential for projects of small cultural institutions and can deal with both planning and project execution. When necessary, they can ask for both internal and external advice, using their connections.

## 2.4 Planning the Project

Having found a suitable source of funding to ensure the economic stability of the project and determine the personnel to be employed, the crucial and most important part of every digitization project initiates. The institution, together with the project manager, has to strategically decide how the objectives of digitization will be achieved. In particular, they have to resolve how they can digitally provide access to their collection and the way to enhance the consultation of their digital objects.

Some strategic decisions have to be implemented to achieve these goals. More specifically, the first decision to be taken concerns the optimization of the fruition experience of digital objects and, therefore, the choice of the hosting platform. Indeed, the repository must adhere to specific criteria and be compliant with specific standards. Operationally, the choice impacts a lot on the digitalization process. The production of images, the type of editing, metadata and the same publication are performed and accordingly adapted to the platform.

Secondly, the institution has to decide on a reliable location where the digitization process will be held. Possible alternatives are two: in-house or outsourced digitization. When planning on whether to digitize within the institution (in-house) or the use of outside services (outsourcing), it is important to consider advantages and disadvantages of each solution and to effectively evaluate them in respect to the limits imposed. An in-house project implicates the examination of technical and logistical aspects.

Thirdly, digitization projects necessarily require the description of digital objects. Metadata allows for identification, access, use, and preservation of digital resources and is therefore directly associated with most of the steps in a digital imaging project workflow. Although it can be costly and time-consuming to produce, metadata represents the identity card of digital objects and allows their searchability and consultation. Images devoid of sufficient metadata are at greater risk of being lost or unfindable.

Finally, the institution must provide a data management plan for the long conservation of its digital collection. This must be implemented on two fronts: on the one hand, by publishing the digital collection on a long-term conservation hosting platform; on the other, by having a correct storage and backup management plan.

### 2.4.1 The Choice of the Long-Term Archival Platform

The digitization process requires the adoption of procedures to guarantee that digital objects remain usable and accessible regardless of future changes in technology. Over time, if correctly carried out, these procedures can contribute to avoid repeating the expensive and time-consuming scanning and cut off the expenses. The first solution for this is an adequate repository.

Digital repositories<sup>102</sup> are essentially digital collections for which:

- content is deposited, whether by the content creator, owner, or third party;
- the repository's architecture manages both content and metadata;
- the repository provides a minimal set of basic services, such as put, get, search, and access control;
- and the repository must be reliable, trustworthy, well-supported, and well-managed.<sup>103</sup>

At an international level, one of the most reliable certifications that define requirements and core characteristics of long-term, trustworthy data repositories is CoreTrustSeal. Launched in 2017,

---

<sup>102</sup> For a complete list of registered data repositories, *cf.* <https://www.re3data.org>.

<sup>103</sup> Denison, T. (2007).

CoreTrustSeal provides rules and guidelines every three years to improve the quality and transparency of repositories and awards certification to institutions that have complied with the specifications. The organization has recently issued a document which establishes requirements for the triennium 2023-2025.<sup>104</sup> They assess, for example, the organisational infrastructure in terms of its goals, legal obligations, and ability to continue providing services as well as the digital object management in terms of its provenance and authenticity, conservation strategy, and quality control. They evaluate IT and security, in terms of the platform's ability to securely store data and maintain its integrity. Therefore, the certification provides a thorough understanding of the many requirements a repository must achieve in order to be deemed durable and accessible. When planning a digitization project, CoreTrustSeal enables the project manager to make decisions regarding the repository.

At an institutional level, the Ateneo Veneto looked for a long-term archival information system that could satisfy these requirements: compliance with FAIR principles; tools to facilitate dissemination and reuse of the deposited digital objects (especially for books); feasibility.

The FAIR principles, recommended by the Association of European Research Libraries in 2018 and by the European Commission in 2020, are guidelines to improve the Findability, Accessibility, Interoperability, and Reuse of digital assets. They were established, using the acronym FAIR, in an article published in the journal *Scientific Data* in 2016 by a group of scientists and organisations.<sup>105</sup> The main objective of these principles is to make data (especially scholarly data) usable by computational technologies. Hence, it is important to define how data, metadata and the IT infrastructure are constructed. In order to re(use) data, the first step is to be able to locate them. Moreover, the user has to know how data can be accessed, including authentication and authorisation. In addition, data need to interoperate with applications or workflows. Finally, data need to a robust metadata schema so that it can be replicated in different settings. FAIR principles are the basis of the adequate construction of a digital platform and the dialogue between data and metadata (inward and outward).

As for the second point, the Ateneo Veneto needed a platform that was suitable for the collection to digitize. The repository needed to have technical experience with library collections, to possibly promote the interconnection between materials of different cultural nature and to be provided with digital high-resolution display and presentation tools for the library heritage.

---

<sup>104</sup> CoreTrustSeal Standards and Certification Board (2022).

<sup>105</sup> Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. (2016). Denison, T. (2007).

Finally, the Ateneo Veneto, like all other small cultural institutes, had to deal with economic and temporal feasibility. It was fundamental to select a repository that would allow for the conclusion of an agreement without compromising the institution's financial resources. The agreement itself was to be concluded as soon as possible, in order to respect the project deadline.

The final choice of the institution fell on Phaidra, the platform used by the University of Padova Library System (SBA) and maintained by the University of Vienna, with whom the University of Padova has been collaborating since 2010. The word Phaidra is an acronym which stands for "Permanent Hosting, Archiving and Indexing of Digital Resources and Assets". There were several points in favour of the stipulation of the agreement. Firstly, the agreement between the two institutions could be signed in a reasonable time and the platform was already used by other institutes in the same geographical area: Ca' Foscari University of Venice and IUAV University (since 2014); University of Bologna (since 2019). From a technical point of view, then, the characteristics of Phaidra responded to the recommendations of the OAIS (Open Archival Information System) and those of CoreTrustSeal. As a comprehensive university digital asset management system with long-term archiving functions, the repository was awarded the certification in 2019<sup>106</sup>. One of its strengths is its embedding in the open-source software environment Fedora, "Flexible Extensible Digital Object and Repository Architecture", dedicated to the construction of digital libraries and archives with native support for Linked Data.<sup>107</sup> Phaidra is also compliant with the FAIR principles. In 2019, Phaidra was in fact selected with other repositories to test innovative developments in the context of the European "FAIRsFAIR" project.<sup>108</sup> The platform assigns to each digital object a permanent link and handles a persistent identifier that certifies to the user its retrievability.

Moreover, the infrastructure uses standard metadata. Phaidra objects are described through a metadata schema called UWmetadata (Universität Wien Metadata).<sup>109</sup> It is an extension of the standard LOM (Learning Object Metadata) schema. Every object is provided with this typology of metadata and the most widely used metadata profile, Dublin Core (DC). Since 1995, DC metadata has represented a concise and simple schema, consisting of only 15 elements, all optional and repeatable, that can be placed in any order. Dialogue between the two different types of metadata is allowed by conversion mapping.<sup>110</sup> The use of DC metadata standardizes Phaidra's objects and facilitates interoperability with external repositories, such as Europeana and Internet Archive. The

---

<sup>106</sup> CoreTrustSeal, *Assessment Information. Phaidra at the Library System of the University of Padova* (2019).

<sup>107</sup> Andreoli, L., Bianchi, E., Cappellato, L. et al. (2019), p. 151.

<sup>108</sup> Bettella, C., Carrer, Y., Turetta, G. (2022).

<sup>109</sup> Cappellato, L. (2023), p. 5.

<sup>110</sup> *Cfr.* Bettella, C. (2018).



direct dialogue with these and other platforms (CulturaItalia, MOVIO, etc.) is possible because of the implementation of APIs and protocols.

With regard to contents, Phaidra stores at present 500 thousand objects that embrace the GLAM (Galleries, Libraries, Archives and Museums) perspective, which means that different cultural objects are available: pictures, documents, books, and educational resources. The result is that it offers the users good possibilities for multidisciplinary connections, and facilitates serendipity. In order to optimize object consultation, many technologies have been developed, including an image viewer, a book viewer, and a collection viewer. The Ateneo Veneto chose the BookViewer option because it suited its needs of publishing. In fact, Book Viewer has specialised features for books, such as the capacity to browse and index them.

In Chapter 4, the platform receives a more thorough study.

## 2.4.2 Budgeting and the Choice of In-House Digitization

After choosing the platform that will host the digital collection, small cultural institutions must decide on the location where the digitalization process will take place. The possibilities are two: in-house or out-sourcing. Most often, this choice can already be influenced by the platform itself. In evaluating one repository, the institute acquires information about practices, requirements of digital objects and their production, and minimum standards. It can even go so far as to exclude in-house digitization and comply with platform technicians' standards that might not be internally guaranteed. The platform may also hold cutting-edge tools or a digitization setup tailored to the kind of cultural items the institute wants to digitise. In both cases, the institute will be inclined to outsource digitization.

More often than not, small institutions are completely aware of the platform's expectations and requirements that directly influence this decision but still have to deal with their own structural limitations and reach a compromise. In particular, they have to weigh three main imposed constraints: lack of economic resources, lack of space and lack of personnel. The availability of finances and their amount is one of the persistent themes of the cultural sector. Institutions are more and more likely to submit sponsorship proposals, funding calls and crowdfunding campaigns because they cannot rely entirely on public resources. If they can benefit from project financing, the agreement itself between the parties will also influence the planning of digitization. The case of the Ateneo Veneto is explanatory. The institution received funding from the Fondazione di Venezia with the conditions of activating "innovative skills and planning on the territory" and

offering “young people opportunities for professional growth”.<sup>111</sup> It is evident that these requirements directly involved the choice of an in-house project. Nevertheless, obtaining funding is not enough to necessarily meet all the requirements or to achieve the quality the institute predetermined. Most of the time, funding is good, but it is not as adequate as it should be to cover the costs of professional equipment, which is out of doubt expensive. The institute needs to budget for all the costs, according to the given amount. Thus, there can be identified a trade-off between cost and professionalism, which often corresponds to a trade-off between cost and quality.

Furthermore, two more constraints have to be considered. The first, the lack of space, directly affects the choice of equipment from a logistic point of view. It is unthinkable that a small institution buys (large) digital machines, because they are state-of-the-art technologies, and does not know where to put them. By the way, once the project is finished, it is not obvious that the institute will run a new digitization project in a short time and this equipment would become an unwieldy purchase. In addition to the equipment, the lack of space directly affects the working environment, which must satisfy some requirements. It should be a single photographic laboratory, uniquely used for the digitization project. It has to be also appropriate to the material being digitized, paying special attention, for example, to light, humidity, vibration, disturbance, and handling of the originals.<sup>112</sup> The working room must be large enough to allow the movement around the equipment and to contain, in addition to the technical equipment, one or more desks to store the material in transit, and compiling relevant records about handling and operations.

Finally, small cultural institutes need to find at least individuals to work on the project. It is not critical that they are both digitization experts. As already mentioned, it is much more important that they have transversal competencies and that overall, the contribution of skills required by digitization is preserved. The methods and profiles to be identified have already been expressed.

A question may arise spontaneously. Is it reasonable to discuss digitization in tiny cultural institutions, and especially to discuss in-house digitization, where neither satisfactory economic resources nor adequate space and the number of employees are present? The response is based on the potential actions and proposals that each institution may formulate. Nevertheless, there are two principles that must inspire digitization. Whether it is in-house or outsourced, digitization must be planned by the institution, which is the depositary and the guarantee of the proper conservation,

---

<sup>111</sup> Fondazione di Venezia, *Bando “Cultura Digitale”* (11 April 2022).

<sup>112</sup> Minerva Working Group 6, ed. (2004), p. 23. Standards required for the digitization environment are stated by ISO 12646. For example, the working room should be “painted with neutral colours.” Monitors should be positioned “to avoid reflections and direct illumination on the screen.” The room illumination should be less than 32 lux between the monitor and the observer, and the light a colour temperature of approximately 5000K and a CRI above 90. For other specifications, *cf.* Federal Agencies Digitization Guidelines Initiative (FADGI) (2023), p. 23.

enhancement, and fruition of its heritage. It is the one that comprehends its cultural value and knows how to promote it more than anybody else. Secondly, good planning has to be done with efficacy, efficiency and, above all, resourcefulness. Resourcefulness may be used to tackle a variety of issues.

The choice of one or the other digitization modality will depend on more specific factors. Outsourced digitization is particularly recommended when originals can be moved without damage, or it is not possible to digitize them within the institution. It is also suggested when the project regards a large number of selected materials and must be performed in a short timeframe or even when there are constraints of space, infrastructure and personnel that exclude any possibility of in-house digitization.<sup>113</sup> The institution will not be concerned about arranging the digitization set; it will simply be responsible for covering the digitization expense to the supplier. Even though it may appear like outsourcing is the best option for tiny organisations, there is one major drawback. Outsourcing totally eliminates one phase of digitization, the production of digital objects, from the control of the institution, so that it does not develop in-depth knowledge of the process. In other words, a stage at which the institution can be renewed by a process of “learning by doing” is completely absent.

In-house digitization is recommended when the collection of originals cannot be moved outside the institution or when digitisation is tendentially easy and can be performed by internal staff. This point depends on standards and on actions to be performed. This internal execution guarantees close control over all procedures, from the handling of materials to the quality of products. “There is no need to send valuable or fragile originals off-site and no worry about working with a vendor who turns out to be incompetent, provides something other than what was required, or goes out of business.”<sup>114</sup> The conservation of the originals, especially when rare and precious, is certainly more guaranteed. Moreover, the institute can equip itself with hardware and software and with a wealth of knowledge and expertise that goes beyond the same digitization and can benefit the entire organisation. However, it must be noted that while purchasing hardware may be a good medium- and long-term investment for the organisation, it may also be problematic due to technical obsolescence.

These were also some limitations, uncertainties, and opportunities considered by the Ateneo Veneto. As mentioned, the choice of an in-house project was almost a mandatory choice for compliance with the conditions of the funding call of the Fondazione di Venezia. Nonetheless,

---

<sup>113</sup> Andreoli, L., Drago, G. (2022), p. 4.

<sup>114</sup> Sitts, M.K. (2000).

there were internal needs. It was impossible to move a hundred early-printed books from the headquarters of the institution to Padua, where the University of Padua's Library System (SBA), which manages Phaidra, has its digital technology labs. Furthermore, the platform provided guidance through recommendations rather than rigid specifications that set quality standards or required the usage of particular technologies.

As a result, the institute had to deal with technical, economic, and logistic aspects. First, a market analysis was done to identify the top equipment models that suited the project needs and the economic resources. All expenses had to be budgeted looking for a compromise between cost and performance. Secondly, it was decided that the "studiolo", a narrow but adequate room, could be used as a dedicated digital working room. The photographic laboratory, made of a PC, monitor and hypothetical planetary scanner, could comfortably be housed there, enabling a digitisation process that could satisfy the principal needs. It was close to the book depository, and this has ensured the preservation of books and the acceleration of their handling. It could also be obscured with curtains, avoiding the interference of external light. It is a single room, where the technical manager and the technical operator could work without disturbance.

### 2.4.3 The Choice of Hardware and its Optimization

Each digitization project that involves the digital reproduction of analogue objects passes through a specific choice of instrumentation. In the circumstance that in-house digitization occurs, this choice is one of the project team's debate topics. The purpose of this task is to find the best hardware and software components that, once optimized, can guarantee the quality required by the project, achieve its objectives and, at the same time, respect the originals' physical integrity.<sup>115</sup> Considering the digitization of two-dimensional documents, the instrumentation typically consists of digital image capture equipment (digital cameras or scanners, light sources) and an appropriate computing platform (computer and visual display unit). Together with hardware, the institute will equip itself with the most appropriate software: an acquisition device management program, a digital asset management (DAM) app, and a raster graphics editor. The decision will be implemented in conformity with the budget, objectives, and the typology of originals. The balance between these three criteria must attain the optimum quality for the digitization process. For the smaller cultural institutions, this will turn into a form of compromise, though. In fact, while it is certainly true that only professional equipment can produce products of the finest quality, small institutions will inevitably be unable to acquire such equipment due to its high cost. The restricted

---

<sup>115</sup> International Federation of Library Associations and Institutions (IFLA) (2015).

budget they can allocate can only permit the purchase of digital acquisition instruments at an estimated expense of a few thousand euros overall. With this in mind, the Ateneo Veneto conducted a specific market analysis and selected the items that might best satisfy both the needs of the originals and those of the project.

#### 2.4.3.1 Planetary Scanner

The choice of the scanning system is firstly determined by the typology of materials. Today, flat scanners, planetary scanners, and digital cameras mounted on tripods are the most frequently utilised pieces of equipment in documentary digitization projects.<sup>116</sup> The available budget (less than one thousand euros) and the characteristics of the books oriented the choice of the Ateneo Veneto specifically towards the typology of the planetary scanners, which are specially designed for the digital reproduction of bound volumes. Basically, they are photo reproduction machines. High-end planetary scanners are equipped with book cradles, their own side and/or front lighting systems, and a fixed arm (planetary scanners with fixed head on a tripod) or a moving arm (planetary scanners with moving arm head) where a camera perpendicular to the support plane of the originals is mounted. They can have their own processor and visual display unit. In contrast to flatbed scanners, books are posed upwards – and therefore are not pressed against a surface, which considerably damage their binding – while the scanner captures a digital image of the document from above. Plastic or glass platens can also be used to guarantee the stability and preservation of the items.

Market research was driven by the awareness that many features have to be surrendered. Automated and semi-automated planetary scanners were overpriced. The choice to exclude automatic products was also dictated by the fragility of the originals and the desire to ensure their safety. Among other things, the presence of inserted leaves and illustrations folded inside the books was not compatible with this type of products. Solutions thus considered were evaluated through a series of technical parameters. They are normally used to describe the quality of the images that are obtained from every process of digitalization, namely raster graphics. Raster graphics are two-dimensional images composed of a tessellation, or “pix-map”, of a definite number of pixels. Pixels are the smallest elements constituting a raster image and their characteristics define the quality of the entire image. The first parameter to take into consideration is resolution, which corresponds to the total number of pixels in an image. The more of these fundamental elements there are, the more defined the image will be. There are two types of resolution. The first, the input resolution,

---

<sup>116</sup> Barbuti, N. (2022).

regards the number of pixels present in the camera sensor that captures the light to create the image. The resulting image will have a number of pixels that is equal to or less than the total number of pixels contained in the sensor image.<sup>117</sup> Pixel density is measured in number of pixels per inch (2,54 cm), the so-called PPI. The second type, DPI (dots per inch), measures the pixel quantity of a raster image when it is intended to be printed. Indeed, modern inkjet printers convert digital pixels into microscopic dots. This is the output resolution.

The quality of a raster image is also the result of the colour rendering. In this case, the quality depends on the number of bits<sup>118</sup> used to indicate the colour of a single pixel. When referring to a pixel, the concept can be defined as bits per pixel (bpp or bpc, bits per channel). 1 bpp assigns only one couple of alternatives for each pixel, producing a binary image made of two colours (frequently black and white). 2 bpp produce a four-colour image, 3 bpp a  $2^3$  (eight)-colour image, and so on. The more bits of information per pixel there are, the more available the colours are and the more accurate the image representation will finally be. The quantity of bits per pixel is called colour depth or radiometric resolution. A common standard in digitization is 8 bpp and it provides an image with 256 different colours.<sup>119</sup> High-quality digitization often uses cameras that capture images with a colour depth of 16 (also known as “High color”), 24 (“True Color”) or even 48 bpp (“Deep color”). These numbers indicate how many colours the camera can capture, not which specific colours they are. Another element defines them and it is called colour model. The most common colour models are RGB (Red - Green - Blue) and CMYK (Cyan - Magenta - Yellow - black)<sup>120</sup>. The representation of colours as tuples of integers, often as three (RGB) or four (CMYK) values

---

<sup>117</sup> Docs Italia, *Linee guida per la digitalizzazione del patrimonio culturale. Versione 1.0* (2022), p. 60. Resolution is typically expressed as a multiplication between pixel grid sizes, for example, 5,000x5,000. The result of this multiplication is in the order of million pixels and therefore is expressed in megapixels (MP). The camera software may have falsely boosted the number of pixels if the final image has a greater number of them. This process is called interpolation.

<sup>118</sup> Bits, or “binary digits”, is the fundamental computational unit and it represents a logical state which can assume one of the two possible values, e.g., “0” or “1”. Considering digital images, bits express the colour gamut associated with each pixel. This is the reason why raster images can be described as “pix-maps” or “bitmaps”.

<sup>119</sup> Barbuti, N. (2022).

<sup>120</sup> Wikipedia. The Free Encyclopedia (2023), under “Color model”. RGB is an additive colour space based on three colours: red, green, and blue. These colours are considered as primary colours. In fact, they can produce a large number of colours by an additive colour mixing. Typical RGB input devices are colour TV, cameras, and image scanners, whereas RGB output devices are TV sets of different technologies (CRT, LCD, plasma, OLED, quantum dots, etc.), computer and mobile phone displays, video projectors, multicolour LED displays and large screens such as the Jumbotron. Colour printers are not RGB devices, but subtractive colour devices typically using the CMYK colour model. CMYK is a subtractive colour model which uses cyan, magenta, yellow, and black. The first three colours are called subtractive primary colours, whereas black is added to improve reproduction of some dark colours. Such a model is called subtractive because inks “subtract” red, green, and blue colours from white light, which is on background: the subtraction of white light and red gives cyan, that one between white and green gives magenta, and finally the subtraction with blue leaves yellow. Cyan, magenta, and yellow components are used for colour printed reproduction and they may be viewed as the inverse of RGB. However, there is not an exact correspondence between RGB and CMYK colour models. So that, colours can change when a digital image defined using the RGB colour model is intended to be printed with a CMYK printer.

or colour components, is described by these abstract mathematical models. The resultant collection of colours, obtained by addition or subtraction of colour components, is known as “colour space” when this model is coupled with a clear description of how the components are to be interpreted, taking into consideration hue, saturation, and brightness.<sup>121</sup> The result is that, for example, RGB images made of three channels (red, green and blue), with 8 bpp each, are rendered with more than 16 million colour values, since every channel has 256 possible values to be represented. The final chromaticities will finally depend on the colour spaces used. Colour spaces like Adobe RGB and sRGB are different, while both rely on the same RGB colour model.<sup>122</sup>

Resolution and colour rendering are important characteristics to evaluate when choosing a scanner and its camera sensor. Digital images are made of light and colour information. The visual result is obtained by the elaboration of this information through the camera processor.<sup>123</sup> The software can apply compression to create the digital image. Image compression is the last variable to consider when choosing a device. It determines the final format of the image which can contain a large number of information, and thus consume a large amount of storage. They can be lossless or lossy. In lossless compression, original data can be perfectly reconstructed from the compressed data with no loss of information. In lossy formats, on the contrary, original data can be reconstructed only with an approximation. An example of a lossless format is TIFF, whereas JPEG format is lossy.<sup>124</sup> These two formats can be automatically produced by the processor of the camera. The result is a digital image. Best digitization practices suggest instead acquiring images in RAW format. This format preserves unprocessed or minimally processed data containing all voltages measured

---

<sup>121</sup> Colour spaces are plane or three-dimensional spaces where colours are represented by three attributes of colour: hue, saturation, brightness. They made it possible to convert the mathematic model of the colour model and to perfect it by using specific rules.

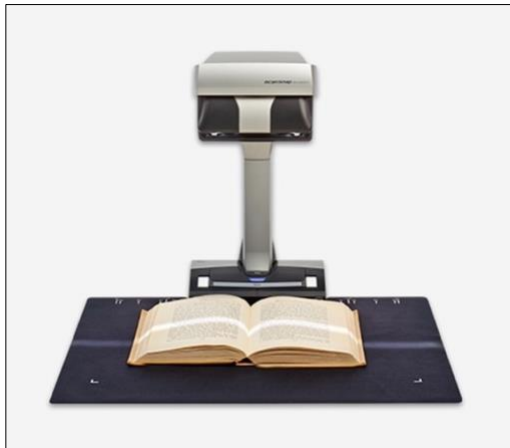
<sup>122</sup> Wikipedia, *Color space* (2023). sRGB is an RGB created cooperatively in 1996 by HP and Microsoft. It is the current defined standard colour space for the web. Adobe RGB colour space was developed by Adobe Inc. in 1998. It was designed to encompass most of the colours achievable on CMYK colour printers, but by using RGB primary colours on a device such as a computer display. It improves upon the gamut of the sRGB colour space, primarily in cyan-green hues. The largest colour space is ProPhoto RGB, developed by Kodak. It encompasses over 90% of possible surface colours in the CIE L\*a\*b\* colour space, and 100% of likely occurring real-world surface colours documented by Michael Pointer in 1980.

<sup>123</sup> Docs Italia, *Linee guida per la digitalizzazione del patrimonio culturale. Versione 1.0* (2022), p. 60. The camera processor elaborates information associated with its sensor. The sensor is made of millions of photodiodes, which are capable of recording the intensity of the captured light but not a specific colour. To accomplish this, special filters (known as Bayer or Bayer Matrix filters) allow photodiodes to record only light in specific wavelength ranges. A quarter of the Bayer filter can read red light, another quarter blue light and the remaining half of the filter reads green light. Using RGB colours, the Bayer filter records three wavelength ranges. Through an algorithm, the processor of the camera will finally convert these analogue electric signals into a digital representation by additively rendering all colours.

<sup>124</sup> Docs Italia, *Linee guida per la digitalizzazione del patrimonio culturale. Versione 1.0* (2022), pp. 58, 60. TIFF is the acronym of “Tagged Image File Format”, a widespread format of raster type, developed by Aldus Corporation and today held by Adobe. It allows considerable flexibility and specifies numerous additional indications such as gamut tables or colour calibration information. As for JPEG, it is the acronym of the Joint Photographic Experts Group, which created the standard in 1992. JPEG encoding greatly reduces the size of a file. The compression is based on the application of suitable averages between adjacent pixels, both with respect to luminance signals and chroma.

by the sensor. This large amount of potentially redundant data needs some additional steps before being displayed (for example, white and colour balance). Then it is saved and converted to other formats. RAW formats make therefore the digitization process reversible, allowing tracing back to the native pre-editing file.

A compromise was reached at the Ateneo Veneto between instrumentation that could satisfy these parameters and the available budget. Among all, two products were considered at the end of the market research: FUJITSU ScanSnap SV600 (*Figure 2*) and CZUR ET-25 Pro (*Figure 3*).



**Figure 2** (on the left) FUJITSU ScanSnap SV600 planetary scanner. ©Ricoh, ScanSnap® SV600 (n.d.)

**Figure 3** (on the right) CZUR ET-25 Pro planetary scanner. ©Scannx, CZUR ET25 Pro Overhead Scanner (n.d.)

A comparison between the two planetary scanners is provided in *Table 1*. These scanners use a different image capture technology. Planetary scanner Fujitsu is equipped with a CCD image sensor on a movable, rotating overhead. CZUR scanner, instead, has a CMOS image sensor on a fixed overhead. Both types of sensors accomplish the same task of capturing light and converting it into electrical signals. Generally, the image quality of a CCD sensor is better than the one provided by a CMOS sensor. However, the latter is better than CCD sensors in aspects such as power consumption and price.<sup>125</sup> Although Fujitsu's CCD sensor was preferable to be chosen, this sensor was mounted on a movable head that was equipped with its own directional LED lamp. This feature was considered since it prohibited the scanner and its lighting system from being further optimized for the project's needs. In terms of image resolution, the CZUR ET-25 Pro scanner was

---

<sup>125</sup> Charged Coupled Device (CCD) and Composite Metal Oxide Semiconductor (CMOS) sensors work on two different imaging technologies of modern cameras. On the one hand, CCD sensors undergo a special production process. Pixels are recorded on the chip and then sent one by one to create the image. Because of this process, they can create high-quality images with a low noise factor. On the other hand, CMOS chips make use of transistors on each pixel and the charge moves through traditional wires. Each pixel receives individual treatment. These sensors have an integrated circuit for reading pixels which is capable of transmitting all data at the same time and therefore less energy is consumed.



likewise seen to be superior: it captures images at a resolution of 25 MP, a pixel density of 330 dpi and a colour depth of 24-bit sRGB (8 bits per channel). On the contrary, FUJITSU ScanSnap SV600 had an inferior, variable resolution (max 285 dpi), which horizontally and vertically changed depending on the scanning angle on the same document. This was due to the fact that this scanner uses a linear, rotating sensor. No information could be found with respect to image colour depth. Furthermore, for both scanners, the size of the scanning area corresponded to that of an A3 paper size (297x420 mm). This was suitable for scanning the project books since the single page of a folio format (maximum height 43 cm) could be also contained inside. However, scanners provided a different scanning speed: the one of Czur was twice as fast as that of Fujitsu.

Finally, the last point of comparison was the master format. Regrettably, neither scanner allowed for the acquisition of images in RAW format. Indeed, the software applied in both cases a lossy compression and digital images were saved using the JPEG format. As a result, there were structural limits to the project that started directly from the digital image capture equipment and did not permit the acquisition of high-quality images in which unprocessed data could be saved in a raw or lossless format. Long-term stored files had already suffered data loss and the digitization process could not be thought to be reversible.

**Table 1** Comparison between FUJITSU ScanSnap SV600 and CZUR ET-25 Pro.

	<b>FUJITSU ScanSnap SV600<sup>126</sup></b>	<b>CZUR ET-25 Pro<sup>127</sup></b>
<b>Sensor type</b>	CCD linear image sensor	CMOS
<b>Scanner type</b>	Movable overhead	Fixed overhead
<b>Resolution</b>	Horizontal scanning: 285-218 dpi Vertical scanning: 283-152 dpi	330 dpi
<b>Colour depth</b>	-	24 bpp
<b>Master format</b>	JPEG	JPEG
<b>Scanning speed</b>	3s	1,5s
<b>Size area</b>	A3	A3

*Source:* author's elaboration from information provided by manufacturers (2023).

<sup>126</sup> Cfr. FUJITSU, *ScanSnap SV600* (2023).

<sup>127</sup> Cfr. CZUR, *CZUR ET Smart Book Scanner* (2023).

Results of the market research and the comparison between the two products were presented to the Ateneo Veneto and after a debate on advantages and disadvantages, the choice was finally made in favour of the planetary scan of the manufacturer CZUR, the CZUR ET-25 Pro scanner.

With reference to guidelines provided by Phaidra<sup>128</sup> and the Ministry of Culture, it is clear that the image quality is lower than the standard. File format required in digitizing rare and precious books is TIFF format (or, if possible, a raw format) with a minimum resolution of 600 dpi (400 dpi if it is a large book) and a 24/48-bit RGB colour space. These guidelines were used to organise the project, but in practice, Ateneo Veneto was unable to buy pricey machinery that could ensure the requirements. As with any internal digitization project, minimal requirements were established internally, and it was agreed to guarantee the project's most crucial goal – the accessibility of the original – at the very least through the text's readability. This is something that was considered appropriate because the originals were in textual form, and the few images present were most often monochrome and large. For the end user, fruition could be ensured.

#### 2.4.3.2 Computer and Monitor

Not a minor device, the computer, with its monitor, must be chosen according to certain criteria. The quality and effectiveness of these elements are essential, in order to maximize the functionality and productivity of the system. The computer should have computing power sufficient to execute several programs simultaneously: for example, acquisition device management program (the software of the planetary scanner), digital asset management (DAM) app, and raster graphics editors. These programs, if professional and in the event that certain operations are performed, require good performance and good RAM memory. National or international guidelines do not list specific computer requirements. In the case of the project developed by the Ateneo Veneto, the computer of the manufacturer HP Inc. had a 2.10 GHz Intel Core i7 CPU and 16.0 GB RAM. These specifications allowed the institute to perform the entire process of digitization without any kind of problem or slowing down.

As for the monitor, it is important that the visual display unit does not filter or alter colour temperature, luminance level, neutral colour balance, and linearity of red, green, and blue (RGB) colours.<sup>129</sup> The process of monitor calibration is meant exactly to visualize digital objects as they were produced and saved by the scanner. Calibration can be conducted through the use of a colour calibrator and its appropriate software. However, it is important to note that this process does not

---

<sup>128</sup> Andreoli, L., Drago, G. (2022), p. 8 and Docs Italia, *Linee guida per la digitalizzazione del patrimonio culturale. Versione 1.0* (2022), p. 62.

<sup>129</sup> Barbuti, N. (2022).

exclude inconsistencies and eventually the monitor may not be balanced properly. The Ateneo Veneto used a HP FHD Monitor. The advantages of this monitor are to be found in the high resolution of the display (a full HD LCD display), which was also anti-reflective, and in the possibility of tilting the monitor both along the longitudinal and the transverse axis. This allowed considerable flexibility and visibility during the scanning and image editing process.

### 2.4.3.3 Prototyping of the Digitization Workstation

When planning an in-house digitization project, functionalities and components have to be tested before the real process of digitization can start. This is owing to the fact that some hardware components have to be especially purchased, thus they require tests in order to see which features are appropriate for the project and which ones are not. In general, the project team must develop a method to combine software with hardware as much as feasible, determining the best optimisation, and providing the instrumentation with the best arrangement. This is particularly true for small cultural organisations, who have to buy individual hardware components spending a limited budget. They assemble them and achieve results that are more reminiscent of homemade instruments than professional ones. This process, which can be called “prototyping of the digitization workstation”, is a characteristic of small institutes that plan projects with restricted budgets and want to improve their equipment.

This prototyping phase begins with the evaluation of the functionalities of the purchased instrumentation and therefore of the exclusion of those that are considered inadequate. Reference is made to the planetary scanner. In the case of the Ateneo Veneto, the Czur scanner consists of a 40cm-high fixed head equipped with a CMOS image sensor and a double bank of LED lights. Along the vertical bar of the scanner arm, another bank of LED lights can be mounted in order to have side illumination. In the base, on the back, a series of connection ports let the user connect a foot pedal which is useful to capture images even if hands are holding the document to be digitized. During the first months of the project, tests took place to check some functionalities provided by the scanner management program, the software “CZUR scanner”. From this, the entire workstation was redesigned and implemented. The *Exercitatio anatomica* by William Harvey and *De formato foetu* by Girolamo Fabrici d’Acquapendente had completely different formats and therefore were used as test books.<sup>130</sup> Two interlinked functionalities were questioned and then discarded: the software’s double-page automatic division and the use of specialized finger cots. Among different processing methods, Czur software offered the possibility of scanning books on two pages,

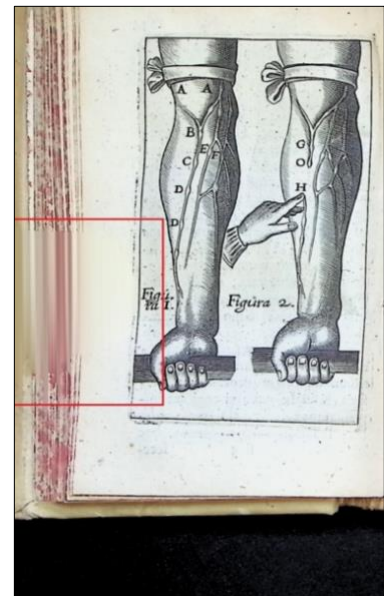
---

<sup>130</sup> For information regarding the books, *cf.* Footnotes [81](#) and [85](#).

obtaining then two distinct images through automatic cropping of the pages in correspondence with the gutter. The mechanism requires the use of two special finger cots, which allow the operator to keep the pages open during the scanning. Moreover, they allow the software to recognize the edges of the pages and to implement the cropping. This method was rejected after some tests for several reasons. First, it brought heavy image processing (such as auto-flattening of curved pages) that contravened the principles of fidelity to the original which are essential to respect in a digitization project. At the same time, it did not work properly, producing wrongly cropped images (Figure 4). The same finger removal was problematic because these cots, in the event that they were correctly identified, were heavily removed and some halos left traces of this process (Figure 5).



**Figure 4** (on the left) The specialized finger cots: no recognition and no cut. ©Ateneo Veneto

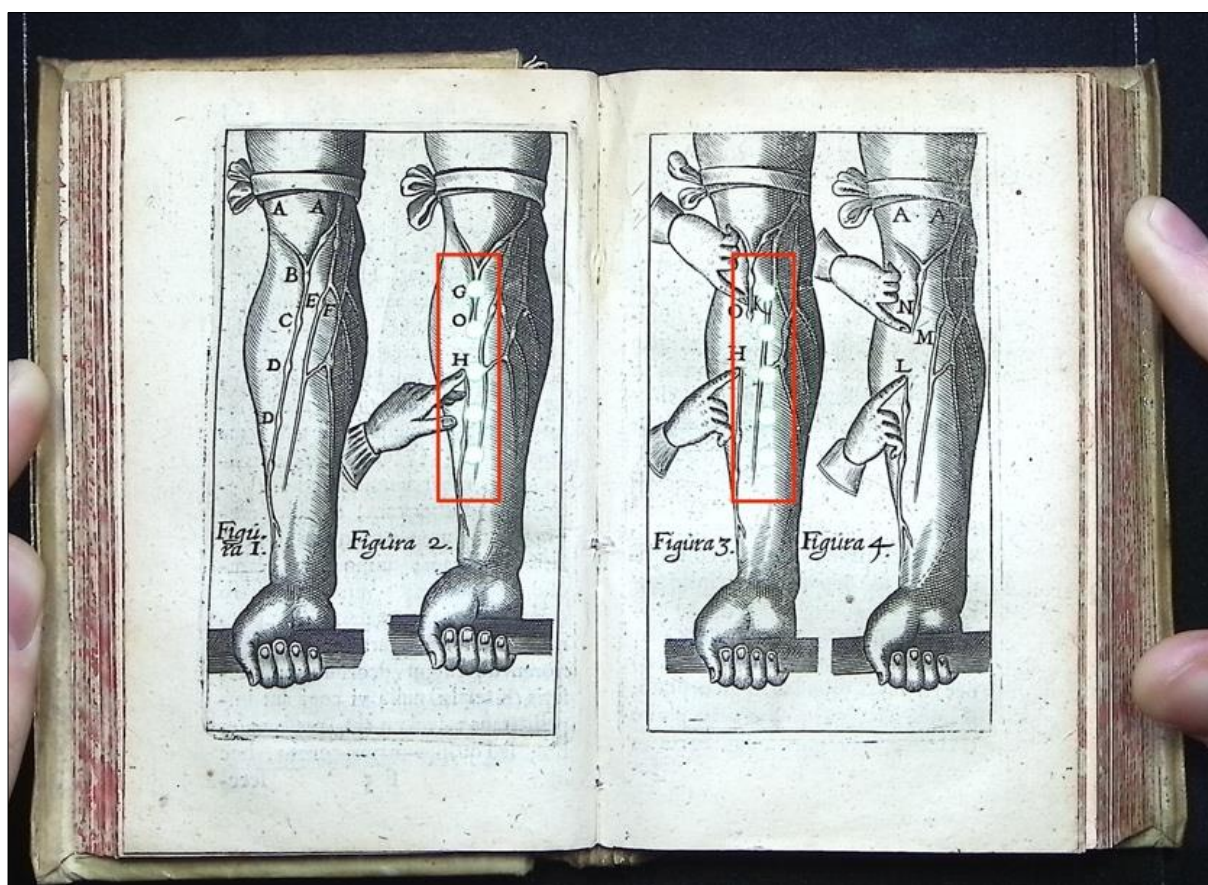


**Figure 5** (on the right) Finger removal and the appearance of a halo. ©Ateneo Veneto

This processing method was discarded, together with others, such as “flat single page”, “combine side” or particular colour modes, such as “auto enhance”, “color”, etc. which process the original image by applying filters or heavy modifications. In order to capture images as cleanly as possible, the scanner’s default shooting settings were also maintained: 25 MP of resolution and 330 DPI. The software lets the user augment the values of these settings, but the final result is an artificial interpolated image and this technique has to be avoided.

The endowment of special finger cots to perform the scans certainly posed a question: how could you keep the book open and execute the digitization without any fingers or other supports appearing in the final digital image? Czur scanner was in fact devoid of any type of book cradle where the book could be placed or any type of plastic or glass platen that could be used to guarantee the stability of items. There was only a black document pad. The project team considered possible

alternatives and solutions. Inspiration came from an examination of high-end planetary scanners<sup>131</sup>, which are equipped with motorized or manual tilting platens, made of glass or plastic, used to stabilise the original and reduce page curvature. Some experiments were conducted using a piece of antiglare glass. The choice of this material satisfied the needs of the project. In comparison to acrylic glasses such as Plexiglas, antiglare glass was entirely transparent, did not reflect or refract light, and cast no shadows. Digital images did not appear distorted as well and colours were preserved. There was just a problem. As it can be seen in *Figure 6*, the upper double bank of LED lights was unusable since their image was perpendicular to the glass and projected onto it.



**Figure 6** The use of a piece of antiglare glass test over Harvey's book and the problem of the scanner's LED lights projection. ©Ateneo Veneto

Thanks to positive test results, an antiglare glass platen was then purchased. Overall, three glass platens were bought by the Ateneo Veneto during the execution of the project. Platens, in fact, took normal wear and tear and needed to be substituted to prevent scratches and streaks from affecting the final digital image rendering. If materials of this type have to be purchased, the suggestion is to first evaluate the resistance and the possible effect they can have on the original to

<sup>131</sup> Cfr. Book2net, *Archive Pro A1* (12 June 2023).

digitize. The thickness of the glass that was evaluated as optimal for the project was 5 mm, thick enough to put pressure on the books and keep them open without compromising them, but not too thin to shatter. The price per square metre amounts to approximately 4 hundred euros, but for the purposes of the project the maximum size was A3 format, and the expense was affordable.

An arrangement of the scanner and the glass platen was necessary. The instrumentation had no precise set-up and there was an unusable lighting system. The geometry of the system was the initial set of requirements for obtaining the proper arrangement of the equipment. It is shown in *Figure 7*. As reported by Lotti, Lunghi, Trumpy, et al., the components must be placed and directed as follows:

- capture plan (object plan): horizontal.
- column of the tripod supporting the camera: vertical.
- arm of the tripod that supports the camera head: horizontal.
- optical axis of the camera: vertical; target point in the centre of the object plane.
- LED spotlights: at the same height from the object plane, in a symmetrical and coplanar position (on the vertical plane) with the centre of the object plane.
- Spotlight axes: oriented 45° from the optical axis or positioned in order to ensure the best uniformity of illumination on the plane.<sup>132</sup>

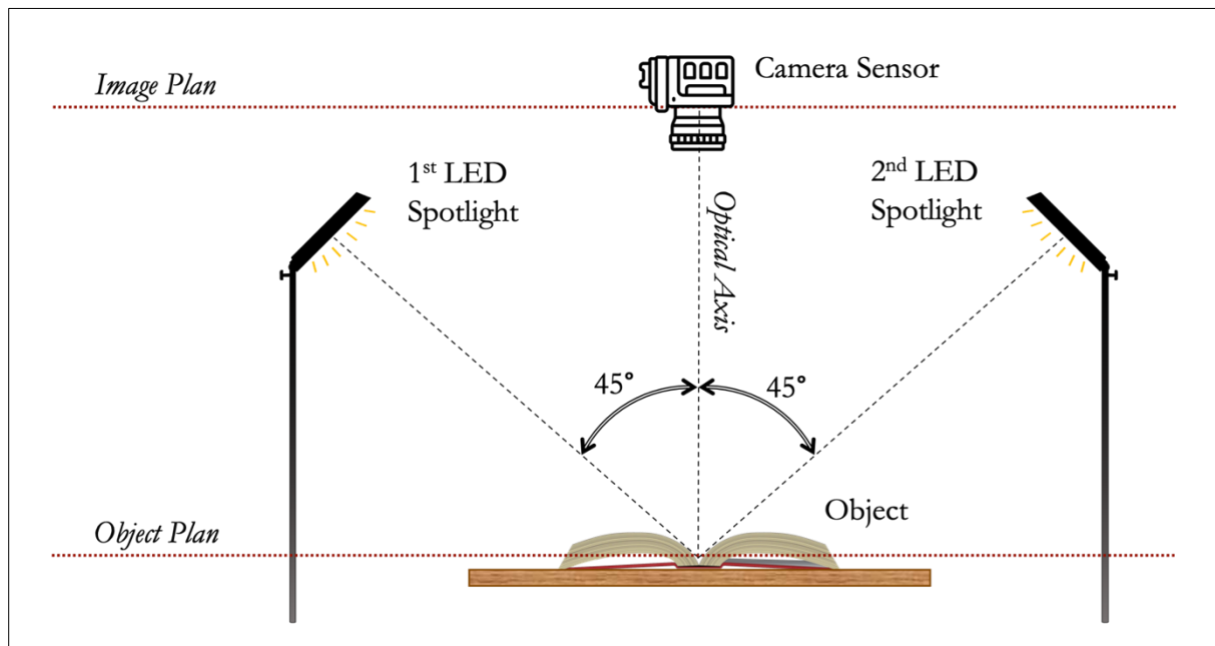
Czur scanner already met the requirements concerning the tripod because it was equipped with a vertical bar on which the top casing was horizontally mounted. On the contrary, it was necessary to purchase an appropriate lighting system, made of two spotlights, that could be placed on the sides according to the guidelines mentioned above. For this purpose, two Neewer 5600K LED spotlights were therefore purchased for a cost of a hundred euros. They consist of a tripod stand, an extension rod, and a top LED panel light.<sup>133</sup> According to guidelines, the lighting system was cold light without emission of UV and IR.<sup>134</sup> The panel could be adjusted 180 degrees, the light had ten dimmable levels of brightness, the tripod was detachable with standard 1/4" thread, and the rod was extensible by almost one metre. Each of these characteristics was advantageous for the proper regulation of the lighting system. The two illuminators were placed on the sides of the capture plane and oriented 45° from the optical axis.

---

<sup>132</sup> Lotti F., Lunghi, M., Trumpy, G. et al., *Digitalizzazione di beni artistici e documentari. Manuale di procedure per un laboratorio fotografico digitale* (2009), p. 7. Author's translation.

<sup>133</sup> Neewer, *NEEWER 2 Packs Dimmable 5600K USB LED Video Light* (n.d.).

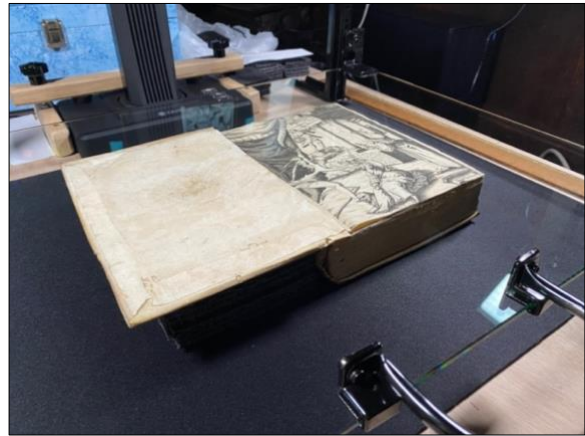
<sup>134</sup> Andreoli, L., Drago, G. (2022). According to Enciclopedia Treccani, under "LED", the possibility of modulating the individual chromatic components of the light and the total absence of ultraviolet (UV) and infrared (IR) components in the spectrum give the LED lights the primacy of less invasive light source on the materials.



**Figure 7** Layout of the shot and lighting arrangement with 45°/45°/0° geometry. ©Author's elaboration on Lotti F., Lunghi, M., Trumphy, G. et al., Digitalizzazione di beni artistici e documentari. Manuale di procedure per un laboratorio fotografico digitale (2009)

What was missing was a real digitization workstation, where the glass platen and the planetary scanner could be fixed, and the book could be placed. For this reason, a workstation was designed and built specifically for the Ateneo Veneto (Figure 8 and Figure 9). The glass panel was attached to hinges on top of a flat wooden panel, which enabled tilting. In particular, the hinges allowed for both transversal rotation and height adjustment in accordance with the book's size. For optimal noise reduction and a continuous, accurate distance between the hardware components, the scanner was fixed to the top of the wooden panel. This prevented the scanner from suddenly shifting during the digitization process, which might have led to issues with book alignment and image rendering. As for the LED spotlights, they were positioned on the sides, oriented towards the object plane, but without fixing them, in order to adjust the distance from the book. They were connected to the USB ports of the computer. Books could finally be placed above the black pad, under the glass panel, and in the centre of the digitizing workstation.

After this phase of prototyping, the Ateneo Veneto had a complete photographic laboratory ready for use. On a desk, to the right, there was the section dedicated to image acquisition, with the digitization workstation and the light sources. To the left, there was the computer workstation where all phases, from the creation of digital objects to their publication, were managed. The phase of prototyping had allowed the institution to have a working environment suitable for the requirements of digitization, practical to use and affordable for a small cultural organization.



**Figure 8** (on the left) The digitization workstation with Czur planetary scanner and LED spotlights. ©Ateneo Veneto

**Figure 9** (on the right) Scarabiccio's book over the workstation and ready for digitization. Note the glass platen and the black pad. ©Ateneo Veneto

#### 2.4.3.4 Other Equipment

In addition to the hardware components mentioned, there are some other additional instruments and accessories.<sup>135</sup> Every photographic laboratory should be provided with the stationery for the office, in order to take annotations and compile record books. A kit of accessories for digitization is also necessary and is composed of latex gloves (for book handling), soft brushes (for book cleaning), labels, screwdrivers, meterstick, spirit level, glass cleaner, cleaning cloths etc. The accessories can be varied according to the specificities of each digitization project.

Three additional items need to be supplied. The first regards the use of thick spongy pieces to be placed beneath the book so that it may be raised or lowered as necessary for scanning. In fact, the workstation and the scanner were not designed to have a V-shaped cradle on which to lay the books and the use of thick spongy pieces seemed the most practical and economical solution. It is recommended to have modular pieces of different sizes to adapt to distinct book formats. 5-mm thick spongy pieces of two different sizes were utilised for the Ateneo Veneto project. The modularity of these objects, together with continuous controls with a spirit level, made it possible to adjust from time to time the height of the book for scanning.

The second item that is required is a colour checker. It is a tool used by photographers and in digitization operations to accurately balance and recreate colours with any lighting source, using any camera or scanner, and under all lighting conditions. A colour checker consists in a 24-squares palette of different colour samples, printed using inks of the highest quality, that were specially developed to replicate a pure hue that can be consistent, intense, and non-reflective. As discussed

---

<sup>135</sup> Barbuti, N. (2022).



in Chapter 3, at the beginning of digitization a photo is taken with this colour rendition chart and it is then used to perform colour calibration in photo editing software. Utilising Calibrite's ColorChecker Passport Photo 2<sup>136</sup> and the Photoshop plug-in Camera Raw, the Ateneo Veneto calibrated the white balance of images. The final picture rendering is influenced by the technological features of the scanner's camera sensor and the light produced by the spotlights mounted on the side, which has a temperature that affects the accurate colour depiction. The preliminary photo with the colour checker is used to take a snapshot of the light conditions present in the photo studio and to make a first verification of the image quality. For a good yield, the light conditions should not change. Otherwise, anytime there are changes, a new calibration is necessary. This calibration process "must aim at the identical reproduction of the chromatic characteristics of the original, not at an arbitrary aesthetic improvement."<sup>137</sup>

With the aim of long-term storage of master and derived files, as well as all project workbooks, the institute needs to store data on different supports. The institution must make two or three copies of these files in order to guard against inadvertent data loss, which would mean losing months and months of arduous effort. It is good practice to save data both in a cloud infrastructure - looking for one specifically or taking advantage of the one already in use at the institution - and on external hard disk drives (HDD).<sup>138</sup> The choice would depend on the size of the storage required by the project. As done by the Ateneo Veneto, the HDDs must be protected in a dust-proof case. Cloud infrastructure storage, which the institution already had, was particularly useful for accessing files from any device connected to the institution's network.

#### 2.4.4 Data Management

After the institution discussed about the choice of hardware components and their optimization, it has to consider a further step of digitization planning, which is data management. This word may encompass all the descriptive passages necessary to identify, access, use, retrieve, and preserve digital resources. Data management is justified from two perspectives that the institution should adopt: an internal (institutional) point of view and an external (user) point of view or, in other words, a macroscopic (project) point of view and a microscopic (digital item) point of view. The institution needs to provide information and therefore to use metadata. From an institutional perspective, data and data management is essential to achieve good organization and transparency of the project, in terms of methodologies and practices used. At the end of digitization, the institute

---

<sup>136</sup> Calibrite, *ColorChecker Passport Photo 2* (17 July 2023).

<sup>137</sup> Andreoli, L., Drago, G. (2022), p. 10.

<sup>138</sup> Cfr. Docs Italia, *Linee guida per la digitalizzazione del patrimonio culturale. Versione 1.0* (2022), p. 49.

should ideally be able to repeat the entire process, relying on descriptive files, working materials, and well-structured directories. It is therefore suggested to adopt a long-term way of thinking that looks beyond the project itself and enables reconstruction of what has been accomplished for the benefit of future digitization initiatives, even if the personnel engaged changes. This is particularly true in small cultural institutes that can often use temporary collaborations for projects. At the Ateneo Veneto, an explanatory readme file explains the nature and type of the project files and folders contained in a single, large directory. This includes lists and mappings of selected books, shared photos of the project, spreadsheets, manuals of equipment, master and derived files of digital objects, guidelines, and external links. All these materials concern data management from the macroscopic point of view of the project.

In particular, descriptive metadata were managed within two Excel spreadsheets. The first included all relevant book information necessary to execute the digitization, whereas the second one dealt with the publication of items. A third spreadsheet was created as a logbook, where it was possible to record hours of work and daily operations. Data about shelfmark, author of the book, title, year of publication, format size, etc. was taken from the national online catalogue OPAC SBN for each book and used to create a first spreadsheet where digitization could be managed (*Figure 10*). Specific nomenclature was assigned to each digital object. This process concerned the microscopic point of view in data management which refers to single digital items. In a small cultural organization, especially when the project is not outsourced, nomenclature may not follow national or international name requirements. What is important is to guarantee the “univocal identification of each file in the context of the digitization project”.<sup>139</sup> The institute can adopt the most appropriate terms with respect to the specificity of its collection. Owing to the library collection, the Ateneo Veneto assigned to each book a folder with a specific name. Nomenclature followed the model “section\_author\_year of publication”, where “section” is one of the eight chronological subdivisions of books, as described in Chapter 2.2. For example, all the files of the aforementioned book *Politica medica per il governo conservativo del corpo humano*<sup>140</sup> written by Bartolomeo Pietrigrassa were placed in the folder “V\_Pietrigrassa\_1650”. As for the individual digital images, filenames followed the nomenclature “collocation.five-digit progressive number.extension”, which means that the same book had a first file, which corresponds to the digital image of the front cover of the book, with the name “8E50.00001.jpg”. Moreover, a progressive numeration was used to differentiate between books written by the same author and released the same year as well as titles

---

<sup>139</sup> Docs Italia, *Linee guida per la digitalizzazione del patrimonio culturale. Versione 1.0* (2022), p. 37.

<sup>140</sup> Cfr. [Footnote 94](#).

contained inside the same book. All this data was managed inside a spreadsheet because the convenience of Excel formulas and cells easily allowed the process of data management even in case of name changes. It made it possible to combine information and prevent nomenclature errors. Furthermore, spreadsheets were used to calculate the total digitization time of each book, in accordance with the number of pages. This was particularly useful to estimate the duration of the process.

Shelfmark	Folder Name	Filename	Section	Card Number	Author's Surname	Year	Image catalogue	Format	cr	OPAC
27A15	II_Averoeres_1511	27A15.	II	12	Averoeres	1511	Immagine		2	31
27A26	II_Avicenna_1508	27A26.	II		Avicenna	1508			2	30
8D44.1	II_Galenus_1625_1	8D44.1.	II	8	Galenus	1625	Immagine		2	35
8D44.2	II_Galenus_1625_2	8D44.2.	II	8	Galenus	1625	Immagine		2	35
8D44.3	II_Galenus_1625_3	8D44.3.	II	8	Galenus	1625	Immagine		2	35
8D44.4	II_Galenus_1625_4	8D44.4.	II	8	Galenus	1625	Immagine		2	35
8D44.5	II_Galenus_1625_5	8D44.5.	II	8	Galenus	1625	Immagine		2	35
27A22	III_Cardano_1554	27A22.	III	31	Cardano	1554	Immagine		2	31
14A49	III_Cardano_1578	14A49.	III	29	Cardano	1578	Immagine		2	31
8E34	IV_Acquapendente_1600	8E34.	IV	32	Acquapendente	1600	Immagine		2	43
8E50	V_Pietrigrassa_1650	8E50.	V	38	Pietrigrassa	1650	Immagine		2	30
8E93	VIII_Andrioli_1711	8E93.	VIII		Andrioli	1711			2	32
13C42	II_Avicenna_1546	13C42.	II	10	Avicenna	1546	Immagine		4	21
13C11	II_Paracelsus_1581	13C11.	II	16	Paracelsus	1581	Immagine		4	20
14B26	III_Baldini_1586	14B26.	III	22	Baldini	1586	Immagine		4	22
13C13	III_Borocchio_1593	13C13.	III	23	Borocchio	1593	Immagine		4	21
13C7	III_Berga_1568	13C7.	III	19	Berga	1568	Immagine		4	22
6D24	III_Bordini_1573	6D24.	III	25	Bordini	1573	Immagine		4	21
14B46.1	III_Bovio_1595	14B46.1.	III	27	Bovio	1595	Immagine		4	20
14B46.3	III_Bovio_1601	14B46.3.	III		Bovio	1601	Immagine		4	20
14B46.2	III_Bovio_1602	14B46.2.	III		Bovio	1602	Immagine		4	20
13C20	III_Cardano_1562	13C20.	III	30	Cardano	1562	Immagine		4	21
13B26	III_DaMonte_1584	13B26.	III		DaMonte	1584			4	24
8B30	III_Oddi_1607	8B30.	III	17	Oddi	1607	Immagine	Immagine2	4	25
13C14	III_Telezio_1590	13C14.	III	26	Telezio	1590	Immagine	Immagine2	4	24
13C17	V_Abraham ben David Arie Portaleone_1584	13C17.	V	37	Abraham ben David Arie Portaleone	1584	Immagine	Immagine2	4	20
13C3	V_Boeckel_1597	13C3.	V		Boeckel	1597			4	21
13C16.4	V_Cremonini_1626	13C16.4.	V	60	Cremonini	1626			4	21
13C16.3	V_Galvani_1620	13C16.3.	V	39	Galvani	1620	Immagine		4	21
13C16.2	V_Gregorio da Rimini_1622	13C16.2.	V		Gregorio da Rimini	1622			4	21
13B2	V_Guidi_1585	13B2.	V		Guidi	1585			4	22.5
13B31	V_Harvey_1651	13B31.	V		Harvey	1651			4	25
34B6	V_Hearne_1600	34B6.	V		Hearne	1600			4	26
13B15.2	V_Mater_1616	13B15.2.	V		Mater	1616			4	21
13B15.1	V_Minderer_1617	13B15.1.	V		Minderer	1617			4	21
13B10	V_Myllus_1618	13B10.	V		Myllus	1618			4	23
13C16.1	V_Romano_1626	13C16.1.	V		Romano	1626			4	21
13C6	V_Scarabiccio_1655	13C6.	V	34	Scarabiccio	1655	Immagine		4	21
13B15.3	V_Stefani_1642	13B15.3.	V		Stefani	1642			4	21

Figure 10 A portion of the spreadsheet with the required information for the digitization and nomenclature. ©Ateneo Veneto

Another spreadsheet with more detailed descriptive information was added, and it was used to characterise digital objects when they were published to the Phaidra repository. Chapter 4 will cover all the details about system ingestion (the procedure for transferring digital images and metadata into online management systems). What is relevant to say here is that all published metadata were organized in a spreadsheet to manage their production, to faithfully copy them during publication and to ensure that in the future the institution can monitor what was published. Data management in this instance was concentrated on the user's perspective, or on information that the user may access online, such as the title, language, description, author, etc. of digital items, which were created in accordance with the repository's specifications.

Some other metadata typologies are produced in a digitization project. Administrative metadata might include information about the project and its team, copyright information and licences, and credit lines.<sup>141</sup> Furthermore, structural metadata are used to describe the internal structure of digital items. For example, they indicate the presence of the frontispiece, chapters, index, etc. in a book.

<sup>141</sup> International Federation of Library Associations and Institutions (IFLA) (2015), p. 13.

Both administrative and structural metadata could be produced and managed on the Phaidra repository.

Finally, technical metadata is usually captured automatically by the acquisition equipment. They include the image's dimensions, compression, resolution, colour depth, colour profile, file size, and date of acquisition. The type and amount of this data depend on the model of the acquisition system. Technical metadata appears in the header of the file.

## 2.5 General Workflow and Timeline of the Project

When planning a project, one of the parameters that must be taken into account immediately is the time factor. Respect for time determines the success of the project and is the result of good planning upstream. This is extremely important in case the institution – that is even more true for small cultural ones – receives an external contribution. If this happens, the institution providing the contribution defines a maximum duration of the project (in some cases extendable), in which the recipient must perform the project. The project must be completed within this time range while adhering to all application-stated criteria and standards, both quantitative and qualitative. In the event that this does not occur, there can be a resolution of the relationship with the consequent revocation of the assigned contribution.

As a result, the planning phase must carefully calculate the correct timing for all phases in the workflow to ensure a good planning of all the activities to be carried out in the given time frame. In the case of the Ateneo Veneto project, project planning took about 20% (more than 130 hours) of the time spent on the entire project. The overall time allowed to the institution was one year from the date of allocation of the contribution by the Fondazione di Venezia and was extendable by additional six months. During this preparatory phase, all decisions regarding the choice to plan an in-house project had to be taken. The activities done in this first phase are: the examination of the technical and logistical aspects; choice of the repository; training of staff and operators involved<sup>142</sup>; selection of originals; preparation of a market study or a tender; choice of equipment; possible arrangement of the digitalisation workstation and creation of a prototype; data management; definition of the project workflow.<sup>143</sup>

---

<sup>142</sup> With reference to the Ateneo Veneto, the training dealt with ancient book cataloguing, digitization and post-capture image editing, metadata, the Phaidra environment, publication, and digital object management.

<sup>143</sup> Andreoli, L., Drago, G. (2022), p. 4.

The second phase regards digitization. This operative step is the one that takes more time – in the case of the Ateneo Veneto, almost 70% of the entire project. It includes: the preparation of originals (handling, inspecting, and cleaning); tests and definition of digitization parameters; creation of master files; quality control, correction of defects and errors; post-capture image processing and colour correction; production of derivative files and data management; relocation of documents.

The third phase in a digitization workflow consists in the system ingestion, namely all the activities that are meant for the long-term conservation of digital objects and – eventually – their publication. It includes: the creation of metadata and other files (such as XML files); publication on the (Phaidra) platform; and revision and correction. In the case of the Ateneo Veneto, this phase took 10% of the overall time.

Ultimately, it is essential to plan a fourth, last phase where, as already mentioned, the project team provides the institution with reports and all the instrumentation is safely replaced and arranged for a new project. This phase involves: data management and storage; arrangement of the photographic laboratory and instrumentation; compilation of reports; provision of documentation; closure of the project.

The digitization and publication process will be respectively presented in Chapter 3 and Chapter 4.



## CHAPTER 3: DIGITIZING RARE AND EARLY-PRINTED BOOKS

Not all digitisation processes are similar from a technical perspective. Books can be digitized by adopting destructive or non-destructive scanning, depending on whether it is important to preserve the integrity of the physical object or to maximize the digital capture process. It is evident that in the case of valuable documents, book scanning is performed according to techniques that are meant to not compromise physical conservation. Book unbinding and cutting are practices to be completely avoided in the context of the digitization of cultural heritage. It should be noted that the concept of preserving the book form during a digitization project is not limited to the adoption or not of extreme techniques such as the ones described above. More generally, scientific digitization implies respecting the features of the books and guaranteeing fidelity to the original. Digital images can be slightly crooked because the binding does not allow an optimal opening. The same usually adopted scanning process, digitization on a flat plane, creates shadows on pages, distortions, and inaccuracies. Where technology cannot be of any help, the digitizer will have to accept these imperfections. In a certain way, digitization does not aim to show a perfect, retouched or embellished digital object. On the contrary, the end user will be able and will have to use that exact digital copy of the valuable analogue one, which possesses imperfections and specificities.

The preservation of the book's original aspect in a digital form was at the heart of the Ateneo Veneto's project. The institute sought to achieve the optimal digitization of its books starting from the equipment that it could dispose of and adopting good practices, as described in Chapter 2. This process involved several operational phases: from the preparation of originals and definition of digitization parameters to the creation of master files by scanning each page of the books. After a quality control, that allowed to trace any defects and errors, the process continued through colour correction and post-capture image processing to terminate in the production of derivative files and their storage.

These steps will now be discussed in detail.

## 3.1 Preliminary Activities

Prior to beginning the digitization, some general precautions may be adopted. The first is to reduce as much as possible the impact of outdoor light within the photographic laboratory. Curtains must be drawn, and the main lights of the working environment must be turned off if they have a direct light on the digitization workstation. LED spotlights must be the only light source inside the room. Even though it can be obvious, this condition ensures uniformity of light in the room during the entire process and helps to accustom the operator's eyes to the right brightness. Moreover, it avoids problems related to the final quality of digital images.

One preliminary activity of every digitization project is the handling and preparation of originals. Depending on the type of collection chosen for the project, there might be good reasons to digitize books according to a certain order. For example, books with the same format and maximum dimensions typically possess similar digitization parameters, such as the distance and height of LED spotlights in relation to the object plane.<sup>144</sup> As a result, it will be convenient to plan the book digitization according to an increasing order of size, passing from the smallest to the largest format. In the case of the “Digitization Project of a Hundred Rare and Valuable Medicine Books”, first the books were digitized starting from sextodecimo format, the smallest one, passing through the duodecimo format, the octavo and the quarto, up to the largest, the folio format, with particular reference to Acquapendente's book.

The size is not the unique consideration to be taken into account. With the purpose of daily planning, it is essential to consider the consistency of books in terms of number of pages. It is preferable to digitize a whole book on the same day in order to ensure constant digitisation parameters and avoid changing the geometry of the workstation. Consequently, the technical operator should consult informative material about the books and properly choose books that can be digitized on the same day. Making beforehand an estimate of the speed employed for each book during the digitization processing can be a valid operational help in this. In the context of the Ateneo Veneto, the digitization speed was calculated for each book by reference to the average time per page on the basis of a number of sample books of the same format. All the estimates were organized in the aforementioned spreadsheet. In this manner the digitizer could know in advance an estimated time of digitization and could choose the appropriate books.

---

<sup>144</sup> *Cfr.* [Subchapter 2.4.3.3 Prototyping of the Digitization Workstation](#).



Furthermore, it is preferable to inquire in advance about the information of the materials to be digitized. Local or national catalogue consultation, such as OPAC SBN, can be used to preventively identify physical characteristics and state of preservation, as well as to trace specificities such as the presence of inserts or engravings.<sup>145</sup>

On the basis of this working organization, books can be handled from the library to the photography lab. It is good practice to note on the project registers the identifiers of the originals moved back and forth and leave a placeholder in the depository in correspondence with the shelfmark of the moved book. It is certainly more effective to take more books at a time for a longer period than to make a daily shift. It is essential to ensure in the laboratory the same conservation standards that are present in the depository, such as temperature and humidity. For conservation but also operational purposes, it is also recommended to keep hands clean and wear latex gloves while handling and digitizing rare books, even if they are not manuscripts.

Before the scanning process, books need to be inspected and cleaned. It is likely that books that are only for limited and internal consultation have a dust layer over the top edge or inside them. This dust and any other dirt must be removed beforehand using soft brushes and at a distance from the digitization workstation as it would dirty the equipment. According to the preservation state of the book, this process will involve only some parts of the original or every single page. Moreover, a preliminary inspection is needed to verify the presence of man-made folds and curvatures of the paper. If any, it is important to gently correct any deformation to guarantee the legibility of the content. This process should in no way compromise the original. Other types of issues that can be traced following a direct analysis are paper oxidation and the presence of holes, uncut edges or particular fragilities, with reference to both paper and binding. Some books may also lack the binding itself.

In other cases, prior inspection of originals can reserve some surprises. For example, some books can result to be totally or partially unopened. Unopened books retain the folds of the original gatherings and contain many pages which cannot be read without first opening the pages from the side or from above with a paper knife.<sup>146</sup> During the digitization project at the Ateneo Veneto, some books needed to be cut in correspondence of single or multiple sheets, as shown in *Figure 11*. In one case, a book was totally unbound and unopened. The book in question is the 15<sup>th</sup>-century octavo *Ioannis Cratonis Vratislaviensis ad artem medicam isagoge*, written by the physician Johann

---

<sup>145</sup> Barbuti, N. (2022).

<sup>146</sup> ABAA, *Glossary of Book Terms* (n.d.), under “unopened”.

Crato von Krafftheim (Breslau, 1519 - 1585).<sup>147</sup> Every signature of the book had eight leaves printed on the paper sheet and had been folded three times (*Figure 12*). In order to divide each leaf, it was necessary to make five cuts for each signature using a paper knife. Once the folds were cut, each page was ready to be digitized and the book was finally readable after more than 450 years.



**Figure 11** (on the left) An unopened book needs some cuts in correspondence of the folds of the original gatherings. ©Ateneo Veneto

**Figure 12** (on the right) Opening and cutting of Ioannis Cratonis Vratislaviensis ad artem medicam isagoge. ©Ateneo Veneto

After a preliminary treatment of the originals, some further basic operations need to be carried out. First, directories and sub-directories should be created inside the data storage device. Each book will have its own folder with the name assigned by the defined nomenclature. Internally, two subfolders will be created: the first one contains derivative files, and the second one contains master files together with another sub-directory called “cc” containing images of the initial calibration with the Colorchecker. Secondly, the book profile is created in the physical or digital logbook or project register. Every book profile is used to annotate important information for digitisation purposes. It consists of the book identifier (shelfmark), its format, the parameters and the relative measures of digitization: spotlight distance  $d$ , spotlight height  $h$ , light intensity  $l$ , book height  $h_{book}$  with respect to the object plane, and digitization time  $t$ .

---

<sup>147</sup> Crato von Crafftheim, J., Da Monte, G.B., Valgrisi, V. (1560). *Ioannis Cratonis Vratislaviensis ad artem medicam isagoge. Additae sunt in libros Galeni de elementis, De natura humana ... Et de temperamentis, & facultatibus naturalibus Periochae Ioannis Baptistae Montani: cum epistola Io. Cratonis, qua recte Galenum legendi ratio breuiter ostenditur.* Venetiis : ex officina Valgrisiana. Shelfmark: 6.C.2.

At this point the digitizer opens the acquisition device management program (CZUR software). Shooting parameters must be left unchanged and correspond to default values. This avoids any artificial interpolation during the digital acquisition. The frequency is set to 60 Hz, the resolution is 25 MP and the pixel density is 330 DPI. Master format corresponds to JPG compression and the image quality is set to the higher one. The subfolder “cc”, the folder of the calibration of the acquisition system, is then configured as the target directory. In fact, before the real digitization process can start, this step, which is discussed in the next section, is required. Files produced during this phase normally include the Colorchecker and they are saved in their folder. Filenames are assigned according to nomenclature, adding the term “cc”.

## 3.2 Calibration of the Acquisition System

Another series of technical operations needs to be conducted with reference to the digital image capture equipment. This has to be properly arranged before the phase of scanning and some shooting parameters have to be determined. As a general premise, it is important to say that final image quality will depend first of all on whether the workstation is kept clean. There may be dust, fragments of paper or dirt on the black document pad or the antiglare glass platen may present halos, marks or fingerprints. When scanning, dirt, however small, can be easily detected in digital images. In some cases, it compromises the quality of reading of the book content. For this reason, it is recommended to properly clean the workstation.

Equipment has to be arranged according to the geometry explained in Chapter 2. Here the dissertation will focus on the lighting system. As already mentioned, LED spotlights have to be oriented 45° from the optical axis. Their arrangement should be expected to be symmetrical. Though their position is variable, it has to be determined in order to obtain the maximum light uniformity possible. With particular reference to the final digital image, what is important to adjust is light exposure. In photography, this term represents the amount of light per unit area reaching the surface of the planetary scanner sensor.<sup>148</sup> Every camera sensor has a dynamic range which requires good light adjustment. In absence of a right exposure, the sensor cannot accurately record it. The final image will be defined as underexposed if it is dark and its details hard to distinguish. On the contrary, it will be regarded as overexposed if there are highlights and faded-looking images.<sup>149</sup> In both cases, significant details will be lost during capture. Therefore, the purpose of

---

<sup>148</sup> Wikipedia. The Free Encyclopedia (2023), under “Exposure (photography)”.

<sup>149</sup> Adobe, *What is an exposure in photography?* (n.d.).

lighting adjustment is to control the physical amount of light the sensor can record, so that “significant” areas of shadow and highlight detail do not exceed the sensor’s exposure range.

A necessary tool for evaluating the exposure of a digital image and deciding on the optimal calibration of the acquisition system is the histogram. The histogram of an image is a graph that reports the number of pixels present for each layer of grey. In other words, it represents the percentage distribution of the image’s brightness values.<sup>150</sup> The histogram’s horizontal axis runs from pure black on the left side to the brightest white on the right, passing through shadows, halftones, and highlights. The vertical axis represents the frequency, or intensity, of each tone. Peaks and valleys indicate, respectively, high and low tone frequency. In order to assess whether a photo is properly exposed, it is therefore essential that its histogram shows a balanced distribution towards the centre and does not touch either the left margin or the right, otherwise, there will be a loss of detail in the image.<sup>151</sup> Underexposed images typically have a left-leaning curve because many pixels are concentrated towards zero. On the contrary, overexposed images have a right-leaning curve.

Consequently, the acquisition system is calibrated in relation to the execution of some tests and the evaluation of the resulting histograms. Firstly, the book is placed at the centre of the digitization workstation. The centre of the object plane can be easily determined through the use of the acquisition software which has guidelines indicating the optical axis. The book is open in correspondence to a sample page. Among a large number of pages, the digitizer chooses the one that visually is the most representative of the dynamics of the book.<sup>152</sup> If necessary, thick spongy pieces can be used to straighten the book. Since the final image will need to go through a calibration procedure to balance the white, all the sample images captured during this arrangement of the lighting system must be made using a Colorchecker. As already mentioned in Chapter 2, the final image will be devoid of its real colours without the use of this tool because of the influence of the specific light temperature produced by the LED spotlights. Consequently, the calibration itself will be inaccurate without the use of Colorchecker. As a result, the palette – in the case of the Ateneo Veneto, Calibrite’s ColorChecker Passport Photo 2 – is placed below the book and aligned to the optical axis (*Figure 13*). Then, the tilting glass platen is lowered on the surface of the book. LED spotlights are accordingly oriented towards the centre of the object plane. It is advisable to wait at least 20 minutes before starting to capture images so that the lights heat up. In the first attempts,

---

<sup>150</sup> Lotti F., Lunghi, M., Trumpy, G. et al., *Digitalizzazione di beni artistici e documentari. Manuale di procedure per un laboratorio fotografico digitale* (2009), p. 11.

<sup>151</sup> Fantin, M., Notolini, G., ed.; Capovilla J., rev. (2018), p. 4.

<sup>152</sup> Lotti F., Lunghi, M., Trumpy, G. et al., *ibid.* The maximum dynamic will have the darkest dark colours and the lightest light colours.

their distance and height are approximately set, especially when it is the first time that the calibration of the acquisition system is executed. Subsequently, the technical operator can arrange equipment starting from measures that have already been used for the previous books. In fact, it was noted that books of similar format possess similar digitization parameters. All measures are taken using a meterstick (*Figure 14*). During these tests, the light intensity has to be determined as well.



**Figure 13** (on the left) Book and Colorchecker placement on the digitization workstation. ©Ateneo Veneto

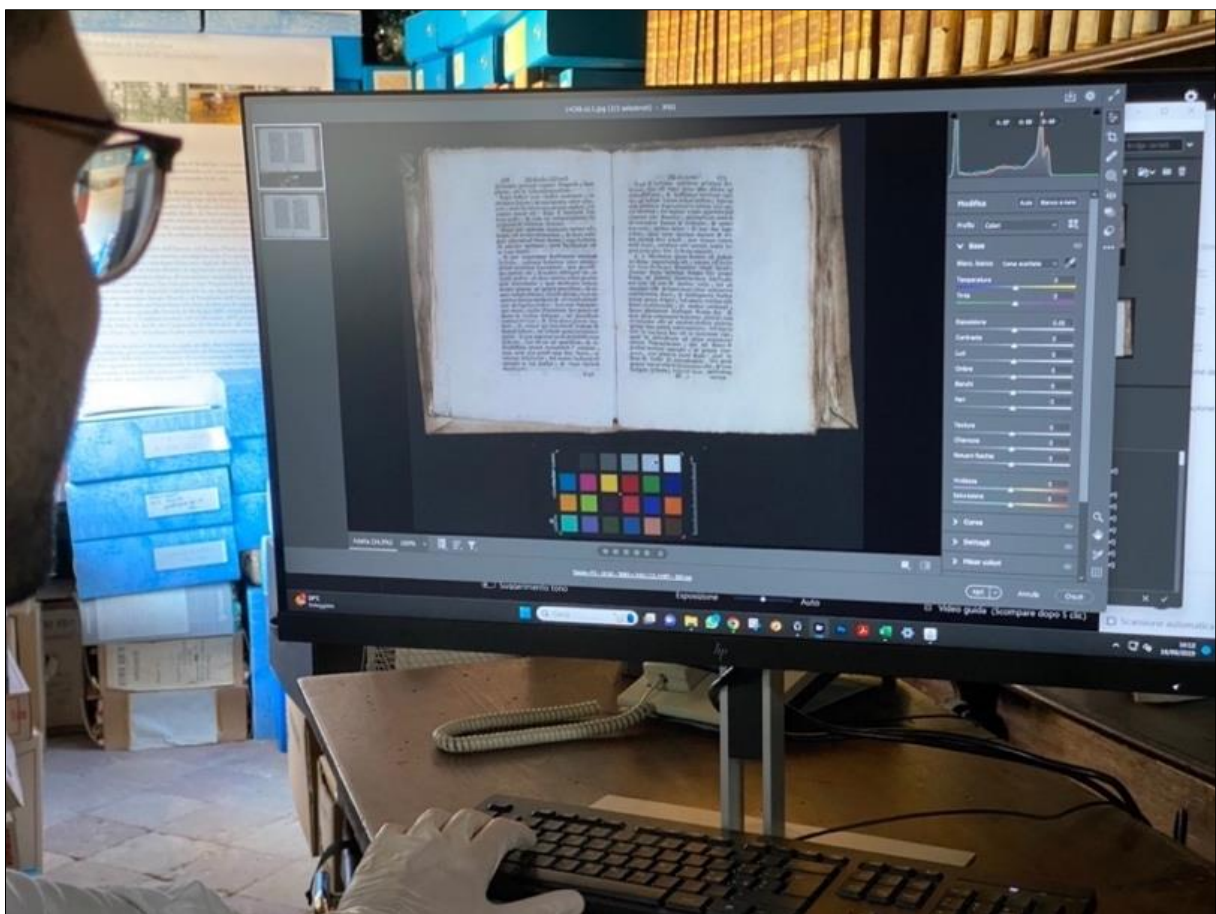
**Figure 14** (on the right) Measuring the distance between the optical axis and the spotlights in the workstation digitization. ©Ateneo Veneto

Using the acquisition software, two images need to be captured. The first one is necessary to calibrate the Colorchecker and the open book. The second one only focuses on the book and, once calibrated, represents the end result. The scanning area is set as required. Images will be saved in the subfolder “cc” of the master folder. After each image acquisition, they are opened in the DAM (digital asset management) app and calibrated. In this regard, a good set of applications and services is provided by Adobe Inc. called Adobe Creative Cloud, the software suite used during the Ateneo Veneto’s project. The workflow runs as follows: the two master images are opened in Adobe Bridge and successively are visualized and processed in Adobe Camera Raw, a Photoshop plugin that easily allows any raster image processing.<sup>153</sup> The calibration operations are shown in *Figure 15*. In Camera Raw, the operator can proceed with white balance by selecting all the images and placing the eyedropper above the second grey of the Colorchecker. In the upper right corner, the histogram shows the result obtained. If the RGB colour curves are well balanced in the centre of the graph,

---

<sup>153</sup> Photoshop® Camera Raw software is usually used to process camera raw images. It analyses raw files and creates and processes a colour image utilising metadata from the image as well as that about the camera. All the modifications in terms of colour, cropping, masks and so on do not compromise the master file. Adjustments are stored as metadata in an accompanying sidecar file or in the file itself (e.g., in a JPG or DNG file). It is required to export the image in a derivative file in order to permanently save these adjustments.

then the image has a correct exposure. Conversely, if the peaks are located near the right or left edges, the image will have a prevalence of dark and light tones with a loss of colour information and image details. In the first case, the calibration of the acquisition system was correctly defined. In the second one, it will be necessary to change the distance, height or light intensity of the LED spotlights in order to make changes in the histogram so that its curves will be placed at the centre. Even if the curves are positioned in the centre, it is still recommended to perform more tests, especially during the first attempts. Small changes may occur, and results optimized. Once the calibration is terminated and the arrangement of the equipment is confirmed, the values are annotated in the project's registers and may serve in the future to re-arrange the workstation for the same book or for those of the same format.



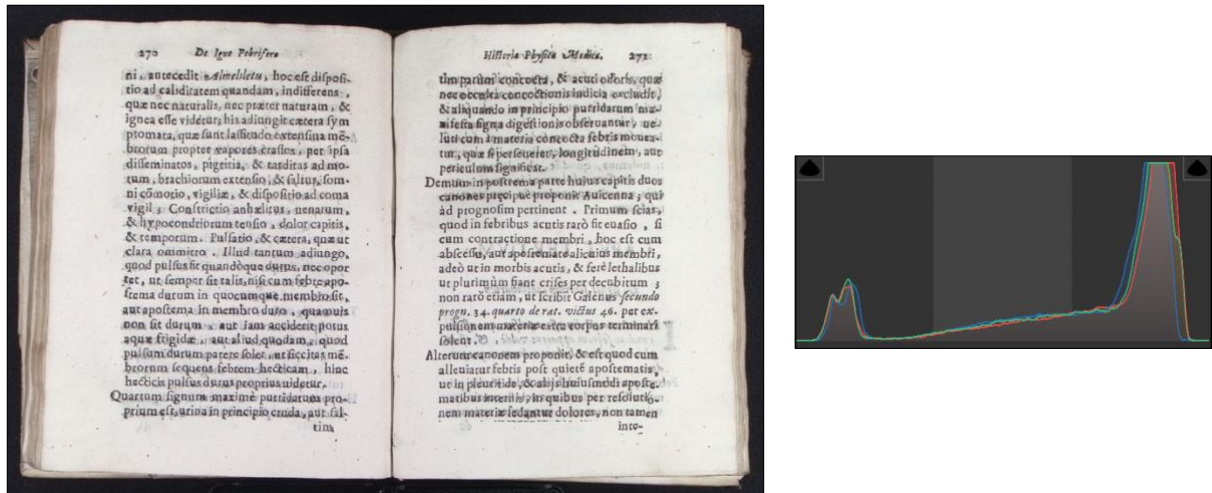
**Figure 15** Image calibration in Camera Raw. ©Ateneo Veneto

Hereinafter, an example of equipment arrangement is proposed taking into consideration two of the tests made for the book *De ortu ignis febriferi historia physica medica ad Avicennae ordinem* by Sebastiano Scarabiccio (Padua, 1609 - 1680).<sup>154</sup> At first, the lighting system was placed at a distance

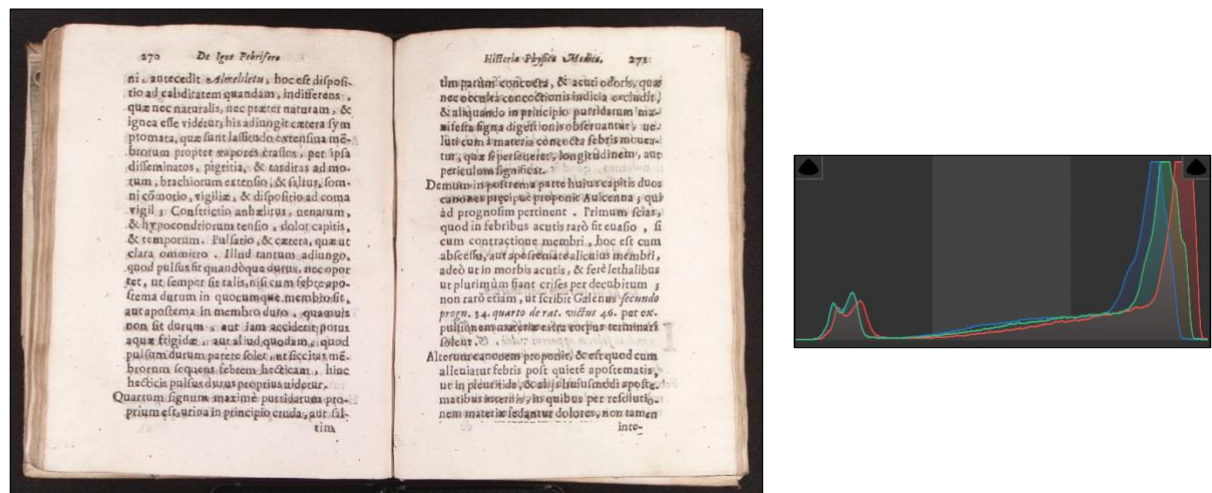
---

<sup>154</sup> Scarabiccio, S., Baruzzi, A. (1655). *De ortu ignis febriferi historia physica medica ad Avicennae ordinem. Auctore Sebastiano Scarabiccio ... Patavii : apud Andream Barutium. Shelfmark: 13.C.6.*

of 55 cm from the optical axis, the spotlight rods were extended to 145 cm from the ground and light panels were set to the fourth brightness level. Once the test images were acquired, they were processed. *Figure 16* shows the digital image before the white balance in Camera Raw, whereas in *Figure 17* the same image is calibrated. After calibration, the RGB colour charts change their curves and the image assumes warmer tones from white-greyish ones. Nevertheless, both the histograms indicated that the image was overexposed, having peaks between the highlight and white ranges. It was therefore necessary to re-arrange the lighting system.



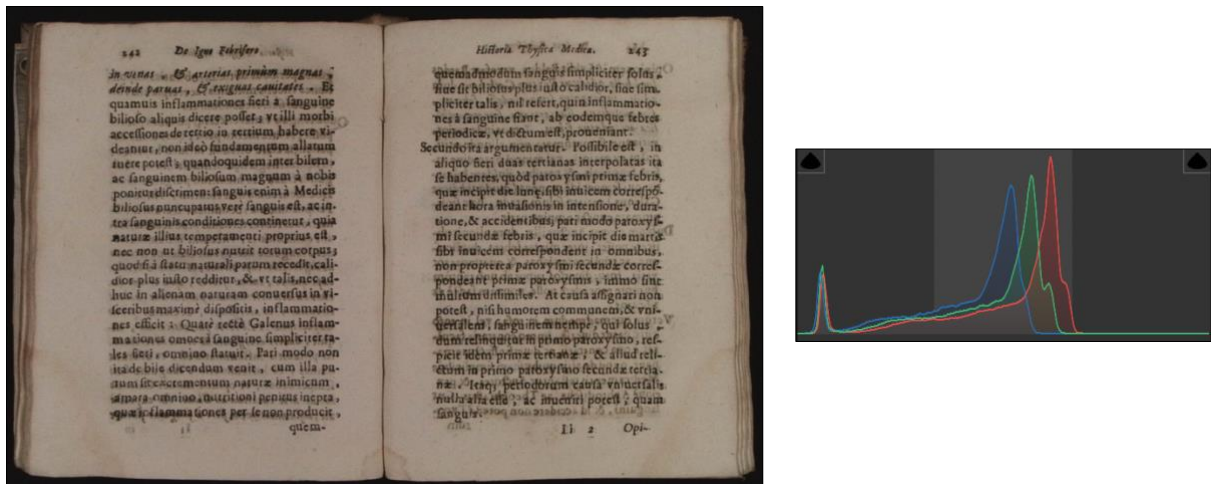
**Figure 16** A sample image of Scarabiccio's book without any calibration of white balance (on the left). LED spotlights are placed with  $d=55\text{cm}$ ,  $h=145\text{cm}$  and  $l=4$ . The corresponding histogram (on the right). ©Ateneo Veneto



**Figure 17** The same sample image of Scarabiccio's book with the calibration of white balance (on the left). The image looks overexposed. The corresponding histogram (on the right). ©Ateneo Veneto

*Figure 18* shows the optimum exposure achieved after several attempts. The distance was finally increased by almost 20 cm and the height by 25 cm. The light intensity itself was reduced to the third brightness level. In this case, the curves were situated in the exposure range and the image effectively reflected the slight yellowing of the book. The image may look darker but the histogram

displays good illumination. As for the small peak in the shadow range, it is generated by all those pixels that form the black outer margin of the image and should not be interpreted as an excess of shadows.



**Figure 18** Another sample image of Scarabiccio’s book with a spotlight arrangement of  $d=78$ ,  $b=170$  and  $l=3$  (on the left). In this case, the calibration of white balance gives the image the optimal exposure. The corresponding histogram (on the right). ©Ateneo Veneto

### 3.3 Scanning

The creation of master files requires multiple technical actions to be conducted at the same time. While scanning, the digitizer is involved in the process both from a manual and a digital point of view. After the preliminary calibration is done and two images with the Colorchecker are captured, the book is placed at the centre of the object plane using the software’s guidelines. The scanning area is set in the acquisition program in order to acquire an image of the document in its entirety. The contours of the document must be also visible, which is why it is necessary to leave a margin of a few millimetres around the document.<sup>155</sup> Actually, the technical operator sets a double scanning area. The left one captures the verso of the previous leaf, while the right one captures the recto of the consecutive leaf. It is important that these two areas are adjacent and have the same height. In this way, there will be two separate master images but, once displayed together, they will perfectly coincide and likely give the impression of reading an open book. This is the view that the user can dispose of in the repository of Phaidra.

The process of image acquisition would take only 1.5 seconds per scan considering the Czur scanner’s speed, but it is inevitable that much more time is needed for the whole process. For each page, it is indeed necessary to raise and lower the glass platen, thumb the page and realign the book

<sup>155</sup> Andreoli, L., Drago, G. (2022), p. 7.



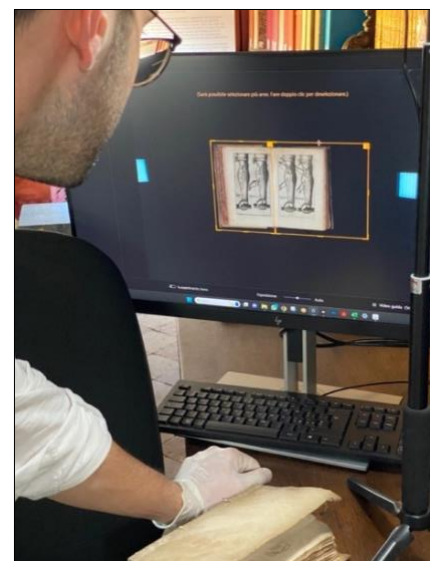
in the centre, control the cleanliness and the presence of peculiarities in the content, verify and adjust the height of the open book through the use of the thick spongy pieces, and regulate the scanning area. The scanning pedal would be pressed only after these controls and settings. On average these operations take 5 seconds per page.

The book digitization proceeds from the first page until the last one. During the acquisition, an image file is captured for each side, recto and verso, of each leaf. The digitization includes not only the textblock but also flyleaves and blank pages and all parts of the binding as well: the front and the back cover, endpapers, together with the fore-edge and the spine if the size of the book allows it. The final number of images will depend on the consistency of the book. In some cases, there can be more than eight hundred scans. All the images are saved in the destination subdirectory of master files.

On a macroscopic level, the most laborious process is the regulation of the height of the open book. The book has its maximum height in correspondence of the front and the back cover and it progressively decreases until the middle of the book. Consequently, the operator will decrease the number of thick spongy pieces put below the front cover for the first part of the book, while they will increase in number until the back cover. This operation is shown in *Figure 19*. In the meantime, the scanning areas undergo an equivalent, digital change in size. In digital images, dimensions are almost subjected to an inverse proportionality. Therefore, a larger image height corresponds to a smaller width and vice versa. For example, the two images captured in the middle of the book have the minimum height, but they laterally present the distortion of the previous and subsequent pages and hence have the maximum width.



**Figure 19** (on the left) Placement of thick spongy pieces next to the book on the digitization workstation. ©Ateneo Veneto



**Figure 20** (on the right) Setting the double scanning area for Harvey's book. ©Ateneo Veneto

This distortion is caused by the bookbinding which exposes the pages located near the covers more externally. For this reason, the scanning areas must be progressively adjusted. *Figure 20* shows an example of a digital adjustment made for the book of William Harvey.<sup>156</sup>

The scanning phase can become a mechanical process over time. The operator has to make it efficient while providing quality and precision. There is still the risk that the operator runs into distractions and errors because of the speed of the process, justified by the amount of originals that the project involves. Quality control of the captured images, considered in the next paragraph, is a further, essential phase in digitization. During the scanning, some basic action can be conducted for this purpose. It is good practice for example to find a correlation between the page and the image numbering, so that the operator can always check that all the pages were digitized. The majority of books have a numbering that begins after a series of unnumbered or numbered pages in Roman numerals. These include the initial binding and flyleaves, the frontispiece, and possibly the index and introduction. The scan numbering typically has the formula  $N_{scan} = n_{page} + k$ , where  $k$  is a constant and expresses the initial part of the book not numbered. If the book numbering is leaf-referred and not page-referred (single recto and verso), the formula would be  $N_{scan} = 2n_{page} + k$ . Once composed, the formula can be used to constantly verify the correctness of the process, avoiding repetitions or omissions of pages.

In the case of the Ateneo Veneto's project, all the master images were automatically compressed in JPG format without the possibility of having raw files. This factor compromised both the preservation of acquisition data and the final quality of images. Nonetheless, as already mentioned, this limit was justified by the fact that the institution had to reach a compromise considered its limited budget. Moreover, pixel density was set and maintained at the default value of 330 ppi in the acquisition program. This would have resulted in obtaining large images in the case of oversized physical documents and small images in the case of small-size documents.<sup>157</sup> It was not possible to regulate the distance between the scanner sensor and the object according to the size of the physical document. The distance was indeed fixed. Nevertheless, some tests were conducted considering small book formats (octavos and smaller ones) in order to avoid this problem. The test was intended to decrease this distance by raising small books of 4 or 5,5 cm. The result was satisfactory and allowed to increase the scanning area of small format books and therefore their readability, without causing blurring. In the case of folios and quartos, this technique was not adopted. In general, this was the biggest difference in the arrangement of the acquisition system. Along with

---

<sup>156</sup> Cfr. [Footnote 92](#).

<sup>157</sup> Andreoli, L., Drago, G. (2022), p. 7.

this, the distance of the LED spotlights also changed progressively. Folios were digitized with lamps at a variable distance of around 90 cm. On the contrary, when sextodecimos were involved, the distance was decreased to almost 60 cm. The height itself of the spotlights slightly decreased for small formats, but for the majority of the books it remained at 170 cm from the ground level. Light intensity was set for all the books at the third level of brightness.

To sum up, with the exception of some cases described below, the scanning did not require a particular differentiation in techniques. Books were almost monochromatic since they had been printed using black carbon-based inks, which was a common feature in early-printed books.<sup>158</sup> In a few cases, some coloured pigments had been used to print the frontispiece, but this factor did not implicate a change in digitization technique.

A further aspect concerns the presence of numerous internal chalcographic illustrations and xylographic initials. An example of these engravings can be found in the opening illustration of the book *Iatrogismorum, seu medicinalium obseruationum Pentecostae...* written by Domenico Panaroli (Rome, 1587 – Sutri, 1657) in 1652 (*Figure 21*).<sup>159</sup> It shows a physician who is educating a young apprentice according to the motto “A Teneris Imprimitur Virtus” (“Virtues are imparted since childhood”). Together with figurative illustrations, there are anatomical engravings that characterize the medicine collection of the Ateneo Veneto. The best practice to digitize engravings is to adjust the sampling density according to the hatching frequency. However, it was not possible to modify it during the project because the scanner sensor was at a fixed distance from the object plane. As a result, the engravings were digitized using the 330 ppi pixel density. In most cases, this methodology did not give particular problems, besides the limited image resolution. In some cases, when the hatching was very dense, it may result in moiré patterns appearing in the images.<sup>160</sup>

Some digitized books were characterized by another peculiar feature, i.e., the presence of handwritten notes. *Figure 22* presents the case of 16<sup>th</sup>-century underlining, corrections and numerous marginal glosses on a page taken from the work *Somniorum synesiorum, omnis generis insomnia*

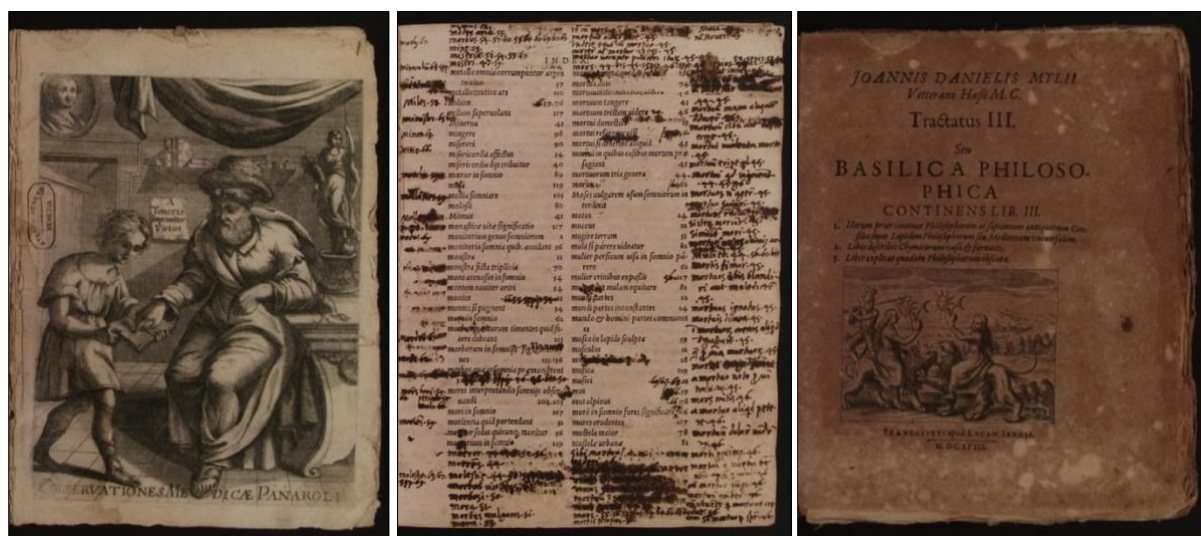
---

<sup>158</sup> Albertin, F., Balliana, E., Pizzol, et al. (2018), p. 148. The article presents a systematic non-invasive investigation of sixty early-printed books kept at the Ateneo Veneto. Data reported was extended to the books of the digitalization project, assuming an analogy in the production technique.

<sup>159</sup> Panaroli, D., Moneta, F. (1652). *Iatrogismorum, seu medicinalium obseruationum Pentecostae quinque vitibus praeceptis, ... quibus diuersa, eaque curiosa (prout aduersa pagina indicabit) in calce adduntur opuscula. Opus certe, non tantum physicis dogmatibus redundans, sed multiplici eruditione iucundum, philosophiae, ac medicinae amatissimis, apprimé vitile ac necessarium. Dominici Panaroli ... Romae : apud Franciscum Monetam. Shelfmark: 13.C.1.*

<sup>160</sup> Lotti F., Lunghi, M., Trumpy, G. et al., *Digitalizzazione di beni artistici e documentari. Manuale di procedure per un laboratorio fotografico digitale* (2009), p. 29. This beating phenomenon consists in the appearance of a kind of “waving” non-existent in the original. It is due to the phenomenon of “aliasing” (the creation of alias) originated the subsampling of the highest spatial frequencies. According to the Nyquist criteria, this sampling frequency should be greater two times than the maximum signal frequency. On the contrary, if the sampling is lower, there are distortions and the appearance of moiré patterns.

*explicantes*... of the famous polymath Girolamo Cardano.<sup>161</sup> In contrast to printed characters, whose inks are black carbon-based, annotations were usually written using Iron-Gall inks. This type of ink, depending on the concentration of metal in the composition, could have important oxidizing properties, causing a colour change towards more or less dark brown and the deterioration of the cellulose of the substrate paper.<sup>162</sup> The results of this oxidation are clearly visible on the page and provoke difficulties in readability. This factor was not a particular technical problem for the project. The resolution of digital images allows the user to zoom in and read handwritten annotations with good definition. The reading appears more complicated where oxidation is in an advanced state. However, this problem also concerns the readability of the original.



**Figure 21** (on the left) The initial chalcographic illustration in Panaroli's book. ©Ateneo Veneto

**Figure 22** (in the middle) 16<sup>th</sup>-century marginal glosses, corrections and underling in Cardano's book. ©Ateneo Veneto

**Figure 23** (on the right) The frontispiece of Mylius' Basilica Philosophica. The whole book is affected by paper oxidation. ©Ateneo Veneto

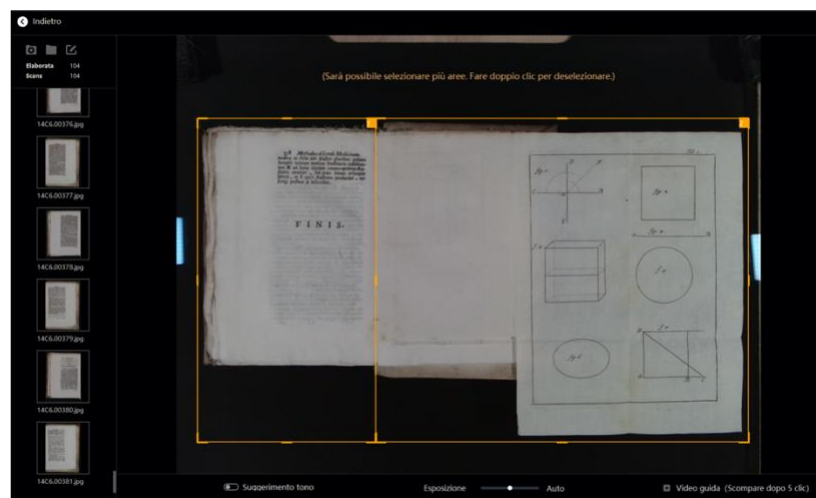
In some other books, oxidation affected the entire substrate paper. A systematic non-invasive investigation of a large corpus of early-printed books conducted in 2017 in collaboration with the Optics Laboratory of the École Polytechnique Fédérale de Lausanne, Ca' Foscari University of Venice and the National Interuniversity Consortium of Materials Science and Technology (Florence) revealed that this deterioration involves especially those volumes of the Ateneo Veneto

<sup>161</sup> Cardano, G., Petri, H. (1562). *Somniorum synesiorum, omnis generis insomnia explicantes, libri 4. Per Hieronymum Cardanum Mediolanensem medicum ac philosophum. Quibus accedunt, eiusdem hæc etiam: De libris proprijs. De curationibus & prædictionibus admirandis* ... Basileæ : per Henricum Petri. Shelfmark: 13.C.20.

<sup>162</sup> Albertin, F., Balliana, E., Pizzol, et al. (2018), p. 148. Iron-Gall inks are made from iron salts, in particular iron(II) sulphate (FeSO<sub>4</sub>) and gallotannic acids. Both the components are diluted in a water-soluble complex that is able to penetrate the paper surface, making it difficult to erase. Over time, Iron-Gall inks are affected by an oxidation phenomenon that gradually changes and darkens the colour to an intense purplish or brownish-black. Oxidation can corrode the paper surface and create holes.

that were printed in Germany and Switzerland.<sup>163</sup> The same tendency can be found in transalpine books digitized during the project. One of the most emblematic books affected by paper oxidation is *Basilica philosophica* written by the German alchemist Johann Daniel Mylius (Wetter, 1583/1585 – 1628 or 1642) and published in Frankfurt in 1618.<sup>164</sup> The frontispiece of this third treaty of the author’s work *Opus medico-chymicum* is shown in *Figure 23*. The sheet is dark brown as a result of a strong oxidation process that leaves here and there only some lighter spots. Digitization, which always has as its primary objective fidelity to the original, could only photograph these conservation conditions without any purpose of page whitening. Pages were fortunately not so brown as to be unreadable, so even in digital images, readability is equally guaranteed. Instead, it had been previously essential to accurately calibrate the lighting system in order to achieve the best light conditions and, therefore, the best image exposure.

A final, technical consideration should be made in reference to the page size, which may not be the same for all the leaves. Firstly, the operator could encounter pages folded on themselves that, once opened, show large inserts. These inserts had usually an illustrative purpose. The digitization technique is usually planned to capture a scan of both the recto and the verso of the open sheet.



**Figure 24** (on the left) Placement of thick spongy pieces under the protruding insert of a book. ©Ateneo Veneto

**Figure 25** (on the right) Setting the double scanning area in the presence of a book insert. ©Ateneo Veneto

<sup>163</sup> *Ibid.* The study is part of *Argeia*, a project “aiming to study early printing technologies, their evolution and, potentially, the identification of physical/chemical fingerprints of different manufactures and/or printing dates”. Sixty Ateneo Veneto’s books were investigated through imaging analysis and X-ray fluorescence (XRF). One of the most important results of the study revealed a potential chemical fingerprint of the differentiation between Venetian and German/Swiss books, especially for papers. The exceptional conservation state of Venetian papers could be hypothetically justified by the formulation of a paper recipe with more gelatine size (animal glue) – and therefore more Ca content – and less Alum ( $KAl(SO_4)_2$ ) – therefore low K amount. The opposite trend is present in transalpine books, causing faster ageing of paper.

<sup>164</sup> Mylius, J.D. (1618). *J.D. Mylii ... Tractatus 3. seu Basilica philosophica continens lib. 3: 1. horum prior continet philosophorum sapientum antiquorum consilia super lapidem philosophorum seu medicinam vniuersalem; 2. liber describit chymicorum vasa & fornaces; 3. liber explicat quaedam philosophorum obscura*. Francofurti : apud Lucam Iennis. Shelfmark: 13.B.10.

For these inserts this process required adjusting the height of the protruding sheet with the rest of the book using the thick spongy pieces (Figure 24). Then, using the acquisition software, the left and right scans were vertically aligned, even if the left one resulted to have a thicker black bottom edge (Figure 25). This solution allows the two pages to perfectly coincide in Phaidra's BookViewer. Some books actually had very large inserts that did not allow for simultaneous digitization of left and right pages. The sheets were not scanned at the same time and a precise work of realignment was necessary, as will be explained later in the section dedicated to post-capture image processing. Rather than increasing, the size of the leaves may decrease. The paper support can indeed be affected by the presence of scratches, wormholes, and moisture damage. These factors irreparably damage the paper support making it fragile. On the one hand, bookworms may bore or chew through volumes seeking food. On the other hand, damage caused to a waterlogged moisture-damaged book can reduce pages to shreds. Two examples of this problem are shown in Figure 26 and Figure 27. In the first case, the frontispiece of Marsilio Ficino's book clearly presents a lacuna in correspondence with the title.<sup>165</sup> This lacuna was caused by humidity stagnation.



Figure 26 (on the left) Ficino's book frontispiece affected by a moisture lacuna. ©Ateneo Veneto

Figure 27 (on the right) Water damage on the fore-edge of the post-incunabulum of Avicenna. ©Ateneo Veneto

<sup>165</sup> Cfr. Footnote 87.

In the second case, moisture damage was far greater and created internal holes in nine leaves, larger holes extending to the fore-edge in two leaves and a whole missing strip of sheet from the top-edge to the bottom edge in five leaves. The lack of text is evident. The book in question is the post-incunabile *Auicenne perhyapatetici philosophi: ac medicorum facile primi Opera in lucem redacta...* published in Venice in 1508.<sup>166</sup> For both books, whenever there were holes in the paper sheet, the digitisation process used a white sheet to mask the underlying content and avoid its capture.<sup>167</sup>

### 3.4 Quality Control

When conducting large-scale digitization, it is important to plan a phase dedicated to quality control. Actually, this is not an independent step. Quality control is performed both during the scanning and the post-captured image processing. In fact, there are some controls that can be made and problems that can be adequately solved in one phase or the other. This is the reason why quality control is discussed at this point in the chapter.

This process requires a preliminary definition of what can be considered acceptable and what is not. Moreover, it needs to establish a verification mode, which can be carried out according to a full or sample inspection.<sup>168</sup> It finally requires the determination of possible corrections and solutions.

Quality control involves four macro-categories. The first one regards the acquisition system and its calibration. The operator has to check the proper arrangement of the digitization workstation. With particular attention to the lighting system, parameters such as distance, angulation, height, and light intensity of LED spotlights must remain unchanged, together with the internal light conditions of the photo laboratory. As already highlighted, cleanliness is an essential factor to respect. The antiglare glass platen has to be repeatedly cleaned in order to prevent halos, marks, fingerprints, and any other typology of dirt. The black document pad itself needs to be cleaned from dust and paper scraps. These impurities can be visible in digital images. In the event that they are detected, it could be a good solution to rescan or digitally remove them during post-captured image processing.

---

<sup>166</sup> Avicenna, Caecilius F., al-Farābī, A.M.Ibn-M., et al. (1508). *Auicenne perhyapatetici philosophi: ac medicorum facile primi Opera in lucem redacta: ac nuper quantum ars niti potuit per canonicos emendata. Logica. Sufficientia. De celo & mundo. De anima. De animalibus. De intelligentijs. Alpharabius De intelligentijs. Philosophia prima.* (Venetijs : mandato ac sumptibus heredum ... Octauiani Scoti cuius ac patritij Modoetiensis ... Per Bonetum Locatellum Bergomensem presbyterum, 1508. Sextodecimo kalendas Maias [16 IV]). Shelfmark: 27.A.26.

<sup>167</sup> Andreoli, L., Drago, G. (2022), p. 7.

<sup>168</sup> Andreoli, L., Drago, G. (2022), p. 16.

A good calibration of the physical environment is also provided by the adjustment in the height of the originals. To check that the surface of the book - and therefore of the glass platen - is straight and does not produce distortions, the technical operator can use a spirit level. 5-mm thick spongy pieces have to be correctly collocated and evenly distributed underneath that textblock that must be raised in height. By browsing the book, these pieces can move to the point where they protrude and can be visible in digital images. Even in this case, it will be advisable to consider whether to repeat the scan or remove the impurities at a later time.

The second macro-category of quality control regards completeness. Master files can be completed both in terms of size and quantity. The scanning area needs to be properly set in size, leaving a 3-4 mm margin around the book. Images of adjacent left and right pages must have the same height and outer edge. The similarity in size will make the double-page view more pleasant once the book is published on the platform. Moreover, digitization must respect quantitative parameters and verify that the book has been scanned in its entirety, without omissions or repetitions. A useful formula (*cfr.* subchapter 3.3) can be used as a constant reference point for a sample checking of the page numbering while digitizing. Nevertheless, the operator has to be attentive to possible changes in page numbering, due to the presence of unnumbered inserts or incorrectly numbered pages. Not infrequently there may be numbering typographical errors that can deceive the operator. There are two types of numbering errors: those related to incorrect writing of page numbers (typos) or those related to an incorrect order of the leaves inside the signature. The first numbering mistake is for example present in the book of Sebastiano Scarabiccio.<sup>169</sup> After page 136, page 173 erroneously follows. The printmaker must have gotten confused and reversed the digits in the number. He dragged this error until page 195 – the number is repeated twice – when he must have realized the fact and returned to the correct page number 161. Even though the numbering is incorrect, the text flows without errors. The same does not happen in the case of errors affecting the order of leaves inside the signature. The text in fact stops on a page and does not continue in the consecutive one. This particular situation is for example shown in the book *Ragionamenti intorno alla nuova medicina dell'acqua...* written by Niccolò Crescenzo.<sup>170</sup> The bookbinder must have incorrectly folded the sheet, and the signature was bound in the following order of pages: 321, 322, 325, 326, 323, 324, 329, 330, 327, 328, 331, 332. Numbers are paired according to the leaves but their order is clearly incorrect.

---

<sup>169</sup> *Cfr.* [Footnote 154](#).

<sup>170</sup> Crescenzo, N., Ramanzini, D. (1741). *Ragionamenti intorno alla nuova medicina dell'acqua, e come la prima volta introdotta ella fosse, difesa, e sostenuta in Napoli ... Opera di Niccolò Crescenzo ...* In Verona : presso Dionisio Ramanzini. Shelfmark: 14.B.59.



During the scanning, the operator needs to pay attention to these peculiarities of books. Most of the time, however, if there is no match between the number of pages and the number of scans according to the formula, it means that a repetition or omission occurred. The correct sequence of master files has to be adjusted. In the case of repetitions, it could be convenient to re-do all the scans until the duplicated one if the error was noticed after a small number of captured images. If not, the operator can use the “Batch Rename” tool in Adobe Bridge.<sup>171</sup> If any omission occurs, the acquisition program usually allows to insert new master files inside those already captured in the chosen position.

The third macro-category of quality control regards readability. To the extent that the conditions of the book permit, the text must always be readable. The operator has to check that there are no folds on the surface of the book and that all the words are distinguishable and are not covered by dust or paper scraps. Textblock might be sometimes inclined because of the binding or the print is not precisely horizontal. It will be necessary to align it as much as possible. In case of holes, the text of the page might appear mixed with the underlying content. As already explained, a blank sheet of paper is a good solution. It is finally important to avoid and reduce shadows present on the surface of the open book, in particular in correspondence of the gutter of the book where there are curvatures.

Quality control consists of a last macro-category, which regards data management. All the master files and directories must respect the defined nomenclature. They must be complete and properly organized according to the principles of retrievability, completeness, and continuity.

### 3.5 Post-Capture Image Processing and Colour Correction

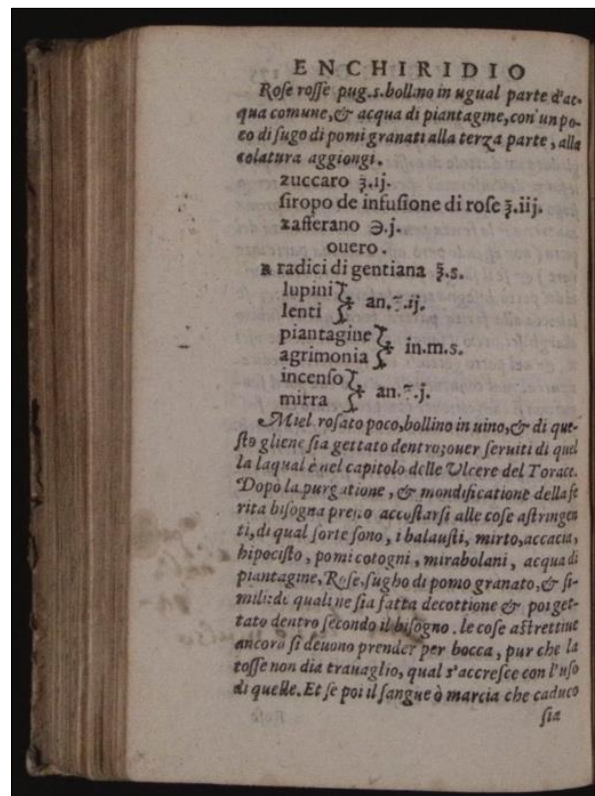
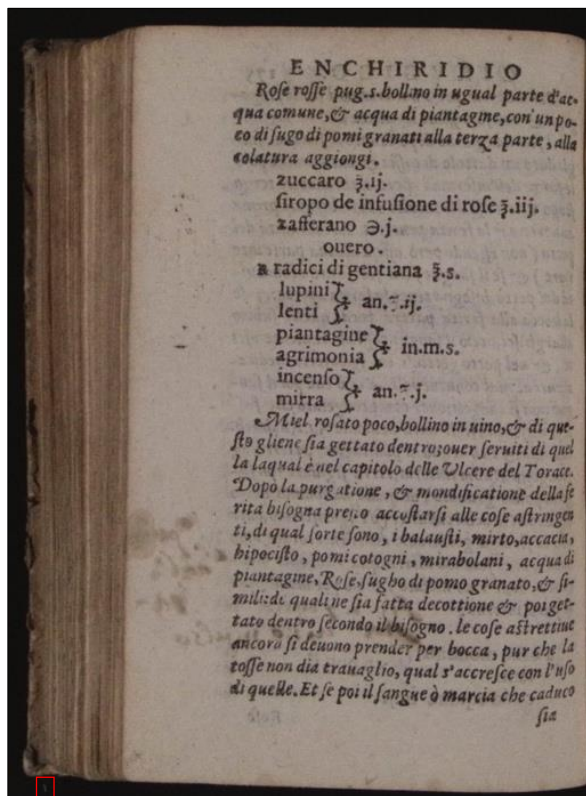
This phase was entirely conducted through the use of Camera Raw, which is the raw image editor provided by Adobe Inc. and part of Adobe Creative Cloud. In some specific cases, this tool was accompanied by the use of Adobe Photoshop. The most important operation that was performed in this sense was colour correction, with special reference to white balance. It is the same process that was explained in Chapter 3.2 but in this case, the calibration is applied to all images. Inside the

---

<sup>171</sup> This tool is strongly recommended in case the latest images are numerous and the book in question has a numbering based on the leaf, which implies a double number of pages. This is the sequence of actions to be carried out. In Adobe Bridge, select the files to rename; open them in Batch Rename; choose the option “Move to other folder” and select the destination folder – to avoid problems, it is better to select a folder different from the one in use. Then, compose the new filename according to the nomenclature: text (the book’s shelfmark with a dot, e.g., “7A259”), sequence number with five digits (the number of the repeated file, e.g., “398”), other possible text (“\_r” for recto and “\_v” for verso), new extension (.jpg). Rename the files, go to the destination folder, select them and move them to the master folder. Import the folder with the new files in the acquisition program and continue scanning.

master file folder, all images, including the one showing the Colorchecker, are opened in Camera Raw. Here the operator proceeds with white balance by selecting all the images and placing the eyedropper above the second grey of the Colorchecker. All the images will change their colours from a shade tending to grey to a warmer yellowish. In the upper right corner, the histogram shows the result obtained for each image. Curves will not be fixed but will change according to the specific hue of the page. In any case, however, they should be placed as far as possible in the centre of the histogram to ensure the optimal exposure of the image.

Inside Camera Raw, quality control and the correction of imperfections were conducted as well. The use of this software prevented the irreversible modification of the image. All the operations in Camera Raw, in fact, are saved in a sidecar file inside the jpg file of the image but become effective only in case of the exportation of the derivative file.<sup>172</sup> The originality of the master file is thus preserved and any editing action is always reversible.



**Figure 28** (on the left) Detection of a trace of dirt on the black pad while digitizing Chaumette's Enchiridio cioè Manuale cirurgico. ©Ateneo Veneto

**Figure 29** (on the right) Removal of the imperfection on the black pad using Camera Raw. ©Ateneo Veneto

<sup>172</sup> Docs Italia, *Linee guida per la digitalizzazione del patrimonio culturale. Versione 1.0* (2022), p. 32. Sidecar files are created by post-captured image processors and are used to store data or metadata that are not supported by the format of a source file. An example of a frequent sidecar file is the format XMP (eXtensible Metadata Platform). In the case of raw or lossless master files, this data is saved as a separate, external file, whereas it is incorporated in the master file if the original master has a lossy compression (e.g., JPG).

Some of the editing operations included the lateral margin cropping and the remotion of traces of dirt on the black document pad. No corrections were ever made to the book. An example of this is shown in *Figure 28* and *Figure 29*, which present the detection of an imperfection (highlighted in red) and its subsequent elimination on the mat. These digital images are taken from the verso of page 175 of the French physician Antoine Chaumette's *Enchiridio cioe Manuale cirugico...* published in Venice in 1605.<sup>173</sup> This and other visible traces were removed using the healing tools provided in Camera Raw: content-aware remove, heal, and clone.

One of the most laborious operations performed during the post-capture image processing is the realignment of adjacent images that could not be scanned using the double scan mode. The edges of the left and the right page did not perfectly coincide and caused a visualization problem once displayed in Phaidra using the *BookViewer*. It was therefore important to digitally realign those pages. This process was performed with particular reference to those books that had large dimensions or had inserts that did not allow for simultaneous digitization of left and right pages. Folio-format books such as Acquapendente's *De formato foetu* or the first five volumes of *Galenii Opera ex nona Iuntarum editione* required single scanning and needed then this realignment process.<sup>174</sup> Moreover, it was not rare to find large, impressive inserts folded up on themselves inside quartos, octavos or smaller format books. In the case of Oddo degli Oddi's *Expositio, in librum artis medicinalis Galeni*, there is a folded 35 cm-wide insert inside the book that had dimensions 17 cm in width and 25 cm in height.<sup>175</sup> It is collocated between pages 24 and 25 and presents the ramifications of the medicine discipline according to Galen's theory. The same insert is interesting for the presence of a central watermark that can be classified into the category of "monts (mountains)".<sup>176</sup> During the post-capture image processing it was necessary to realign page 24 with the recto of the insert and page 25 with the verso. The pages were opened in pairs in Photoshop, put side by side and visualized in real dimensions. As shown in *Figure 30*, two guidelines were then placed for each image at the point where the right and left pages meet above and below. The pages are aligned in height by regulating the dimensions of the upper and lower external margins. The advantage of

---

<sup>173</sup> Chaumette, A., Poli, H. (1605). *Enchiridio cioe Manuale cirugico, di Antonio Calmeteo ... Nel quale si insegna la ragione, e la pratica di curare tutti i mali esterni, co' loro rimedij si uniuersali, come particolari. Nuouamente tradotto di latino in volgare per opera del sig. Horatio Poli, detto Montagnana. Aggiuntoui in fine La ragione di guarir il mal francese, sicura & approuatissima*. In Venetia : appresso Giacomo Vincenti. Shelfmark: 7.A.259.

<sup>174</sup> Cfr. Footnotes [81](#) and [85](#).

<sup>175</sup> Cfr. Footnote [89](#).

<sup>176</sup> The first and still the most extensive study on watermarks is the mammoth four-volume work *Les Filigranes* (Genève, 1907) written by the Swiss filigranologist Charles-Moise Briquet. The work analyses more than 16 thousand types of watermarks, each with its reproduction and description. He was the first scholar to suggest the use of watermarks for dating paper.

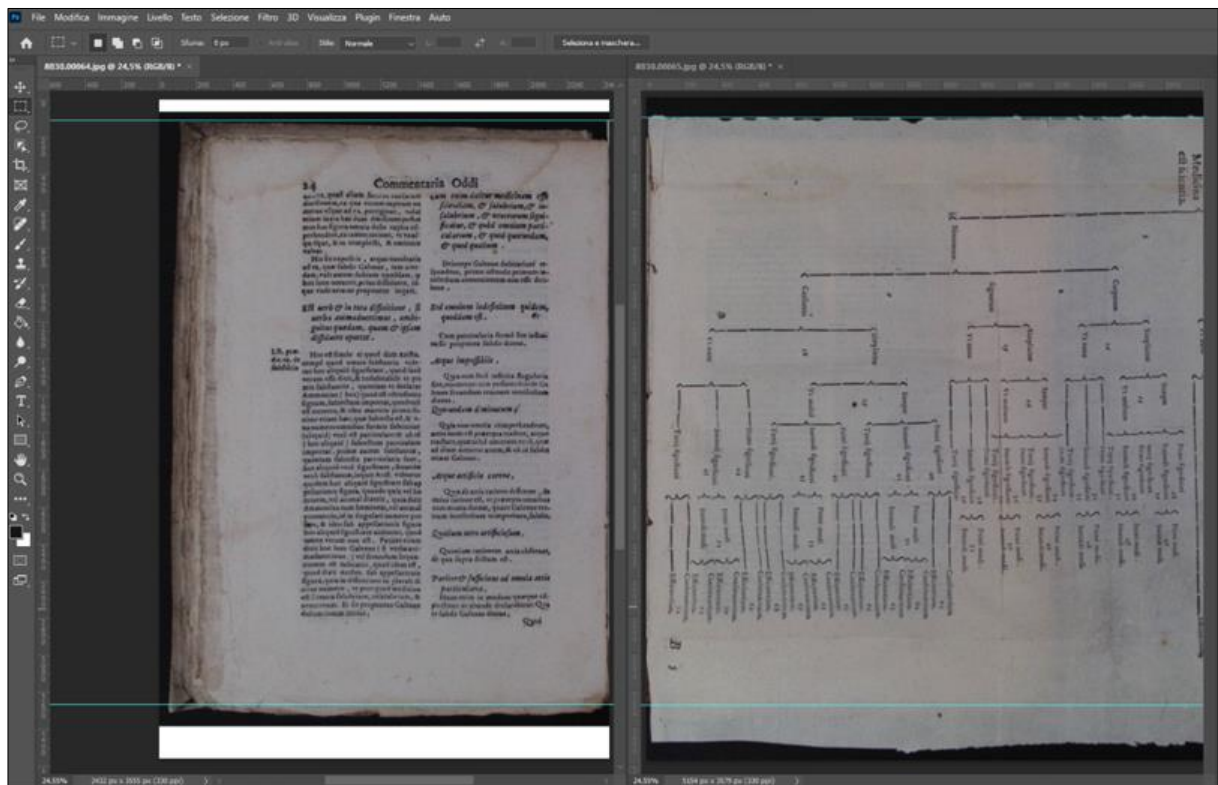


Figure 30 Realignment of left and right master page in the presence of the insert inside Oddo degli Oddi's book. ©Ateneo Veneto

this technique is to preserve the original shape of the page and to intervene only on the external margin whose integrity does not matter for the purposes of digitisation.

It is likely that a portion of blank edge has to be added above or below one of the pages. These portions are then filled with a pattern that is created from the document pad captured in that specific image so that there is uniformity in the background. Once this process is completed, the two images can be displayed using an image visualization software in order to check if the final

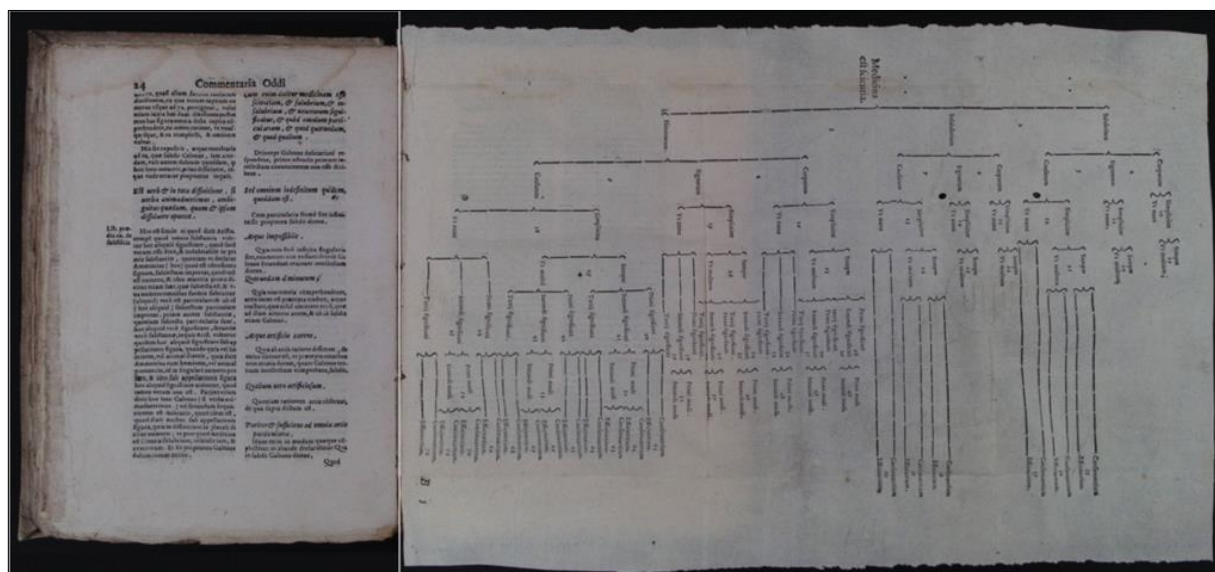


Figure 31 The final result after the process of realignment in Oddo degli Oddi's book. ©Ateneo Veneto

result is satisfying. If not, the process must be re-established. *Figure 31* shows the final result of the alignment done for the insert inside Oddo degli Oddi's book. The top and bottom edges of the two pages perfectly coincide. Changes made to the master files are saved and only at a later stage are made permanent through the creation of the derivative files.

### 3.6 Production of Derivative Files

Files meant to be published are produced from post-processed images. In Camera Raw, all modified masters are selected and converted into new JPG files. Image quality is usually reduced when creating derived files because lossless formats are converted into lossy formats. In the case of the Ateneo Veneto's project, however, the master and the derived files have the same compression, the JPG format. The quality of the images published in Phaidra, therefore, is not much different from that of the original files. Derivative files were saved by setting the colour depth to 24-bit (8 bits per channel), using the same sRGB colour profile of acquisition and the same resolution of 330 dpi. No reduction of quality or resolution was used since the masters' dimensions were not so large. Derivative files were then saved in a separate subdirectory, called "jpg".

Subsequently, it is important to create backup copies in external hard disk drives or on Cloud infrastructure storage. This has to be done for each book folder. Finally, registers are compiled, calculating digitization times, annotating the number of files created and other interesting information and practices implemented for each book.

### 3.7 A Particular Digitalization Case: Acquapendente's *De formato foetu*

Since the beginning of the Ateneo Veneto's project, the book that represented a significant challenge to digitalization was *De formato foetu* written by Girolamo Fabrici d'Acquapendente.<sup>177</sup> The book kept by the institution – the edition published by Francesco Bolzetta in Venice in 1600 – has two aspects worthy of attention: the dimensions and the content. With regard to dimensions, the book was printed in a large, folio format and the pages have the following measures: 28 cm in width and 43 cm in height. As a result, the scanner sensor could capture only one page at a time

---

<sup>177</sup> Cfr. [Footnote 81](#).

and the scanning had to be done without the use of the double scan function. It was then necessary to re-align every page with the adjacent one in the post-capture image processing. Secondly, the book contains 33 large engraved anatomical tables, some of which cover both the left and the right pages. During the image processing, it was, therefore, important to align not only the edges of the pages but also the lines of the engravings.

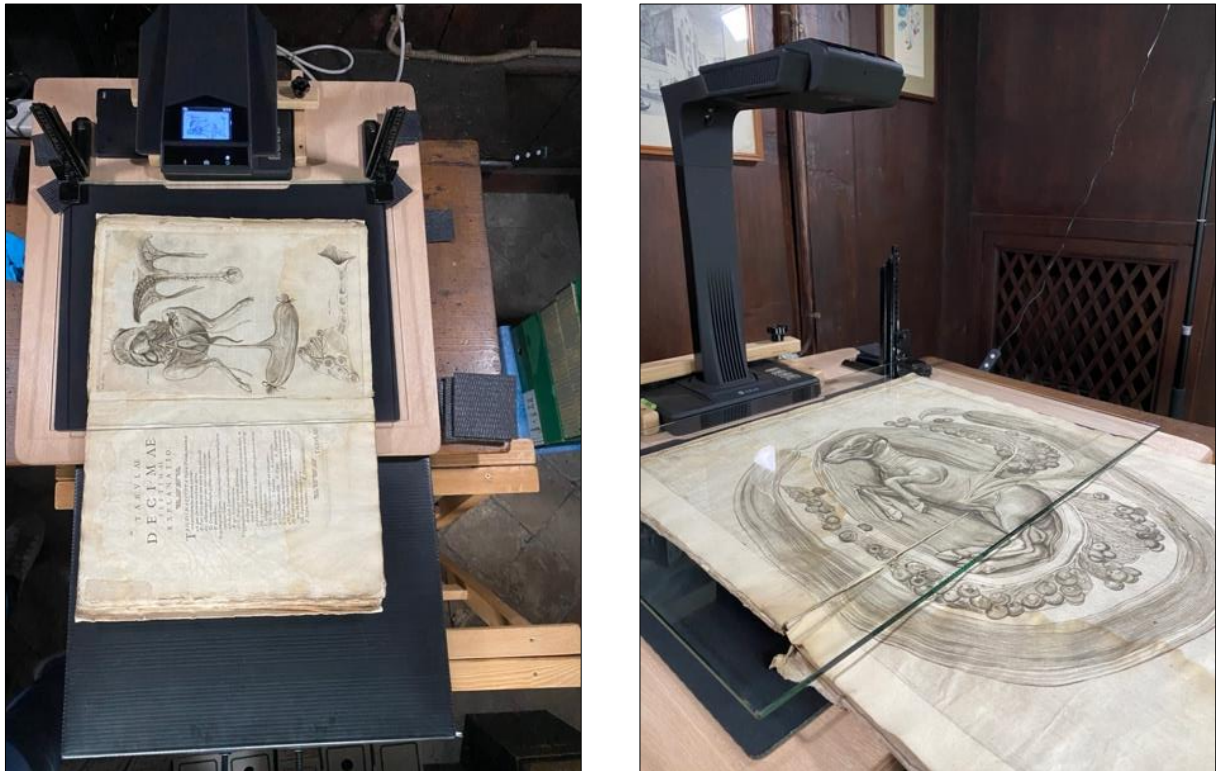
First of all, the book was preliminarily inspected. Each page was dusted using a fine bristle brush (*Figure 32*). Curvatures and little folds of the paper were gently stretched.



**Figure 32** *Dusting the pages of Acquapendente's De formato foetu.* ©Ateneo Veneto

The workstation was calibrated following the parameters used for large-format books. LED spotlights were placed at a distance of 90 cm from the optical axis and at a height of 170 cm from the ground. The book was then scanned page by page, from the front to the back cover (*Figure 33*). It was convenient to first scan all the rectos and then all the versos. In this way, the height of the book increased progressively, avoiding the continuous change of the thick spongy pieces underneath or the rotation of the book. This allowed for better scanning accuracy. However, once the rectos had been digitized, it was important to note and realise the scanning of the adjacent versos using the same height of the book. In fact, the subsequent digital re-alignment could be

effective only with scans that had the same proportions and, therefore, of the same distance from the scanner sensor. The capturing phase produced a total number of 174 scans. Rectos were saved in one folder and versos in another one. It was then necessary to inspect the completeness of the scans, by cross-checking rectos and versos and comparing digital images with the physical pages.



*Figure 33 (left and right) Arranging Acquapendente's book on the digitization workstation. ©Ateneo Veneto*

Post-capture image processing was conducted in two phases. For the first time, images were calibrated and rotated using Camera Raw and then saved together in a unique subdirectory. Rectos were renamed using the suffix “\_r”, whereas “\_v” was used for versos. In a second time, these first processed images were opened in pairs in Photoshop, cropped, re-aligned, purified of imperfections and saved as final derivative files. Although this method may seem laborious, it does not have a permanent impact on master files and was therefore preferred. On the contrary, performing all post-editing passages (rotation, cropping, alignment, elimination of impurities) at the same time would have compromised the master files.

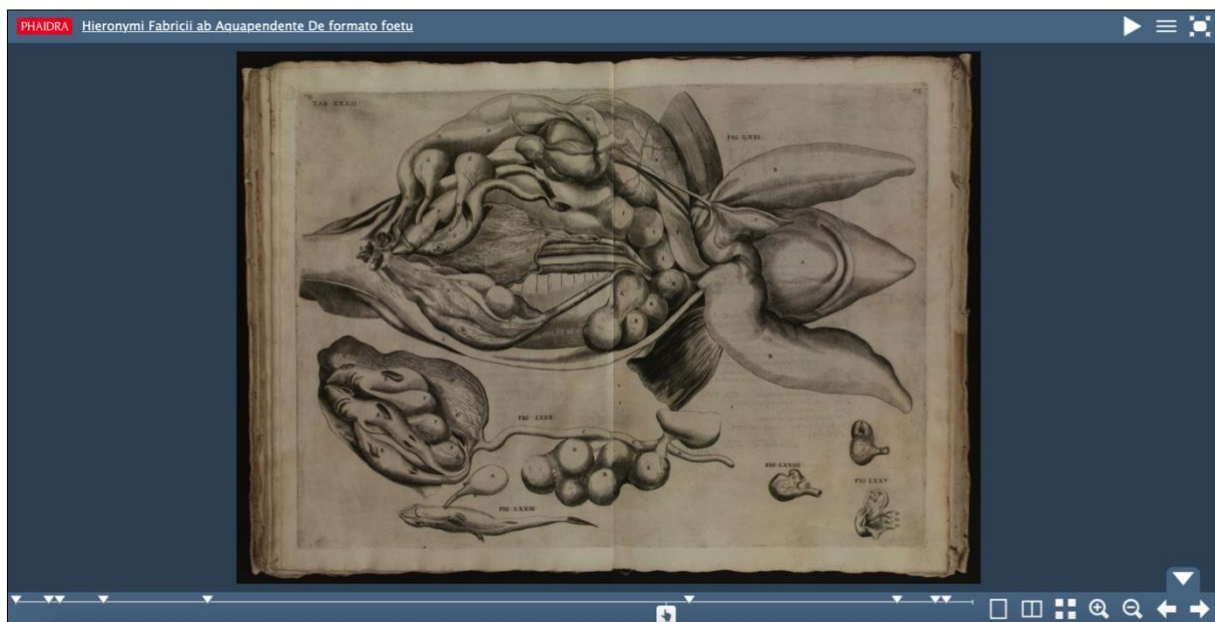
With regard to the re-alignment of adjacent pages, more than two couples of guidelines were set in case of continuous engravings. It was essential to set at least a couple of guidelines for each figure in order to make the lines match. In *Figure 34*, for example, there were six guidelines since two continuous engravings illustrate the anatomy of a fish, with particular attention to the oviduct. The guidelines show that the left page is higher than the right. By adding upper and lower margins, the guides were matched and the images were aligned, as shown in *Figure 35*. The final result of the re-

alignment process can be seen in *Figure 36*, which represents the open book as it can be read in Phaidra using the *BookViewer*.



**Figure 34** (on the left) Realignment (before) of the left and right pages of the 30<sup>th</sup>, second part, table of Acquapendente's *De formato foetu*. Guidelines do not correspond. ©Ateneo Veneto

**Figure 35** (on the right) Realignment (after) of the left and right pages of the 30<sup>th</sup>, second part, table of Acquapendente's *De formato foetu*. Guidelines now correspond. ©Ateneo Veneto



**Figure 36** Final result of the 30<sup>th</sup>, second part, table of Acquapendente's *De formato foetu* after the realignment as it can be seen in Phaidra BookViewer. ©PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023)

*De formato foetu* was the book that required the longest digitization time: more than 15 hours of work, a time that included both the scanning and the post-capture image processing. This means almost five minutes of digitization for each page on average. The process was laborious but it was totally worth it.



### 3.7 Some Considerations

The “Digitization Project of a Hundred Rare and Valuable Medicine Books” produced more than 30 thousand digital images taken from 104 books. Almost 70% of the overall time of the project was allocated for the digitization. In performing the process, there was an evident difference between large and small books. The digitization of quartos and smaller formats required almost half a minute per image, whereas folios needed more than double. In fact, the scanner’s capturing area limited the digitization to one page at a time and implied laborious post-capture image processing. Nevertheless, large formats could be handled and browsed with much more convenience. Small books had to be raised in height in order to augment the scanning area and their binding exerted a greater resistance in keeping them opened with the use of the glass platen. As a result, small formats were more unstable during digitization.

The process entails benefits and issues that can be highlighted for every project of this type with particular reference to small cultural institutes. Besides the principal benefits that result from the objectives of the project – the accessibility of cultural heritage –, there are some additional advantages connected to the process. The most evident one regards the monitoring and control of the originals. Inspection and cleaning are activities not frequently carried out by small cultural institutions due to the loss of time and the presence of alternative tasks. Nonetheless, they are part of the conservative operations that a cultural organization has to conduct in favour of its heritage. During the digitization process, technical operators directly assess the status of the books and can plan the necessity of restorative interventions. Furthermore, they inspect the conservation environment, evaluate the arrangement of the books and may notice problems related to inventory and cataloguing. Not infrequently errors with the shelfmark can be discovered or some books may have been catalogued as a single title, ignoring that the volume is a collection of several titles. The close comparison with the text, which is one of the main advantages of digitization, can highlight hitherto unknown titles, which must be separated and catalogued as independent. Inserts and particular features can be discovered or rediscovered by institutions and used for public promotion through social media and institutional communication. In other cases, there can be a direct benefit to the online catalogue. The book’s corresponding online OPAC inventory sheet may not have been created, or the book may have been associated with an erroneous sheet. It will be then necessary to make the appropriate interventions and provide the proper inventory accessibility.

Time and cost represent the major limits to digitization. Each activity must be conducted within the prefixed time in order to meet the deadline of the project. Many operations may therefore be performed quickly or mechanically. Others, on the contrary, must be excluded beforehand from

the project in view of their implementation in the future. This, for example, happened in the case of Ateneo Veneto's project with regard to OCR<sup>178</sup>. Much depends on the capabilities of technical operators, e.g., how they calibrate the workstation, adjust the parameters of digitization, scan, and perform the post-editing. In fact, in small digitization projects, most of the process is manually managed and executed. The institution's tight budget precludes in fact the possibility of buying semi-automatic or fully automatic high-resolution scanners. As a result, some of the excluded functions must be performed manually by technical operators, not without the use of ingenuity and brilliant solutions. For example, the size of the scanning area of the CZUR scanner used by the Ateneo Veneto did not allow smooth digitization of large format volumes, forcing the institute to a laborious post-capture image processing. There may indeed be a waste of time and energy in the case of small digitization projects. In addition to this, a lower quality level may be achieved in comparison to high performances offered by more advanced machines.

One of the most serious problems that can occur during digitization is damage to the originals. Reference is not made to the carelessness or ineptitude of the project team, but rather to the damage that the originals may suffer due to the structural limitations of the digitization workstation. The whole process weighs on the binding of the book. According to its rigidity, the binding can be exposed to more or less pressure while the book is kept open. Due to this fact, new fissures can appear on the spine, the cover or the fore-edge of the book. The rigid spine, in particular, can bend inward highlighting a vertical fold on it. This type of problem can be solved by holding and scanning the book in a V-shaped book cradle. Pressure on the book spine is reduced and at the same time, the internal curvature in the gutter is less pronounced in this way. When the workstation is not designed to have a V-shaped book cradle – as happened in the case of the Ateneo Veneto – meticulous attention must be paid to the forces acting on the book lying on a flat plane. The operator should seek a compromise between book preservation and effective scanning. Due to the binding and the particularity of the book, pages can be undoubtedly slightly oblique and the text cannot be perfectly horizontal. Some distortion can arise. Nevertheless, by taking advantage of the maximum performance offered by the hardware and software, technical operators must always proceed with the utmost respect for the originals. They have the privilege to handle valuable books with their hands, to browse them and give them a new, digital life. They somehow temporally own the originals during digitization. Consequently, they have to feel responsible and implement all the actions necessary for their preservation.

---

<sup>178</sup> Cfr. [Footnote 15](#).

## CHAPTER 4: REACHING OUT TO THE PUBLIC. THE PHAIDRA REPOSITORY

The digitization process produces digital objects obtained from the conversion of analogue assets. Yet, another step is required to reach the objectives of the project – the online publication. The publication represents indeed the final strategic passage that effectively makes digital objects available to everybody interested in them. The importance of a reasoned choice of the project's repository was discussed in Chapter 2. As already stated, a good repository has two main functions. As a digital asset management system, it enables long-term archiving of digital objects, adhering to specific requirements and standards. Furthermore, as a digital collection, it provides accessibility and fruition to the objects through tools that facilitate the dissemination and reuse of the deposited digital objects. In the case of cultural institutions of big dimensions, digital objects are frequently published on specific platforms that were internally developed.<sup>179</sup> These platforms are usually linked to the main institutional website where a section presents the collection to the public and offers the user the possibility to search, browse and discover it. In the case of small cultural organizations, on the contrary, it is likely that an internal digital repository is not affordable or not convenient. The institution will therefore take advantage of external platforms. This was also the Ateneo Veneto's choice in using the Phaidra repository managed by the University of Padova Library System (SBA). Some of the features of Phaidra that were generally considered in the strategic planning have already been explained. In this chapter, characteristics and tools are closely discussed together with the system that allows ingestion, which is the process of transferring digital images and metadata into the platform. With specific reference to metadata, online publication requires the production of some descriptive categories that are essential to identify the objects. The Ateneo Veneto's books were accordingly described, uploaded and made accessible in Phaidra. Nor all the typologies of metadata were produced nor all tools were used. In accordance with the deadline of the project, the team decided which features were important to provide from the beginning and which ones could be postponed. Presentation possibilities and multidisciplinary connections offered by the platform are numerous and the project did not envisage to make use of all of them. Some additional steps were indeed planned in the future.

---

<sup>179</sup> Minerva Working Group 6, ed. (2004), p. 48.

## 4.1 The Repository

“Permanent Hosting, Archiving and Indexing of Digital Resources and Assets”, also known as PHAIDRA, was designed and built in 2008 by the University of Vienna.<sup>180</sup> The project was born through the collaboration of the university library system and computer centre (ZID - Zentraler Informatikdienst), which chose to base the repository on the open-source software Fedora.<sup>181</sup> From the beginning, Phaidra was conceived as a DAMS (Digital Asset Management System) platform where specific categories of the academic community, such as professors, students, and researchers, could permanently archive, describe and access their publications, making them globally available over the Internet. Today Phaidra represents a European, certified long-term repository where digital cultural objects are archived and disseminated. Parallel to the development of PHAIDRA, the University of Padua was looking for a platform that could host the digital objects that different digitisation campaigns of its collections were producing.<sup>182</sup> From the point of view of the university, digitalization was understood as part of the so-called “Third Mission”, the set of activities conducted by universities to enter into direct interaction with society. Those digitization campaigns started in the 1990s but it was only in 2014 that a vast program was launched and included all the SBA poles. As a result, the University of Padova considered and then chose the Phaidra repository of the University of Vienna. This choice was formalized in 2010 with a collaboration and development agreement between the two universities. The collaboration has taken into account the presence of a consolidated technology for the storage of digital objects, cost containment, innovative aspects of the platform and the possibility of modular developments from a European perspective. Moreover, it has allowed the sharing of the platform and the exchange of experiences and competencies, although Phaidra is formally maintained by the University of Vienna. The section of Phaidra managed by the Library System of the University of Padova has its own domain of access.<sup>183</sup> Here are archived rare documents or other items of historical, artistic, cultural or academic interest directly digitized by the University of Padova or by all those institutions that have decided to use the Phaidra platform for publishing their digital collections. In particular, the Ca’ Foscari and IUAV universities in Venice have been collaborating with Phaidra since 2014 and the University Museum System of the University of Bologna since 2019. Minor cultural organizations joined Phaidra as well, such as the Library of the Episcopal Seminary of Padua, the Theological

---

<sup>180</sup> University of Vienna, *Phaidra* (2023).

<sup>181</sup> Montanaro, D. (2017), p. 10.

<sup>182</sup> Andreoli, L., Bianchi, E., Cappellato, L. et al. (2019), p. 149.

<sup>183</sup> PHAIDRA - University of Padua, *PHAIDRA – Digital Collections* (2023).

Faculty of Triveneto, the Aloisianum Philosophical Institute, the Marciana National Library of Venice, the University Library of Padua (Ministry of Culture) and, from 2023, the Library of the Ateneo Veneto. The launch of collaborations with external cultural institutions has greatly enriched the number of digital objects hosted on the platform and created virtual collections, overcoming the physical and administrative barriers due to the differentiated location of analogue objects. The widespread Phaidra community now comprises the following countries: Austria, Bosnia and Herzegovina, Italy, Montenegro and Serbia. Phaidra now keeps more than 500 thousand objects which include early-printed and modern books, manuscripts, photographs, wall didactic tables, maps, museum objects, archive materials, and parchments. Three languages are available: Italian, English, and German. In contrast with the corresponding Viennese platform, which provides a wide range of university users the opportunity to publish digital content, the Paduan Phaidra allows only trained members of university staff and external institutions to do it. The Phaidra staff is composed of librarians and IT staff of the Digital Library Unit of the Library System and guarantees the maintenance and development of Phaidra. At the same time, the navigation to the site and the digital objects access is opened to everybody free of charge.

As shown in *Figure 37*, Phaidra's digital environment consists of three main compartments, which are ingestion, archival, and dissemination<sup>184</sup>:

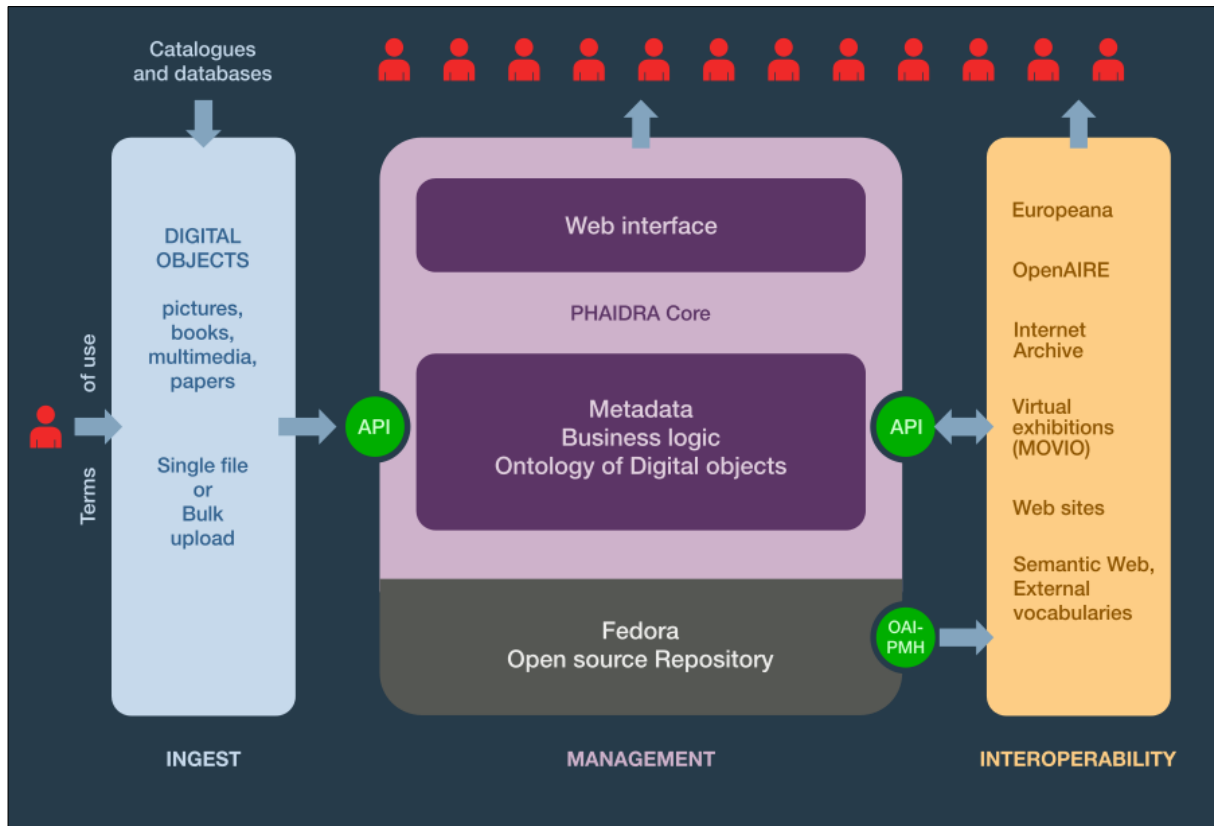
- Ingestion: the uploading of digital objects is done through Phaidra's web interface or the application called "Phaidra Importer" by authorized users. Public API may be used to make massive uploads.
- Archival management: digital objects with their metadata and relationships, together with all the descriptive tables, are stored in the Fedora platform<sup>185</sup> and organized according to the logic of Phaidra's application in order to be available and searchable through the web interface. Different visualization tools are available: BookViewer, ImageViewer, and CollectionViewer.
- Dissemination (Interoperability): the use of APIs allows the transfer of digital resources from Phaidra to other platforms (Europeana, openAIRE, Internet Archive) and the integration of objects and collections in virtual websites and exhibitions (MOVIO),

---

<sup>184</sup> Andreoli, L., Bianchi, E., Cappellato, L. et al. (2019), p. 151.

<sup>185</sup> Fedora, *Home* (2023). The original research that led to the first Fedora prototype was conducted in the Digital Library Research Group in the Computer Science Department at Cornell University (USA) in 1997. The Fedora project is now led by the Fedora Governance Group and is maintained by the DuraSpace Community-Supported Programs Division of Lyris. This open-source repository system is especially suited for digital libraries and archives, both for access and preservation of single digital objects or complex digital collections. In 2022 the last software version 6.2.0 was released.

encouraging the dissemination of content. Metadata is standardized to the Dublin Core schema and exposed through the OAI-PMH (Open Archive Initiative Protocol for Metadata Harvesting)<sup>186</sup> for the exchange of data in a distributed environment.



**Figure 37** The Phaidra digital environment. ©PHLAIDRA - University of Padua, PHAIDRA – Digital Collections (2023)

Phaidra is a repository compliant with FAIR (Findable, Accessible, Interoperable, Reusable) principles.<sup>187</sup> These principles were recommended by the European Commission Horizon 2020 guidelines to encourage the sharing and reuse of open data. The platform uses established standards, unique and persistent identifiers, and a large amount of metadata (Findable). Clear access policies to archived objects are presented to users (Accessible). Data and metadata use a formal, accessible, shared and widely applicable language for the representation of knowledge (Interoperable). Digital objects are finally equipped with a set of properties which allow their reuse (Reusable). Phaidra’s “FAIRness”, in other words, the compliance of the platform with FAIR principles, satisfies thirteen of the fifteen requirements, as measured by Phaidra’s staff.<sup>188</sup> For this reason, as already mentioned, in 2019 Phaidra was selected with other repositories to test innovative

<sup>186</sup> Open Archives Initiative, *OAI-PMH* (2023). The protocol is “a low-barrier mechanism for repository interoperability”. OAI-PMH Data Providers are repositories that expose structured metadata via the protocol. Service Providers then make OAI-PMH service requests to harvest metadata.

<sup>187</sup> Cfr. [Subchapter 2.4.1 The Choice of a Long-Term Conservation Platform](#).

<sup>188</sup> Andreoli, L., Bianchi, E., Cappellato, L. et al. (2019), p. 153.

developments in the context of the European “FAIRsFAIR” project.<sup>189</sup> Phaidra was technically designed following the OAIS recommendations. In the same 2019, the platform undertook the evaluation process as long-term DAMS through the prestigious certification – finally awarded – issued by CoreTrustSeal.<sup>190</sup>

Interoperability is undoubtedly one of the main strengths of Phaidra. The platform is designed to be interoperable both inbound and outbound. In order to attract digital resources belonging to different GLAM (Galleries, Libraries, Archives, Museums) organizations, the platform was provided with diversified metadata schemes. EAD standard<sup>191</sup> is used for archival documents, ICCD standard<sup>192</sup> for museum assets, and UNIMARC standard<sup>193</sup> for library resources. To this end, exhaustive mapping between source and internal metadata schemes has been developed. The Phaidra metadata model, called UWmetadata, is functional for the conversion and integration of all these different types of specific metadata. To be interoperable, metadata is also converted and displayed according to the Dublin Core schema, which represents the worldwide metadata standard.

Furthermore, the repository is designed to be in dialogue with other national and international DAMS platforms. The basis of this dialogue is the creation of APIs and the use of the aforementioned OAI-PMH. Nevertheless, Phaidra lacks a resource model based on the Resource Description Framework (RDF) standard. The use of non-interoperable or restricted vocabularies represents a significant obstacle to interoperability as well. The staff is developing a new semantic model and implementing a new vocabulary management system to overcome this limitation. Considering APIs (Application Programming Interfaces), they are used to extend the digital environment built in Fedora and implement other services such as search, content management, metadata extraction and reuse. Specific APIs were designed to bilaterally import and export digital collections and related metadata from and to Internet Archive. Between 2012 and 2013 and in recent years, some of Phaidra’s digital collections were published in Europeana thanks to the European project Linked Heritage. Phaidra joined also CulturaItalia, the Italian Cultural Aggregator managed by the Ministry of Culture, where some collections were published. The joining of

---

<sup>189</sup> Bettella, C., Carrer, Y., Turetta, G. (2022), p. 120. FAIRsFAIR involved 22 partner institutions, from eight Member States of the European Union, with the primary objective of fostering FAIR data practices in Europe. The project was developed along four thematic axes – *data practices; data policy; certification; training, education and support* – comprising seven working packages.

<sup>190</sup> *Cfr.* [Chapter 2.4.1 The Choice of a Long-Term Conservation Platform](#).

<sup>191</sup> Istituto Centrale per gli Archivi – ICAR, *EAD* (27 July 2022).

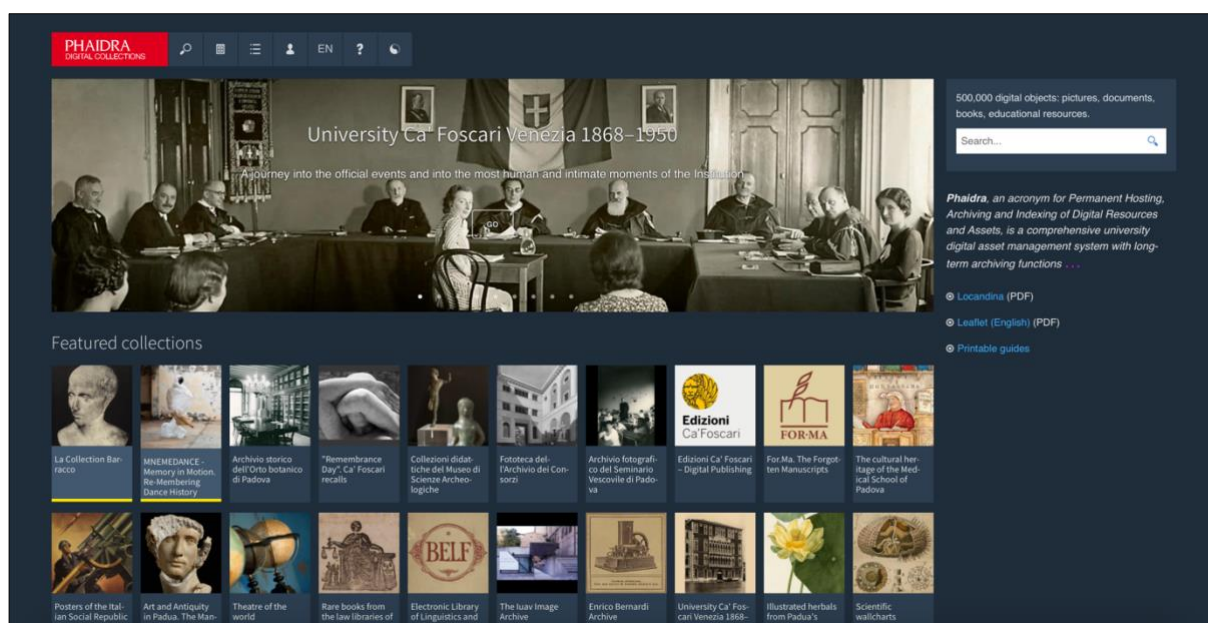
<sup>192</sup> Istituto Centrale per il Catalogo e la Documentazione – ICCD, *Standard catalogfici* (2023).

<sup>193</sup> International Federation of Library Associations and Institutions (IFLA), *UNIMARC formats and related documentation* (2023).

CulturaItalia is strategically important because this platform represents a trusted partner and national aggregator for Europeana.

Finally, one of the most interesting external collaborations is represented by the integration between Phaidra and MOVIO, the ICCU (Central Institute for the Union Catalogue of Italian Libraries and Bibliographic Information) open-source software for creating virtual exhibitions. In this regard, the University of Padova's SBA developed a plug-in to allow the insertion of Phaidra's objects and collections in MOVIO's virtual exhibitions through a simple mask. Virtual exhibitions allow the inclusion of digital objects within real and more interacting narrative paths and foster their dissemination.

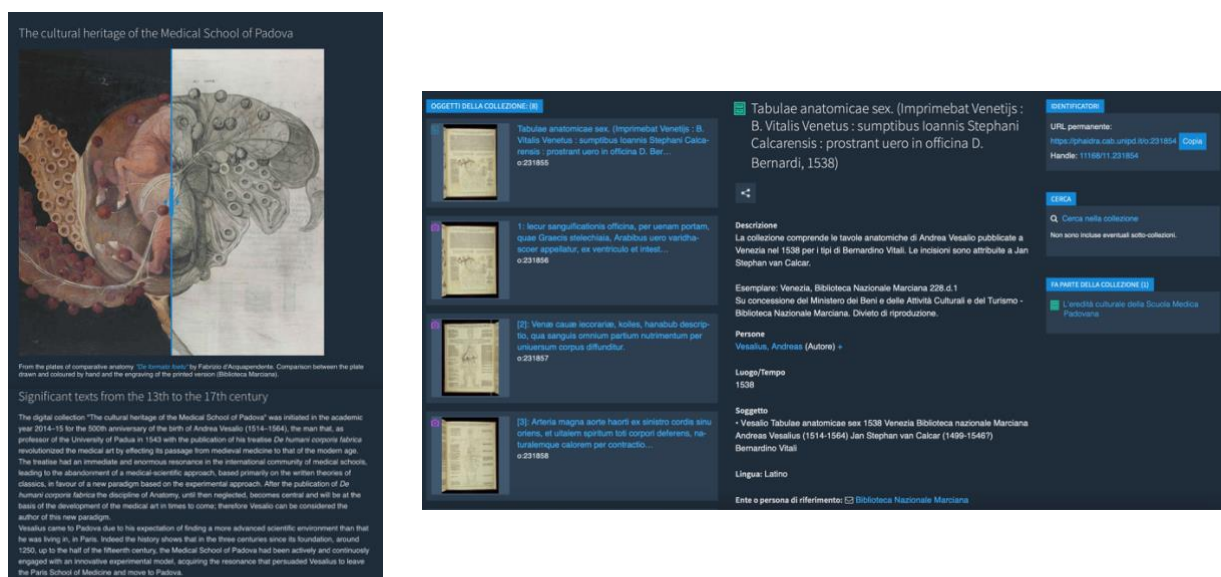
By typing "phaidra.cab.unipd.it" on a search engine, users can access the homepage of the section of Phaidra that is managed by the Library System of the University of Padova. As shown in *Figure 38*, the Phaidra homepage displays a slideshow of images representing some of the most interesting items in the Phaidra collections. Shortcuts of the most recently added collections and a link to a complete list of them are accessible beneath the slideshow. The heterogeneity of the contents archived in Phaidra can be immediately grasped from this view of the homepage. Collections range from both the humanistic and scientific areas. For example, the collection called "The cultural heritage of the Medical School of Padova" highlights the contribution of the Paduan Medical School to the development of modern medicine. "Posters of the Italian Social Republic at CASREC" is instead a historical collection that investigates the Fascist propaganda between 1943 and 1945. Differently, the "Electronic Library of Linguistics and Philology" collection documents Italy's rich panorama of dialects.



**Figure 38** Phaidra's homepage. ©PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023)



On the webpage's right side, a simple search form is shown as well as important links to Phaidra's printable guides, which clarify the platform's functioning and the process of ingestion and management of digital objects. At the top, the horizontal navigation menu allows the user to browse and search the platform's content. The user can access here the search function and browse the repository by setting the object type or the institution as search keys. Besides other distinctive icons, the one of a drawer metal file cabinet provides access to All collections. A collection is a set of relationships between several digital objects. It can contain objects part of the same cultural collection that were digitized in the context of the same project or it can aggregate objects belonging to different cultural organizations. Collections are frequently presented through an informative cover page. From here, the user can explore the collections' sections and subsections, up to view the individual digital objects. A visual example of a digital collection is shown in *Figure 39*. Like any digital object, collections and each component are identified by unique and persistent identifiers. According to FAIR principles, they also possess their own description and metadata.



**Figure 39** An example of digital collection in Phaidra: “The cultural heritage of the Medical School of Padova”. On the left, the cover page of the collection. On the right, an internal section. ©PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023)

Creating a collection is a time-consuming process, but it has the great advantage of grouping and enhancing digital objects, which would otherwise remain isolated. The Ateneo Veneto is planning to create as soon as possible a digital collection with all the books digitized during the project. This has not been possible before precisely because of the amount of operations that the project already required. However, the user can search and visualize single digital objects. *Figure 40* shows Phaidra's page of one of the books digitized and published by the Ateneo Veneto, Garcia da Orta's

*Dell'istoria de i semplici aromati...*<sup>194</sup> The digital object is a book and for this reason, an icon representing a book is displayed alongside the title (and the subtitle, if present). Object types are: image, book, PDF document, the aforementioned collection, video, data, resource, and audio.

The screenshot shows the PHAIDRA digital collections interface. At the top left, the PHAIDRA logo and navigation icons are visible. The main content area is divided into several sections:

- Image:** A thumbnail of an old, worn book cover with a small icon of a book and a magnifying glass.
- Metadata:** A section labeled 'METADATA' containing 'Dublin Core' and 'Phaidra metadata'.
- Title:** 'Dell'istoria de i semplici aromati, et altre cose che vengono portate dall'Indie Orientali pertinenti all'vso della medicina. Di don Garzia Dall'Horto medico portughese, con alcune breui annotazioni di Carlo Clusio. . . .
- Description:** A detailed text block describing the book's content, including authors like Garzia Dall'Horto and Carlo Clusio, and publication details from 1616.
- People:** A list of individuals associated with the work, such as Orta, Garcia da (Author), L'Ecluse, Charles de (Author), Monardes, Nicolas (Author), and Briganti, Annibale (Technical Translator).
- Format:** 'application/pdf (153.60 MB)'
- Subject:** 'Dewey Decimal Classification -- Technology (600) -- Medicine & health (610)'
- Object languages:** 'Italian'
- Sources:** 'View in catalogue'
- Rights:** 'This work is licensed under a Creative Commons CC BY-NC-SA 4.0 License.'
- Contact:** 'Biblioteca dell'Ateneo Veneto'
- Identifiers:** A section on the right containing a 'Permalink' and a 'Handle'.

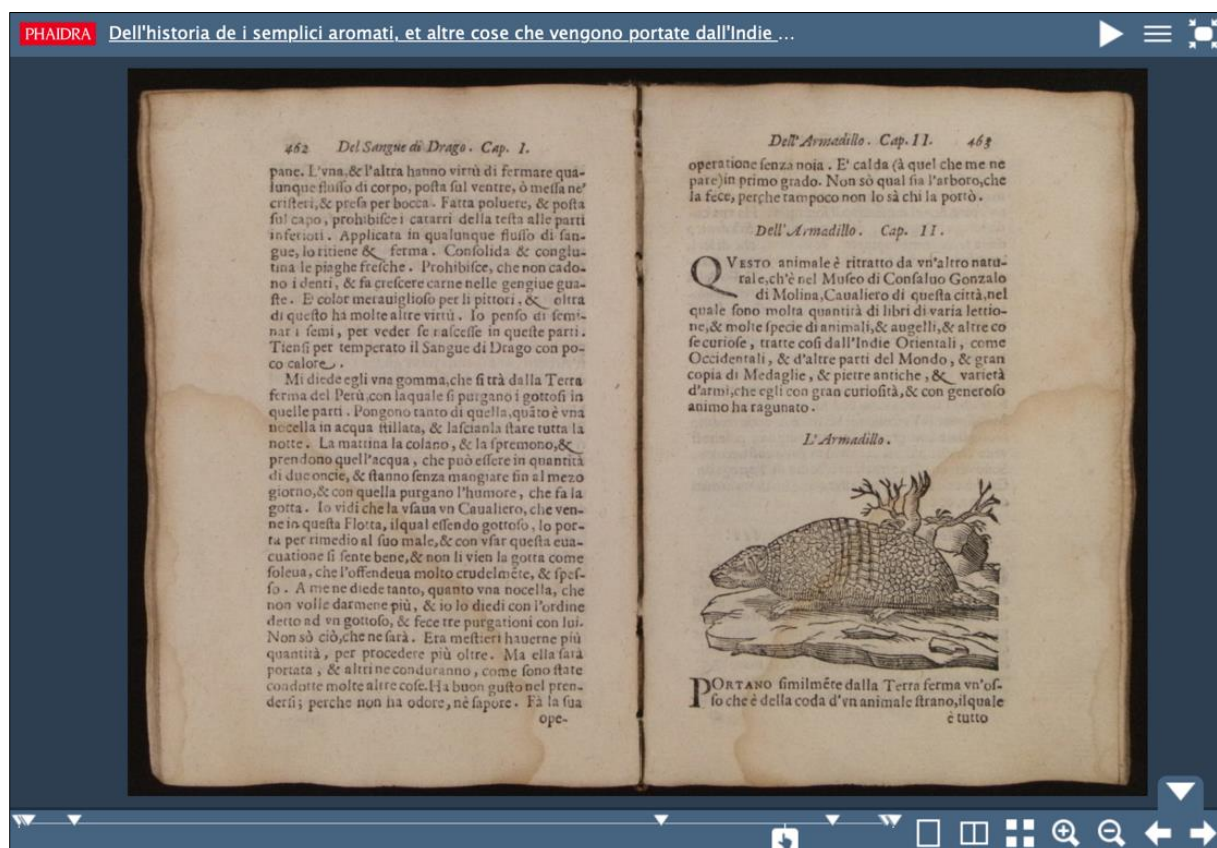
**Figure 40** An example of a digital object and metadata description in Phaidra: Garzia da Orta's *Dell'istoria de i semplici aromati...* digitized by the Ateneo Veneto. ©PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023)

The object's preview is shown on the left. Identifiers, such as permanent links and handles, are available on the right, together with cross-reference to other objects if available. In the middle, under the title, the Dublin Core metadata is presented under a series of labels. The description is a field that freely describes the subject according to the intention of the publisher. People label groups any person or entity related to the life of the object, such as author, translator, engraver, operator, or curator, with the exception of the editor, who has a separate label. Place and Time indicate the spatial and temporal coverage of the object. The Date label can be also displayed when one of the date fields in the metadata editor's calendar format has been checked. Format shows the file format, whereas Subject groups keywords, classifications and chronological subject.

<sup>194</sup> Cfr. [Footnote 95](#).

Language and Sources follow in the list. The licence attributed to the object is highlighted in the Rights section. Contacts of the institution or the person who edited the page are shown at the end. Underneath, full metadata can be accessed. In addition, XML metadata files are accessible through a series of links on the left. These include Phaidra Metadata, based on the UW/metadata model, but also Dublin Core Metadata, created through a conversion of the first type. A more detailed study of Phaidra's metadata is conducted in the following subchapter.

According to the typology, the object can be investigated by clicking on the preview on the left. The BookViewer is the best application to visualize books in Phaidra. All the books are automatically visualized using this tool. *Figure 41* shows the same book written by Garcia da Orta displayed in the BookViewer. The tool provided several functionalities. Buttons on the top right are used respectively to activate the automatic scroll of the book, to open the tab of information, index and downloads, and to visualize if needed the book in full-screen size. Among the large quantity of information accessible here, it is important to highlight the table of contents (if present), the permalink of the page and the possibility of downloading it both as a JPG image or a PDF file.



**Figure 41** Garcia da Orta's *Dell'istoria de i semplici aromati...* displayed in the Phaidra BookViewer. ©PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023)

This means that each page of the book is archived and searchable independently in Phaidra with its own metadata. The book, therefore, is a collection of pages that is bound together from a digital

point of view. Returning to the description of the features offered by the BookViewer, the scroll bar at the bottom allows the user to directly reach internal subdivisions of interest, such as the frontispiece, chapters and bookbinding (if this division was previously created). Finally, the user can browse the book using the buttons on the bottom right, choosing the view (single page, double page, thumbnail view) and zooming in and out.

With regard to images, these are displayed through the ImageViewer, a viewer that is specifically designed to investigate the image's details by progressively zooming in and out. The tool is particularly useful for high-resolution images or with minute details, such as maps and paintings. The ImageViewer uses the IPIImage software which supports the IIIF protocol.<sup>195</sup>

## 4.2 Metadata

In Chapter 2 metadata was included among the key points of digitization planning and in the subchapter 2.4.4 “Data Management” the importance of data transparency and organization was discussed. For convenience, it was decided to deal with metadata at this point with the awareness, however, that it is a topic that concerns digitalization from the beginning of the project and that goes beyond its closure. Much metadata was evident and was collected during the selection of originals, the scanning and the post-capture image processing. Nevertheless, it was only at the time of the publication of the books in Phaidra that data was collected in its entirety and was standardized and uploaded online according to the descriptive tabs arranged by Phaidra's Metadata Editor. That is why metadata is discussed at this point.

In general terms, the word metadata means “data that provides information about other data”. The study about metadata conducted by Murtha Baca, head of Digital Resources Management and the Vocabulary Program at the Getty Research Institute in Los Angeles, provides a more detailed definition in this regard. According to it, metadata can be defined as “a structured description of the essential attributes of an object”.<sup>196</sup> An item provided with metadata is indeed a (digital) object described with several typologies of data which are organised according to a clear and defined structure. Without these assumptions, metadata is practically useless. The definition is frequently

---

<sup>195</sup> IPIImage, *IPIImage* (2023). IPIImage is “an advanced high-performance feature-rich image server system for web-based streamed viewing and zooming of ultra high-resolution images”. IPIImage supports the IIIF (International Image Interoperability Framework) protocol, which is nowadays considered the most interoperable image protocol for standardized image retrieval. It was created in 2012 by a community of the world's leading research and national libraries and non-profit repositories in order to design a technology capable of interoperable image delivery. IPIImage is one of the Image Servers of IIIF. Moreover, IIIF is also supported client-side within IIPMooViewer, which is the IPIImage's presentation API, allowing it to be used with any IIIF-compliant server.

<sup>196</sup> Baca, M., ed. (2008), Chapter 2 “Metadata and the Web”.

illustrated by the metaphor of a library card catalogue, in which the individual entries describe the books. This metaphor is particularly effective in depicting the history of the use associated with this term. In fact, metadata is not only about the digital world. Before the advent of the Internet, it was recorded in analogue formats such as card catalogues and file labels until the 1980s, when libraries started to convert their catalogue data to digital databases. Metadata has been created since human beings have been organizing collections of items. Such metadata was then incorporated into digital information systems from the 1990s and more and more from the 2000s.

The primary function of metadata is indeed the description and consequently the retrieval of digital objects in a repository or collection. Descriptive tabs can be automatically or manually organized into the structure of DAMS. Additional data may refer to registration, cataloguing, and indexing processes. Moreover, the development of the Internet can justify the increasingly considerable use of metadata and other type of functions. Metadata is a tool that can allow better exploitation of the Internet's opportunities while addressing some of its issues. First of all, metadata can help users to effectively search and find relevant information and discover resources. Rich, consistent and interoperable descriptive tabs allow searchability and retrievability on search engines or online databases. Metadata recorded in the same format or mapped across various formats also enables the consultation and access of digital objects archived in several repositories. Each scientific discipline has its own technical language and therefore its specific metadata. Nonetheless, in recent years the debate among IT operators has stressed the importance of having common standards concerning the descriptive categories and their structure in order to guarantee interoperability between information systems.

Furthermore, metadata has the advantage of being able to describe not only the intrinsic content of digital objects but also the extrinsic data that form the context in which the object was created and belongs, as well as the set of relationships that form its structure. The possibility of including this information together with the digital object is a great advantage. The fluidity and speed of the digital world can endanger the integrity and authenticity of objects. It often happens that objects are decontextualized or copies of the originals are spread over the network. In these cases, metadata has the same effect as an identity card for digital objects and allows users to corroborate their availability and authenticity. At the same time, from the point of view of the institution, metadata promotes the dissemination of digital objects without posing a significant danger. In other cases, its attribution makes it possible to identify both the original digital objects and the multiplicity of versions created. In the case of digitization projects, as mentioned, different types of files are produced, from master files to derivative files up to those for online publication. They are files

with different quality levels and technical specifications. As a result, metadata allows their distinction.

Likewise, information about rights, reproduction licences, and legal restrictions can be attributed through metadata to digital objects.

Although there is no absolute agreement in the literature, there can be identified four major types of metadata that digitisation projects should take into account<sup>197</sup>:

- Descriptive metadata provides information needed for the identification, authentication and description of digital objects. This typology includes title, author, date, place, keywords, cataloguing records, curatorial information, and details about the original.
- Administrative metadata helps to manage the resource or the collection of which it is part, like resource type, location information, information about the digital acquisition, legal permissions, licences and the digitization project.
- Structural metadata describes the manner in which the components of an object are organized, and types, versions, relationships, and other characteristics of digital materials.
- Technical metadata includes all the information about the generation and the management of the file, such as software and hardware acquisition documentation, format, compression, resolution, and dimensions.

After this brief introductory explanation about the main purposes and advantages of the use of metadata, hereafter will be discussed the type of categories of metadata implemented in Phaidra to describe digital objects and their organization. The repository's Metadata Editor manages all four aforementioned typologies of metadata. Depending on the type of digital object, descriptive and administrative metadata can be exported in an XML file, called "UWmetadata", during the publication process, as explained in the following subchapter. In the case of museum objects, there may be an additional special metadata profile, called Museum metadata.<sup>198</sup> With reference to structural metadata, Phaidra allows the operator to create the book's internal structure and export it in an XML file, called "phaidraimporterstructure". For all types of objects, then, it is possible to create relationships, subsections, and sections, up to a complete collection. Technical metadata is finally shown in the section dedicated to "Technical Data" and in an EXIF (Exchangeable Image File Format) viewer if data about image acquisition was provided.

---

<sup>197</sup> Baca, M., ed. (2008), Chapter 1 "Setting the Stage"; Docs Italia, *Linee guida per la digitalizzazione del patrimonio culturale. Versione 1.0* (2022), p. 29; Wikipedia. *The Free Encyclopedia* (2023), under "Metadata".

<sup>198</sup> Cappellato, L. (2023), p. 6.

Phaidra objects are described through the UWmetadata (Universität Wien Metadata) schema, the conceptual data model created by the University of Vienna specifically for the repository.<sup>199</sup> As explained by the university, UWmetadata was created through an expansion of the IEEE 1484.12.1-2002 – Standard for Learning Object Metadata (LOM).<sup>200</sup> Elements of different metadata namespaces belonging to other schemes were combined with this model.<sup>201</sup> Similarly to LOM schema, UWmetadata has a hierarchical tree structure composed of thirteen descriptive categories, which include distinct attributes and sub-attributes. The metadata set used by the University of Padova has been customized over the years according to needs and local specificities. It consists of nine categories, which are: *General*, *LifeCycle*, *Technical Data*, *Rights & Licences*, *Classification*, *Association*, *Contextual Allegation*, *Provenience*, and *Bibliographic Data*.

Thanks to a mapping that converts the different UWmetadata fields, a second, Dublin Core metadata profile is automatically generated for each digital object in Phaidra.<sup>202</sup> The Dublin Core, also known as the Dublin Core Metadata Element Set (DCMES), was inspired and created during the “OCLC/NCSA Metadata Workshop” held in 1995 in Dublin, Ohio (USA).<sup>203</sup> The term “core” refers to the features of broadness and generality of the constituting elements, making the set “usable for describing a wide range of resources”. The set consists of only fifteen elements that can be optionally and repetitively used to describe both digital and physical objects.<sup>204</sup> Due to the flexibility and breadth of applications, DCMES has become a widespread, interoperable metadata standard and the preferred schema for metadata mapping and harvesting.<sup>205</sup> The Dublin Core Metadata Initiative (DCMI) is currently responsible for formulating and updating the metadata set. Phaidra uses DCMES both for the Web visualization of its digital objects and for the publication

---

<sup>199</sup> Andreoli, L., Bettella, C., Cappellato, L., et al. (2019), p. 2.

<sup>200</sup> Learning Object Metadata (LOM) is a metadata model used to describe learning objects and similar resources. According to the Institute of Electrical and Electronics Engineers (IEEE), a learning object is “any entity, digital or non-digital, that may be used for learning, education or training”. The IEEE 1484.12.1-2002 LOM Standard defines relevant attributes of learning objects to be described and for this reason, is internationally widespread. 70 attributes are tree-structured and form nine descriptive areas, which are: *General*, *LifeCycle*, *Meta-metadata*, *Technical*, *Educational*, *Rights*, *Relation*, *Annotation*, and *Classification*. LOM categories are usually encoded in XML files.

<sup>201</sup> The descriptive tabs are: *Contextual Allegation*, *Provenience*, and *Bibliographic Data*.

<sup>202</sup> Cfr. Bettella, C. (2018).

<sup>203</sup> DCMI, *Dublin Core™ User Guide* (n.d.). The “OCLC/NCSA Metadata Workshop” owes its name to the two promoting organisations, the OCLC (Online Computer Library Center) library consortium based in Dublin, and the National Center for Supercomputing Applications (NCSA). At this event, it was discussed how a core set of metadata for Web-based resources would be extremely useful for categorizing the Web and allow searchability and retrievability. This marked the birth of the “Dublin Core™ metadata”.

<sup>204</sup> The fifteen DC metadata elements are based on multiple Resource Description Framework (RDF) vocabularies. They are: *Contributor*, *Coverage*, *Creator*, *Date*, *Description*, *Format*, *Identifier*, *Language*, *Publisher*, *Relation*, *Rights*, *Source*, *Subject*, *Title*, and *Type*.

<sup>205</sup> DC was formally standardized as ISO (International Organization for Standardization) 15836, as IETF (Internet Engineering Task Force) RFC 5013 and in the USA as ANSI/NISO (National Information Standards Organization) Z39.85.

and exposure of metadata to other platforms. The integration of Dublin Core metadata inside Phaidra has promoted interoperability with external digital aggregators, such as Europeana and Internet Archive, and enhanced the exposure of Phaidra content to web search engines.<sup>206</sup> In this regard, better indexing by search engines was provided in 2021 when each Phaidra digital object incorporated metadata exposed in Schema.org's vocabulary.<sup>207</sup> Data exposure in a distributed digital environment is also fostered through the aforementioned OAI-PMH. A final strength of UWmetadata is its ability to convert and integrate specific metadata profiles, such as ICCD museum standards, EAD encoding for archives, and MARC bibliographic metadata.

As regards the UWmetadata set, the University of Padova uses nine descriptive tabs that are managed in Phaidra's Metadata Editor. Metadata cannot be inserted using special formatting such as bold or italics style or using HTML code in order to ensure data interoperability. Some elements, in particular title, subtitle, description, keywords, and coverage, can be duplicated and expressed in multiple languages (Italian, English and German). If present, they are visualized according to the language chosen by the user. All the other elements are instead automatically translated by the platform. The first metadata category is *General Data* and groups general information that identifies the digital object. In this regard, the permanent and unique identification of the object is guaranteed by the automatic assignment of an identifier and a handle. The identifier is built on the serial number that the object has within the Phaidra platform, while the handle is an identifier managed by external servers to Phaidra and remains unchanged even if the object is moved to another repository.<sup>208</sup> Other identifiers may be assigned to the object in addition to the Phaidra system's own identifier, such as AC-Number, DOI, HTTP/WWW, ISBN and others. General Data's obligatory fields are title, language, and description. To these, the subtitle, alternative title, keywords, and space-time coverage may be added.

The second descriptive category is called *LifeCycle*. The history and current state of the resource are illustrated through those entities who had or continue to have any relation with the creation and maintenance of the analogue object and the corresponding digital one. There are three main elements. Firstly, the section called "Role" specifies the type of contribution that the person made to the object. It may be indicated the creator such as the author, the cartographer, the composer, the photographer, etc. but also the editor, the publisher, the printer or the translator, as well as

---

<sup>206</sup> Cappellato, L. (2023), p. 5.

<sup>207</sup> Schema.org, *Homepage* (n.d.). The community promotes schemas for structured data on the Internet. These schemas are a set of Types and Properties arranged in a hierarchy. The vocabulary is used by over 10 million sites and currently consists of 803 Types, 1465 Properties, 14 Datatypes, 87 Enumerations and 463 Enumeration members.

<sup>208</sup> Cappellato, L. (2023), p. 11.



those people who contributed after the creation of the cultural item and those who took part in the current digitization project. Secondly, personal data that identifies the author on the dedicated content can be associated, for example, name and surname, relevant titles or institution. Thirdly, Date may be used to express the date of contribution in relation to the indicated role.

The third UWmetadata descriptive tab, *Technical Data*, is mostly automatically recorded by the system and includes the format and size of the object and permanent links. In *Duration*, the publisher can indicate the time taken to watch or listen to a multimedia file.

*Rights and Licences* constitute a fourth category of metadata and describe the intellectual property rights and conditions of use for the resource. The institution can specify the type of licence that authorizes the reproduction, distribution, broadcasting, supply and editing of protected objects by way of a Creative Commons licence. Contrary to “All rights reserved”, which declares that the object is not reproducible without express authorization, Creative Commons licences, now at their 4.0 version, grant the public permission to use it under copyright law.<sup>209</sup> The publisher can choose among seven different licence types. Starting from the most permissive, they are:

- No Rights Reserved (CC0). The object is completely in the public domain and everyone can copy, modify, distribute and perform the work, even for commercial purposes and without asking permission.
- Attribution (BY). Licensees can share and adapt the object, even for commercial use, only if they give the author or licensor appropriate credits (attribution) in the manner specified by these.
- Attribution-ShareAlike (BY-SA). This licence allows reusers to share and adapt, even for commercial use, the digital object, so long as attribution is given to the creator and under a licence identical to the licence that governs the original work.
- Attribution-NonCommercial (BY-NC). This licence allows reusers to share and adapt the digital object only for non-commercial purposes and only so long as attribution is given to the creator.
- Attribution-NonCommercial-ShareAlike (BY-NC-SA). This licence allows reusers to share and adapt the digital object under the same rights as the previous licence and under identical terms of licence that govern the original work.

---

<sup>209</sup> Creative Commons, CC Licences (16 August 2023). Creative Commons is an American non-profit organization that works to overcome legal obstacles to the sharing of knowledge and creativity. The organization has released several copyright licences, known as Creative Commons licences. They represent a free, simple, and standardized way to grant copyright permissions for creative and academic works.

- Attribution-NoDerivatives (BY-ND). This licence allows reusers to share the material in unadapted form only so long as attribution is given to the creator. The licence allows for commercial use, whereas prohibits the distribution of modified materials.
- Attribution-NonCommercial-NoDerivatives (BY-NC-ND). This licence allows reusers to copy and distribute the object in unadapted form only for non-commercial purposes and by giving the author or licensor appropriate credits.

Phaidra encourages institutions that use the platform to publish digital objects to allow a broader, but regulated circulation of materials by avoiding the choice of “All rights reserved” and preferring one of the Creative Commons licences. The BY-NC-SA licence is particularly recommended. The platform gives the possibility to change the type of copyright notice but only in a wider sense of distribution and use.

The fifth descriptive category is *Classification*. This tab can be used to classify digital objects according to international thesauruses and classification systems. This type of metadata is not mandatory, yet it fosters the definition of the object and its retrieval during thematic research.

Phaidra provides four different classification models:

- “The ACM Computing Classification System” (1998 version), which is a standard classification system for the computing field.<sup>210</sup>
- “EuroVoc Thesaurus” (version 4.2), a multilingual, multidisciplinary thesaurus covering the activities of the EU.<sup>211</sup>
- “Dewey Decimal Classification” (DDC) – only in the English edition –, which is a decimal library classification used to organize knowledge through ten classes, each divided into ten divisions.<sup>212</sup>
- “Physics and Astronomy Classification Scheme” (PACS), developed by the American Institute of Physics (AIP), which represents a hierarchical classification for themes related to physics, astronomy, and similar sciences.<sup>213</sup>

Phaidra offers a “Search Help” tool to the publisher. It is sufficient to enter a search term and the system returns one or more classification strings.

*Association* is the sixth UWmetadata category. It allows the publisher to associate the digital object with the institution where the analogue object is kept or the one that created the digital one.

---

<sup>210</sup> Association for Computing Machinery, *The 2012 ACM Computing Classification System* (n.d.).

<sup>211</sup> Publication Office of the European Union, *EuroVoc* (n.d.).

<sup>212</sup> OCLC, *Dewey Services* (n.d.).

<sup>213</sup> Physical Review Journals, *Physics and Astronomy Classification Scheme (PACS)* (11 January 2008).

The seventh category, *Contextual Allegation*, records metadata for the original object. In particular, it is possible to express its material and size and the geographical coordinates. Additional information can be provided in the next category, *Provenience*, particularly useful to indicate an open or uncertain date or when the author, publisher or printer are unknown.

The last category indicates *Bibliographic Data* about the object, such as the title of the series, pagination, place and date of publication, editor and the permanent link to the online local catalogue.

As explained in the next subchapter, the Ateneo Veneto decided to select and provide some of these metadata categories during the publication process, in view of an extension in the future.

### 4.3 The Publication Process

The process of archiving digital objects in a long-term repository requires some preliminary activities. First of all, it is necessary to decide which typologies of the forementioned metadata the institution wants to provide, where to find them, and how they can be organized and arranged in view of the system ingestion. Secondly, the project team has to inspect all the digital material archived and check that formats satisfy the recommendations for the online publication. It will then study the specific loading mechanisms of the platform and perform, if possible, some tests. Finally, the publication will be realized for each digital object, providing selected categories of metadata, structural one included. A final online quality control is necessary to avoid errors, make corrections and last changes.

Metadata was selected and imported from OPAC SBN books' description. The project team decided to provide those categories of metadata that are mandatory in Phaidra, together with some other types. In fact, the project's tight timeframe did not allow the loading of a complete set of metadata and for this reason, it was preferred to select metadata that did not take too much time, while always ensuring a clear and accessible description of digital objects. The obligatory metadata categories in Phaidra are title and title language, language, description, role, and copyright licence. All this information was conveniently taken from the respective descriptive fields in OPAC. For a visual example, reference is made to above *Figure 40*. With regard to the description category, each book was provided with the International Standard Bibliographic Description (ISBD), which was considered a complete, consolidated standard description form.<sup>214</sup> The "role" field was completed

---

<sup>214</sup> International Federation of Library Associations and Institutions (IFLA), *International Standard Bibliographic Description (ISBD)* (n.d.). ISBD is a set of bibliographic description rules created by the already-mentioned International Federation of Library Associations and Institutions (IFLA). The first consolidated edition was published in 2011. In

with data referring to the author(s) of the work. For several books, this field was duplicated or tripled to include secondary authors, engravers and translators, which are all included by OPAC SBN in the category of “author”. Differently, the specific “publisher” metadata entity is going to be entered in the future. The Ateneo Veneto decided then to choose the licence Attribution-NonCommercial-ShareAlike (BY-NC-SA) to provide access and reuse of its digital object. This choice accepted Phaidra’s suggestion to encourage a wider and free, albeit regulated, dissemination of cultural items on the platform. Moreover, the project team decided to associate each object with the institution profile. Though not mandatory, this metadata category allows the user to perform a search on the portal by setting as a filter the belonging institution. Thus, all the objects digitized during the project can be found and selected, even if a collection has not been created so far. Finally, other two metadata elements were published. The first one is the thematical classification according to the Dewey Decimal Classification (DDC) system, which is the same for all the books: “600: Technology → 610: Medicine & health”. The second one is the permanent link to the online local catalogue, as part of *Bibliographic Data*. It was considered important to provide the link from the very beginning so that the user could access other information for any doubt or clarification.

All metadata entries were organized in a dedicated spreadsheet. Here they were associated with the corresponding book, identified by the corresponding shelfmark, folder name and file name. The spreadsheet was particularly useful for three reasons. First, once the metadata in Excel were well defined, they could be easily ingested in the system using the “copy-paste” technique, thus preventing any type of typing error. Secondly, it was possible to plan the process, deciding the books’ order of publication and recording those that were already published. Thirdly, the publisher could estimate the time spent for the entire publication process, considering that the system takes from three to four seconds per image.

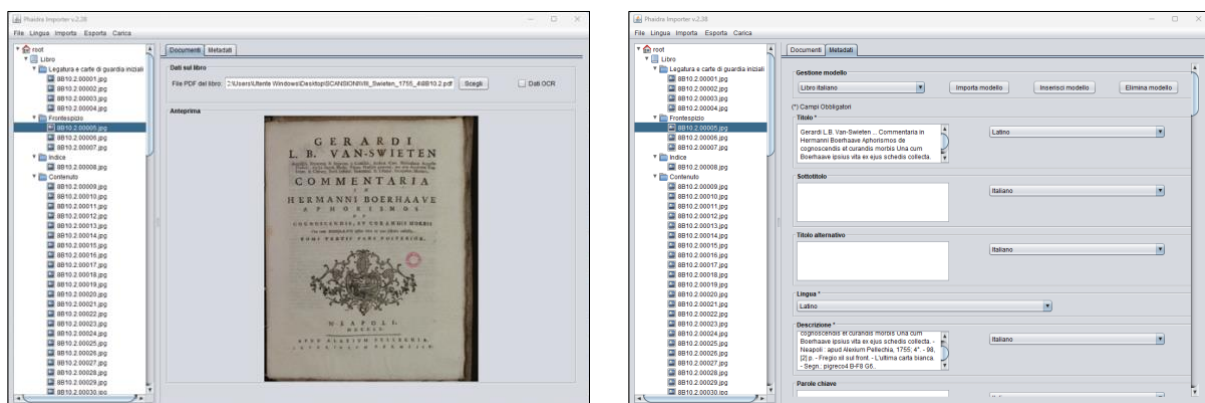
Afterwards, the project team inspected all archived folders created during digitization. It was important to check that folders, and especially those of derivative files that would be published, were complete, ordered, and without external elements. Phaidra required two typologies of files: JPG images and PDFs. The former ones were ready from the end of digitalization. The latter ones, instead, were subsequently created through the use of the book loading program, called “Phaidra Importer”. This app can be downloaded and installed on the publisher’s computer only after authentication. It is a Java app needed to upload books, collections and videos in Phaidra.

---

2022 it was updated with a new, expanded edition. The standard consists of nine descriptive areas, which are: *Content form and media type area*, *Title and statement of responsibility area*, *Edition area*, *Material or type of resource specific area*, *Publication, production, distribution area*, *Material description area*, *Series area*, *Notes area*, *Resource identifier and terms of availability area*.

The program gives the opportunity to upload digital objects to one of two destinations (servers): Phaidra Production, where objects are accessible to everyone, and Phaidra Test, a server not indexed by search engines where tests can be performed. A third server is reserved for Phaidra’s developers. In particular, Phaidra Test was used by the Ateneo Veneto to test the Phaidra platform and its functions, loading some test books and evaluating the results.

As regards the publication process, it is carried out through this app, Phaidra Importer, and then completed online using Phaidra’s Metadata Editor. The two main operations performed in the program are the creation of the object’s internal structure and the assignment of metadata. Both operations are visually exemplified by *Figure 42* and *Figure 43*. The examples are taken from the publication of Herman Boerhaave’s *Praelectiones academicae, tomus II*.<sup>215</sup>



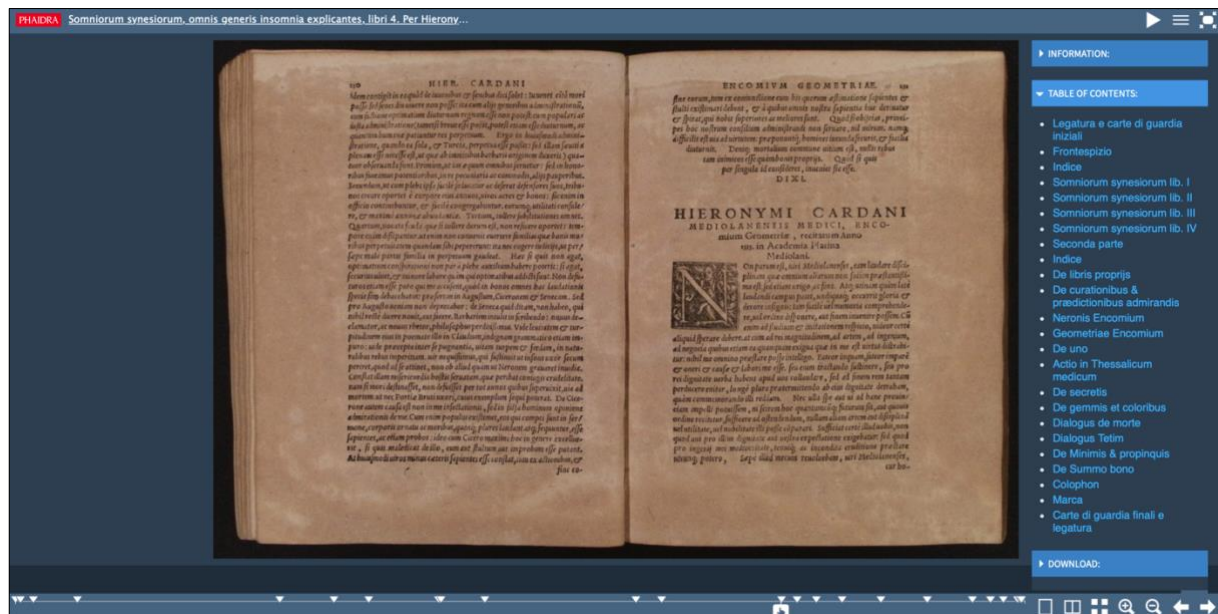
**Figure 42** (on the left) Phaidra Importer: creating the structure of Boerhaave’s *Praelectiones academicae, tomus II*. © PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023)

**Figure 43** (on the right) Phaidra Importer: entering metadata of Boerhaave’s *Praelectiones academicae, tomus II*. © PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023)

The book’s internal structure is a structural type of metadata and is visualized in the BookViewer’s scroll bar at the bottom, which allows the user to directly reach internal subdivisions of interest. Phaidra Importer enables the publisher to create this structure by dividing internal files into folders, each of which will be an internal subdivision of the object. In the case of the Ateneo Veneto’s early-printed books, it was decided to indicate the following sections: initial binding and flyleaves; frontispiece; content; table of contents and index (if any); colophon (if any of particular interest); final flyleaves and binding; book fore-edge and spine (if digitized). During this process, attention must be paid to the files’ order, otherwise the book, once loaded, will show the pages in erroneous order. In some cases, the book’s internal structure can be so complicated that several sections need to be created. It is the case of Girolamo Cardano’s *Somniorum synesiorum, omnis generis insomnia*

<sup>215</sup> Cfr. [Footnote 97](#).

*explicantes*..., which served in the previous chapter as an example to explain the non-rare presence of handwritten notes in these books.<sup>216</sup> In this book, the publication of the main work indicated by the title was accompanied by a list of minor works, such as *De libris proprijs*, *De curationibus & praedictionibus admirandis*, *Neronis Encomium*, etc. It was therefore essential to highlight the presence of these minor works during the creation of the internal structure. In the end, 19 additional sections were created. The complexity of the structure is visible in the *Figure 44* both in the scroll bar underneath the open book and in the table of contents on the side.



**Figure 44** An example of complex internal structure: Cardano’s *Somniorum synesiorum, omnis generis insomnia explicantes*. ©PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023)

In Phaidra Importer’s Metadata panel, it is possible to enter those categories of data that were previously described. In this phase, the creation of one or more metadata models could be a good solution to simplify this operation. A metadata model is a precompiled form that contains those entries that are repetitive among digital objects. In the case of the Ateneo Veneto’s project, two models were created, different only for the language field: Italian and Latin. The pre-compiled fields are: role (author), licence (BY-NC-SA), and the Dewey classification (“600: Technology → 610: Medicine & health”). Those models can be imported from time to time, accelerating the insertion time. The remaining fields are completed as indicated. The app does not offer a thesaurus from which to draw the standardized data of indicated persons, such as authors. For this reason, data was entered in the same form present in OPAC SBN.

<sup>216</sup> Cfr. [Footnote 161](#).

During the ingestion process, it is important to pay attention to two factors. Once published, books cannot be cancelled by the institution, while instead it will be necessary to request the cancellation to Phaidra's technicians. As a result, before giving confirmation to the upload, the publisher should check the correctness of the internal structure and metadata. All types of metadata, with the exception of the licence, can also be modified successively, while the structure will be unchanged. Nevertheless, the publisher has to remember that Phaidra's ingestion system considers as a digital object not only the book in its entirety but also all the pages that compose it. For this reason, metadata entered in Phaidra Importer is also assigned to each image and, unlike that of the book, it is regrettably not editable later using the Metadata Editor. For this reason, it is always better to avoid inserting temporary or partial metadata.

Before uploading, the publisher will need to import the PDF of the book (it will be the file that the user can download from Phaidra). If not already available, it may be created directly using Phaidra Importer. It is finally good practice to have the XML files of the internal structure and the metadata entered. These files should be saved together with the folders of the master and derivative files, considering that they constitute a trace of the process. Depending on the image size and connection speed, the process will take around three to four seconds per image.

Once published, the book is accessible to anyone online. A final check and correction phase is expected. In the BookViewer, the publisher will control the visual rendering of the book and its internal structure. Temporary metadata will then have to be replaced and missing metadata inserted.

The screenshot displays the Phaidra Digital Collections Metadata Editor interface. The main content area shows the book title "Praellectiones academicae in proprias institutiones rei medicae. 2: Arteria, cor, pulmo, sanguis, glandula, cerebrum" and the author "Boerhaave, Herman (Author)". The description includes details about the edition: "2: Arteria, cor, pulmo, sanguis, glandula, cerebrum. Hermann Boerhaave ... Praelectiones academicae in proprias Institutiones rei medicae. Edidit, et notas addidit Albertus Haller [...] Tomus primus [- sextus]. - Editio prima veneta post secundam Gottingae. - Venetiis : apud Simonem Oochi sub signo Italiae, 1743. - [4], 360 p.; 4\* - Segn.: [pl greco]2 A-2Y4. - Var. B: [2], 360 p. - Segn.: r1 A-2Y4. - Var. C: p. 13 numerata 1...". The interface also shows sections for "IDENTIFIERS" (Permalink: https://phaidra.cab.unipd.it/o.529205, Handle: 11108/11.529205), "THIS OBJECT REFERS TO THE FOLLOWING OBJECTS (1)", and "FOLLOWING OBJECTS ARE REFERENCING THIS OBJECT (1)". The "METADATA" section lists "Dublin Core" and "Phaidra metadata". The "Rights" section indicates the work is licensed under a Creative Commons CC BY-NC-SA 4.0 License. The "Contact institution or person" field is "Biblioteca dell'Ateneo Veneto". A "Show full metadata" button is visible at the bottom.

**Figure 45** Updating Boerhaave's *Praellectiones academicae*, tomus II, in Phaidra's Metadata Editor. ©PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023)

The publisher will use the Phaidra's Metadata Editor, the one that is shown in *Figure 45*. Specifically, Phaidra Importer does not accept special characters, so it will be necessary to insert them only at a later time. Additionally, missing metadata, such as the association with the hist institution, will be provided.

Finally, Phaidra offers the possibility to create cross-references between digital objects that have aspects in common, regardless of whether the objects are kept by the same institution. It is a tool that allows the creation of interesting thematic or multidisciplinary links between objects belonging to different institutions, increasing the fruition and enhancement of cultural items within the platform. For the moment, the Ateneo Veneto decided to establish cross-references between digital objects whose corresponding analogue ones are part of the same printing edition. For instance, the seven tomes of the already mentioned work *Praelectiones academicae* by Herman Boerhaave were interconnected. *Figure 46* shows the result of this operation.

**Figure 46** Boerhaave's *Praelectiones academicae*, tomus II, in Phaidra. On the right, under permanent identifiers, purple cross-references indicate the previous and following tome of the same printing edition. ©PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023)

## 4.4 Some Considerations

The publication took 10% of the time of the entire project: more than 70 hours of work and 104 published books, for a total number of 31,853 images uploaded in Phaidra. In contrast to the digitization process, this phase seems to be less influenced by the assets that the institution has.



While economic and technological parameters have a direct influence on the scanning process, the quality of image acquisition and the treatment of cultural items, these factors are less evident in the case of the publication process, where the images produced are published as they are inside the platform. The same choice of platform will be dictated by the type of digital object and the needs of the institution, rather than the economic possibilities. In this sense, the gap between small and big cultural institutions is less noticeable in the publication phase. On the contrary, what can represent a dividing line between digitization projects performed by small and big cultural institutions is the time factor. It is likely that large digitization projects can enjoy a more consistent project team than small organizations, such as the Ateneo Veneto. With equal time, the former will be able to program and implement a more complete and varied publication of digital objects, while the latter will provide a more basic one. As a result, small cultural institutions will have to adapt their action according to the potential time limit, making cuts also on publication. This is the reason why the Ateneo Veneto allocated only 10% of the overall time spent on the project on publication and now plans to extend it with some additional steps. Some may be performed shortly, while others in the foreseeable future. Firstly, the metadata set is going to be expanded to include bibliographic data and data relating to the book's lifecycle. Furthermore, the digital collection will be realized in Phaidra. Books may be divided into those sections that were described in subchapter 2.2 "Selection of Documents". An initial descriptive cover page will present the collection and the digitization project. Cross-references will provide relationships between digital objects of the same typology and extend the scope of dissemination.

Outside the Phaidra environment, a series of operations are also planned in order to give accessibility to digital objects. Specifically, the institute will request to admit the links of its digital resources archived in Phaidra within the national (OPAC SBN) and local (Polo-VEA SBN) catalogues. Users who search for one of the digitized items within these catalogues will then find links to Phaidra's platform where they can enjoy digital resources. This operation has the advantage of increasing the accessibility of digital objects through national and local catalogues but also of promoting the name and image of the institute and the digitization project itself.

Some other steps towards the extension of the actual online publication will be taken in the future. They will take advantage of Phaidra's resources for content diffusion and enhancement towards other platforms, first of all Internet Archive.<sup>217</sup> This digital library gives the possibility to broaden

---

<sup>217</sup> Internet Archive, *Homepage* (n.d.). The Internet Archive is an American digital library founded in 1996 to archive and preserve the contents that were published on the Internet. The Wayback Machine provides access to more than 26 years of web history. The platform archives other cultural items as well and it is also an activist organization, advocating a free and open Internet. According to the numbers reported on its website, today the Internet Archive

the fruition of digital objects by the creation of searchable PDFs, full-text transcriptions or EPUB files. It will therefore be necessary to request the technicians of the University of Padua the use of this function. The dialogue between Phaidra and Internet Archive is possible through APIs that convey digital collections and related metadata in Internet Archive while gaining the forementioned files.<sup>218</sup> Furthermore, it would be interesting to benefit from the collaboration that Phaidra has with MOVIO, the ICCU open-source software for creating virtual exhibitions. As already explained, the University of Padova's SBA developed a plug-in to allow the insertion of Phaidra's objects and collections in MOVIO's virtual exhibitions through a simple mask. The creation of virtual exhibitions would be the chance to further promote the Ateneo Veneto's digital collection and foster its dissemination.

Publication, unlike digitalization, is a process that can never be definitively closed. The technology change, the availability of new functionalities in the archiving platform and new bridges to other platforms determine possible future interventions aimed at a better dissemination of digital objects. This is true both for small cultural institutes and for the big ones. The choice of a good repository will also depend on this aspect. The repository has to adequately fulfil the role of a long-term platform where digital objects can be securely archived. It must be jointly designed as an open and interconnected digital environment. Interoperability and the use of metadata models based on standards are the two keywords for the widespread dissemination of cultural items. It will ensure that the platform can evolve and connect with other DAMSs in the future. In this sense, it is significant what Phaidra's technicians are doing for improving the resource model and vocabulary of the platform according to the RDF model and the schema.org vocabulary. These operations will further facilitate the dialogue and interconnection with other repositories on the Internet.

---

contains 735 billion web pages, 41 million books and texts, 14.7 million audio recordings (including 240,000 live concerts), 8.4 million videos (including 2.4 million Television News programs, 4.4 million images, and 890,000 software programs.

<sup>218</sup> Andreoli, L., Bianchi, E., Cappellato, L. et al. (2019), p. 155.

## CONCLUSION

This dissertation has set out to understand whether it is possible for small cultural institutes to plan digitization projects and the type of strategies they can develop. It examined the requisite passages of digital project planning, investigating the pivotal decisions that these institutes must undertake in order to transcend the limitations that structurally characterize them. The discussion was stimulated by the project and the related practices of digitization and publication planned by the Ateneo Veneto in Venice. This specific case served as test case of what small-sized organizations can effectively program and carry out. Whilst the Ateneo Veneto's experience represented the sample project of this dissertation, it could be considered at the same time its main limitation. In fact, taking into consideration several technological possibilities and a variety of institutes and practices, it would have been possible to obtain a wider range of results.

Overall, the research reveals that limitations do not vanish with the institution's planning of the project; they are persistent before, during and after the planning. Restricted financial resources and logistical limits influence the decision of outsourcing the project or performing it in-house. They also determine the careful choice and purchase of a series of hardware and software components. In most the cases, the budget will not be as adequate as it should to cover the costs of the cutting-edge equipment and the institute will arrive at a compromise between cost and quality. This has evident consequences in the execution of the digitization, where technical operators will be required to compensate for some missing functionalities offered by automatic or semi-automatic planetary scanners with a greater manual and time-consuming effort, demanded in the most exquisitely scanning phase and in post-capture image processing as well. The feature of limitation that characterizes small cultural institutes inevitably echoes in their instrumentation, their processes and the quality of the outputs. The final digital objects will reflect the technical specifications of the digital image capture equipment. Derivative files that possess unsatisfactory technical parameters, such as image resolution, pixel density, color depth, and file compression, would be unlikely perfectible afterwards.

Strategies formulated by small cultural institutes indicate that the most important variable in planning a digitization project lies in the operators' capabilities rather than in the available budget. It would be undoubtedly advantageous for such institutions to have at their disposal a large sum of money to be able to buy advanced equipment and easily achieve results above the standards. If money is the starting point, what makes the difference between a mediocre or successful

digitization project is budgeting, together with other crucial decisions. The institution will decide the composition of the project team, which will then be responsible for the budget management, the choice of the long-term hosting platform, the optimization of the equipment and all the operation of digitization, publication and metadata management. A good planning is worth more than a generous sum of money and represents the trump card especially for small cultural institutions, which must surmount a labyrinth of structural impediments of economic, logistical and HR dimensions. In this case, a compromise must be accepted. Most importantly, best practices will be able to conciliate feasibility with quality during the planning phase, purchasing and arranging an instrumentation system that reach the minimum digitization standards proposed nationally and internationally – with reference to FADGI's guidelines<sup>219</sup> and the National *Guidelines for the digitalisation of cultural heritage*<sup>220</sup>. The choice seems even more difficult since some technical parameters may be sacrificed in favour of others and commensurate with the project objectives, which greatly vary from those of a big cultural institution. In small-sized projects the optimization of the digitization workstation can resemble a real prototyping, meant to increase hardware efficiency and performance by testing and arranging singular components. During this phase, operators' resourcefulness is of great importance and can produce considerable improvements in the final quality of the digital objects. Moreover, it can be extremely beneficial for the preservation of cultural items during the execution of the operations, as the features of the workstation design and the digitization process can irreparably damage the originals.

The small cultural institutes that will take greater advantage from a digital project are the ones that paradoxically do not think only about the project but instead consider the digital as part of their cultural strategy. The adoption of a digital strategy allows the institute to achieve greater objectives and seize more opportunities thanks to a long-term logic that takes into consideration several projects. From a planning point of view, the institution will be still required to complete its project before the deadline but already owning an existing instrumentation and competences acquired in the previous project will save time. Moreover, the choice of the archival platform will be done in view of the construction of a digital collection that will be enlarged, preserved and enhanced over time. Some additional publication steps can be executed hereto inside the platform: digital objects' metadata can be extended, connections with other interoperable platforms can be created and improved and new presentation tools can be used. At the same time, the adoption of a digital strategy requires the project team to accomplish the objectives of completeness and transparency,

---

<sup>219</sup> Federal Agencies Digitization Guidelines Initiative (FADGI) (2023).

<sup>220</sup> Docs Italia, *Linee guida per la digitalizzazione del patrimonio culturale. Versione 1.0* (2022).

by writing internal reports and guidelines so that practices and methodologies are codified and “decrypted”. Data and documents should also be organized for the sake of future accessibility.

The creation of a digital strategy will give advantages to small cultural institutes which go beyond direct benefits of digitization practices – more accessibility, monitoring and inspection of the originals, the promotion of conservative and restorative interventions and advancement in the physical and online cataloguing. A long-term vision will encourage the organization to plan the digitization project in-house, investing in higher quality acquisition equipment and sharpen practical and specialist skills. Secondly, a digital strategy will enable the institution to stay up-to-date with the digital transformation and perfect digitization practices which by their nature are in evolution. It is not excluded that the conception of such a long-term, potentially ambitious strategy could serve as a strong point during the search for sources of funding, just because some financing institutes could consider it forward-looking and more reliable in comparison to single, self-referential digitization projects. The digital world can represent a driving force for a wider visibility and new forms of engagement with the public. The most ambitious aspiration of a digital strategy will be to gradually help small institutions to reinvent themselves. The assumptions are not lacking, since the cultural significance of these institutions can be used to fill gaps in the digital world, where too often contents are homologated, inaccurate and not certified. Those organizations will be asked to explore the prospect possibilities in the context in which they are inserted and use the digital as a means for its institutional activities. In this sense, no dichotomy exists between the physical and the digital.

At the same time, digitization should not be conceived as a miraculous solution for all the problems of cultural institutions. Other interventions are necessary on several fronts to reach a regenerative tide. It will be important to evaluate common operational strategies with other small cultural institutions, so that individual forces can assemble and achieve more visible results. Concerning cultural operators, it will be essential to invest in processes of upskilling and reskilling, while from a national point of view the legislator should rethink the regulations about internships in order to encourage greater participation of external figures in institutional activities. More generally, the struggle of small institutions for a digital transformation should be accompanied by a greater attention of the regional, national and European leadership towards them. The objectives to build a national and European digital cultural heritage – that is not just limited to replicate the cultural artefacts but also includes “stories and experiences” and the “cultural and socio-historical context”<sup>221</sup> – require time, dedication and widespread participation by all cultural institutions, small

---

<sup>221</sup> *Cfr.* [Footnote 29](#).

ones included. As a result, it seems more appropriate than ever that public funding should go to the benefit of organizations that through digitization can develop their potential. In this sense, it will be interesting to understand how the remaining funds of the NPD's sub-investment 1.1.5 – related to digitization campaigns, for a total amount of 57.4 million euros<sup>222</sup> – will be allocated. But first, the financing action should be anticipated by the systematic study on the phenomenon of small cultural institutions and the collection of specific statistical data, in order to encourage targeted policies towards them.

---

<sup>222</sup> *Cfr.* [Footnote 66](#).

# REFERENCES

## BIBLIOGRAPHY

- Agosti, M., Esposito, F., Meghini, C., et al. (2011). *Digital Libraries and Archives. 7th Italian Research Conference, IRCDL 2011, Pisa, Italy, January 20-21, 2011. Revised Papers*. Heidelberg (DE), Springer Berlin. Available at: <https://link.springer.com/book/10.1007/978-3-642-27302-5#toc>. [Accessed 14 July 2023]
- Albertin, F., Balliana, E., Pizzol, et al. (2018). *Printing materials and technologies in the 15th–17th century book production: An undervalued research field*, "Microchemical Journal", 138, pp. 147-153. Available at: <https://doi.org/10.1016/j.microc.2017.12.010>. [Accessed 28 July 2023]
- Andreoli, L., Bettella, C., Cappellato, L., et al. (2019). *University of Vienna Metadata UWmetadata*. Padova, Università di Padova. Available at: <https://phaidra.cab.unipd.it/static/phaidra-uwmetadata.pdf>. [Accessed 2 July 2023]
- Andreoli, L., Bianchi, E., Cappellato, L. et al. (2019). *Phaidra, un archivio digitale FAIR per la disseminazione e l'accesso integrato a testi, testimonianze, immagini e storie del patrimonio culturale*, "DigItalia. Rivista del digitale nei beni culturali", 14, no. 1, pp. 147–157. Available at: <https://digitalia.cultura.gov.it/article/view/2281>. [Accessed 13 June 2023]
- Andreoli, L., Drago, G. (2022). *Guidelines on digitisation of two-dimensional documents*. Padova, Università di Padova. Available at: <https://phaidra.cab.unipd.it/static/linee-guida-digitalizzazione-EN.pdf>. [Accessed 28 June 2023]
- Andreoli, L., Visentin, M. (2019). *Guida a Phaidra Importer*. Padova, Università di Padova. Available at: <https://phaidra.cab.unipd.it/static/guida-phaidra-importer.pdf>. [Accessed 11 July 2023]
- Andreoli, A., ed. (2022). *Beni privati e interessi pubblici: storia e attualità delle biblioteche accademiche italiane*, Atti del Convegno di Rovereto 2019. Verona, Scripta edizioni.
- ATHENAWP3, ed. (2009). *Digitisation Standard Landscape for European Museums, Archives, Libraries*. ATHENA. Available at: <https://phaidra.cab.unipd.it/o:6785>. [Accessed 11 July 2023]
- Baca, M., ed. (2008). *Introduction to metadata*. Los Angeles (USA), The Getty Research Institute. Available at: [https://www.getty.edu/research/publications/electronic\\_publications/intrometadata/index.html](https://www.getty.edu/research/publications/electronic_publications/intrometadata/index.html). [Accessed 14 July 2023]
- Barbuti, N. (2020). *Ripensare i dati come risorse digitali: un percorso difficile?*, in "La svolta inevitabile: sfide e prospettive per l'informatica umanistica", Marras, C., Passarotti, M., Franzini, G. et al. ed., pp. 19-23. Milano, Università Cattolica del Sacro Cuore. Available at: <https://ricerca.uniba.it/handle/11586/303938.4#>. [Accessed 25 July 2023]

- Barbuti, N. (2022). *La digitalizzazione dei beni documentali. Metodi, tecniche, buone prassi*. Milano, Editrice Bibliografica.
- Bergamin G., Messina M. (2010). *Magazzini digitali- dal prototipo al servizio*. “DigItalia. Rivista del digitale nei beni culturali”, 5, no. 2, 144–153. Available at: <https://digitalia.cultura.gov.it/article/view/246/157>. [Accessed 14 July 2023]
- Besser, H., Hubbard, S (2003). *Introduction to Imaging. Revised Edition*. Los Angeles (USA), The Getty Research Institute. Available at: [https://www.getty.edu/research/publications/electronic\\_publications/introimages/intro.html](https://www.getty.edu/research/publications/electronic_publications/introimages/intro.html). [Accessed 12 July 2023]
- Bettella, C. (2018). *PHAIDRA\_DC Metadata Element Set*. Padova, Università di Padova. Available at: [https://phaidra.cab.unipd.it/static/phaidra\\_dc-metadata-element-set.pdf](https://phaidra.cab.unipd.it/static/phaidra_dc-metadata-element-set.pdf). [Accessed 11 July 2023]
- Bettella, C., Carrer, Y., Turetta, G. (2022). *La valutazione FAIRness di un archivio digitale certificato: tra principi teorici e azioni pratiche*. “DigItalia. Rivista del digitale nei beni culturali”, 17, no. 1, pp. 113–133. Available at: <https://doi.org/10.36181/digitalia-00043>. [Accessed 13 July 2023]
- Bloomberg, J. (2018). *Digitization, Digitalization, And Digital Transformation: Confuse Them At Your Peril*. “Forbes”, April 29. Available at: <https://www.forbes.com/sites/jasonbloomberg/2018/04/29/digitization-digitalization-and-digital-transformation-confuse-them-at-your-peril>. [Accessed 8 July 2023]
- Calibrate, *ColorChecker Passport Photo 2* (17 July 2023). Available at: <https://calibrite.com/it/product/colorchecker-passport-photo-2>. [Accessed 21 July 2023]
- Cappellato, L. (2023). *Guida all'archiviazione*. Padova, Università di Padova. Available at: <https://phaidra.cab.unipd.it/static/guida-all-archiviazione.pdf>. [Accessed 13 July 2023]
- Cecotti, H. (2022). *Cultural Heritage in Fully Immersive Virtual Reality*. “Virtual Worlds”, 1, no. 1, pp. 82–102. Basel (CH), MDPI. Available at: <http://dx.doi.org/10.3390/virtualworlds1010006>. [Accessed 12 September 2023]
- Cerullo, L., Negri, A. (2023). *L'infrastruttura software per il patrimonio culturale (ISPC) come abilitatore di un Ecosistema digitale nazionale del patrimonio culturale*. “DigItalia. Rivista del digitale nei beni culturali”, 18, no. 1, pp. 38–50. Available at: <https://doi.org/10.36181/digitalia-00059>. [Accessed 4 September 2023]
- Cohen, D.J., Rosenzweig, R. (2005). *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web*. Philadelphia, University of Pennsylvania Press. Available at: <https://chnm.gmu.edu/digitalhistory>. [Accessed 11 July 2023]
- CoreTrustSeal Standards and Certification Board (2022). *CoreTrustSeal Requirements 2023-2025 (V01.00)*. Zenodo. Available at: <https://doi.org/10.5281/zenodo.7051012>. [Accessed 13 July 2023]



- Denison, T. (2007). *Chapter 10 - Library and information systems: a work in progress*, in Ferguson, S., ed., "Topics in Australasian Library and Information Studies. Libraries in the Twenty-First Century". Chandos Publishing, pp. 165-177. Available at: <https://doi.org/10.1016/B978-1-876938-43-7.50010-7>. [Accessed 13 July 2023]
- Drago, G., ed. (2019). *Formati dei file raccomandati per l'archiviazione a lungo termine e per la disseminazione web in Phaidra*. Padova, Università di Padova. Available at: <https://phaidra.cab.unipd.it/static/formati-file-compresso.pdf>. [Accessed 16 July 2023]
- European Commission, Directorate-General for Research and Innovation, Brunet, P., De Luca, L., Hyvönen, E., et al. (2022). *Report on a European collaborative cloud for cultural heritage: ex – ante impact assessment*. Brussels, Publications Office of the European Union. Available at: <https://data.europa.eu/doi/10.2777/64014>. [Accessed 14 September 2023]
- European Union (2014). *Council conclusions of 21 May 2014 on cultural heritage as a strategic resource for a sustainable Europe*. In "Official Journal of the European Union", C183, no.8. Available at: [https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52014XG0614\(08\)](https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52014XG0614(08)). [Accessed 12 September 2023]
- Fantin, M., Notolini, G., ed.; Capovilla J., rev. (2018). *Vademecum digitalizzazione*. Padova, Università di Padova.
- Federal Agencies Digitization Guidelines Initiative (FADGI) (2023). *Technical Guidelines for Digitizing Cultural Heritage Materials. Third Edition*. Available at: <https://www.digitizationguidelines.gov/guidelines/digitize-technical.html>. [Accessed 11 July 2023]
- Fondazione Symbola, Unioncamere (2023). *Io sono Cultura. Rapporto 2023. L'Italia della qualità e della bellezza sfida la crisi*. Roma, Symbola. Available at: <https://www.symbola.net/ricerca/io-sono-cultura-2023/>. [Accessed 4 September 2023]
- Gandin, K. (2021). *Cambiamenti nella comunicazione social dei musei nel corso della pandemia Covid-19. Il caso della Scuola Grande di San Rocco* [Bachelor's Thesis]. Venezia, Università Ca' Foscari di Venezia, supervisor: Erri, A.
- Gonzales, R.C., Woods, R.E. (2018). *Digital Image Processing*. Harlow (GB), Pearson Global Edition.
- Gruppo di lavoro sulla digitalizzazione del materiale cartografico, a cura di (2006). *Linee guida per la digitalizzazione del materiale cartografico*. Istituto Centrale per il Catalogo Unico delle Biblioteche Italiane. Available at: <https://www.internetculturale.it/getFile.php?id=44395>. [Accessed 14 July 2023]
- Gruppo di lavoro sulla digitalizzazione del materiale fotografico Gennaio 2004, a cura di (2004). *Linee di indirizzo per i progetti di digitalizzazione del materiale fotografico*. Roma, Istituto Centrale per il Catalogo Unico Available at: <https://www.internetculturale.it/getFile.php?id=44396>. [Accessed 2 July 2023]

- International Federation of Library Associations and Institutions (IFLA) (2004). *IFLA Principles for the Care and Handling of Library Material*. Bari, Biblioteca Nazionale di Bari. Available at: <https://repository.ifla.org/handle/123456789/1269>. [Accessed 2 July 2023]
- International Federation of Library Associations and Institutions (IFLA) (2015). *Guidelines for Planning the Digitization of Rare Book and Manuscript Collections*. Den Haag (NL), IFLA. Available at: <https://repository.ifla.org/handle/123456789/458>. [Accessed 2 July 2023]
- Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche (ICCU), *Linee guida e standard*. Available at: <https://www.internetculturale.it/it/1131/linee-guida-e-standard>. [Accessed 23 July 2023]
- Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche (ICCU), *Linee guida per la digitalizzazione e metadata*. Available at: <https://www.iccu.sbn.it/it/normative-standard/linee-guida-per-la-digitalizzazione-e-metadata/>. [Accessed 23 July 2023]
- Kama, A.F.L. de F., Manini, M.P., Baptista, D.M. (2016). *Analysis of criteria and requirements to access in Rare Books digital libraries – a longitudinal study*. Em Questão. Available at: <https://doi.org/10.19132/1808-5245223.227-244>. [Accessed 14 July 2023]
- Kremers, H. ed. (2020). *Digital cultural heritage*. Cham (CH), Springer.
- Lavoie, B. (2014). *The Open Archival Information System (OAIS) Reference Model: Introductory Guide (2nd Edition)*. The Digital Preservation Coalition. Available at: <http://dx.doi.org/10.7207/twr14-02> [Accessed 13 July 2023]
- Lotti F., Trumpy, G., CNR IFAC (Istituto di Fisica Applicata "Nello Carrara"), Fondazione Rinascimento Digitale (2007). *La digitalizzazione documentaria*. Available at: <https://phaidra.cab.unipd.it/o:469492>. [Accessed 17 July 2023]
- Minerva Working Group 6, ed. (2004). *Good Practices Handbook*. Available at: <https://www.minervaeurope.org>. [Accessed 11 July 2023]
- Ministero della cultura (2022). *Decreto Ministeriale n. 298 26/07/2022 “Assegnazione delle risorse alle Regioni e alle Province Autonome per la Missione 1 - Digitalizzazione, Innovazione, Competitività e Cultura, Componente 3 - Cultura 4.0 (M1C3), Misura 1 Digitalizzazione del patrimonio culturale (PNRR)”*. Available at: [https://pnrr.cultura.gov.it/wp-content/uploads/2022/08/Assegnazione-Risorse\\_DM-298\\_26.07.22.pdf](https://pnrr.cultura.gov.it/wp-content/uploads/2022/08/Assegnazione-Risorse_DM-298_26.07.22.pdf). [Accessed 15 September 2023]
- Ministero della cultura, Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library, ed. (2022). *Piano nazionale di digitalizzazione del patrimonio culturale 2022 – 2023* (Version 1.1). Available at: <https://digitallibrary.cultura.gov.it/il-piano>. [Accessed 11 July 2023]
- Montanaro, D. (2017). *Digital Ecosystems for Open Science - Il caso Università di Vienna, l'esempio di Phaidra* [Master's Thesis]. Venezia, Università Ca' Foscari di Venezia, supervisor: Raines D.; co-supervisor: Ridi, R. Available at: <http://hdl.handle.net/10579/11805>. [Accessed 14 July 2023]

- Pirrelli, M. (2022). *Corte dei Conti: la digitalizzazione della cultura ancora in ritardo*. “Il Sole 24 Ore”, October 25. Available at: <https://www.ilsole24ore.com/art/corte-conti-digitalizzazione-cultura-ancora-ritardo-AEfgcUBC>. [Accessed 6 July 2023]
- Raines, D., catalogo della mostra a cura di (2007). *Anatomia di una biblioteca. Cinquanta volumi di medicina dalla collezione storica dell'Ateneo Veneto*. Venezia, Ateneo Veneto.
- Sahle, P. (2016). 2. *What is a Scholarly Digital Edition?* In “Digital Scholarly Editing: Theories and Practices”, Driscoll, M.J., Pierazzo, E., eds. Cambridge (UK), Open Book Publishers. Available at: <http://dx.doi.org/10.11647/OBP.0095>. [Accessed 11 September 2023]
- Simionato, G. ed. (2007). *Cultura e tradizione accademica. Il ruolo degli atenei tra passato e future*, Atti Convegno di Studio degli Atenei di Venezia, Treviso, Brescia e Bergamo 2006. Treviso, Ateneo di Treviso.
- Sitts, M.K. (2000). *Handbook for Digital Projects: A Management Tool for Preservation and Access*. Andover (USA), Northeast Document Conservation Center. Available at: <https://www.nedcc.org/assets/media/documents/dman.pdf>. [Accessed 11 July 2023]
- Testa, S. (2015). *Italian Academies and their Networks, 1525-1700. From Local to Global*. New York (USA), Palgrave Macmillan. Available at: <https://doi.org/10.1007/978-1-137-43842-3>. [Accessed 11 September 2023]
- The Association for Library Collections and Technical Services, Preservation and Reformatting Section (2013). *Minimum Digitization Capture Recommendations*. Available at: <https://www.ala.org/alcts/resources/preserv/minimum-digitization-capture-recommendations>. [Accessed 4 July 2023]
- Trček, D. (2022). *Cultural heritage preservation by using blockchain technologies*. “Heritage Science” 10, no. 6. Available at: <https://doi.org/10.1186/s40494-021-00643-9>. [Accessed 12 September 2023]
- Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. (2016). *The FAIR Guiding Principles for scientific data management and stewardship*, “Scientific Data”, 3. Available at: <https://doi.org/10.1038/sdata.2016.18>. [Accessed 11 July 2023]
- Woods, A., Wilcox, D., Bernstein, D. (2019). *Fedora Project and Community Update. The 14th International Conference on Open Repositories*. Hamburg (DE), Zenodo. Available at: <https://doi.org/10.5281/zenodo.3554157>. [Accessed 14 July 2023]

## WEBSITES AND DIGITAL RESOURCES

- ABAA, *Glossary of Book Terms* (n.d.). Available at: <https://www.abaa.org/glossary>. [Accessed 26 July 2023]
- Adobe, *What is an exposure in photography?* (n.d.). Accessible at: <https://www.adobe.com/uk/creativecloud/photography/discover/exposure-in-photography.html>. [Accessed 26 July 2023]

Agenzia per la Promozione della Ricerca Europea (APRE), *Horizon Europe* (2021). Accessible at: <https://horizoneurope.apre.it/he-in-breve>. [Accessed 8 September 2023]

Association for Computing Machinery, *The 2012 ACM Computing Classification System* (n.d.). Accessible at: <https://www.acm.org/publications/class-2012>. [Accessed 23 August 2023]

Ateneo Veneto, *Ateneo Veneto - il più antico istituto culturale attivo per tutti a Venezia* (27 February 2023). Available at: <https://ateneoveneto.org/ateneo-veneto-venezia>. [Accessed 9 July 2023]

Ateneo Veneto, *La Biblioteca*. (8 July 2023). Available at: <https://ateneoveneto.org/la-biblioteca>. [Accessed 10 July 2023]

Ateneo Veneto, *Statuto* (20 March 2023). Available at: <https://ateneoveneto.org/statuto>. [Accessed 9 July 2023]

Ateneo Veneto, *Storia* (21 April 2023). Available at: <https://ateneoveneto.org/la-storia-dellateneo-veneto>. [Accessed 7 July 2023]

Book2net, *Archive Pro A1* (12 June 2023). Available at: <https://book2net.net/it/produkte/a1-it,scanner-darchivio,scanner-per-musei,repssysteme-it/archive-pro-a1>. [Accessed 21 July 2023]

Commissione Tecnologie Digitali di ICOM Italia, *Glossario Museologia digitale e nuove tecnologie* (2020). Available at: <https://www.icom-italia.org/wp-content/uploads/2020/10/ICOMItalia.CommissioneTecnologie.2020.GlossarioTecnologie.pdf>. [Accessed 4 September 2023]

CoreTrustSeal, *Assessment Information. Phaidra at the Library System of the University of Padova* (2019). Available at: <https://www.coretrustseal.org/wp-content/uploads/2019/11/Phaidra-at-the-Library-System-of-the-University-of-Padova.pdf>. [Accessed 14 July 2023]

Creative Commons, *CC Licences* (16 August 2023). Available at: <https://creativecommons.org/about/cclicenses>. [Accessed 23 August 2023]

CZUR, *CZUR ET Smart Book Scanner* (2023). Available at: <https://www.czur.com/product/ETscanner>. [Accessed 19 July 2023]

DCMI, *Dublin Core™ User Guide* (n.d.). Available at: <https://www.dublincore.org/resources/userguide>. [Accessed 22 August 2023]

Docs Italia, *Linee guida per la digitalizzazione del patrimonio culturale. Versione 1.0* (2022). Available at: <https://docs.italia.it/italia/icdp/icdp-pnd-digitalizzazione-docs/it/v1.0-giugno-2022>. [Accessed 9 July 2023]

Enciclopedia Treccani (2023). Available at: <https://www.treccani.it>. [Accessed 12 July 2023]

European Commission, *Curation of digital assets and advanced digitisation* (2020). Available at: <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/dt-transformations-12-2018-2020>. [Accessed 12 September 2023]

- European Commission, Italian Government, *National Recovery and Resilience Plan* (2021). Available at: <https://www.italiadomani.gov.it/content/sogei-ng/it/en/home.html>. [Accessed 4 September 2023]
- European Commission, *Italy in the Digital Economy and Society Index* (2022). Available at: <https://digital-strategy.ec.europa.eu/en/policies/desi-italy>. [Accessed 4 September 2023]
- European Union, *NextGenerationEU* (n.d). Available at: [https://next-generation-eu.europa.eu/index\\_en](https://next-generation-eu.europa.eu/index_en). [Accessed 7 September 2023]
- Fedora, *Home* (2023). Available at: <https://fedora.lyrasis.org/>. [Accessed 21 August 2023]
- Fondazione di Venezia, *Bando "Cultura Digitale"* (11 April 2022). Available at: <https://www.fondazionedivenezia.org/archivio-attivita/bando-cultura-digitale>. [Accessed 14 July 2023]
- Fondazione Istituto Internazionale di Storia Economica F. Datini - Minerva Project, *The Minerva Project* (07 January 2014). Available at: <https://www.istitutedatini.it/biblio/eng/minerva.htm>. [Accessed 6 July 2023]
- FTA Online News (2020). *Industria 4.0: la Quarta Rivoluzione Industriale*. "Borsa Italiana", June 24. Available at: <https://www.borsaitaliana.it/notizie/sotto-la-lente/rivoluzione-252.htm#:~:text=Finoad%20oggi%20le%20rivoluzioni,nascita%20della%20macchina%20a%20vapore>. [Accessed 6 September 2023]
- FUJITSU, *ScanSnap SV600* (2023). Available at: <https://www.fujitsu.com/uk/products/computing/peripheral/scanners/soho/sv600>. [Accessed 19 July 2023]
- Gartner, *Gartner Glossary* (2023). Available at: <https://www.gartner.com/en/glossary>. [Accessed 1 September 2023]
- GO FAIR, *FAIR Principles* (21 January 2022). Available at: <https://www.go-fair.org/fair-principles>. [Accessed 13 July 2023]
- IBM Blog, *What Is Optical Character Recognition (OCR)?* (5 January 2022). Available at: <https://www.ibm.com/blog/optical-character-recognition>. [Accessed 13 August 2023]
- IIPImage, *IIPImage* (2023). Available at: <https://iipimage.sourceforge.io>. [Accessed 21 August 2023]
- International Federation of Library Associations and Institutions (IFLA), *International Standard Bibliographic Description (ISBD)* (n.d.). Available at: <https://www.ifla.org/references/best-practice-for-national-bibliographic-agencies-in-a-digital-age/resource-description-and-standards/bibliographic-control/international-standard-bibliographic-description-isbd>. [Accessed 24 August 2023]

- International Federation of Library Associations and Institutions (IFLA), *UNIMARC formats and related documentation* (2023). Available at: <https://www.ifla.org/publications/unimarc-formats-and-related-documentation>. [Accessed 20 August 2023]
- Internet Archive, *Homepage* (n.d.). Available at: <https://archive.org>. [Accessed 25 August 2023]
- Istituto Centrale per gli Archivi – ICAR, *EAD* (27 July 2022). Available at: <https://icar.cultura.gov.it/standard/standard-internazionali/ead>. [Accessed 20 August 2023]
- Istituto Centrale per il Catalogo e la Documentazione – ICCD, *Standard catalogafici* (2023). Available at: <http://www.iccd.beniculturali.it/standard-catalogafici>. [Accessed 20 August 2023]
- Istituto Centrale per il Catalogo Unico, *Linee guida per il PND* (n.d.). Available at: <https://www.iccu.sbn.it/it/normative-standard/linee-guida-per-la-digitalizzazione-e-metadati/linee-guida-per-il-pnd/index.html>. [Accessed 14 July 2023]
- Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library, *Investimento 1.1* (2023). Available at: <https://digitallibrary.cultura.gov.it/investimento-1-1>. [Accessed 17 September 2023]
- Istituto Nazionale di Statistica (ISTAT), *L'accessibilità di musei e biblioteche* (2022). Available at: <https://www.istat.it/it/files//2022/12/accessibilita-luoghi-cultura-dic2022.pdf>. [Accessed 7 July 2023]
- Istituto Nazionale di Statistica (ISTAT), *Statistiche culturali – Anno 2021* (2022). Available at: <https://www.istat.it/it/archivio/279105>. [Accessed 6 July 2023]
- Lotti F., Lunghi, M., Trumpy, G. et al., *Digitalizzazione di beni artistici e documentari. Manuale di procedure per un laboratorio fotografico digitale* (2009). Available at: <https://phaidra.cab.unipd.it/o:469494>. [Accessed 13 July 2023]
- Neewer, *NEEWER 2 Packs Dimmable 5600K USB LED Video Light* (n.d.). Available at: <https://it.neewer.com/collections/led-panel-light/products/tabletop-led-light-66601363>. [Accessed 21 July 2023]
- OCLC, *Dewey Services* (n.d.). Available at: <https://www.oclc.org/en/dewey/resources.html>. [Accessed 23 August 2023]
- Online Etymology Dictionary (2023). Available at: <https://www.etymonline.com>. [Accessed 9 September 2023]
- Open Archives Initiative, *OAI-PMH* (2023). Available at: <https://www.openarchives.org/pmh>. [Accessed 20 August 2023]
- Osservatorio Tecnologico per i Beni e le Attività Culturali, *Schema di capitolato per attività di digitalizzazione* (2008). Available at: <http://www.otebac.it/index.php?it/127/capitolato-tecnico-digitalizzazione>. [Accessed 20 July 2023]

- PHAIDRA - University of Padua, *PHAIDRA – Digital Collections* (2023). Available at: <https://doi.org/10.25430/phaidra>. [Accessed 16 August 2023]
- Physical Review Journals, *Physics and Astronomy Classification Scheme (PACS)* (11 January 2008). Available at: <https://journals.aps.org/PACS>. [Accessed 23 August 2023]
- Publication Office of the European Union, *EuroVoc* (n.d.). Available at: <https://op.europa.eu/s/yXC3>. [Accessed 23 August 2023]
- Re3data.org, *REgistry of REsearch Data REpositories* (n.d.). Available at: <https://www.re3data.org>. [Accessed 13 July 2023]
- Regione del Veneto, *Digitalizzazione del patrimonio culturale - Avviso pubblico PNRR* (2023). Available at: <https://www.culturaveneto.it/it/la-tua-regione/digitalizzazione-del-patrimonio-culturale-avviso-pubblico-pnrr>. [Accessed 17 September 2023]
- Schema.org, *Homepage* (n.d.). Available at: <https://schema.org>. [Accessed 23 August 2023]
- University of Vienna, *Phaidra* (2023). Available at: <https://phaidra.univie.ac.at>. [Accessed 17 July 2023]
- Wikipedia. The Free Encyclopedia (2023). Available at: <https://en.wikipedia.org>. [Accessed 11 July 2023]

## ICONOGRAPHIC REFERENCES

- Lotti F., Lunghi, M., Trumpy, G. et al., *Digitalizzazione di beni artistici e documentari. Manuale di procedure per un laboratorio fotografico digitale* (2009). Available at: <https://phaidra.cab.unipd.it/o:469494>. [Accessed 13 July 2023]
- PHAIDRA - University of Padua, *PHAIDRA – Digital Collections* (2023). Available at: <https://doi.org/10.25430/phaidra>. [Accessed 21 August 2023]
- Ricoh, *ScanSnap® SV600* (n.d.). Available at: <https://www.pfu-us.ricoh.com/scanners/scansnap/sv600>. [Accessed 24 July 2023]
- Scannx, *CZUR ET25 Pro Overhead Scanner* (n.d.). Available at: <https://www.scannx.com/products/overhead-scanners/czur-et25-pro>. [Accessed 24 July 2023]





# LIST OF TABLES AND FIGURES

<b>Table 1</b> Comparison between FUJITSU ScanSnap SV600 and CZUR ET-25 Pro.....	65
<b>Figure 1</b> Some of the Ateneo Veneto's books digitized during the project. ©Ateneo Veneto.....	46
<b>Figure 2</b> (on the left) FUJITSU ScanSnap SV600 planetary scanner. ©Ricoh, ScanSnap® SV600 (n.d.).....	64
<b>Figure 3</b> (on the right) CZUR ET-25 Pro planetary scanner. ©Scannx, CZUR ET25 Pro Overhead Scanner (n.d.).....	64
<b>Figure 4</b> (on the left) The specialized finger cots: no recognition and no cut. ©Ateneo Veneto .....	68
<b>Figure 5</b> (on the right) Finger removal and the appearance of a halo. ©Ateneo Veneto .....	68
<b>Figure 6</b> The use of a piece of antiglare glass test over Harvey's book and the problem of the scanner's LED lights projection. ©Ateneo Veneto .....	69
<b>Figure 7</b> Layout of the shot and lighting arrangement with $45^\circ/45^\circ/0^\circ$ geometry. ©Author's elaboration on Lotti F., Lunghi, M., Trumpy, G. et al., Digitalizzazione di beni artistici e documentari. Manuale di procedure per un laboratorio fotografico digitale (2009) .....	71
<b>Figure 8</b> (on the left) The digitization workstation with Czur planetary scanner and LED spotlights. ©Ateneo Veneto .....	72
<b>Figure 9</b> (on the right) Scarabiccio's book over the workstation and ready for digitization. Note the glass platen and the black pad. ©Ateneo Veneto .....	72
<b>Figure 10</b> A portion of the spreadsheet with the required information for the digitization and nomenclature. ©Ateneo Veneto.....	75
<b>Figure 11</b> (on the left) An unopened book needs some cuts in correspondence of the folds of the original gatherings. ©Ateneo Veneto. 82	82
<b>Figure 12</b> (on the right) Opening and cutting of Ioannis Cratonis Vratislaviensis ad artem medicam isagoge. ©Ateneo Veneto .....	82
<b>Figure 13</b> (on the left) Book and Colorchecker placement on the digitization workstation. ©Ateneo Veneto.....	85
<b>Figure 14</b> (on the right) Measuring the distance between the optical axis and the spotlights in the workstation digitization. ©Ateneo Veneto .....	85
<b>Figure 15</b> Image calibration in Camera Raw. ©Ateneo Veneto .....	86
<b>Figure 16</b> A sample image of Scarabiccio's book without any calibration of white balance (on the left). LED spotlights are placed with $d=55\text{cm}$ , $h=145\text{cm}$ and $l=4$ . The corresponding histogram (on the right). ©Ateneo Veneto .....	87
<b>Figure 17</b> The same sample image of Scarabiccio's book with the calibration of white balance (on the left). The image looks overexposed. The corresponding histogram (on the right). ©Ateneo Veneto .....	87
<b>Figure 18</b> Another sample image of Scarabiccio's book with a spotlight arrangement of $d=78$ , $h=170$ and $l=3$ (on the left). In this case, the calibration of white balance gives the image the optimal exposure. The corresponding histogram (on the right). ©Ateneo Veneto .....	88
<b>Figure 19</b> (on the left) Placement of thick spongy pieces next to the book on the digitization workstation. ©Ateneo Veneto.....	89

<b>Figure 20</b> (on the right) Setting the double scanning area for Harvey's book. ©Ateneo Veneto.....	89
<b>Figure 21</b> (on the left) The initial chalcographic illustration in Panaroli's book. ©Ateneo Veneto .....	92
<b>Figure 22</b> (in the middle) 16 <sup>th</sup> -century marginal glosses, corrections and underling in Cardano's book. ©Ateneo Veneto .....	92
<b>Figure 23</b> (on the right) The frontispiece of Mylius' <i>Basilica Philosophica</i> . The whole book is affected by paper oxidation. ©Ateneo Veneto	92
<b>Figure 24</b> (on the left) Placement of thick spongy pieces under the protruding insert of a book. ©Ateneo Veneto.....	93
<b>Figure 25</b> (on the right) Setting the double scanning area in the presence of a book insert. ©Ateneo Veneto .....	93
<b>Figure 26</b> (on the left) Ficino's book frontispiece affected by a moisture lacuna. ©Ateneo Veneto.....	94
<b>Figure 27</b> (on the right) Water damage on the fore-edge of the post-incunable of Avicenna. ©Ateneo Veneto.....	94
<b>Figure 28</b> (on the left) Detection of a trace of dirt on the black pad while digitizing Chaumette's <i>Enchiridio cioe Manuale cirurgico</i> . ©Ateneo Veneto .....	98
<b>Figure 29</b> (on the right) Removal of the imperfection on the black pad using Camera Raw. ©Ateneo Veneto.....	98
<b>Figure 30</b> Realignment of left and right master page in the presence of the insert inside Oddo degli Oddi's book. ©Ateneo Veneto .	100
<b>Figure 31</b> The final result after the process of realignment in Oddo degli Oddi's book. ©Ateneo Veneto .....	100
<b>Figure 32</b> Dusting the pages of Acquapendente's <i>De formato foetu</i> . ©Ateneo Veneto.....	102
<b>Figure 33</b> (left and right) Arranging Acquapendente's book on the digitization workstation. ©Ateneo Veneto.....	103
<b>Figure 34</b> (on the left) Realignment (before) of the left and right pages of the 30 <sup>th</sup> , second part, table of Acquapendente's <i>De formato foetu</i> . Guidelines do not correspond. ©Ateneo Veneto .....	104
<b>Figure 35</b> (on the right) Realignment (after) of the left and right pages of the 30 <sup>th</sup> , second part, table of Acquapendente's <i>De formato foetu</i> . Guidelines now correspond. ©Ateneo Veneto .....	104
<b>Figure 36</b> Final result of the 30 <sup>th</sup> , second part, table of Acquapendente's <i>De formato foetu</i> after the realignment as it can be seen in Phaidra BookViewer. ©PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023).....	104
<b>Figure 37</b> The Phaidra digital environment. ©PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023)...	110
<b>Figure 38</b> Phaidra's homepage. ©PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023).....	112
<b>Figure 39</b> An example of digital collection in Phaidra: "The cultural heritage of the Medical School of Padova". On the left, the cover page of the collection. On the right, an internal section. ©PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023) .....	113
<b>Figure 40</b> An example of a digital object and metadata description in Phaidra: Garcia da Orta's <i>Dell'istoria de i semplici aromati...</i> digitized by the Ateneo Veneto. ©PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023).....	114
<b>Figure 41</b> Garcia da Orta's <i>Dell'istoria de i semplici aromati...</i> displayed in the Phaidra BookViewer. ©PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023).....	115
<b>Figure 42</b> (on the left) Phaidra Importer: creating the structure of Boerhaave's <i>Praelectiones academicae, tomus II</i> . © PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023).....	125

**Figure 43** (on the right) *Phaidra Importer: entering metadata of Boerhaave's Praelectiones academicae, tomus II.* © PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023)..... 125

**Figure 44** *An example of complex internal structure: Cardano's Somniorum synesiorum, omnis generis insomnia explicantes.* ©PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023)..... 126

**Figure 45** *Updating Boerhaave's Praelectiones academicae, tomus II, in Phaidra's Metadata Editor.* ©PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023)..... 127

**Figure 46** *Boerhaave's Praelectiones academicae, tomus II, in Phaidra. On the right, under permanent identifiers, purple cross-references indicate the previous and following tome of the same printing edition.* ©PHAIDRA - University of Padua, PHAIDRA – Digital Collections (2023)..... 128

