Think Big: Digital Scholarly Editions as / and Infrastructure

Anna-Maria Sichani
DiXiT fellow - Huygens KNAW
PhD University of Ioannina-GR
amsichani

DiXiT Workshop “The educational and social impact of Digital Scholarly Edition”
AIUCD 2017 Conference
Roma, January 24, 2017
"Houston, We've Had a Problem"

Digital scholarly editing (DSE) projects

- project-based
- in all shapes, sizes, and flavours
- one specific research question - editorial school
- edition → THE end-output
- development/planning of the project: deliverables-oriented
- well-defined deliverables: well-defined audience
How do we want to develop & use our digital editions?
“Think Big” approach

- Infrastructure - system logic
- Modular design approach

*Golden rules*

Digital editions should not be planned or equated to the one final product

But as an open, extendable, repurposable systems or processes, components with unexpected reuses, applications and diverse audiences.
Infrastructure logic

✓ a material and experiential presence that is embedded in the practices and experience of research, which builds on and enhances that which already exists, that unites creators, stakeholders and users

✓ structures, systems, and facilities serving and supporting for a community - project to function

✓ facilities, resources and related services used by the scientific community to conduct top-level research in their respective fields (research infrastructure)

✓ a thing that shifts into the background and is invisible

✓ facilitates cooperation

✓ enable to share knowledge and resources
Modular design

✓ an approach that subdivides a system into *smaller parts* called modules or ‘skids’, that can be *independently created* and then *used in different systems*.

✓ a modular system can be characterized by functional partitioning into discrete scalable, *reusable modules*; rigorous use of well-defined modular interfaces; and, making use of industry *standards* for interfaces.

✓ reduction in cost (due to less customization & shorter learning time)

✓ flexibility in design

✓ *n*-augmentation (adding new solutions by merely plugging in a new module)

✓ exclusion
*Golden rules*

**Tech choices**

- Standards! Standards! (TEI, metadata)
- Platform & technology independent workflows

**Operational choices**

- Blue-sky thinking: imagine the unconventional, the unexpected
- Occam’s razor: preference for simplicity in the scientific method
- Agile planning: focus on processes (VS Fountain – focus on deliverables)

**Robust documentation**

- Open access and open source policies

- Open-ended deep access model → reuse
- Clear credit attribution & respect licensing choices of creators
- Keeping track / communicate your reuse
How it would look like?
ways of reuse → aspects of value:

1. base/raw material for new kind of research and scholarship

2. ideal test-bed for new computational/analytical approaches (text analysis, topic modelling)

3. easy way to produce derivatives and varying outputs (ePub, PDF) ⇒ add-ons / on demand services ⇒ multiple uses ⇒ specialized audiences

4. reference material for teaching purposes

5. great potential for cost and time-avoidance

6. distributed archiving solution
What could be seen as modular component?

- digitized facsimiles & metadata
- source files
  - encoded XML
  - related schemas & ODDs
  - scripts for processing-visualization
  - transformation scripts and query algorithms
  - entity relational models
- project’s documentation
  - (successful or not) funding proposals
  - feasibility studies
  - editorial conventions
  - budget plans etc.
Examples

an online work environment for uploading facsimiles, transcription, annotation, publishing of digital editions

[ build upon the model of a decentralised, distributed architecture of openly available sources, services and functionalities (virtual research environment) ]
a sustainable distributed research environment offering open source tools and services to support humanistic scholars during the entire process of research, especially in digital scholarly editing.

the vision of a digital ecosystem is based on the open source idea, allowing free exchange of tools and data, whereby adaptation concerning discipline-specific needs is made possible.
Sounds challenging - Let’s work on this!

Scenario 1

National Library → National Literature corpus

Scenario 2 ???
“Think Big” approach

- flexible layered model of interlinked components to create a robust workflow for the creation, publication, and reuse of scholarly resources

- a modularised structure not only in visualisation/processing, but also in creating and using resources and components
Any ideas, suggestions?

Thank you!

anna-maria.sichani@huygens.knaw.nl

amsichani