Tracking the life of the book: an interview with Cristina Dondi

Cristina Dondi is a leading scholar of bibliography and book history who joined Sapienza University of Rome in 2024 after 30 years at Oxford. Her key contributions include creating the Material Evidence in Incunabula (MEI) database in 2009 and leading the ERC-funded "15cBOOKTRADE" Project (2014–2019) [1], both focused on incunabula and the printing revolution. She has been the Secretary of the Consortium of European Research Libraries (CERL) since 2009 and is currently coordinating a new project on the migration of cultural heritage.

The interview focuses on her perspective regarding digital transformation, specifically exploring the impact of digitisation and the potential of artificial intelligence (AI), such as image recognition, on research and curriculum development within the field of the History of the Book.

Q1. Your foundational work, notably the 15th century book (incunabula) project, pioneered the application of quantitative and digital methodologies to bibliographical analysis, fundamentally shifting our understanding of the distribution, movement and ownership history of early printed books. Could you elaborate on the significance of this methodological shift?

The methodological shift, applied to incunabula already described at the bibliographical level in the Incunabula Short Title Catalogue (ISTC) database [2], consisted of three components:

(1) Detailed copy-specific cataloguing: Combining internal evidence (ownership inscriptions, manuscript annotations, decoration, binding, etc.) and external evidence (historical library catalogues; sale and auction catalogues; sequestration lists, etc.).

(2) Creation of the international database MEI [3]: Designed to keep the data together, allowing researchers to reconstruct dispersed collections and track the movement of books from the 15th to the 21st century.

(3) Promotion of collaboration and shared knowledge.

Q2. As secretary of the consortium of European research libraries (CERL), you steer an organisation central to heritage scholarship. What are the organization's strategic priorities in addressing the challenges of digital scholarship, particularly concerning data governance and integrity?

The continuous development of CERL's digital resources is our top priority and the focus of our financial investment. The strategic priorities in digital scholarship are:

- *Data integration:* Promoting and ensuring the integration of all bibliographical and copy-specific data for European and American printed heritage up to circa 1850 into a specialised network.
- *Enhancing access:* Supporting access to this data through the CERL Thesaurus (which includes variant spellings of imprint places, personal and corporate names) to enable high-quality and comprehensive scholarly research.

Digital Library Perspectives
Vol. 42 No. 1, 2026
pp. 214-218

2059-5816
DOI 10.1108/DLP-02-2026-216

- *Preserving bibliographic heritage:* Maintaining the bibliographic heritage of Europe, including incorporating large international resources such as the ISTC, STCN and ESTC, with the goal of improving access and interoperability.
- *Extending the MEI model:* Extending the model created for copy-specific data collection and analysis with MEI to other research projects (e.g. HPB-Provenance, Patrimonit, Matmed) to ensure long-term sustainability and improve interoperability within the digital environment [4].

To address the need for global accessibility and interoperability, CERL is progressively making the data in its digital resources available as linked open data, beginning with the Thesaurus.

Q3. The intersection of physical artefacts and their digital representations is vital. Given the "material turn" in textual scholarship, what are the most exciting technological developments that you see enhancing the scholarly study of manuscripts and incunabula? How can these digital surrogates foster new forms of academic and public engagement that go beyond simple facsimile?

Freely available digitisations have been immensely useful, advancing incunabula research in multiple scholarly ways. Furthermore, I have made extensive use of digitisations for the creation of many videos displayed in the exhibition "Printing R-Evolution 1450–1500" [5]. Since then, all the digital material has been made available online and is regularly used for teaching and presentations, with libraries also running them in their exhibition spaces (for example Monastery of Santa Scolastica in Subiaco).

Q4. With your international profile and teaching role at La sapienza university, how are you innovating your curriculum to integrate advanced digital perspectives, specifically linked data principles and the emerging capabilities of artificial intelligence (AI)? what core competencies must today's students in bibliography and textual scholarship acquire to meet the demands of modern research and library management?

In Oxford, during the 15cBOOKTRADE Project [6], we experimented with image matching of illustrations in incunabula to help find when woodcuts were reused and copied by different printers in different places and time. This year we plan to apply image recognition to the images of provenance contained in CERL's Provenance Digital Archive (PDA) [7]. But to achieve high quality advancement in research every implementation of AI requires a vast amount of (new) data collection and data organization. At the core of any exciting possibility stands good quality cataloguing, and I am very concerned by the pockets of important material not yet catalogued in our libraries.

At present I teach, in English, History of the Book (specifically the introduction of printing in Europe and around the world, in Latin and non-Latin alphabets) to a very diverse and large group of first year undergraduates within the Global Humanities BA course; they come from all over the world. I regularly use the videos created for the Venice exhibition and for other projects I have been involved in, as well as the digital resources we created (MEI, PDA, 1516, Text-inc) to explain the historical objectives which we are working towards, with a strong emphasis on methodology. But I also take them to libraries, to the Tipoteca Italiana and to Fabriano to experiment with the materiality of printing.

Students should have a direct experience of the object of their studies, that is books and libraries; contributing to ongoing research projects or large cataloguing campaigns and internships are very important; these experiences should reinforce their knowledge of the

existing main digital resources as provided by teaching; and they should speak English, to benefit from international mobility.

Q5. Open science principles advocate for transparency and accessibility, which can be challenging when dealing with fragile and geographically dispersed historical book materials. What is your perspective on the greatest opportunities and ethical challenges associated with promoting true open science, including data sharing and licencing, within the field of historical bibliography?

Everything CERL hosts and curates is made freely available; for the political aspects we conform to European legislation.

For us at CERL, another concern is version control. Scholars annotate and enrich the data they download from libraries/large repositories. When, in the not too distant future, scholars begin to make use of the Cultural Heritage Cloud, and themselves upload data there, eventually it will become very hard to understand which version of the data is available, who created it and precisely what the intellectual status of that data is.

Q6. Your projects have a broad impact. How does your team define and measure the success and outreach of large-scale digital initiatives, moving beyond simple website metrics? Specifically, what strategies have been most effective in translating complex bibliographical data (e.g. provenance information, watermarks) into user-friendly and impactful research tools for both specialist academics and the wider library/archival professional community?

One measure of success of our digital resources is the willingness of libraries and scholars to contribute to them – an investment of time, expertise and money. Other measures include the frequency of citation, invitations to speak at both scholarly events and those for a more general public, and scholarly publications and articles written for a wider audience. Then, of course, there is success in acquiring funding for research and special projects (for example, Dante1481 [8]; the digitisation of incunabula in Italian monastic libraries [9]; the digital reconstruction of the library of San Giorgio Maggiore of Venice, which is ongoing).

In terms of strategy, high-quality cataloguing – not in a silo but as part of a connected network – is always the core component of our research projects. However, in research projects, high-quality cataloguing should always be accompanied by a scholarly output and effective outreach, such as videos and exhibitions. Communication at multiple levels is essential.

Q7. The scale of digital bibliography demands intense international and cross-disciplinary collaboration, spanning fields from history and library science to computer science and data modelling. What are the most effective collaboration models you have implemented to bridge these disciplinary divides and ensure coherence, longevity and sustainability across massive, multi-institutional projects?

When I started MEI in 2009, I had the good sense of asking the Data Conversion Group of Göttingen State and University Library to design it with me. They have designed, managed and further developed most of CERL's digital resources, in collaboration with the librarians in CERL. The collaboration with, and support of, CERL libraries is part of the success of MEI and other CERL digital resources. I worked closely with a number of engineers specialised in data visualisation and image recognition (Min Chen, Andrew Zissermann and the Visual Geometry Group in Oxford; Alfie Abdul Rahman and the team of computer scientists at King's College London). Clear and close communication with the scientists is

vital. It is also very important, in my experience, to have clear historical objectives to begin with.

Q8. Beyond the initial creation phase, the long-term preservation and sustainability of complex digital bibliographical data are paramount. What specific technological protocols and organisational frameworks – for data standards, metadata management and archiving – must CERL and similar consortia prioritise to ensure these valuable datasets remain intact, authentic and usable for future generations of scholars and digital systems?

CERL adheres to internationally defined data structures and cataloguing practices, applying them to historical material. It ensures consistency and interlinking by promoting the use of standards and persistent identifiers, thereby gathering rich, specialised information and avoiding, as much as possible, the duplication of effort and loss of valuable data. Furthermore, clearly defining the data sets is becoming increasingly important so that researchers are fully aware of their contents and limitations.

All data added to CERL datasets are derived from trusted sources (rigorous cataloguing undertaken at CERL member libraries, to agreed standards). The data are carefully analysed to harmonise them from various cataloguing traditions and are ultimately supported by a coherent model and stable database environment that offers tailored indexing, searching and browsing. This illustrates the many steps and processes involved in transforming information and knowledge into usable data.

Q9. Reflecting on your significant contributions, is there one project, finding or methodological breakthrough that you consider the most personally rewarding or transformative for the field of bibliography as a discipline? Could you share a brief anecdote illustrating the challenge you overcame and the resulting scholarly insight?

The most rewarding project was undoubtedly the creation of MEI, following ten years of detailed incunabula cataloguing at the Bodleian Library. Despite the wealth of data, the work to extract information on the European circulation of Venetian books from the six-volume catalogues for an article (Dondi, 2008) proved extremely arduous. This difficulty led to the conclusion that the use of incunabula as historical sources could only be improved by reorganising copy-specific information (such as provenance, binding and decoration). From this, the model of the "life cycle of the book" was born, tracing its movement as a sequence of geolocated provenance blocks.

Q10. Looking ahead, how do you foresee the roles of the physical book, the digital surrogate and the bibliographer evolving over the next decade? What is the single most critical piece of advice you would offer to young scholars and library professionals who are currently entering or specialising in the field of digital library studies and heritage collections?

The digital resources which we have or are creating have to be comprehensive to sustain historical research. This is a long term aim. Therefore, I would like to stress again the all important awareness that not everything has been catalogued and even less has been digitized. And if it has been catalogued, it may not have been catalogued to the level of detail that you require for your research. Therefore research needs high quality cataloguing of the heritage collections, and the availability of images of the books, preferably linked to the records. Young scholars should be aware of the vast quantity of books in our libraries; if they are looking for a topic for their MA or PhD they should go and talk to their local librarian. Provenance research (for which high quality cataloguing is required) is essential for the

understanding of our entangled and highly mobile heritage collections; although increasingly wide-spread, there is so much still to do, including the integration of provenance data pertaining to GLAM heritage material.

Cristina Dondi
University of Rome La Sapienza, Rome, Italy, and

Anna Maria Tammaro
Department of Information Engineering, University of Parma, Parma, Italy

Notes

[1.] https://15cbooktrade.ox.ac.uk/project/project-team/

[2.] British Library. Incunabula Short Title Catalogue (ISTC): https://data.cerl.org/istc/_search

[3.] CERL Material evidence in incunabula (MEI): https://data.cerl.org/mei/_search

[4.] CERL Databases: data.cerl.org/all

[5.] Printing R-Evolution 1450-1500. Fifty Years that Changed Europe. www.printingrevolution.eu/

[6.] 1516 A Visually Searchable Database of Printed Illustration from the 15th and 16th centuries: www.robots.ox.ac.uk/∼vgg/research/1516/

[7.] CERL Provenance Digital Archive: https://pda.cerl.org/

[8.] Dante 1481: www.printingrevolution.eu/dante-1481/

[9.] The digitization of the incunabula of monastic libraries in Italy: the Polonsky Foundation's project with BNCR and CERL: http://digitale.bnc.roma.sbn.it/progettopolonsky/en/1/home

Reference

Dondi, C. (2008), "La circolazione europea degli incunaboli veneziani documentata dalle edizioni conservate alla biblioteca bodleiana, oxford", in Pon, L. and Kallendorf, C. (Eds), *The Books of Venice/Il Libro Veneziano, Miscellanea Marciana, v. 20 (2005-2007). Venezia: La Musa Talìa: Biblioteca Nazionale Marciana*, Oak Knoll Press, New Castle, DE, pp. 179-190, available at: www.oakknoll.com/resources/tableofcontents/100392.pdf