CorpusTracer: A CIDOC database for tracing knowledge networks

Florian Kräutli and Matteo Valleriani
Max Planck Institute for the History of Science, Germany

Abstract

In our research, we study mechanisms of knowledge dissemination based on the structural and social networks surrounding the edition history of a single text: the *Tractatus de sphaera* by Johannes de Sacrobosco. By applying methods from network analysis, we investigate how specific commentaries on the text circulated, which actors were responsible for them and what factors supported or hindered the spread of specific kinds of knowledge. The basis of this investigation is represented by CorpusTracer, a database that stores the required data in a suitable format and with the required level of expressivity. In this article, we present the design of our database and our data model based on CIDOC-CRM and FRBRoo. We discuss the implementation and suitability of the conceptual and technical realization for our research question. We conclude that FRBRoo fits well to the task at hand. We found that the comparatively complex data structure it requires can be sufficiently abstracted through current implementation methods. As the research continues, our data model will have to grow and we expect that the presented methods will be sufficient to accommodate our future requirements.

Correspondence:
Florian Kräutli,
Boltzmannstraße 22,
14195 Berlin,
Germany.
E-mail:
fkraeutli@mpiwg-berlin.mpg.de

1 Introduction

Historical research is increasingly confronted with an ever-expanding, worldwide process of digitization of historical sources. In the frame of this development, the analysis of text corpora, as has always been the case for instance in medieval studies, is becoming an important approach in many other areas of historical research.

Concerning the medieval era, text corpora compiled during the early modern period can also be mostly defined as corpora of commentaries, usually on ancient works but also on late medieval ones.

Early modern scientific commentaries are often regarded as works of less relevance when compared, for instance, to the original works of authors such as Andreas Vesalius (1514–64) or Guidobaldo del Monte (1545–1607), or, a little later, René Descartes (1596–1650). In truth, the differences between scientific commentaries and other works in most cases concern only those characteristics that are needed to define a literary genre, and they do not express any value in reference to their scientific relevance. Although such books contain the so-called 'original work', on which comments are written, the commentary texts—whether in form of annotations or addenda or appendices, just to mention a few approaches—were genuine occasions for the authors of the commentaries to present their own ideas and to spread knowledge innovations also and often in direct and explicit opposition to the knowledge expressed in the original work. Moreover, as Markus Asper has clarified, the authors of commentaries followed both specific agendas and practices. While the agenda, eventually formally expressed in the openings of the books, could show for instance a benevolent approach to the original work, the practice—the method used

doi:10.1093/llc/fqx047 Advance Access published on 18 August 2017

to compile the commentary and content-related issues—could in fact radically diverge from the agenda (Asper, forthcoming). This approach represented a potential for innovation that could be exploited, also when institutional conditions were not optimal.

2 Commentaries and Knowledge Evolution

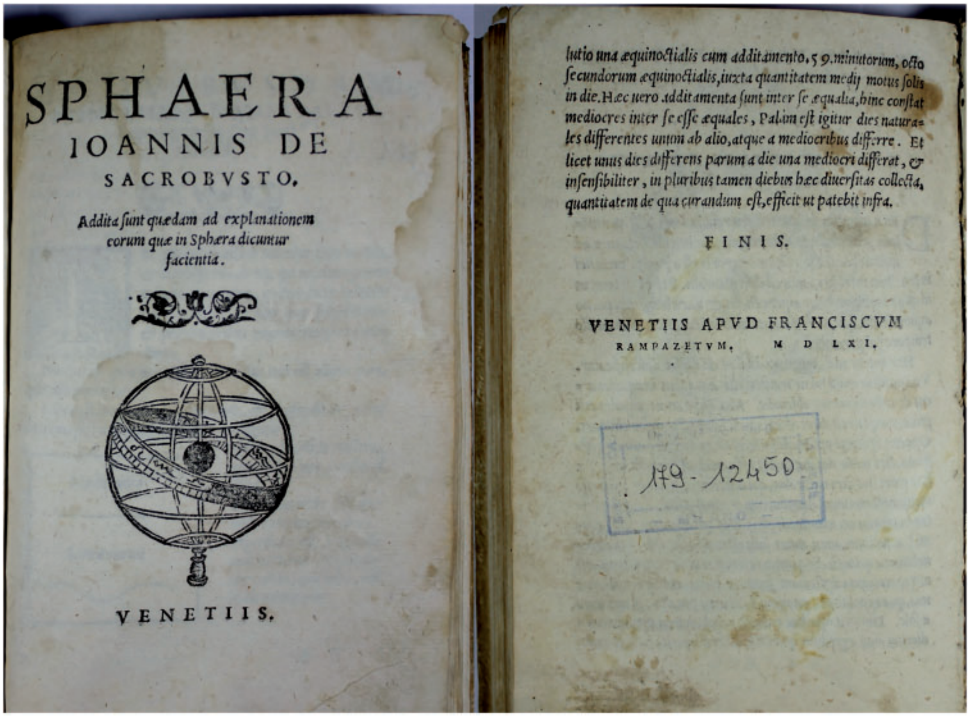
In the frame of the research project 'The Sphere. Knowledge System Evolution and the Shared Scientific Identity of Europe' (sphaera.mpiwg-berlin.mpg.de), we investigate mechanisms for the spread of knowledge innovations by making use of models developed in the area of network analysis.

We focus on a university textbook on geocentric cosmology compiled during the thirteenth century that gave rise to an extremely successful commentary tradition that lasted until the seventeenth century: *Tractatus de sphaera* by Johannes de Sacrobosco. The print history of the *Tractatus* begins in 1472; the printed material books containing the original work amount to over 300 different editions produced by printers all over Europe. Beginning with the analysis of the printed books, we investigate how those books evolved. Authors prepared new texts and commentaries, including novelties such as new mathematical procedures or descriptions of new mathematical instruments (Valleriani, 2017). Printers, on the other hand, selected which commentaries had to belong to the same edition; perhaps they ordered new engravings or decided to add extracts from pertinent but older works, for instance, from ancient ones. Therefore we analyze the evolution of a book on two levels: (1) the conceptual one, by looking for the relations between 'text parts' built between the first occurrence of a knowledge innovation in one commentary and its later adoptions in other editions; (2) the social one, by looking at relations, such as contracts concerning the sale of rights of reproductions or the loan of woodblocks and engravings, between printers of subsequent editions that show similarities in terms of the adoption of the same innovations. We create in this way a multilevel network on whose basis we can investigate (1) the profile of the knowledge that was most successful in terms of adoptions (for instance, whether most adopted innovations belonged more to the technical, natural philosophical, or mathematical knowledge domain); (2) which social mechanisms and actors were mostly responsible for the phenomena of spread (or hindrance) of innovations.

The example below shows a commentary printed in Venice in 1561 (Sacrobosco, 1561) (Fig. 1). In this case, there is not even an explicit author of the commentary. The book can nevertheless be considered a commentary because of three differences from the original text. The first is an addition at the beginning of the text concerning a few easy definitions belonging to Euclidean geometry (Sacrobosco, 1561, A 1r–[7r]) (Fig. 2).[1] The original text begins with the geometric definition of a sphere, and this addition is to be understood as a propaedeutic to Sacrobosco's geometric definition of a sphere, with which he opens his treatise. In this case, the printer has decided to follow a very clear trend as a great number of existing commentaries begin in exactly the same way with the same text and images. Sacrobosco's original work is reproduced but entirely uncommented (Sacrobosco, 1561, A [7r]–D [7v]). After this work, there is, first, an image that explains the process of the moon phases and, second, a very short portion of text, which is only seven lines long, extracted from the encyclopedic *Naturalis historia* of Pliny.[2]

Strangely, the printer sets the end of the treatise (of Sacrobosco) only after Pliny's extract. However, this might be due to the fact that the extract also served to fill the page. Finally, the printed edition concludes with the addition of Proposition 22 of the Third Book of Regiomontanus' *Epitoma in «Almagestum» Ptolomaei* on the causes of the inequality of length of the solar day during the year (Fig. 3).[3]

The portions of texts added to Sacrobosco's treatise in this edition are absolutely pertinent to the subject matter and indeed they touch upon specific detailed issues discussed in the original tract. They are all extracts from works older than 1561 and therefore cannot be called innovations as if they were reflecting upon recent discoveries. However, they definitely could have been innovations in the



SPHAERA
IOANNIS DE
SACROBVSTO.

Additæ ſunt quædam ad explanationem eorum quæ in Sphæra dicuntur facientia.

VENETIIS.

lutio una æquinoctialis cum additamento, 59. minutorum, octo ſecundorum æquinoctialis, iuxta quantitatem medij motus ſolis in die. Hæc uero additamenta ſunt inter ſe æqualia, hinc conſtat mediocres inter ſe eſſe æquales, Palam eſt igitur dies naturales differentes unum ab alio, atque a mediocribus differre. Et licet unus dies differens parum a die una mediocri differat, & inſenſibiliter, in pluribus tamen diebus hæc diuerſitas collecta, quantitatem de qua curandum eſt, efficit ut patebit infra.

FINIS.

VENETIIS APVD FRANCISCVM
RAMPAZETVM. M D LXI.

Fig. 1 Frontispiece and colophon of the treatise published in 1561 by the printer Franciscum Rampazetum. Courtesy of Biblioteca di Filosofia dell'Università degli Studi di Milano.

frame of the text history of this corpus as the printer could have found something relevant in the ancient literature. The innovation consists in this case in the conceptual connections between works that until that moment had belonged instead to different traditions.[4]

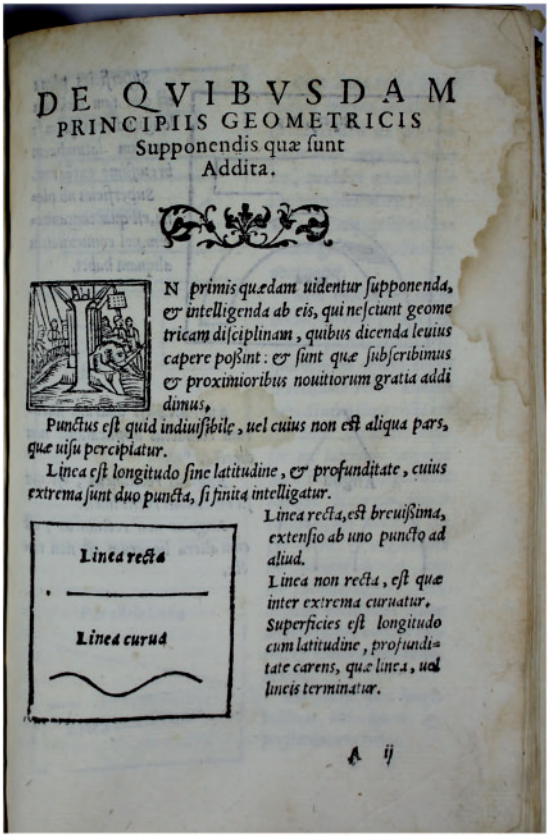
Many of these portions of texts as well as specific commentaries and annotations are not announced in the frontispiece of the work. In other words, they do not belong to what we call bibliographic data. Nevertheless, the appearance and subsequent adoptions of these text portions determine content-related relations between the treatises of this text corpus: they determine the network that, in its dynamic, shapes the evolution of a knowledge system.

3 The Database

An analysis of a text corpus that heads toward the application of mathematical models of network theory requires a space where all these data can be stored in form of relations. In fact, we go as far as storing the data in a format where it all consists of relations. We are able to do so by using RDF language (Resource Description Framework) that describes resources as a set of statements. These statements, called triples, express one kind of relation between two entities. A single book is then described by a multitude of triples, which are stored by the database.

By storing every aspect of the book as a relation, we do not need to decide in advance what ultimately constitutes the nodes in our network. Using the same data set, we can analyze the conceptual relations of books and the texts they contain, the social network of books and their publishers, or other types of networks that prove to be of interest as the research continues.

The atomic relations that define the triples are specified by the data model. To model the data, we use CIDOC-CRM (Crofts *et al.* 2011) together with



DE QVIBVSDAM
PRINCIPIIS GEOMETRICIS
Supponendis quæ ſunt
Addita.

IN primis quædam uidentur ſupponenda, & intelligenda ab eis, qui neſciunt geometricam diſciplinam, quibus dicenda leuius capere poßint: & ſunt quæ ſubſcribimus & proximioribus nouitiorum gratia addidimus,

Punctus eſt quid indiuiſibile, uel cuius non eſt aliqua pars, quæ uiſu percipiatur.

Linea eſt longitudo ſine latitudine, & profunditate, cuius extrema ſunt duo puncta, ſi finita intelligatur.

Linea recta, eſt breuißima, extenſio ab uno puncto ad aliud.

Linea non recta, eſt quæ inter extrema curuatur.

Superficies eſt longitudo cum latitudine, profunditate carens, quæ linea, uel lineis terminatur.

A ij

Fig. 2 First page of the introduction to Euclidean geometry as a propaedeutic to Sacrobosco's treatise. From Sacrobosco, 1561, A [7r]. Courtesy of Biblioteca di Filosofia dell'Università degli Studi di Milano

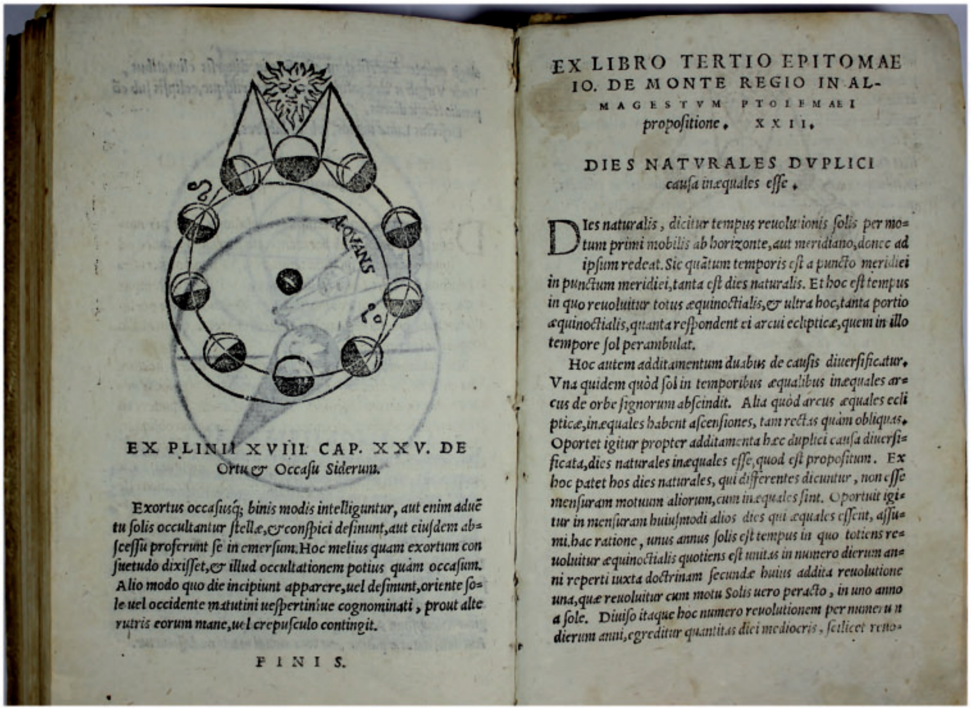
the FRBRoo (Bekiari *et al.* 2015) extension. CIDOC-CRM is a reference model that enables unified access to data, while allowing the representation of individual and subject-specific characteristics. As a conceptual model, it does not demand a specific technical implementation but is usually applied following linked data approaches and using RDF (CIDOC/RDF).

FRBRoo is the CIDOC implementation of the Functional Requirements for Bibliographic Records—FRBR (IFLA, 2009). At the heart of FRBR are the so-called Group 1 entities: Work, Expression, Manifestation, and Item. They each represent a different level of abstraction, ranging from the conceptual, immaterial notion of a Work, to the specific and physical instantiation of an Item. These entities enable us to be more specific when we refer to a 'book'. Do we mean the specific copy that we own (Item) or any copy of a specific book (Manifestation)? What constitutes the translation of a book? In FRBR, it is the Expression that is being translated. The Work, finally, is 'a distinct intellectual or artistic creation' (ibid. p. 17). Works enable us to refer, for example, to Homer's *Iliad*, without having to reference one particular instantiation. Works, however, are also the most abstract of the four entities, and it might not always be possible—or indeed practical—to declare something as a Work.

FRBRoo adopts these entities as subclasses of CIDOC entities and further specifies them. To comply with the CIDOC nomenclature of using letters and numbers to identify entities—Exx for classes, Pxx for entities—FRBR entities are named and numbered: Fxx for classes, Rxx for properties.

3.1 Our data model

We apply the classes and properties provided by CIDOC-CRM and FRBRoo to develop a data model that is meaningful for our research questions and is likewise supported by available evidence.[5] Developing a data model is an iterative approach that requires us to oscillate between the specifications of FRBRoo, our own sources, and our research questions. Initially, we started from the concept of a Work, Sacrobosco's *Tractatus de sphaera*—the original tract—as a 'F15 Complex Work', which is a constellation of interrelated Works. All of the material books we study would then constitute members of this Complex Work, as well as being Works in their own right. Using FRBRoo's terminology, we can then further specify if a Work only reproduces Sacrobosco's original tract (F14 Individual Work), if it contains added and edited commentary texts (F17 Aggregation Work) or if it combines existing expressions of the original tract and commentaries in a single treatise (F19 Publication Work). However, while developing the data model, we questioned our ability to discriminate between different types of Works. We would only know the specific type of a Work after having completed



EX PLINII XVIII. CAP. XXV. DE Ortu & Occaſu Siderum.

Exortus occaſusq; binis modis intelliguntur, aut enim aduẽtu ſolis occultantur ſtellæ, & conſpici deſinunt, aut eiuſdem abſceſſu proferunt ſe in emerſum. Hoc melius quam exortum conſuetudo dixiſſet, & illud occultationem potius quàm occaſum. Alio modo quo die incipiunt apparere, uel deſinunt, oriente ſole uel occidente matutini ueſpertiniue cognominati, prout alterutris eorum mane, uel crepuſculo contingit.

FINIS.

EX LIBRO TERTIO EPITOMAE IO. DE MONTE REGIO IN ALMAGESTVM PTOLEMAEI propoſitione. XXII.

DIES NATVRALES DVPLICI cauſa inæquales eſſe.

Dies naturalis, dicitur tempus reuolutionis ſolis per motum primi mobilis ab horizonte, aut meridiano, donec ad ipſum redeat. Sic quãtum temporis eſt a puncto meridiei in punctum meridiei, tanta eſt dies naturalis. Et hoc eſt tempus in quo reuoluitur totus æquinoctialis, & ultra hoc, tanta portio æquinoctialis, quanta reſpondent ei arcui eclipticæ, quem in illo tempore ſol perambulat.

Hoc autem additamentum duabus de cauſis diuerſificatur. Vna quidem quòd ſol in temporibus æqualibus inæquales arcus de orbe ſignorum abſcindit. Alia quòd arcus æquales eclipticæ, inæquales habent aſcenſiones, tam rectas quàm obliquas. Oportet igitur propter additamenta hæc duplici cauſa diuerſificata, dies naturales inæquales eſſe, quod eſt propoſitum. Ex hoc patet hos dies naturales, qui differentes dicuntur, non eſſe menſuram motuum aliorum, cum inæquales ſint. Oportuit igitur in menſuram huiuſmodi alios dies qui æquales eſſent, aſſumi. hac ratione, unus annus ſolis eſt tempus in quo totiens reuoluitur æquinoctialis quotiens eſt unitas in numero dierum anni reperti iuxta doctrinam ſecundæ huius addita reuolutione una, quæ reuoluitur cum motu Solis uero peracto, in uno anno a ſole. Diuiſo itaque hoc numero reuolutionem per numerum dierum anni, egreditur quantitas diei mediocris, ſcilicet reuo-

Fig. 3 After Sacrobosco's treatise, an image is added that explains the moon phases and finally a pertinent passage from Pliny's *Natural History*. Here, the printer sets the end of the original treatise and adds Proposition XXII from the third book of Regiomontanus' *Epitoma in Almagestum Ptolomaei* on the causes of the inequality of the length of the days. From Sacrobosco, 1561, D [8r]. Courtesy of Biblioteca di Filosofia dell'Università degli Studi di Milano

our analysis of the texts they contain. We cannot use the Work entity as a basis for gathering data.

We therefore developed the data model on the basis of what was at hand: the digitized versions of physical copies of books. Instead of building the semantic network from an abstract notion of a Work, we begin with its physical instantiation, or Item.

The current state of the data model is reproduced here (Fig. 4). It captures bibliographic data of each treatise as well as data on the individual texts (Expressions) they contain.

At the center of the semantic network is the physical book (F5 Item). A book 'carries' a F24 Publication Expression, which is essentially its content. The Publication Expression then 'incorporates' a F22 Self-Contained Expression. This is the main text of a book, while the Publication Expression also includes additions such as page numbers, tables of content, indices, formatting, and so on. Both expressions have been created through events that were carried out by actors (the author or publisher) at a certain time (the date of publication) in a certain place (the place of publication). Other bibliographic data are modeled in a similar fashion. We were able to model most of these aspects based on available examples (Europeana, 2013, Norton, 2014).

The ostensibly simple property of the number of pages in a book required further consideration. Initially, we modeled this as a property of the physical book. Strictly speaking, this is incorrect as the number of pages in a bibliographic context refers to the number of pages a book is supposed to have. Relating it to a physical book means we refer to an individual copy, which might have pages missing. While entering the data, we realized that we cannot make a reliable statement about either. All

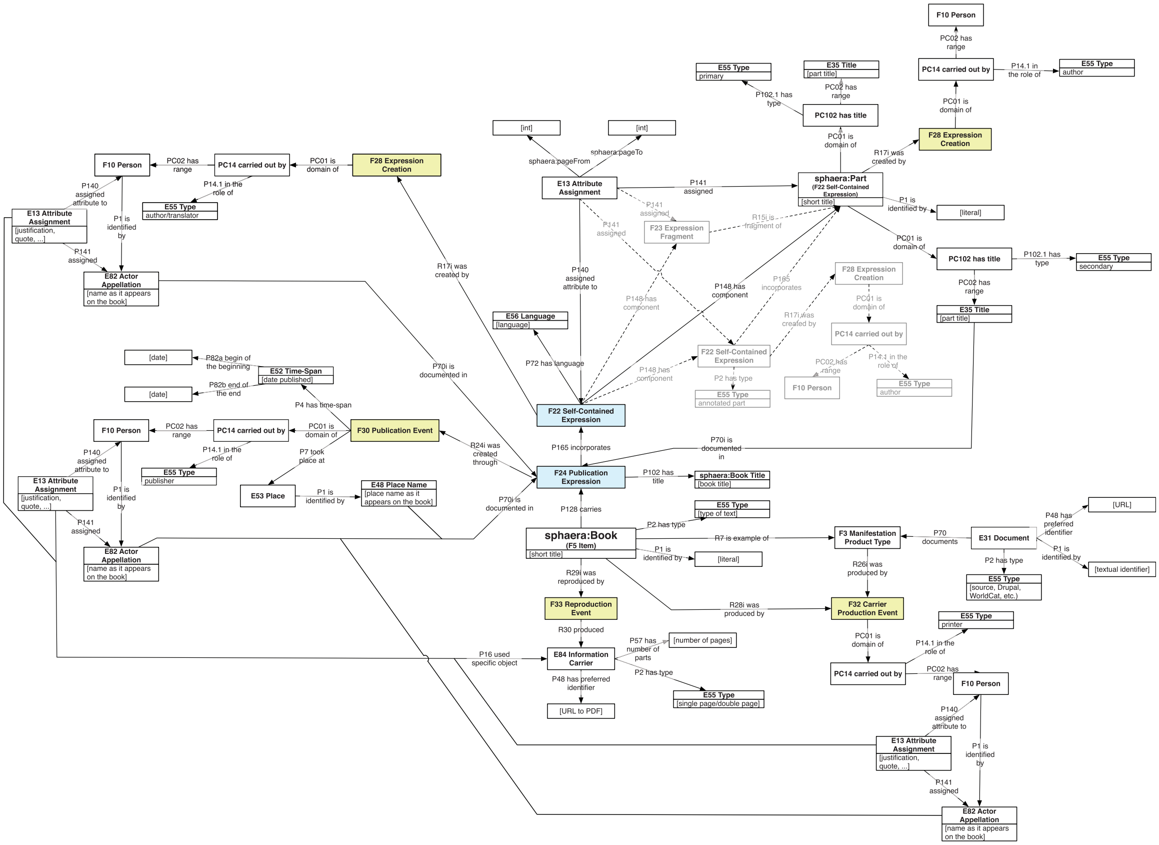


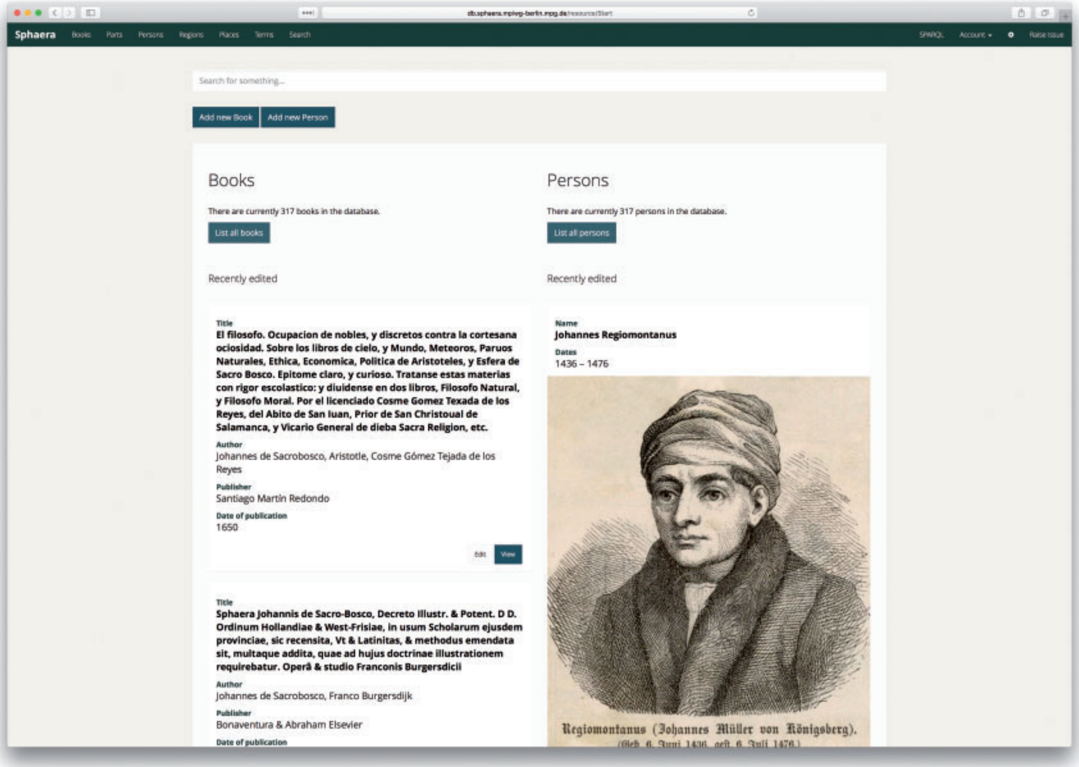
Fig. 4 The FRBRoo data model that specifies the modeling of different aspects of the books. We continuously adapt and extend the model to accommodate new findings.

we can state with certainty is how many pages are represented in the PDF file. The data model enables us, but also forces us, to be specific about what we know and cannot know from our sources.[6]

The different kinds of commentary texts can all be considered Expressions in their own right, which make up the Expression that is a book's text. We can further specify whether an Expression has been included in its entirety, or—as in Pliny's example—only a portion of it. We model this through a book including an Expression Fragment, which is a fragment of the original text (F23 Expression Fragment → R15i is fragment of → F22 Self-Contained Expression). Some books contain annotated texts, such as the original tract of Sacrobosco with another author's commentaries set graphically next to it. In these instances, we want to be able to identify a book as containing Sacrobosco's text, but nevertheless treat the annotated texts as expressions in their own right. We model this as a book including an expression that itself incorporates Sacrobosco's original text (F22 Self-Contained Expression → P165 incorporates → F22 Self-Contained Expression). Currently, we discriminate between three ways a text can appear in a treatise: in its complete form, as an extract, or complete and annotated. As we encounter more forms, we will extend our model.[7]

3.2 Implementation of the database

Digital tools that enable data entry directly in CIDOC/RDF and in a user-friendly manner are currently few and far between. For our project, we make use of a platform that is being developed in



Books

There are currently 317 books in the database.

Recently edited

Title
El filosofo. Ocupacion de nobles, y discretos contra la cortesana ociosidad. Sobre los libros de cielo, y Mundo, Meteoros, Paruos Naturales, Ethica, Economica, Politica de Aristoteles, y Esfera de Sacro Bosco. Epitome claro, y curioso. Tratanse estas materias con rigor escolastico: y diuidense en dos libros, Filosofo Natural, y Filosofo Moral. Por el licenciado Cosme Gomez Texada de los Reyes, del Abito de San Iuan, Prior de San Christoual de Salamanca, y Vicario General de dieba Sacra Religion, etc.

Author
Johannes de Sacrobosco, Aristotle, Cosme Gómez Tejada de los Reyes

Publisher
Santiago Martin Redondo

Date of publication
1650

Title
Sphaera Johannis de Sacro-Bosco, Decreto Illustr. & Potent. D D. Ordinum Hollandiae & West-Frisiae, in usum Scholarum ejusdem provinciae, sic recensita, Vt & Latinitas, & methodus emendata sit, multaque addita, quae ad hujus doctrinae illustrationem requirebatur. Operâ & studio Franconis Burgersdicii

Author
Johannes de Sacrobosco, Franco Burgersdijk

Publisher
Bonaventura & Abraham Elsevier

Date of publication

Persons

There are currently 317 persons in the database.

Recently edited

Name
Johannes Regiomontanus

Dates
1436 – 1476

Regiomontanus (Johannes Müller von Königsberg).
(Geb. 6. Juni 1436, geſt. 6. Juli 1476.)

Fig. 5 The landing page of CorpusTracer, which was built based on the ResearchSpace platform.

the context of ResearchSpace, a project led by the British Museum and funded by the Mellon Foundation.

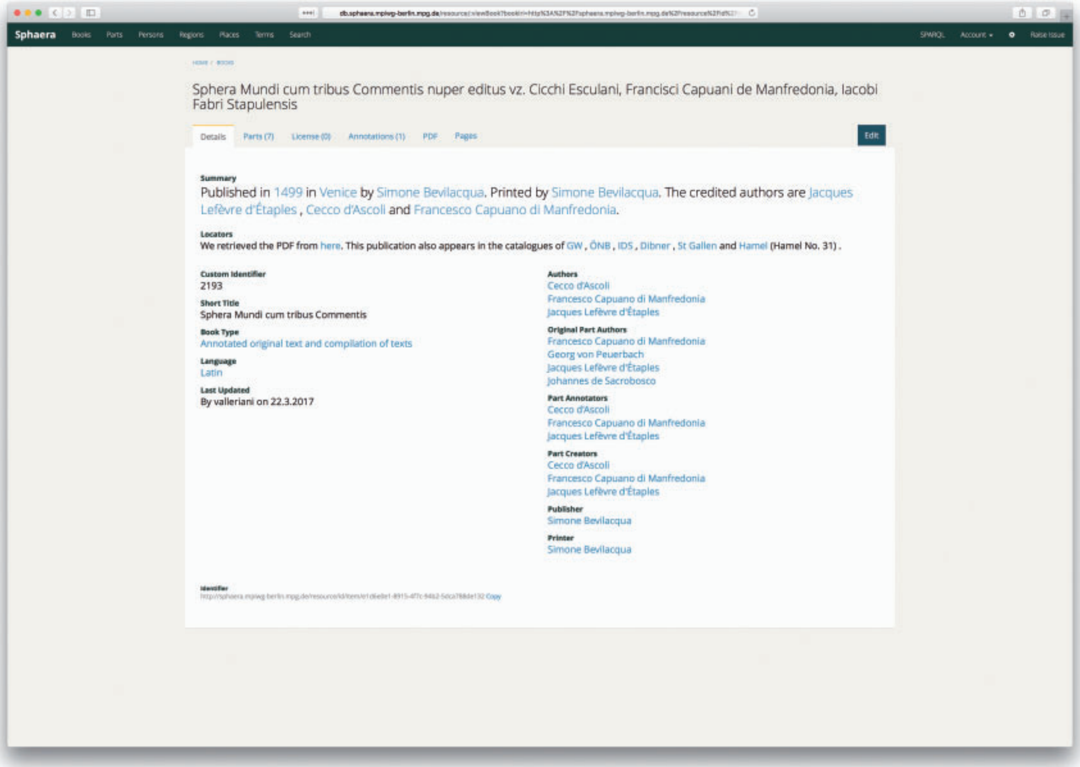
ResearchSpace aims to make the knowledge that is increasingly created through and within digital systems usable and accessible. Central to these efforts is the development of a software environment that enables researchers to browse, query, annotate, and—most importantly for this project—create semantic data.

We create our data set using a front-end interface we built that is based on the core platform and modules of ResearchSpace. Data are stored in a Blazegraph database and can also be queried directly via a SPARQL end point.

The landing page of the database environment shows recently edited book records in the left column, and recent person records on the right (Fig. 5). The illustrations depicting the persons as well as their dates of birth and death are pulled from Wikidata. When creating a new record for a person, users have the option of linking a person's record in the database to the corresponding person on Wikidata.

By linking entities to Wikidata, we can make a clear separation between our own knowledge that we derive from our collection, and canonical knowledge, such as the geographic location of cities that comes from external sources. We use Wikidata rather than other reference databases, as it provides a heterogeneous interface for many kinds of entities: people and places, but also professions and terms. However, for our purposes, it is incomplete (roughly half of the people we find in our sources do not exist in Wikidata) but is also editable, and we can add new entities as we identify them.[8]

Selecting a book in our database brings up a view containing a summary of its details (Fig. 6). In this



Sphera Mundi cum tribus Commentis nuper editus vz. Cicchi Esculani, Francisci Capuani de Manfredonia, Iacobi Fabri Stapulensis

Summary
Published in 1499 in Venice by Simone Bevilacqua. Printed by Simone Bevilacqua. The credited authors are Jacques Lefèvre d'Étaples , Cecco d'Ascoli and Francesco Capuano di Manfredonia.

Locators
We retrieved the PDF from here. This publication also appears in the catalogues of GW , ÖNB , IDS , Dibner , St Gallen and Hamel (Hamel No. 31) .

Custom Identifier
2193

Short Title
Sphera Mundi cum tribus Commentis

Book Type
Annotated original text and compilation of texts

Language
Latin

Last Updated
By valleriani on 22.3.2017

Authors
Cecco d'Ascoli
Francesco Capuano di Manfredonia
Jacques Lefèvre d'Étaples

Original Part Authors
Francesco Capuano di Manfredonia
Georg von Peuerbach
Jacques Lefèvre d'Étaples
Johannes de Sacrobosco

Part Annotators
Cecco d'Ascoli
Francesco Capuano di Manfredonia
Jacques Lefèvre d'Étaples

Part Creators
Cecco d'Ascoli
Francesco Capuano di Manfredonia
Jacques Lefèvre d'Étaples

Publisher
Simone Bevilacqua

Printer
Simone Bevilacqua

Identifier
http://sphaera.mpiwg-berlin.mpg.de/resource/id/item/e1d6e8e1-8915-4f7c-9462-5dca788de132 Copy

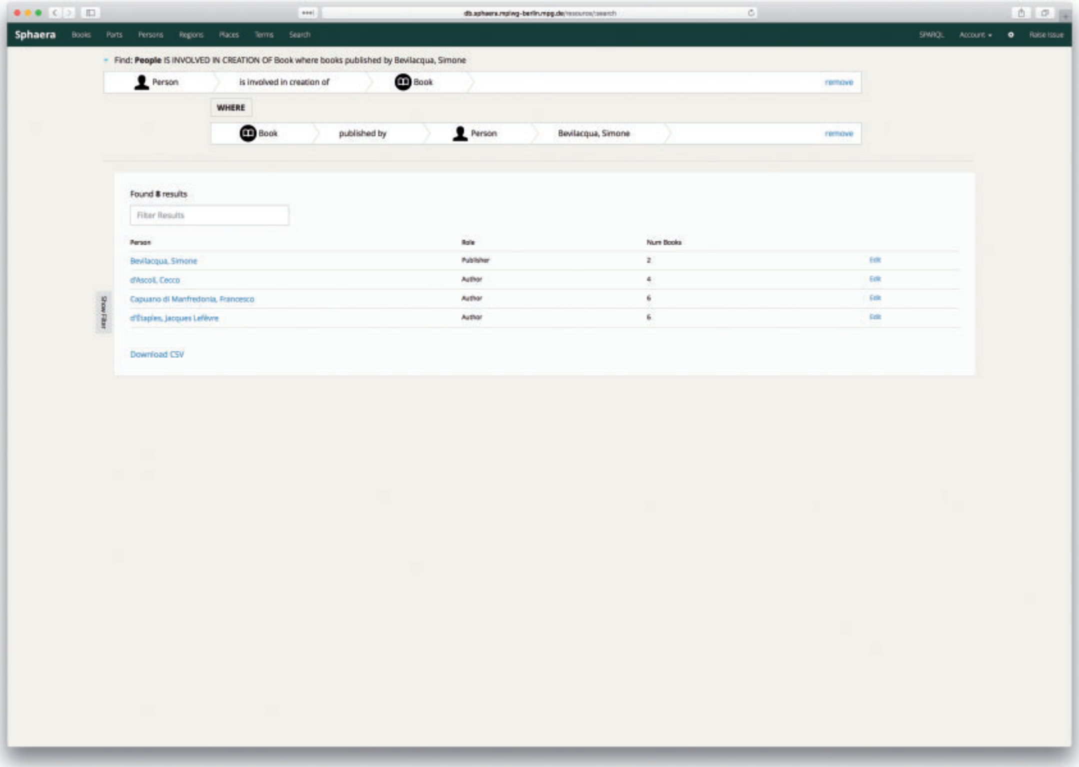
Fig. 6 Selecting a book calls up a summary of the data along with annotations and a PDF viewer in a separate tab.

example of a treatise published in 1499 by Simone Bevilacqua, we see how we are able to discriminate between different types of authors: the authors announced in the title (Cecco d'Ascoli, Francesco Capuani da Manfredonia, Jacob Fabri d'Etaples) and those who do not belong to the bibliographic data in the strict sense. The latter includes Johannes de Sacrobosco, whose *Tractatus* is actually reproduced three times in the book, but in versions annotated by the credited authors.

The ability to not only state the fact that a text appears in a book but also to represent within the data how a certain part of the book came about, which text it is based on, and who contributed to it, is a key requirement for our later analysis.

The editing screen of the books is organized into individual sections for data concerning the book as a whole, its publication, authoring and printing, links to references of the book in external catalogues, the texts it contains, and finally, annotations added by the researchers while compiling the data set.

The front end of the database includes views that enable a basic level of analysis of the data. We adopt the Semantic Search interface from ResearchSpace that allows users to query the database for relations without having to write queries by hand and, crucially, without having to know the underlying data model. It applies the concept of Fundamental Categories and Relationships (Tzompanaki and Doerr, 2011), which translates a complex data model to a set of simple, 'fundamental' paths. Through the interface we can find, for example, all the books that share a certain collection of texts, or find all persons related to a specific publisher by the books they published (Fig. 7).



Find: **People** IS INVOLVED IN CREATION OF Book where books published by Bevilacqua, Simone

Person | is involved in creation of | Book | remove

WHERE

Book | published by | Person | Bevilacqua, Simone | remove

Found 8 results

Filter Results

Person	Role	Num Books	
Bevilacqua, Simone	Publisher	2	Edit
d'Ascoli, Cecco	Author	4	Edit
Capuano di Manfredonia, Francesco	Author	6	Edit
d'Étaples, Jacques Lefèvre	Author	6	Edit

Download CSV

Fig. 7 The interface includes a visual query builder that enables researchers to explore relations between books, parts, persons, places, and time frames.

4 Discussion and Future

We found that the abstraction of a book's content introduced by FRBR aligns well with our research question. Using FRBRoo to model our data, we are able to meet the required level of expressivity. With the chosen front-end implementation, we can do so while keeping the database maintainable for non-expert users.

The entities and relations that are implemented in CIDOC and FRBRoo have so far sufficed for our purposes, and we do not expect to exhaust them as we continue adding to and refining our model. The challenge will be to maintain the usability of the interface when we add different types of data later on to investigate more levels of our network.

According to our epistemic approach (Renn *et al.*, 2016), we investigate the diffusion of knowledge innovation by considering three levels. The first, already configured in the current database, concerns content-related matters. The second concerns the relations between treatises that are of a social, institutional, and economic nature. At this layer, we look at the actual carriers of the transaction costs, which are the publishers and the social micro-region represented by all actors gravitating around their workshops: printers, engravers, translators, sellers, teachers, courtiers, and finally, authors. In this respect, we want to trace relations between treatises that are based on events, such as the contracts between publishers to sell or buy rights of reproduction, to lend or borrow woodblocks, engraving plates or entire printing plates, or agreements between publishers and educational institutions to prepare the required textbooks in advance (Pantin, 1998). The third layer is a kind of meso-layer (Lazega, 2016) represented by the material books produced.

In this framework, the social layer is defined as the structure of the network, the material books are the vehicles circulating along the edges of the structure, and the conceptual layer represents the dynamic aspect of the network. We intend to extend our database to make it capable of saving relational data concerning the structure of the network and data for the investigation of three further networks concerned with the dynamics of knowledge innovation along that structure.

References

Asper, M. (Forthcoming). Doing commentaries: Ancient and modern. In Amirav, H. and Markschies, C. (eds), *Catenae*. Leuven: Peeters Publishers.

Bekiari, C., Doerr, M., La Boeuf, P., and Riva, P. (eds) (2015). Definition of FRBRoo: A Conceptual Model for Bibliographic Information in Object-Oriented Formalism. https://www.ifla.org/files/assets/cataloguing/FRBRoo/frbroo_v_2.4.pdf (accessed 27 April 2017).

Crofts, N., Doerr, M., Gill, T., Stead, S., and Stiff, M. (eds) (2011). Definition of the CIDOC Conceptual Reference Model. ICOM/CIDOC CRM Special Interest Group. http://www.cidoc-crm.org/sites/default/files/cidoc_crm_version_5.0.4.pdf (accessed 27 April 2017).

Sacrobosco, J. (1561). *Sphaera Ioannis de Sacrobvsto: addita sunt quaedam ad explanationem eorum quae in Sphera dicuntur facientia*. Venetiis: apud Franciscum Rampazetum.

Europeana. (2013). Final Report on EDM – FRBRoo Application Profile Task Force. http://pro.europeana.eu/taskforce/edm-frbroo-application-profile (accessed 6 April 2016).

IFLA. (2009). Functional Requirements for Bibliographic Records: Final Report. IFLA. http://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf (accessed 27 April 2017).

Lazega, M. (2016). *Multilevel Network Analysis for the Social Sciences. Theory, Methods and Applications*. Heidelberg: Springer.

Norton, B. (2014). Book of the Dead Project. http://www.cidoc-crm.org/Resources/book-of-the-dead-project-0 (accessed 6 April 2017).

Pantin, I. (1998). Les problèmes de l'édition des livres scientifiques: l'exemple de Guillaume Cavellat. In **Bibliothèque Nationale** (ed.), *Le Livre dans l'Europe de la Renaissance: Actes du XXVIIIe Colloque International d'Etudes Humanistes de Tours*. Paris: Promodis, Editions du Cercle de la Librairie, pp. 240–52.

Renn, J., Wintergrün, D., Lalli, R., Laubicher, M., and Valleriani, M. (2016). Netzwerke als Wissensspeicher. In Mittelstraß, J. and Rüdiger, U. (eds), *Die Zukunft der Wissensspeicher. Forschen, Sammeln und Vermitteln im 21. Jahrhundert*. Munich: UVK Verlagsgesellschaft, pp. 35–79.

Tzompanaki, K. and Doerr, M. (2012). *A New Framework for Querying Semantic Networks*. Greece: Institute of Computer Science, F.O.R.T.H. Crete.

Valleriani, M. (2017). The tracts of the sphere. Knowledge restructured over a network. In Valleriani, M. (ed.), *The Structures of Practical Knowledge*. Dordrecht: Springer, pp. 421–73.

Notes

[1] No author can be univocally identified for this addition. However, this text circulated already in the fourteenth century in manuscript form, also as a commentary on *The Sphere* of de Sacrobosco.

[2] For comparison, see *C. Plinii Secundi naturalis historiae libri trigintaseptem. A Paolo Manutio multis in locos emendati. Castigationes Sigismundi Gelenii*, Venetiis, Paulum Manutium, 1559, p. 483, lines 5–12.

[3] For comparison, see Johannes Regiomontanus, *Epytoma Joannis de Monte Regio in almagestum Ptolomei, per Johannem Hamman de Landoia*, Venetiis, 1496, Liber tertius, Prop. XXII, 3v.

[4] Additions of text extracts from Pliny's *Natural History* as well as from the works of Regiomontanus were definitely already long used in 1561 when the treatise discussed here was published. In this respect, this treatise is merely an example of adoptions of innovations that emerged in previous editions. Many of the treatises that belong to this corpus are compounded by a number of texts that can easily go up to twelve or even more.

[5] On the one side, the knowledge retrieved should be specific and detailed and, on the other, no abstract inference, uncorroborated by the analyzed sources, should surreptitiously enter the data sets. Developing a data

model means striking a balance between representing the knowledge accurately, but not overstating our claims.

6 The electronic copies of the books often include additional pages with copyright and licensing information. Nevertheless, the length of the actual PDF file turns out to be the most reliable source concerning the length of the analyzed treatises.

7 Since the time of writing, we have already extended the model to include translations of texts and, consequently, annotated translations, translation fragments, and partial translations.

8 In a later upgrade, we expanded the database to enable linking to further electronic resources such as the CERL Thesaurus (https://thesaurus.cerl.org).