Deutsche DH, Passau 26.03.2014

Brauchen die Digital Humanities eine eigene Methodologie?

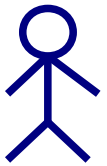
Übersicht

- Aspekte der Operationalisierung geistes- und sozialwissenschaftlicher Fragestellungen
- Beispiel „ePol – Postdemokratie und Neoliberalismus“
- Zusammenfassung

Aspects of operationalization

roles

humanities



computer science

DH project character

	DATA MANAGEMENT	**DATA PROCESSING**
humanities	data modeling: defining entities (E) and their relations (R)	requirements: iterated process of (1) devising research needs, (2) approaches to fulfil them (3) assessment of results
computer science	service provision / translation of ER-models into research infrastructures: • DBMS + GUI • MAXQDA et al. • ...	research: (1) identification, creation, adaption and (2) iterated application, improvment of suitable technologies

Aspects of operationalization

WRONG *(technology-driven)*

- We use *(incidentally)* available / ad-hoc compiled text collections
- We apply generic *(black box)* tools (well) known to us
- *(Afterwards)* We try to establish a meaningful connection between our initial research question and the automatically generated output data

RIGHT *(requirements-driven)*

- We select / compile a text collection by carefully specified criteria beforehand
- We apply procedures best matching our data / research interest. If not available, we develop them based on requirements we identify systematically.
- We use our text collection as basis for a profound validation of our initial research question

Aspects of operationalization

How can we support the process of „operationalization“?

For social sciences

- *Apply / adapt established methodology of qualitative and quantitative research*
  - But, use (large) written text corpora instead of survey data
- Understanding: Identifying / extracting meaningful entities
  - *Complement qualitative research based on texts (e.g. MAXQDA)*
- Explaining / Quantitative paradigm: measuring (causal) relations
  - *Hypothesis development* + *validation* by empirical testing on text collections
- A reconciliation of qualitative and quantitative methods?

Aspects of operationalization

How can we support the process of „operationalization“?

For social sciences and humanities in general

- *Apply requirements engineering and modeling as part of a software engineering process*
  - *Identify domain model and workflows*
  - *Identify key functions and define their implementation*
- Requires **„mutual understanding“ / „common language“** of humanities scholars and computer scientists

Example: *ePol* – Tracking down economization

- Joint research project in the eHumanities; cooperation with Prof. Gary Schaal, department for political theory (HSU)
- Period: June 2012 – May 2015
- 2 x 1,5 academic employees, + student assistants
- licences for longitudinal retro-digitized German newspaper corpora (Zeit, FAZ, Süddeutsche, TAZ)
  → 3.5 Million documents

ePol – Tracking down economization

Research question:

In the debate on post-democracy **political theory** claims that contemporary western democracies tend to justify politics in a **neoliberal manner**. This manner, characterized by means of economization, has become dominant over the past decades.

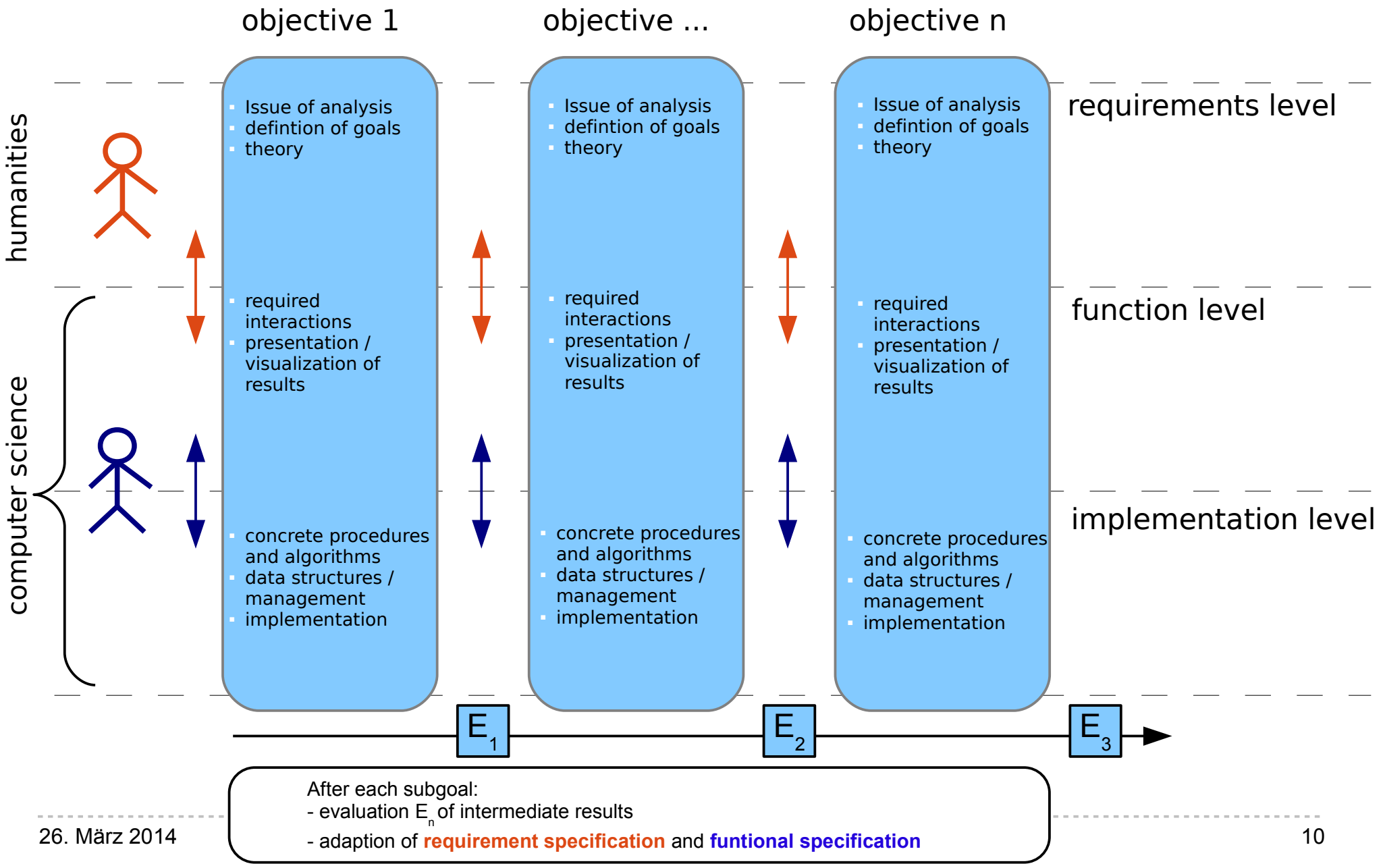
→ ePol-projects: empirical evaluation of this hypothesis for Germany

ePol – der „Ökonomisierung“ auf der Spur

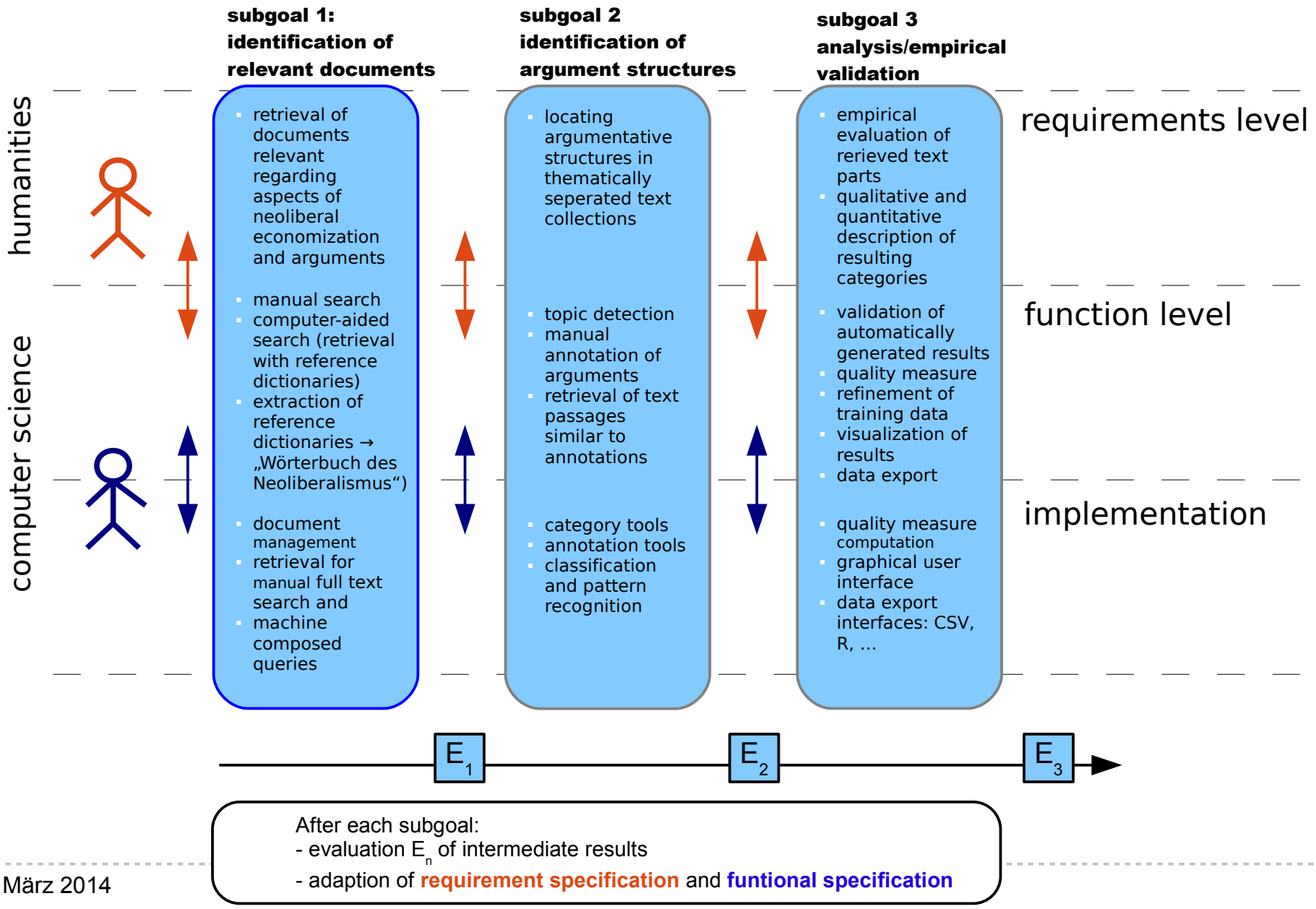
Research issues from the computer science perspective

1) Operationalization of political science research questions for computer-assisted text analyses
2) Development of a technical research infrastructure for large scale text corpora adaptable to heterogenous research interests
3) development of specifically adapted analyses procedures
4) procedures to evaluate computer generated results

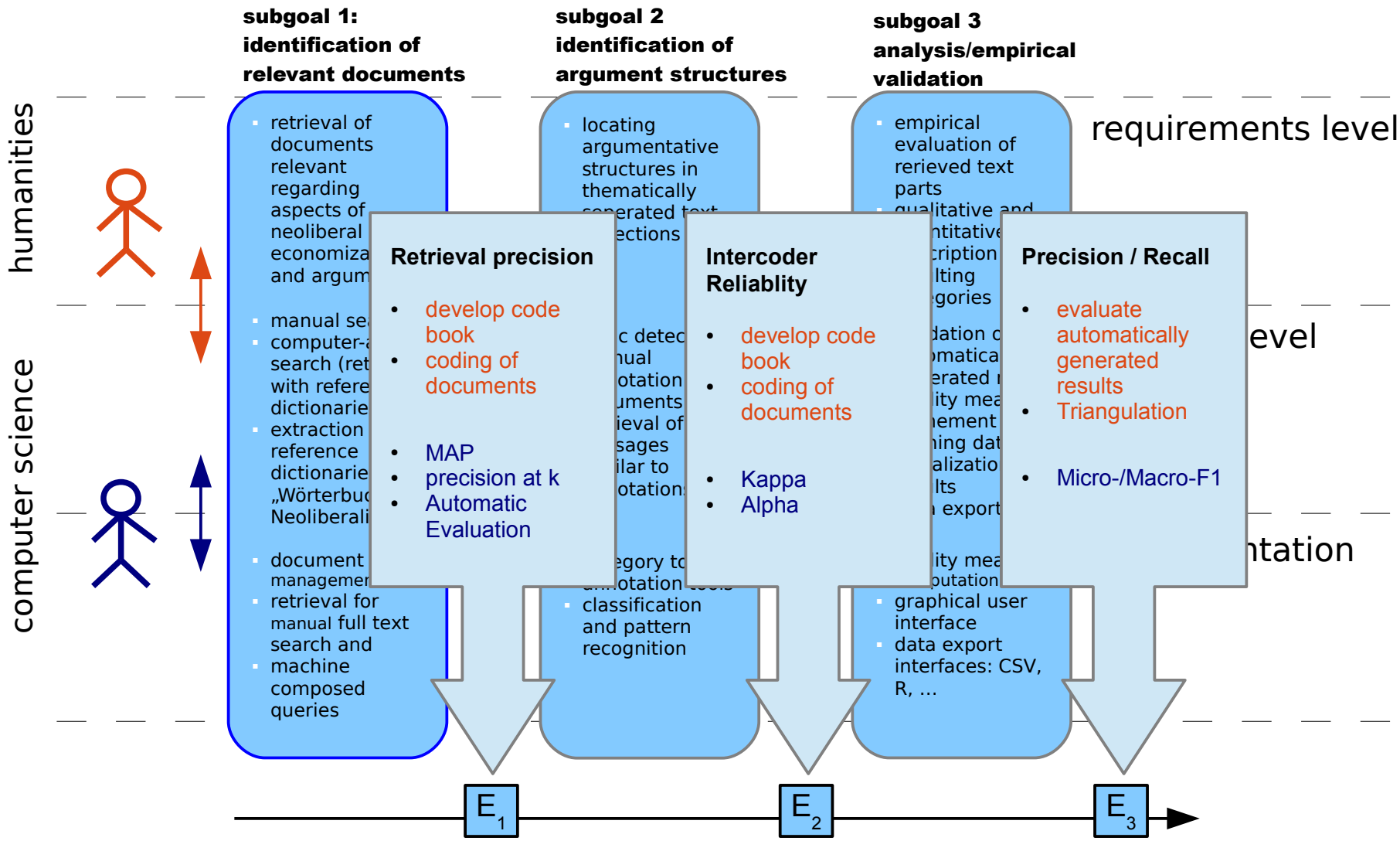
Requirements Engineering



Requirements Engineering



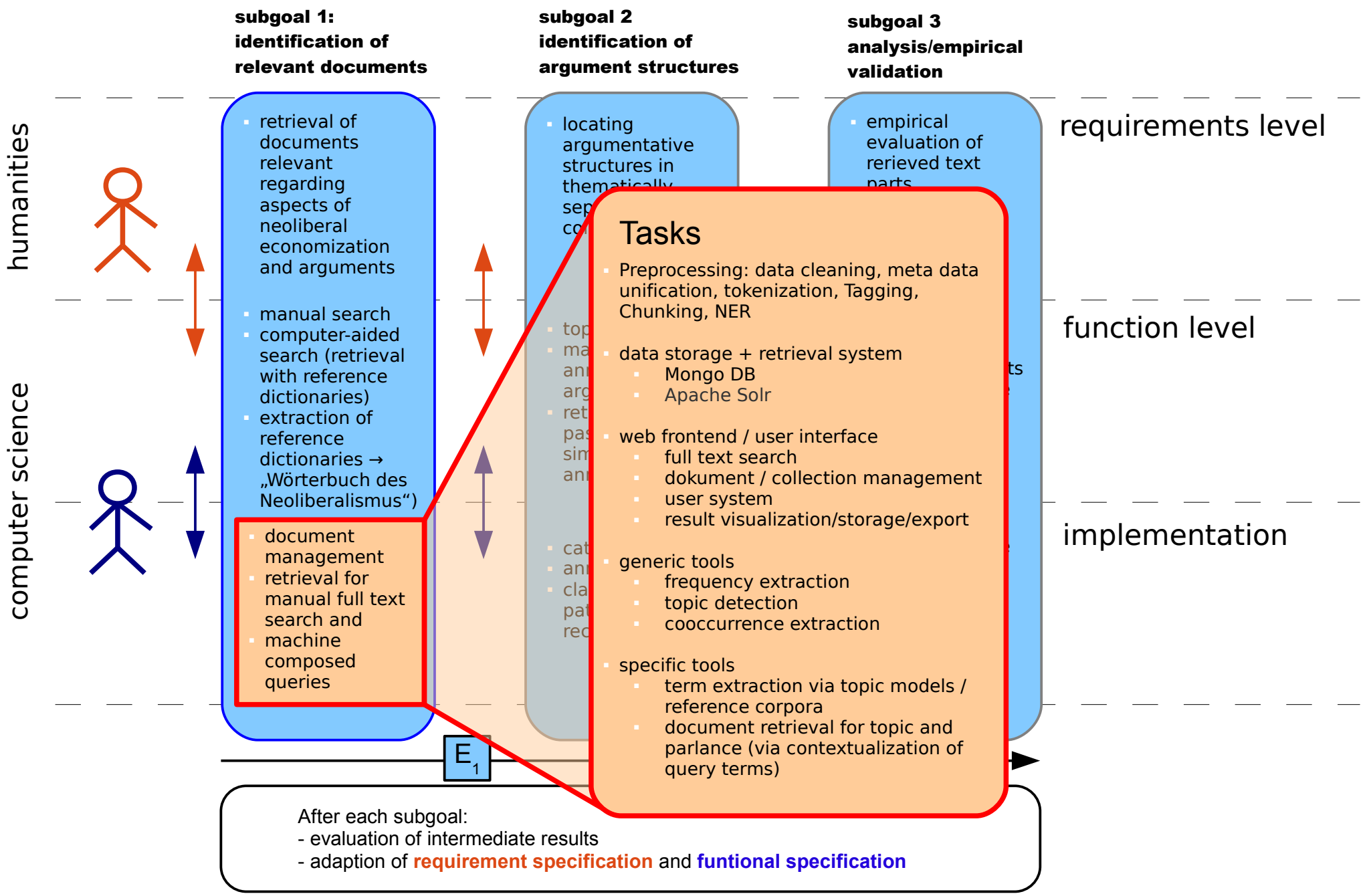
Requirements Engineering



Requirements ePol-project

UNIVERSITÄT LEIPZIG

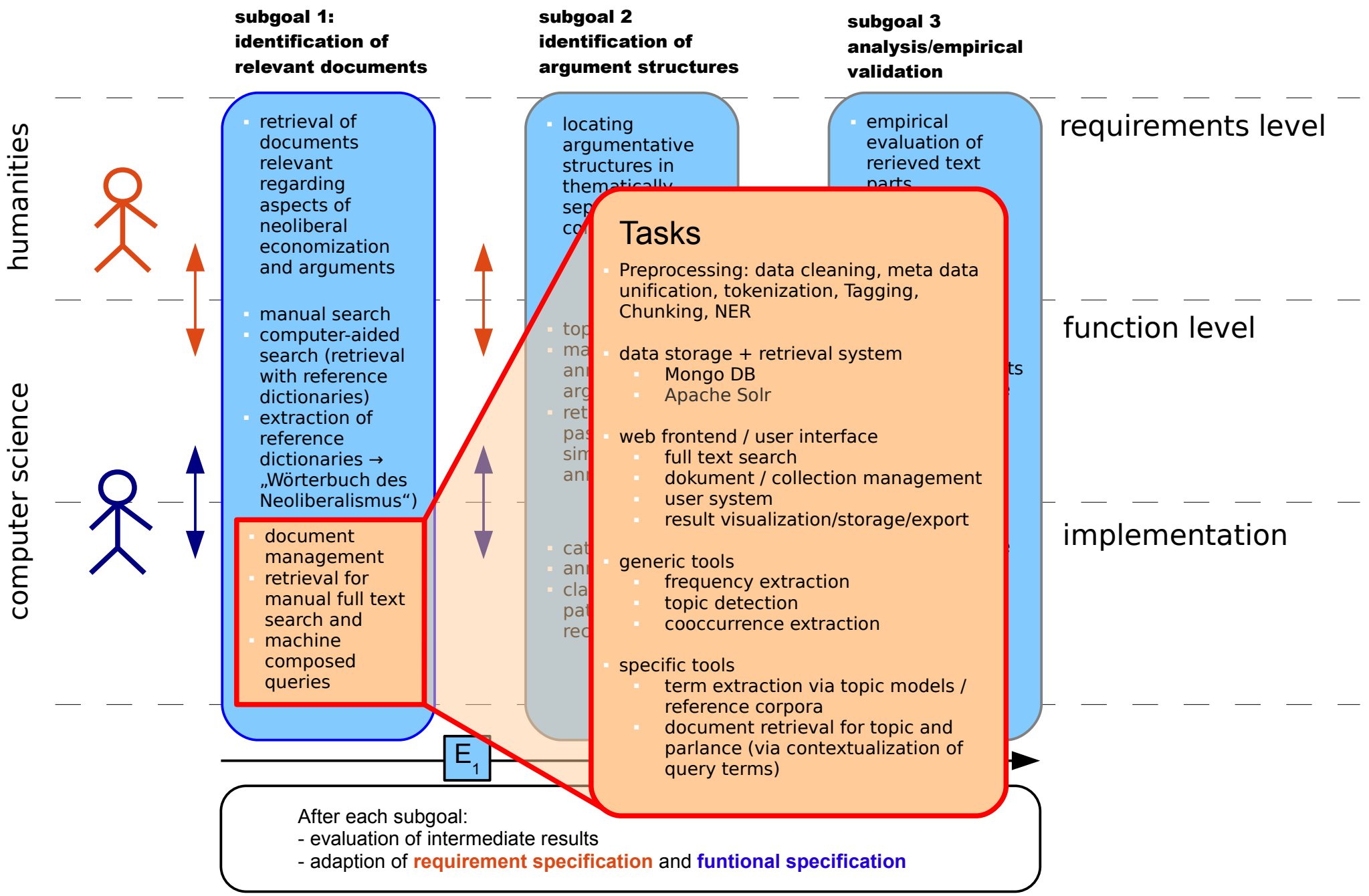
ASV Automatische Sprachverarbeitung



Requirements ePol-project

UNIVERSITÄT LEIPZIG

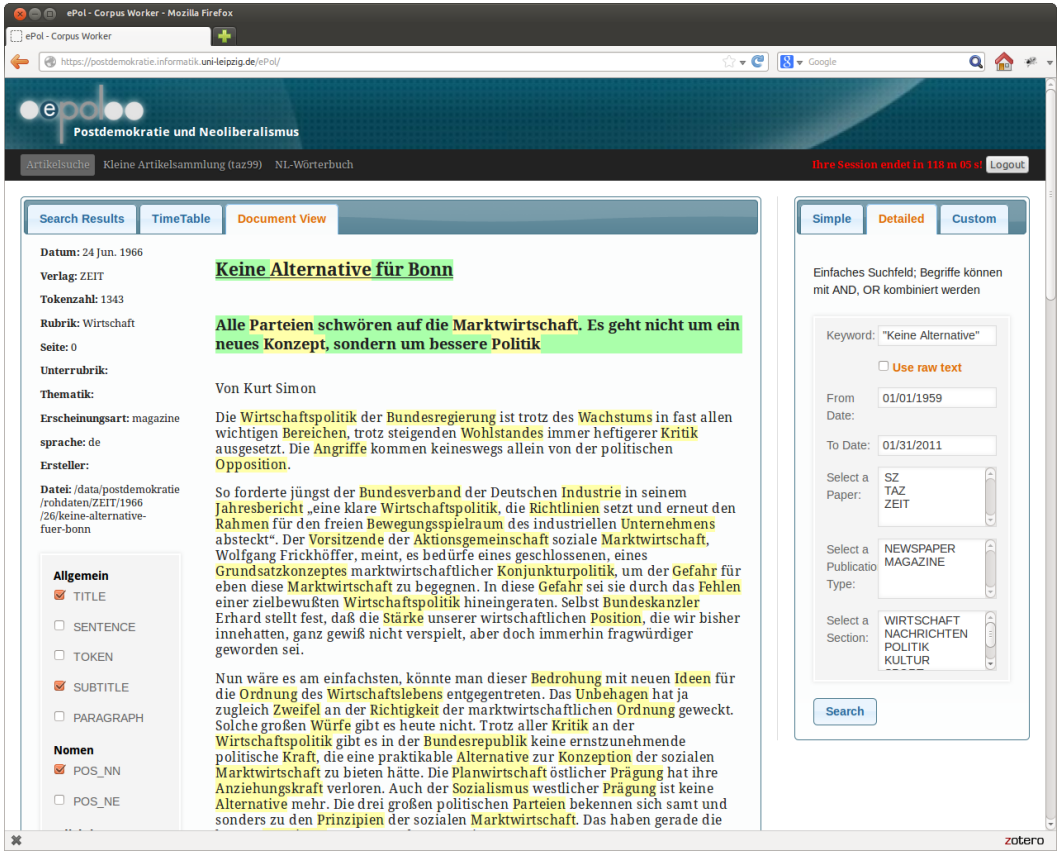
ASV Automatische Sprachverarbeitung



User interface for text analyses (Zeitungstexte)

Leipzig Corpus Miner:
A web interface for

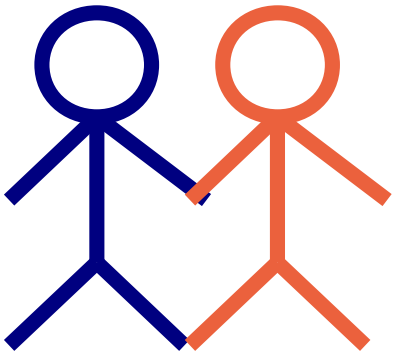
- corpus exploration
- document collection management
- control of analysis processes
- visualization of results



Conclusion

- What we need
  - modeling processes for systematic analysis of requirements from both perspectives (humanities and computer science)

    → open issue for future research



- What we recommend
  - adapted procedures / individual solutions rather than generic tools
  - iterated refinement of *requirement specifications* as well as *feature specification* during the overall analysis process
  - strong emphasis on evaluation / validation of computer-generated results