A Research Tool for the ERC-Funded EMoBookTrade Project

Giliola Barbero[1(✉)] and Luigi Tessarolo[2]

[1] Università degli Studi, Udine, Italy
giliola.barbero@uniud.it
[2] Venice, Italy

Abstract. The ERC-funded EMoBookTrade project, led by professor Angela Nuovo and based at the University of Udine, addresses the issues of book prices and of book privileges in early modern Europe (1540–1630c.). To achieve the goals set, a Web application is under development by the research team. This paper describes technical solutions found to ensure usability and effectiveness in recording complex historical data, and automatisms which allow data analysis.

Keywords: Digital humanities · Digital history · History of the book
Price history · Economic history

1 The ERC-Funded EMoBookTrade Project

The Early Modern Book Trade project, funded by the European Research Council, aims to understand how the European book market worked during the sixteenth century and the first decades of the seventeenth [1]. The project team is analyzing prices of books, privileges granted by Western governments and any other evidence concerning bookshops management and commercial networks.

Until this moment the history of the book has taken into account cultural and technical issues related to the production and circulation of books, while economic and juridical aspects received little attention. Nevertheless, several innovative researches show how important the booksellers' economic organization was in spreading texts and ideas, not only in Europe but also in the new world [2, 3].

To understand how economic and juridical issues influenced the cultural weight of printed books, the EMoBookTrade project is developing five main activities. (1) One research activity is devoted to investigate prices set by Italian publishers and booksellers (Manuzio, Giolito, Scoto, Compagnia bresciana) and to study the inventory of Gian Vincenzo Pinelli's private library, as it was assessed for sale during an auction that took place in Naples in 1608. (2) A second research activity on European prices employs Christophe Plantin's archive in Antwerpen as well as French publishers' and booksellers' trade lists. (3) A third task focuses on the management of a Venetian bookshop

L. Tessarolo—Independent Scholar.

G. Serra and C. Tasso (Eds.): IRCDL 2018, CCIS 806, pp. 201–208, 2018.
https://doi.org/10.1007/978-3-319-73165-0_20

and the pricing of books in the early seventeenth century, and consists in the transcription and analysis of Bernardino Giunti's stock book. (4) The census, edition and interpretation of privileges granted by the Republic of Venice and preserved in the Archivio di Stato in Venice are underway. (5) The correspondence between Venetian wholesale bookseller Giovanni Bartolomeo Gabiano and his European business partners, which describes the technique of building and managing a transnational network of book distribution is in the process of being edited.

To manage these sources and related data, a database application has been developed, in collaboration with the project team and Luigi Tessarolo, an IT specialist with a long standing experience in the field of digital humanities [4].

The database application consists of a backend and a frontend. Through the backend (which has already been working for six months) the EMoBookTrade research team enters data and manage controlled vocabularies and authority files. Researchers can also formulate complex queries to help data interpretation. The frontend is still under development: it is intended to publish data and make them accessible to the scholars outside the project. Both the backend and the frontend are in English.

The reliability of prices analysis depends not only on the accuracy of research activities and on a correct historical interpretation of the sources, but also on the semantic evaluation of data and of their relations inside the database (DB) conceptual model. In a web application for an innovative historical research, the semantic value of each entity must be defined precisely, because there are not always existing standards to adopt.

But implementing a common language and a correct data classification is not the only task when designing a DB application for humanistic researchers. Also usability and effectiveness of the Database Management System (DBMS) are fundamental, because researchers do not want to spend time in doing what they do not need to do while working on traditional paper notebooks or on word processors.

Finally, the EMoBookTrade project and the related DB application far are proving that the effectiveness of a digital repository of historical sources does not rely only on digital images but also on their transcription and interpretation, a task that only a research project can afford.

This paper aims to describe the DB (Sect. 2) and highlight the problems–not only technical–encountered when designing a Database Management System (DBMS) for historical research (Sect. 3); while–as the frontend is still in progress–only few information will be given on it (Sect. 4).

2 Entities in the DB

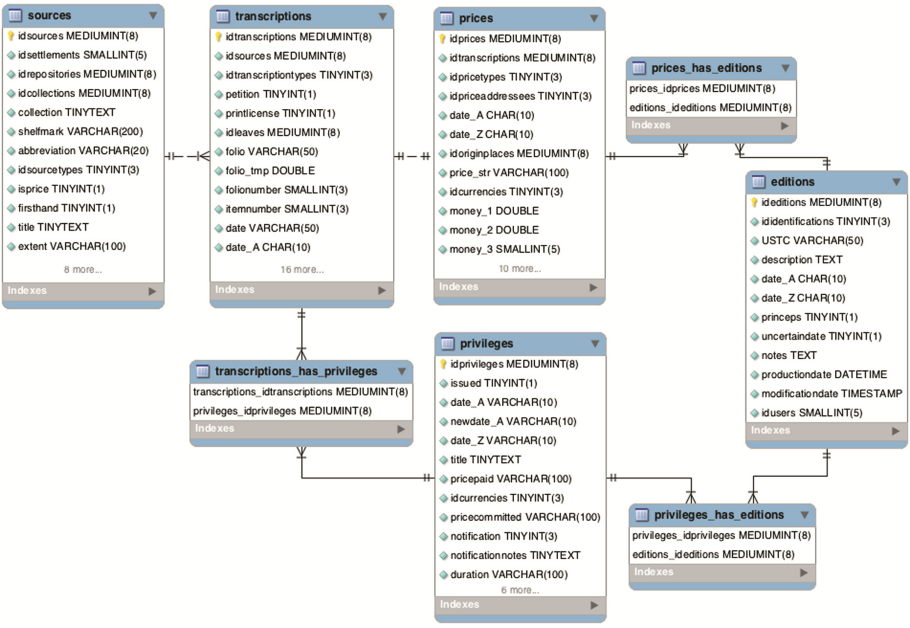
The EMoBookTrade DB and its corresponding DBMS aim to publish sources and to support their interpretation and analysis. The team needs to discover and enquiry whether there were constants in pricing books, what the average prices were, whether and how prices change over time and which elements could prompt publishers and booksellers to set different prices from the average (lower or higher).

To evaluate these issues, first of all a great effort was made to design a valid conceptual model.

In the EMoBookTrade DB the main entities are the following:

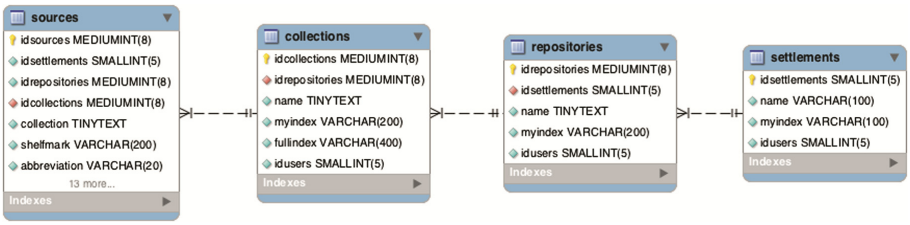
- sources: publisher catalogues, bookshop and bookstore inventories, archive documents;
- transcriptions: texts transcribed from the sources that provide information on prices and privileges;
- prices: an amount of money indicated in one or more sources;
- privileges: a public act described in one or more sources;
- editions: a series of printings of the same book priced and/or privileged by one or more sources;
- names: personal, corporate or geographic names, e.g. authors, publishers, places of publication, booksellers;
- currencies in prices, e.g. Lire veneziane, Ducati napoletani, Livres tournois.

Items of this list represent the historical objects and events considered by the EMoBookTrade project at an abstract level, while the logical structure of the tables is more complicated. Here is a simplified illustration of the fundamental tables and their relationships:



Other DB tables contain price-specific information which is one of the two cores (the other one is about privileges) of the EMoBookTrade project: prices, currencies and exchange rates are linked to editions as follows:

The attributes of each entity have been designed to respect standard, at least when existing. Sources identifier data (source location) for example respect the TEI standard for manuscript description (Settlement, Repository, Collection, Shelfmark) [5]:



And even if data about editions do not respect Marc [6] and Unimarc [7] models (because they would be too complicate for a non-bibliographic DB), each edition is identified by a public ID derived from institutional repositories such as the *Incunabula Short Title Catalogue* (ISTC) [8] and the *Censimento delle edizioni italiane del XVI secolo* (EDIT16) [9].

To design the conceptual model (entities) and the data structure (tables and relationships), as a first step, the team discussed and listed 'what' they wanted to record and, in some cases, what they had already listed in previous DBs. In fact a DB of Plantin's archive and a DB of privileges were already in use by some members of the new team. From these previous DBs many fundamental ideas came to the EMoBookTrade project, but afterwards a great effort has been made to restore those DBs structures and to design new tables and their relationships.

DBs developed for individual research in the field of humanities in fact often carry some recurring mistakes, that are tolerable in a private context but cannot be managed in a shared public digital repository.

One of the most common mistakes is to consider as one single entity's attributes sets of data that do not have a one-to-one ratio. E.g. librarians and IT specialists working for libraries with bibliographic formats (Marc, Unimarc) know very well that the relationship between authors' names and bibliographic descriptions is a many-to-many one, but

in several individual research contexts bibliographic descriptions and authors' names are stored in two fields of the same table and when authors are more than one, different names are copied together in the same field. The consequence is that each record can be read by human eyes as on a page on paper, but data as a whole cannot be ordered and managed automatically.

A too much scant attention to the relational structure can be observed also in other common practices. For example few pre-set tables with controlled vocabularies are designed and free fields are preferred when any sort of classification (classification of source, editions…) has to be recorded, that does not help in retrieval.

Moreover another common mistake in individual DBs is to design a unique field to store ontologically different data. For example it can be a discouraging experience to explain to humanistic researchers that the 'responsibility' a person has in a specific context (e.g. the responsibility to be the copyist of a document) is not to be considered a fixed 'qualification' of that person, because the same person can be a copyist in a context and an author in another one.

Last but not least also recording different kind of information in a field designed for a specific typology of data, just because 'there is a blank to fill up', is a typical misunderstanding of individual research tools. This is again a sort of typical confusion between a DB record and a traditional page, where the responsibility of understanding the semantic value of data and their internal links falls on the human reader.

On the other hand, from the previous DBs already in use, the new EMoBookTrade DB derived a great set of specialist data collected during the previous research, which saved the team from the task of speculating an abstract model and the entities which could be proved useful in the future. As the design of a DB devoted to an historical research project cannot follow existing standardization - for its innovative nature -, previous experience and data have been fundamental to design the EMoBookTrade DB.

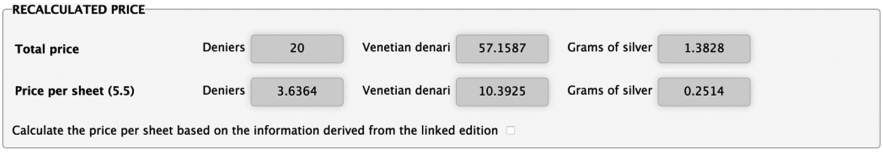
3 Usability and Effectiveness in the DBMS

After designing the logical model and creating the data structure, in developing the DBMS, the main task was to put the team in the condition to operate efficiently through the backend. The backend was organized in the safest and most expedient possible way, even if these two needs can often conflict. To this regard, balanced solutions were found, and at present the EMoBookTrade team members are working in an suitable condition.

The DBMS complies with the principles outlined below.

- No data has to be saved through the existing 'save' buttons, but when an empty record is closed, a pop-up message appears asking to the client if that record should survive or not.
- No data should ever be entered twice. E.g. when a researcher is transcribing a series of prices from the same inventory, and each price is intended to populate a single record, a basic set of data shared by all the prices–e.g. the date, the typology of price, the place where that price was set - is inherited (although they can also be changed manually).

– When information is already available online in free access repositories, such information can be imported automatically minimizing manual interventions. For this reason old books are not being catalogued by the EMoBookTrade team. Instead bibliographic descriptions are imported automatically from existing national and international catalogues and bibliographies such as USTC [8], EDIT16 [9], and VD16 [10].
– Numeric series (such as the order number of pages of a single catalogue) are entered automatically, although a manual change option is also available.
– Researchers are not required to perform manual calculations, because multiple automated functions are made available instead. For example, for each edition considered, the EMoBookTrade DBMS calculates the price per printing sheet. This is made possible by a function that automatically divides the total price found in the source by the number of leaves, divided in turn by the volume's format. For example, if a given book carries the total price of 1560 denari and it is known to be an in 4° of 360 leaves (for a total of 90 printing sheets), the price per printing sheet will be 17.33 *denari*. Moreover each price is automatically calculated in the minimum sub-multiple of the original currency used by the source, then converted in Venetian *denari* (when different) and in the correspondent grams of silver:



The conversion from foreign currencies to Venetian *denari* and from *denari* to grams of silver also takes into account the date when pricing was set, because the chronological fluctuation in exchange rates are also recorded in the DB.
– Any data resulting from complex serial calculations is a dynamic entity and is not to be permanently fixed and stored within the DB. Any manual intervention made on a single numeric value will result in a consequent automatic readjustment of all the interlinked values. E.g. if a recorded total price needs to be corrected, because at a first glance the researcher misread the original source, the correspondent price per printing sheet will be automatically recalculated, and the same goes for exchange rates and so forth.

Moreover several data, as they result from the interpretation of historical sources, cannot always be taken as fully reliable. For example sources may be damaged, misleading or carry incoherent information.

At times, data can even be 'unidentified.' This, far from being a paradox, is a commonplace of historical bibliography. This is the case when a source cites a work without providing sufficient bibliographic data: e.g. 'Rime di Petrarca'. The given case does not allow to establish an univocal relation between the cited - and priced work - and a known edition. In order to store information that are still valued by the researcher, the information regarding the price of the cited work will be linked to a fictitious bibliographic record (unidentified edition) based on the scanty information provided by the source: e.g. 'Francesco Petrarca, Rime'.

Appropriate models of data, usability and effectiveness of the DBMS and the freedom to record also doubtful data, help researchers engaged in digital humanities to accept the introduction of new shared digital tools in their traditional environment. Without care of such issues, it is very difficult to develop and employ DBs and DBMSs in the field of humanities and in humanistic projects [11].

4 The Future: Data Analysis

While the DB and the backend system described here are already working, the frontend - the public web site - is being designed. Nevertheless, in the backend the possibility to manage data for individual research already exists. At present, prices can be ordered and sorted by total price, price per printing sheets or quantity of printing sheets, and, most importantly, editions can be searched to verify if they had received privileges and to find out what prices were set by publishers or paid by customers for them. Data can also be downloaded by the team researchers.

This module of the backend has already been exploited to elaborate data for the congress *Selling & Collecting. Printed Book Sale Catalogues and Private Libraries in Early Modern Europe* held in Cagliari in September 2017 [12].

In the meantime the frontend is under construction. The EMoBookTrade team is discussing different solutions in order to create a research engine suitable to the scholars and their scientific needs. The starting point of the main public search will be from editions, to obtain prices and privileges connected with them. Scholars will be able to answers questions like "How much did Petrarca's *Canzoniere* cost in Venice duting the 15th century?", or "Which edition of Ariosto's *Orlando furioso* received a privilege in Italy?". Moreover they will be able to obtain and compare prices in any currency and/or in grams of silver and to order prices and privileges by different criteria.

Due to the analytical conceptual model designed, and to the amount of data collected by researchers, we trust in a coherent and suitable future development of the frontend. Just because the EMoBookTrade researchers have striven to share a common language and a common data classification, scholars outside the project will be able to share the same logic and consequently to discuss and exploit scientific results of the project.

References

1. EMoBookTrade, An Evidence-based Reconstruction of the Economic and Juridical Framework of the European Book Market, http://emobooktrade.uniud.it/
2. Nuovo, A., Ammannati, F.: Investigating book prices in early modern Europe: questions and sources. JLIS.it **8**, 3, September 2017. https://doi.org/10.4403/jlis.it-12365

3. Musisque Deoque, Un archivio digitale di poesia latina, PRIN 2005 e 2007, progetto scientifico di P. Mastandrea, R. Perrelli, G. Biondi, L. Zurli, V. Viparelli, progetto informatico di L. Tessarolo. www.mqdq.it. Pede certo - Metrica Latina Digitale, Università di Udine, progetto scientifico di L. Mondin, progetto informatico di L. Tessarolo, 2013. www.pedecerto.eu. Pietro Metastasio, Drammi per musica, PRIN 1998, progetto di A.L. Bellina e L. Tessarolo. www.progettometastasio.it. Carlo Goldoni, Drammi per musica PRIN 2001, progetto di A.L. Bellina e L. Tessarolo. www.carlogoldoni.it. Varianti all'opera, FIRB 2008–2012, Università di Padova, di Milano e di Siena. www.variantiallopera.it. Miguel de Cervantes, Novelas ejemplares, PRIN 2004, progetto scientifico di D. Pini; progetto informatico di L. Tessarolo. www.cervantes.cab.unipd.it. AXON - Iscrizioni storiche greche, Università Ca' Foscari Venezia, coordinatrice S. De Vido, supervisore C. Antonetti, progettazione e realizzazione S. Palazzo, M. Socal, L. Tessarolo (progetto informatico) (2015). http://virgo.unive.it/venicepigraphy/axon
4. Squassina, E.: Authors and the System of Publishers' Privileges in Venice. Gutenberg Jahrbuch, 42–74 (2016)
5. TEI –Consortium. Text Encoding Initiative, http://www.tei-c.org/index.xml
6. MARC standards. Library of Congress: Network Development and MStC standards Office, https://www.loc.gov/marc/
7. UNIMARC Bibliographic, https://www.ifla.org/publications/unimarc-bibliographic-3rd-edition-updates-2012?og=33
8. USTC Universal Short Title Catalogue, http://www.ustc.ac.uk/
9. EDIT16 Censimento nazionale delle edizioni italiane del XVI secolo, http://edit16.iccu.sbn.it/web_iccu/ihome.htm
10. VD 16 Verzeichnis der im deutschen Sprachbereich erschienenen Drucke des 16. Jahrhunderts, https://opacplus.bib-bvb.de/TouchPoint_touchpoint/start.do?SearchProfile=Altbestand&SearchType=2
11. Vitali, S.: On DB and historical research, Passato Digitale. Bruno Mondadori, Milano (2004)
12. Selling & Collecting. Printed Book Sale Catalogues and Private Libraries in Early Modern Europe, http://convegni.unica.it/icsc/