

Digital Scholarly Editions and Libraries*

Szilvia Maróthy (ELTE University)

mthy.szilvi@gmail.com

ILIDE Conference, Jasná, 16–18. April, 2018.

„...we are still at an early stage of Humanities Computing in terms of the development of methodology. People speak about revolutions and the immediate changes that we can see on the surface of things. But deep changes require, I will not say tens of years, but hundreds.”

Tito Orlandi in *Computation and the Humanities. Towards an Oral History of Digital Humanities*, eds. J. Nyhan, E. Flinn, 2016.

<https://link.springer.com/content/pdf/10.1007%2F978-3-319-20170-2.pdf>

* Supported by the ÚNKP-17-3 New National Excellence Program of the Ministry of Human Capacities 



About the Project

- *New National Excellence Program for doctoral candidates*
- September 1. 2017. – June 30. 2018.
- collect, analyze, describe Hungarian DSEs

Aims:

- popularize DSEs, make them more visible
- give a historical overview (from 1993 to present)



Catalogues

Patrick Sahle's annotated bibliography (420 items)

<http://www.digitale-edition.de/>

Greta Franzini's catalogue on GitHub (planned: 350 items +HU)

- detailed (49 field)
- opened for contributors
- simple CSV on GitHub

https://github.com/gfranzini/digEds_cat/



What is DSE?

Digital?

- “A digitized edition is not a digital edition.”
- “A digital edition cannot be given in print without significant loss of content and functionality.”
- “Scholarly digital editions are scholarly editions that are guided by a digital paradigm in their theory, method and practice.”

Sahle, “What is a Scholarly Digital Edition?”

<https://www.openbookpublishers.com/htmlreader/978-1-78374-238-7/ch2.xhtml>

Digital?

- digitized not equals born digital – BUT:
 - how do we manage reborn digital?
 - 1) a printed edition digitized, OCR-d, corrected
 - 2) get a new structure, encoding
 - 3) enriched with LOD
 - 4) get a multi-functional representation (e. g. flexible, synoptic, stochastic)
 - born digital (some WYSIWYG editor)
 - born digital (e. g. XML+PHP)



Scholarly?

- 1) "historical documents": editing is concerned with documents which are already there.
- 2) "representation": covers (abstract) representation as well as presentation (reproduction). As I use to say: transmedialization (representation by data) and medialization (presentation by media). Publishing descriptive data (e.g. metadata) without reproduction is not critical editing.
- 3) "critical / scholarly" (erschließend): reproduction of documents without critical examination is not scholarly editing. A facsimile is not a scholarly edition.

Sahle, <http://www.digitale-edition.de/vlet-about.html>



Scholarly?

Catalogue of DSEs of G. Franzini:

Account of textual variance (0, 0.5, 1)

1 = This edition is “based on full-text transcription of original texts into electronic form”.

Value of witnesses (0, 0.5, 1)

1 = The witnesses are traditional philological primary sources (including manuscripts, inscriptions or papyri).

Edition?

“... in digital editions the encoded texts themselves are the most important long-term outcome of the project, while their initial presentation within a particular application should be considered only a single perspective on the data. [...] Furthermore, all initial presentations are also ephemeral, bound to be either modified over time as technologies and forms of digital publishing change, or languish in obsolescence on a forgotten server.”

M. Turska, J. Cummings, S. Rahtz:

Challenging the Myth of Presentation in Digital Editions <https://journals.openedition.org/jtei/1453>

“... completely decouple the edition data from the visualisation mechanism.”

Roberto Roselli del Turco: <http://books.openedition.org/obp/3423>



The TEI Principles

- 1) TEI XML focuses on the meaning of text, rather than its appearance
- 2) TEI XML is independent of any particular software environment
- 3) TEI XML was designed by and for the scholarly research community, which is also responsible for its ongoing development

Lou Burnard, “What is the Text Encoding Initiative?”

<http://books.openedition.org/oep/679>



Edition?

- What is an edition?
 - the encoded text (a critical representation of a historical resource, e. g. in TEI XML)
 - the encoding + presentations (e. g. for WWW)
 - the encoding + presentations + software environment for presentations



Edition?

Publishing

- a single XML e. g. on GitHub is a publication?
<https://github.com/marothyszilvi/oms>
 - publisher: GitHub
 - presentation: e. g. web browser

Archiving

- what to preserve?



Who reads DSEs?

- Scholars?
 - not reliable
 - always changing
 - difficult to cite
- General public?
 - too much information
 - difficult to read
 - ?responsive design
 - ?search engine optimization



“The Battle We Forgot to Fight”

[Roberto Rosselli Del Turco]

Scientific criteria for DSEs:

- sustainability
- usability: e. g. responsive design, search engine optimization
- success and acceptance: reliable, easy to cite, easy to find and identify (e. g. DOI or other PID, Canonical Text Services, SignPosting)



About the Hungarian Corpora

- total: 48 editions (at the moment)
- 1993–2004 HTML
- after 2004 TEI XML (P4 and P5)
- all of them are Open Access
- some of them are Open Source
- mostly Hungarian, but Latin, German, Slovakian etc. also presents
- period: early modern, long 19th century, modern



Best practices in sustainability

- Textology Research Group
HAS—Debrecen University
<http://textologia.unideb.hu/>
 - the infrastructure is provided by the University Library
- DigiPhil
Petőfi Literary Museum
<http://digiphil.hu/>
 - the infrastructure is part of the Museum's

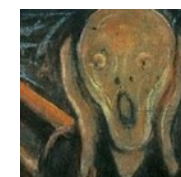
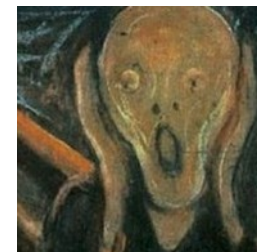
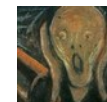


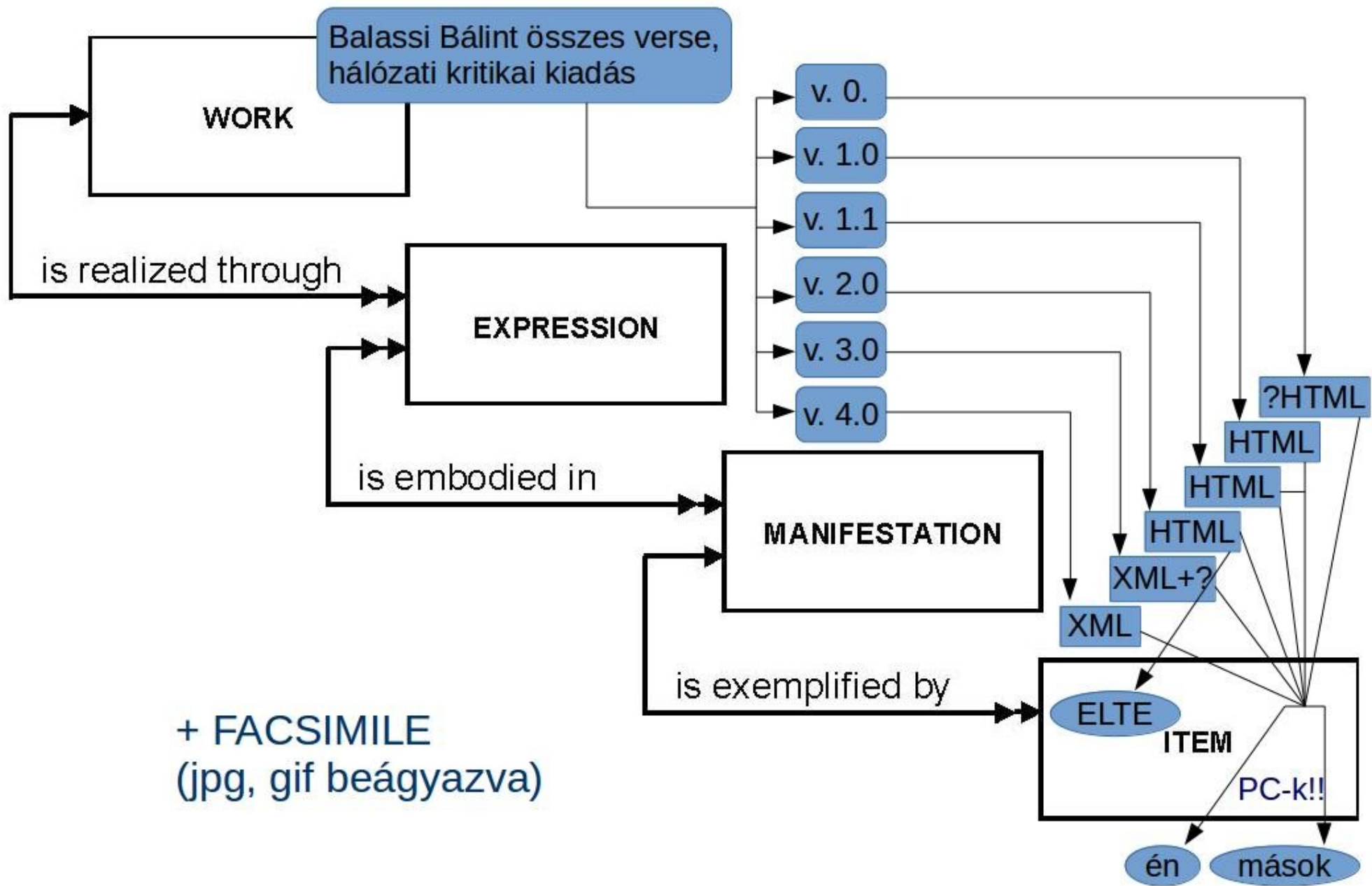
An example...

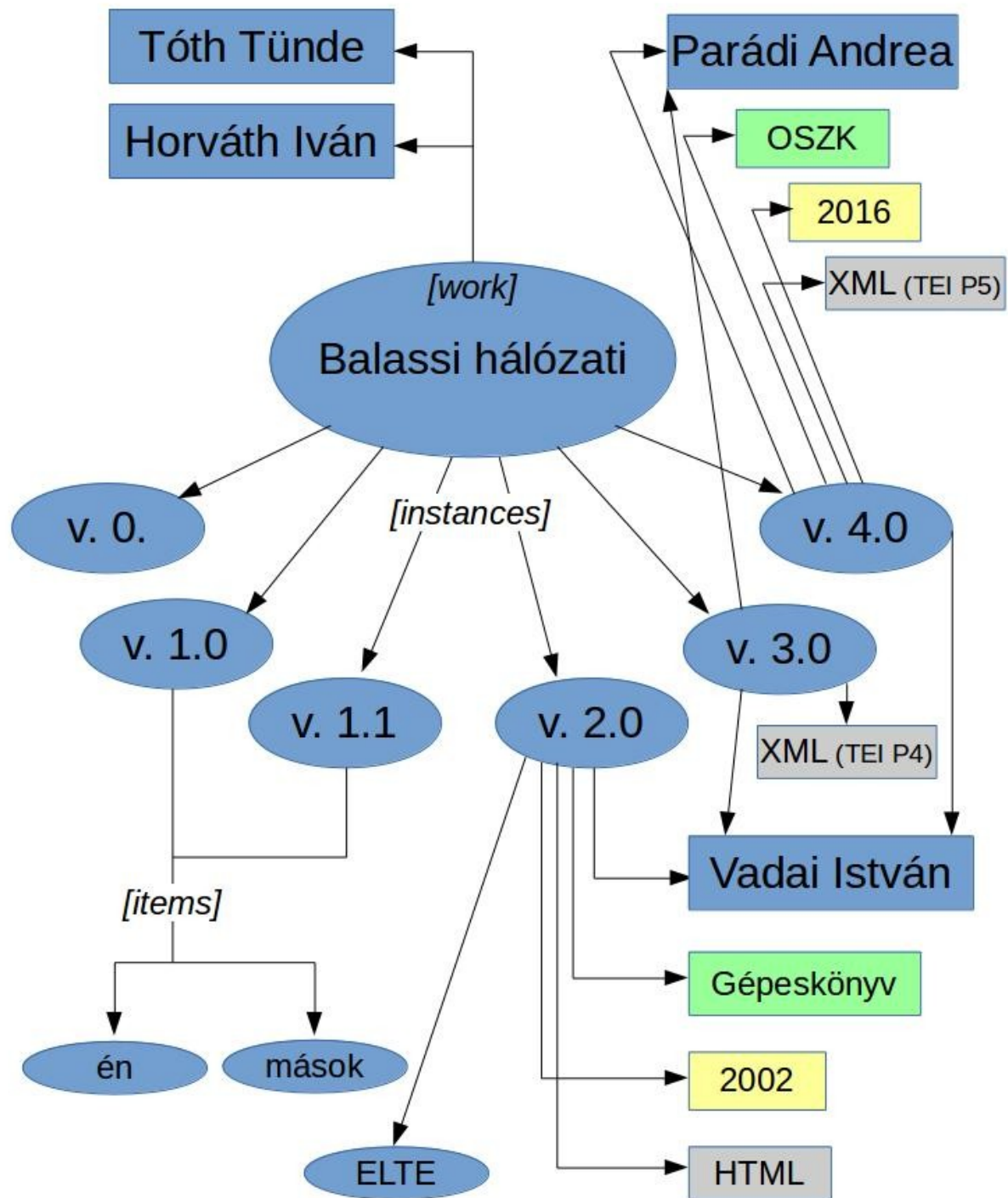
Versions and/or Editions?

“Balassi Bálint összes versei”

- v. 0. – 1993–1996 → Gopher then HTML,
(I found the HTML from 1996 on Internet Archive)
- v. 1.0 – 1998 → HTML, ?files owned by authors
- v. 1.1 – 1999 → HTML, files: authors and me :)
rather overwritten than versioned
- 2.0 – 2002 HTML, online (ELTE-server)
- 3.0 – 2004 → TEI XML files (TEI P4),
WWW representation lost (?)
- 4.0 – 2016 → TEI XML file (TEI P5)
no representation yet







The Catalogue



International:

https://github.com/gfranzini/digEds_cat

Search and visualizations:

<https://dig-ed-cat.acdh.oeaw.ac.at/>

Hungarian:

https://github.com/marothyszilvi/digEds_hun



What's next?

- publish the descriptions of the catalogue using library standards (RDA? BIBFRAME?)
- convince institutions (e. g. ELTE DH Centre, National Széchényi Library) to preserve these “books” and find out archival strategies together
- popularize editions for scholars, students and the general public
- help creators in sustainability (e. g. with recommendations)

