MODELLING SCHOLARLY PUBLICATIONS FOR SUSTAINABLE WORKFLOWS

KLAUS THODEN, MAX PLANCK INSTITUTE FOR THE HISTORY OF SCIENCE

ELPUB, 23RD INTERNATIONAL CONFERENCE
2-4 JUNE 2019, MARSEILLE
OUTLINE

• Edition Open Access
• Document Modeling
• Use Case: Recovering the Back-Catalogue
Edition Open Sources

Edition Open Sources (EOS) pioneers a new paradigm in publishing, devoted to open access and high-quality peer review. This collaborative scholarly endeavor publishes academic editions of primary sources in the history of science in online, digital and print formats and offers post-doctoral fellowships in the digital humanities.

Max Planck Research Library

This platform provides open access to the publications of the Max Planck Research Library for the History and Development of Knowledge. The aim is to disseminate peer-reviewed results of scholarly work to a broad audience rapidly and at low cost. The volumes presented are directed at scholars and students in a wide range of disciplines and are available online, as print-on-demand books and open-access publications.
MAX PLANCK RESEARCH LIBRARY FOR THE HISTORY AND DEVELOPMENT OF KNOWLEDGE

Shifting Paradigms
Thomas S. Kuhn and the History of Science
Alexander Blum, Konstanze Gauvorgu, Christian Jost, Jürgen Renn (eds.)

The Globalization of Knowledge in History
Jürgen Renn (ed.)

Algebra in Cuneiform
Introduction to an Old Babylonian Geometrical Technique
Jean Hoyrup

mprl-series.mpg.de
Kapitel 1
Alvaurus Thomas und sein Liber de triplici motu

1. Forschungssübersicht zur Literatur über Alvaurus Thomas

contrary?

This remark can be extended to other periods and intellectual shifts. For instance, in an article on Kuhn's employment of his epistemological categories to the emergence of quantum theory ("a scientific revolution par excellence"), Jochen Büttner, Jürgen Renn and Matthias Schemmel argued, against the gestaltic-switch thesis, that "crisis" was the outcome rather than the source of theory discontinuity and that crisis might even count as a feature of "normal science" (Büttner et al. 2003, 56).

**Controversy versus Linguistic / Conceptual Misunderstanding**

Kuhn's conversion-like treatment of the Copernican paradigm shift downplays the argumentative strategies employed in the controversy over the heliocentric and geokinetic theories. Scholars' choice between terrestrial mobility and immobility was indeed complex and depended on the weight they attached to special aspects at the expenses of others, as well as on their philosophical and cultural choices, and their political and religious bias: e.g. the lack of observable stellar parallax and the physical and scriptural difficulties were enough for Brahe to reject terrestrial motion but not for Bruno, Galileo and Kepler who developed very different counterarguments depending on their philosophical backgrounds and convictions.

By contrast, Kuhn tended to treat controversies as mere misunderstandings. For instance, he linked Copernicus's preference for mathematical-astronomical harmony, as opposed to Aristotelian natural philosophy, to his lack of comprehension of the reasons against terrestrial motion resting on terrestrial physics:

But an excessive concern with the heavens and a distorted sense of values may be essential characteristics of the man who inaugurated the revolution in astronomy and cosmology. The blinders that restricted Copernicus' gaze to the heavens may have been functional. They made him so perturbed by discrepancies of a few degrees in astronomical prediction that in an attempt to resolve them he could embrace a cosmological heresy. (CR, 184)

Kuhn's statement has no historical evidence since Copernicus's education, his concerns about scholastic and theological opposition to his hypotheses as well as his first-hand knowledge of Aristotle bear witness to the contrary. Thus, Copernicus's
A NEW PLATFORM

- Why re-develop the platform?
- Overcome shortcomings of existing platform:
  - Hard to extend functionality
  - Database flaws, hard-coded elements
  - Cumbersome document processing workflow
- Enhance re-usability of publications
FORMAT CHANGES THROUGH THE YEARS
SUSTAINABLE?

• Little thought about the sustainability of the sources
• Little compliance with existing standards
THINK DATA, NOT PAPER!

• Find ways to exploit the contents and structure of publications

• Prepare for further evaluation of publications, other than reading
  • Citation networks
  • Natural Language Processing
  • Named Entity Recognition
ABSTRACT MODEL OF A SCHOLARLY PUBLICATION

• Create a blueprint for scholarly monographs
• What are the atoms of scholarly literature?
• And in which configuration should they be allowed to occur?
ABSTRACT MODEL OF A SCHOLARLY PUBLICATION

apple: <graphic url="/inline/A.jpg"/>, or a penguin: <graphic height="8mm" url="/inline/tux.png"/>.

This is a regular paragraph. Words can appear in <hi rend="italic">italics</hi>, - and some characters are <hi rend="superscript">superscript</hi> or <hi rend="subscript">subscript</hi>. There is also <hi rend="math">Math font</hi> for additional symbols: <hi rend="math">⊕</hi>. In some cases, you can use <hi rend="bold">E0A</hi>bold</hi> directly. <hi rend="smallcaps">Small caps</hi> are also possible.

If you want to include a longer quote in a text, use the block quote feature. And inside that we also introduce you to how you can insert footnotes (<bibl><ref target="#Piaget_1985" type="authoryear">Piaget 1985</ref></bibl>).<p>
<quote>A rather short block quote.<note n="1" place="bottom" xml:id="chap03_ftn1"/><bibl><ref target="#Kaulbach_1960" authoryear"/> citedRange from="320" to="322"/></bibl></note></quote>

Poems<note n="2" place="bottom" xml:id="chap03_ftn2"/>If you
ABSTRACT MODEL OF A SCHOLARLY PUBLICATION

```xml
linebreak =
    element lb {empty}

markup =
    hi

hi =
    element hi {
        attribute rend { "italic" | "superscript" | "subscript" | "math" | "bold" | "struck" | "spaced" | "smallcaps" },
        ( text | index | indexanchor | ref )+
    }

theorem =
    element ab {
        attribute type { "theoremdeclaration" },
        attribute xml:id { xsd:NCName },
        text
    }
```
I believe that the time is ripe for significantly better documentation of programs, and that we can best achieve this by considering programs to be works of literature. Hence, my title: »Literate Programming«

Donald E. Knuth
Nel 1450, l’anno del Giubileo, si reca a Roma, riceve da Niccolò V il cappello cardinalizio, viene nominato Vescovo-Principe di Bressanone e legato pontificio per la Predicazione del Giubileo in Germania. Il 1450 è un anno molto proficuo dal punto di vita della produzione letteraria del cardinale. Durante quest’anno scrive un opuscolo matematico, il *De quadratura circuli*. E un’altra opera importante, inaugurato dal dialogo, vale a dire i tre libri *De sapientia divina*, *De sapientia humana* e *De sapientia divina*; IV. *De sapientia divina*. Nel 1451 e 1452 Cusano viaggia insieme ad un apostolico insieme a una trentina di chilometri, per lo più a dorso di un cavallo, dalla Germania alla città dell’Austria.
DOCUMENT PROCESSING

- Legacy Data
- EOATeX
- Office Documents

TEI-P5 XML

- PDF
- EPUB
- Online
RECOVERING THE BACK-CATALOGUE
### RECOVERING THE BACK-CATALOGUE

<table>
<thead>
<tr>
<th>Title</th>
<th>Serie</th>
<th>Number</th>
<th>Datepublished</th>
<th>Published</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guidobaldo del Monte's Mechanicorum Liber</td>
<td>Sources 1</td>
<td></td>
<td>Dec. 16, 2010</td>
<td>(</td>
</tr>
<tr>
<td>Herausforderung Energie</td>
<td>proceedings 1</td>
<td></td>
<td>June 3, 2011</td>
<td>(</td>
</tr>
<tr>
<td>The Globalization of Knowledge in History</td>
<td>studies 1</td>
<td></td>
<td>July 20, 2012</td>
<td>(</td>
</tr>
<tr>
<td>Non-Isothermal Kinetic Methods</td>
<td>textbooks 1</td>
<td></td>
<td>April 24, 2013</td>
<td>(</td>
</tr>
<tr>
<td>The Equilibrium Controversy</td>
<td>sources 2</td>
<td></td>
<td>Jan. 24, 2012</td>
<td>(</td>
</tr>
<tr>
<td>Catalysis for the Conversion of Biomass and Its Derivatives</td>
<td>proceedings 2</td>
<td></td>
<td>Feb. 28, 2013</td>
<td>(</td>
</tr>
<tr>
<td>Research and Pedagogy</td>
<td>studies 2</td>
<td></td>
<td>Sept. 10, 2013</td>
<td>(</td>
</tr>
<tr>
<td>Algebra in Cuneiform</td>
<td>textbooks 2</td>
<td></td>
<td>Dec. 8, 2017</td>
<td>(</td>
</tr>
<tr>
<td>Bernardino Baldi’s In mechanica Aristotelis problematica exercitatioes</td>
<td>sources 3</td>
<td></td>
<td>July 11, 2011</td>
<td>(</td>
</tr>
<tr>
<td>New Vistas on Old Problems</td>
<td>proceedings 3</td>
<td></td>
<td>April 29, 2013</td>
<td>(</td>
</tr>
<tr>
<td>Wissensgeschichte der Architektur</td>
<td>studies 3</td>
<td></td>
<td>Dec. 2, 2014</td>
<td>(</td>
</tr>
<tr>
<td>Bernardino Baldi &quot;In mechanica Aristotelis problematica exercitatioes&quot;</td>
<td>sources 4</td>
<td></td>
<td>Aug. 8, 2011</td>
<td>(</td>
</tr>
<tr>
<td>Guidobaldo del Monte (1545–1607)</td>
<td>proceedings 4</td>
<td></td>
<td>May 23, 2013</td>
<td>(</td>
</tr>
<tr>
<td>Wissensgeschichte der Architektur</td>
<td>studies 4</td>
<td></td>
<td>Dec. 2, 2014</td>
<td>(</td>
</tr>
<tr>
<td>The Role of Gravitation in Physics</td>
<td>sources 5</td>
<td></td>
<td>Feb. 14, 2011</td>
<td>(</td>
</tr>
</tbody>
</table>
CONCLUSION

• Create an abstract document model
  • Documentation for editors
  • Automatic validation of XML documents
• Integrate model into publication workflow
• Apply model to legacy data for a uniform document base
THANK YOU

- edition-open-access.de
- edition-open-sources.org
- mprl-series.mpg.de
- mail: eoa@mpiwg-berlin.mpg.de
- code: github.molgen.mpg.de/EditionOpenAccess
- twitter: @EditionOA
- blog: eoa.hypotheses.org