

# A Workflow Analysis Perspective to Scholarly Research Tasks

Marijn Koolen - KNAW Humanities Cluster, Netherlands  
Sanna Kumpulainen - Tampere University, Finland  
Liliana Melgar-Estrada - Utrecht University, Netherlands

Virtual DARIAH Event 2020 - 10-11-2020





Common Lab Research Infrastructure  
for the Arts and Humanities

# Research question: what and how to evaluate digital infrastructures?

- Why evaluate systems?
  - Know that we're actually supporting research, gradually improve tools
- What do we evaluate?
  - Individual tools? Typical combinations of tools? Entire infrastructure?
- How do we evaluate?
  - Traditional LIS evaluation frameworks take information system viewpoint:
    - laboratory models (functionality), testing (usability)
  - Our approach: focus evaluation on **scholarly processes/workflows**
- Why focus on scholarly primitives?
  - Operationalise the activities in digital research
  - Understand workflow as sequences of activities
  - And link them to tools

# Operationalizing Evaluation

What does/should “evaluate” in a workflow perspective mean?

## 1. Simplistic approach

- a. Is activity X supported by the infrastructure? (what do we mean by “supported”)

## 2. Slightly more elaborate approach

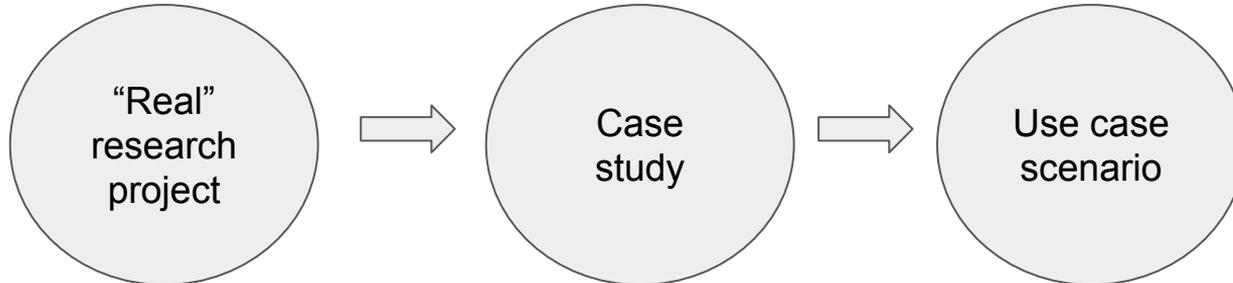
- a. How well is activity X supported? (what do we mean by “how well”?)
- b. How well = How many interactions are required for activity X?
- c. How well = How many scholars use the infrastructure for activity X?

## 3. Layered view

- a. How well is activity X supported?
- b. How well are sequences/combinations of activities that include X supported?
- c. How well are workflows of type Y supported?

# Method

- Qualitative approach
- Interviewing DH researchers, focus on specific research
  - What they did, why, how and who did what (collaboration aspects)
- reductionistic/simplistic approach to very rich and complex processes
  - isolate main tasks and model them in typical workflows.
- From “case studies” to “use case scenarios”
  - We depart from real research projects carried out by scholars
  - Our unit of analysis is a small/focused “research project”



# Research Questions

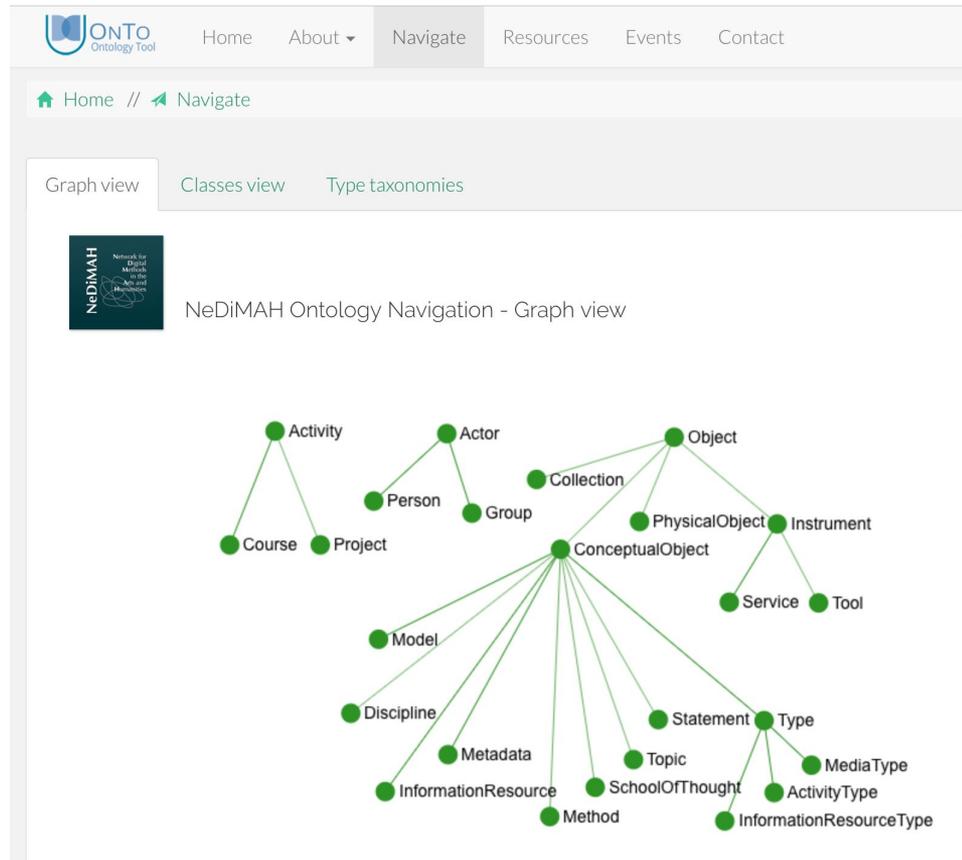
- RQ1. What are the characteristics of research workflows in the selected cases of digital humanities research projects?
- RQ2. Is it possible to identify commonalities in the sequences of, or transitions between, research activities across similar digital humanities research projects?
- RQ3. What are sequences or transitions that need further investigation?

# Case Studies

- Two cases studies:
  - RP1:
    - **Topics:** management of migration flows in Europe, 1950-1990
    - **Aim:** demonstrate bi-directional influence: politics <-> science
    - **Data:** 200 pubs of international committees and journals on European migration
    - **Approach:** identify themes and network of actors over time
  - RP2:
    - **Topic:** maritime careers in Dutch East India Company, 1680-1800
    - **Aim:** analyse social mobility of Dutch native and migrant workers
    - **Data:** 800,000 contracts in pay ledgers from 3400 journeys NL <-> East Indies
    - **Approach:** reconstruct careers, promotions, demotions
- Focus on historical research
  - We analysed research processes in other humanities disciplines in Melgar et al. (CHIIR 2017)

# Previous work on modelling scholarly processes

- Scholarly Primitives (Unsworth, 2000)
- Methodological Commons (McCarty & Short 2002)
- Scholarly Information Practices (Palmer et al. 2009)
- TaDiRAH (Borek et al. 2017)
- NeDiMAH Methods Ontology (Hughes et al. 2015)
- Scholarly Ontology (Pertsas & Constantopoulos 2017)



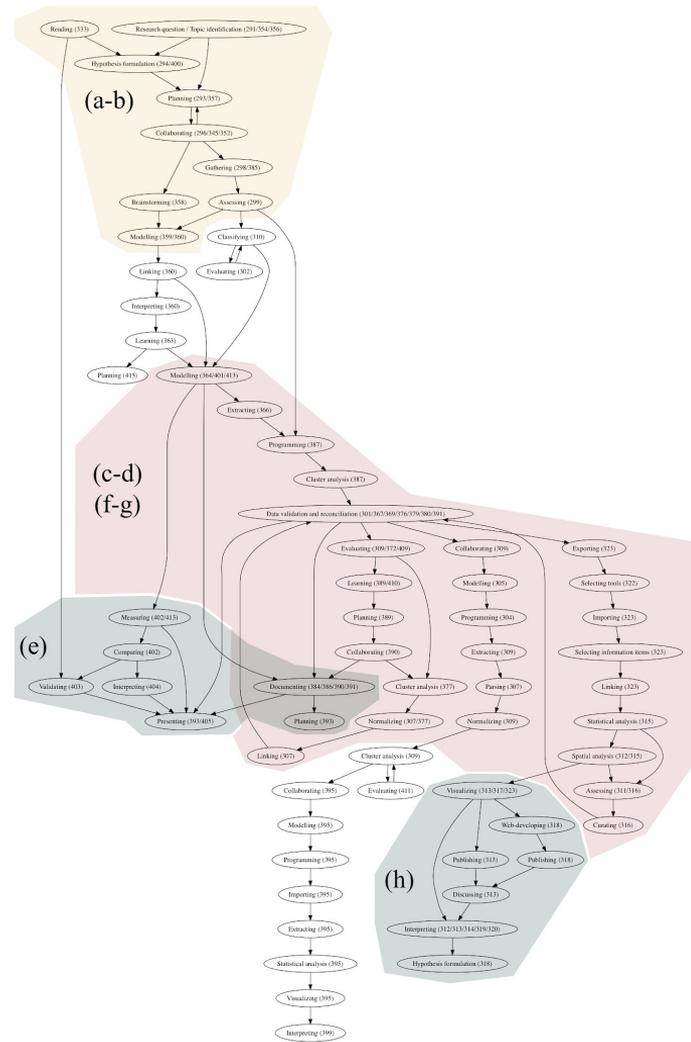
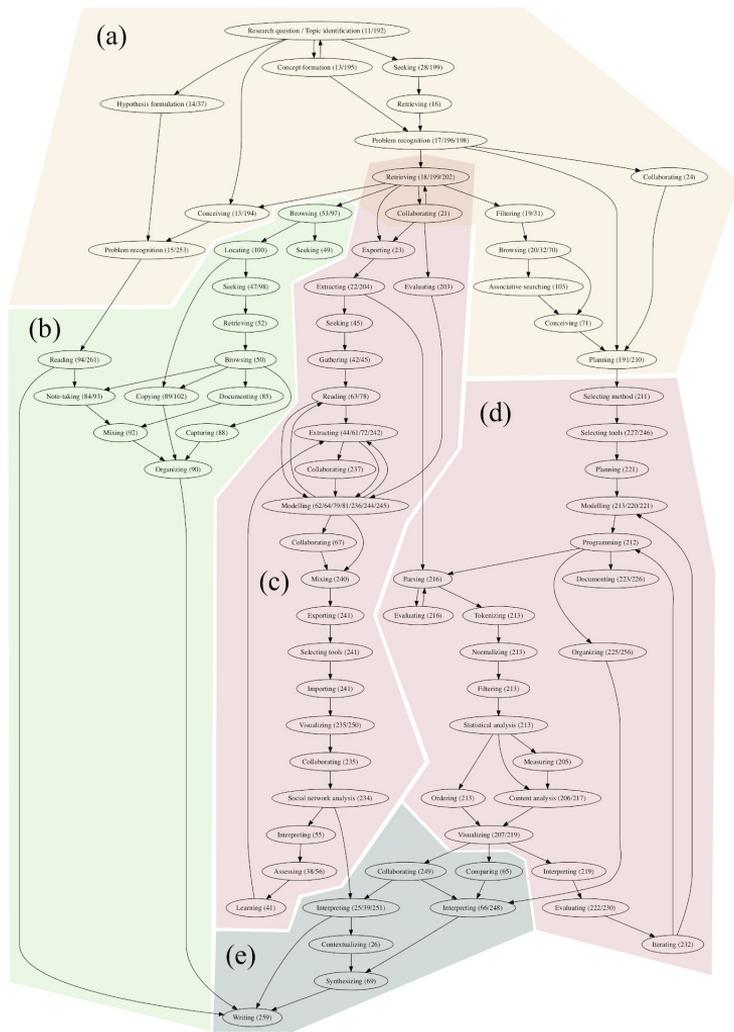
# Conceptualizing Workflows

- *Activity*: a unique and distinct **type** of human mental or physical action
- *Transition*: switch from one instance of an activity to an instance of another activity
- *Workflow*: a coherent set of *activities* and *transitions* leading to a certain research outcome (a paper/presentation/report)

# Coding Process

- Transcribe and segment interviews
  - Bottom-up coding by all three authors individually
  - Merge codes and map to NeMO where possible
- Result:
  - 80 codes, of which 69 same as NeMO activity (see <https://zenodo.org/record/3626814>)
  - 11 new activities: copying, importing, normalizing, note-taking, reading, sampling, ...
- Identify transitions between coded instances of activities
  - Use GraphViz to construct workflow visualization
  - Checked workflows with interviewees for correctness









# Implications for System Design and Evaluation

- Identify potential hurdles for users to switch between activities
  - Design and evaluate support for such transitions
- In sciences, transitions are typically modeled as **pipeline**
  - Many parts of humanities workflows are **exploratory**, don't fit this metaphor
  - Many activities are highly interactive and rapidly alternate
    - esp. between high and low cognitive load activities: **reading**, **extracting**, **modeling**
- In the Humanities, a **Palette** metaphor is more appropriate:
  - Desire to have 'everything' in single screen for making associations
  - Multiple functionalities and views in single screen,
  - **Exploratory**: researcher picks next activity from palette based on all available information
- Future direction: focus on support at the intersection between activities

# Thank You!

Questions?

More extensive slides: <http://bit.ly/chiir-2020-workflows>

Auxiliary materials: <https://zenodo.org/record/3626814>

# References (1 of 2)

Borek, L., Perkins, J., Schöch, C. and Dombrowski, Q., 2017. TaDiRAH: A case study in pragmatic classification.

Pia Borlund and Peter Ingwersen. 1997. The development of a method for the evaluation of interactive information retrieval systems. *Journal of documentation* 53, 3 (1997), 225–250.

Pia Borlund and Jesper W. Schneider. 2010. Reconsideration of the simulated work task situation: a context instrument for evaluation of information retrieval interaction. In *Proceedings of the third symposium on Information interaction in context*. ACM, New York, NY, USA, 155–164.

Flanagan, J.C., 1954. The critical incident technique. *Psychological bulletin*, 51(4), p.327.

Hughes, L., Constantopoulos, P. and Dallas, C., 2015. Digital methods in the humanities: understanding and describing their use across the disciplines. *A new companion to digital humanities*, pp.150-170.

Järvelin, K., Vakkari, P., Arvola, P., Baskaya, F., Järvelin, A., Kekäläinen, J., Keskustalo, H., Kumpulainen, S., Saastamoinen, M., Savolainen, R. and Sormunen, E., 2015. Task-based information interaction evaluation: The viewpoint of program theory. *ACM Transactions on Information Systems (TOIS)*, 33(1), p.3.

# References (2 of 2)

Kuhlthau, C.C., 1991. Inside the search process: Information seeking from the user's perspective. *Journal of the American society for information science*, 42(5), pp.361-371.

Liliana Melgar, Marijn Koolen, Hugo Huurdeman, and Jaap Blom. 2017. A Process Model of Scholarly Media Annotation. In Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval (CHIIR '17). ACM, New York, NY, USA, 305–308.  
<https://doi.org/10.1145/3020165.3022139>

McCarty, W. and Short, H., 2002. Mapping the field. Report of ALLC meeting held in Pisa, April 2002. See <http://www.allc.org/content/pubs/map.html>.

Palmer, C.L., Teffeau, L.C. and Pirmann, C.M., 2009. Scholarly information practices in the online environment. *Report commissioned by OCLC Research. Published online at: [www.oclc.org/programs/publications/reports/2009-02.pdf](http://www.oclc.org/programs/publications/reports/2009-02.pdf).*

Van Zundert, J., 2012. If you build it, will we come? Large scale digital infrastructures as a dead end for digital humanities. *Historical Social Research/Historische Sozialforschung*, pp.165-186.

# Method

- Unit of analysis:
  - research project with focused RQ, with at least one publication
- Semi-structured interviews using Critical Incident Technique
  - Focused on **experienced tasks** and **task structure**
  - Inspired by **Task-Based Information Interaction** evaluation framework (Järvelin et al. 2015)
  - Focused on activities as characterized by the interviewees (Bernardou et al. 2013)
    - What they did
    - How they did it
    - Why they did it (aims and goals of the activity)
    - Who did what? (collaboration aspects)
  - Also asked **demonstrations** of data, tools, scripts, visualizations
- Two interviewees per project
  - Researcher triangulation (Garijo et al. 2014) + data triangulation (Kumpulainen 2017)

# Method

- Unit of analysis:
  - research project with focused RQ, with at least one publication
- Semi-structured interviews using Critical Incident Technique
  - Focused on **experienced tasks** and **task structure**
  - Inspired by **Task-Based Information Interaction** evaluation framework (Järvelin et al. 2015)
  - Focused on activities as characterized by the interviewees (Bernardou et al. 2013)
    - What they did
    - How they did it
    - Why they did it (aims and goals of the activity)
    - Who did what? (collaboration aspects)
  - Also asked **demonstrations** of data, tools, scripts, visualizations
- Two interviewees per project
  - Researcher triangulation (Garijo et al. 2014) + data triangulation (Kumpulainen 2017)

# Interviews

- Interview scholars using *critical incident technique* (Flanagan 1954)
  - Pick a recent finished project and elaborate on the process
  - Semi-structured, using interview guide (see <https://zenodo.org/record/3626814>)
  - Interviews lasted 90-120 minutes, recorded with both audio and video
  - Consent form for recording interview, storing data and publishing about it
- Interviewees
  - RP1:
    - Political historian + historical source editor
    - Digital historian + data scientist + information specialist
  - RP2:
    - Social-economic historian
    - Social-economic historian + data curator + information specialist

# What have we done until now...

Idea  
formulation

Method  
selection

Design of  
interview  
guide

Conducted  
3 Interviews

Transcribed  
2 interviews

Reading  
research  
project papers

Bottom up coding  
(each one of us)

Agreement on  
codes (Code  
book)

Recoding

# Codebook

Normalized code	NeDiMAH Methods Ontology equivalent	Scope note
Analyzing	Analyzing	Use more specific analysis method if possible
Annotating	Annotating	
Archiving	Archiving	
Assessing	Assessing	
Associative searching	Associative searching	
Bookmarking	Bookmarking	
Brainstorming	Brainstorming	For collaborative idea generation. For solitary idea generation, use conceiving.
Browsing	Browsing	
Capturing	Capturing	
Classifying	Classifying	
Cluster analysis	Cluster analysis	
Collaborating	Collaboration	
Collecting	Collecting	
Commenting	Commenting	
Comparing	Comparing	
Conceiving	Conceiving	
Concept formation	Concept formation	
Consulting	Consulting	
Content analysis	Content analysis	
Contextualizing	Contextualizing	
Copying	has no NeDiMAH equivalent	manually by writer / typing / copy-pasting
Curating	Curating	
Data recognition	Data recognition	
Data validation and reconciliation	has no NeDiMAH equivalent	Includes data mapping
Developing	Developing	Add target (algorithm, tool, website, ...)
Discussing	Discussing	
Documenting	Documenting	

Available at:  
<https://zenodo.org/record/3626814>

# Methodological Reflection

- Issues with our approach
  - Differences in background knowledge of coders
    - Coder with NLP background interprets activities differently than information scientist
  - Levels of activities:
    - **analyzing** vs.
    - **importing, parsing, tokenizing, filtering, statistical analysis**
  - Connection between low and high level activities might be insightful for design
- How to use
  - Conduct evaluation
  - Use transitions as focus in simulated works tasks (Borlund & Ingwersen 1997, Borlund & Schneider 2010)