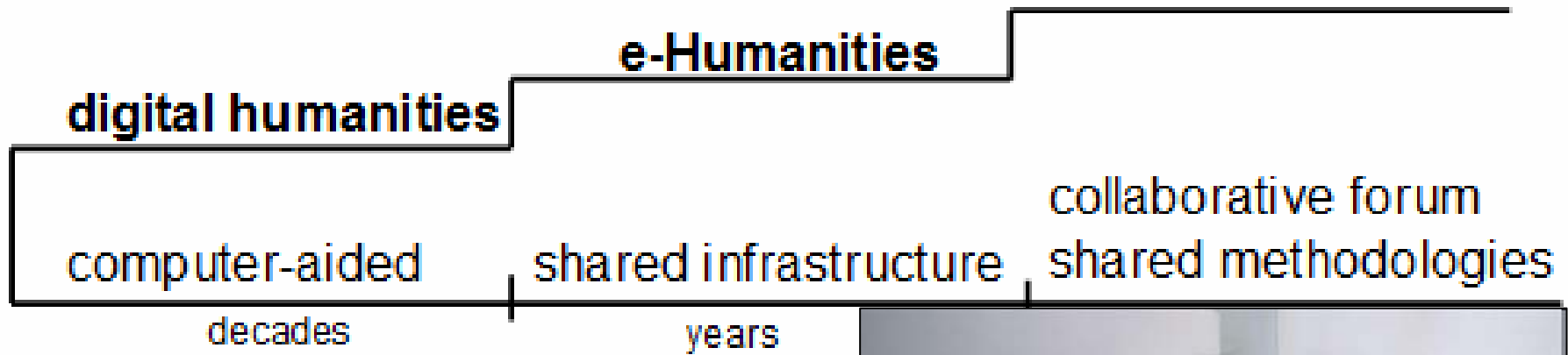


the quality in quantity - enhancing text-based research

Bernie Ács, National Center for Supercomputing Applications, UIUC, USA
Andreas Aschenbrenner, State and University Library Goettingen, Germany
Tobias Blanke, Centre for e-Research, King's College London, UK
Patrick Harms, State and University Library Goettingen, Germany
Mark Hedges, Centre for e-Research, King's College London, UK
Felix Lohmeier, State and University Library Goettingen, Germany
Wolfgang Pempe, State and University Library Goettingen, Germany
Angus Roberts, University of Sheffield, UK
Kathleen Smith, State and University Library Goettingen, Germany





<http://www.sixdifferentways.com/photos/spamalot-stairs.jpg>

quantitative



comparative [breadth]

- (statistical) evaluation
- information extraction
- re-representation /
visualisation

qualitative

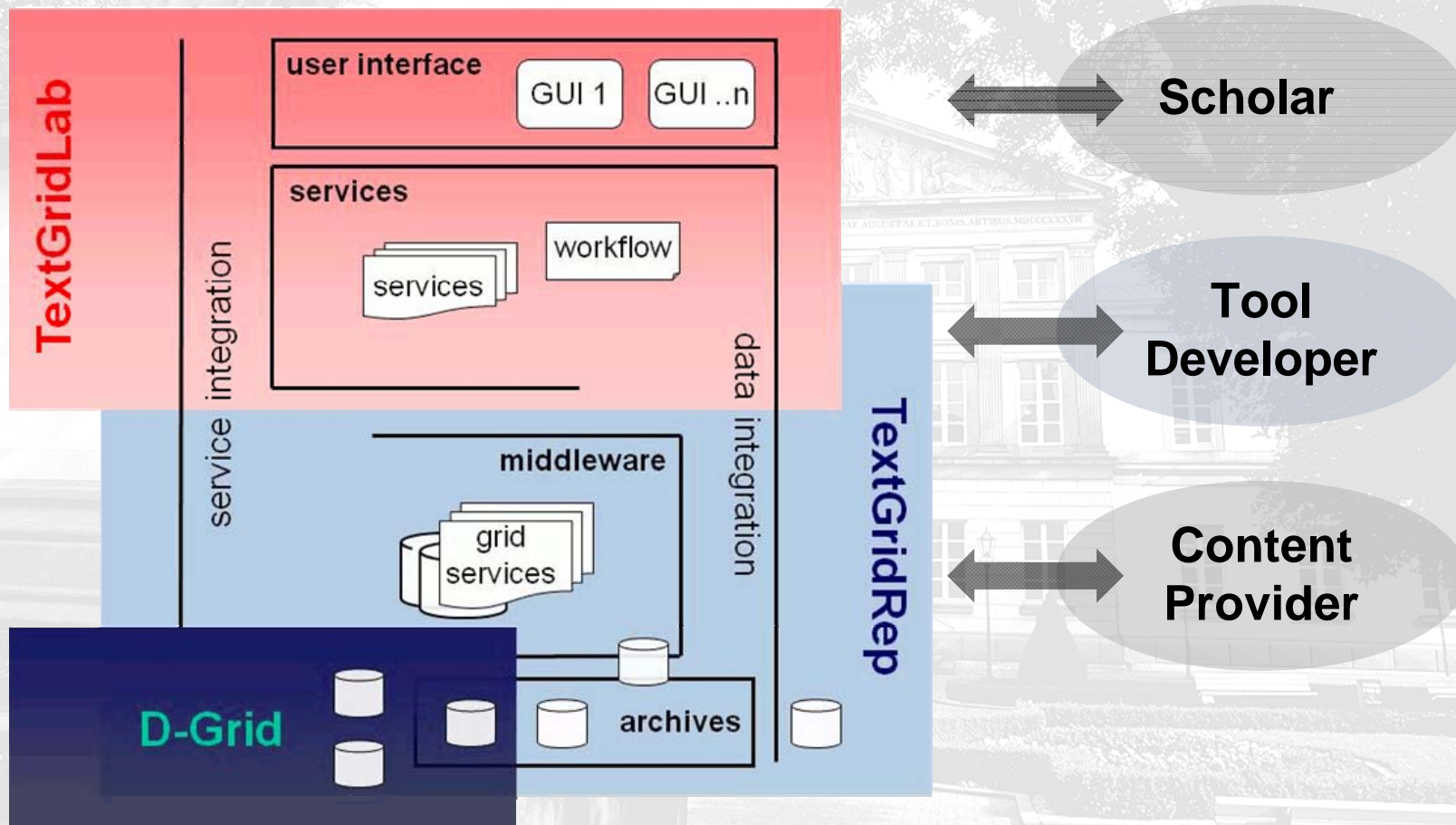


source as such [depth]

- observing
- analyzing,
understanding
- annotating

complimentary

TextGrid Architecture



TextGrid Services and Tools



XML-Editor



Graphical Link Editor



Workflow Editor



Search Tool



Dictionary Search Tool



Collationer



User and Project
Management



Metadata Annotator



Streaming Editor



Lemmatizer



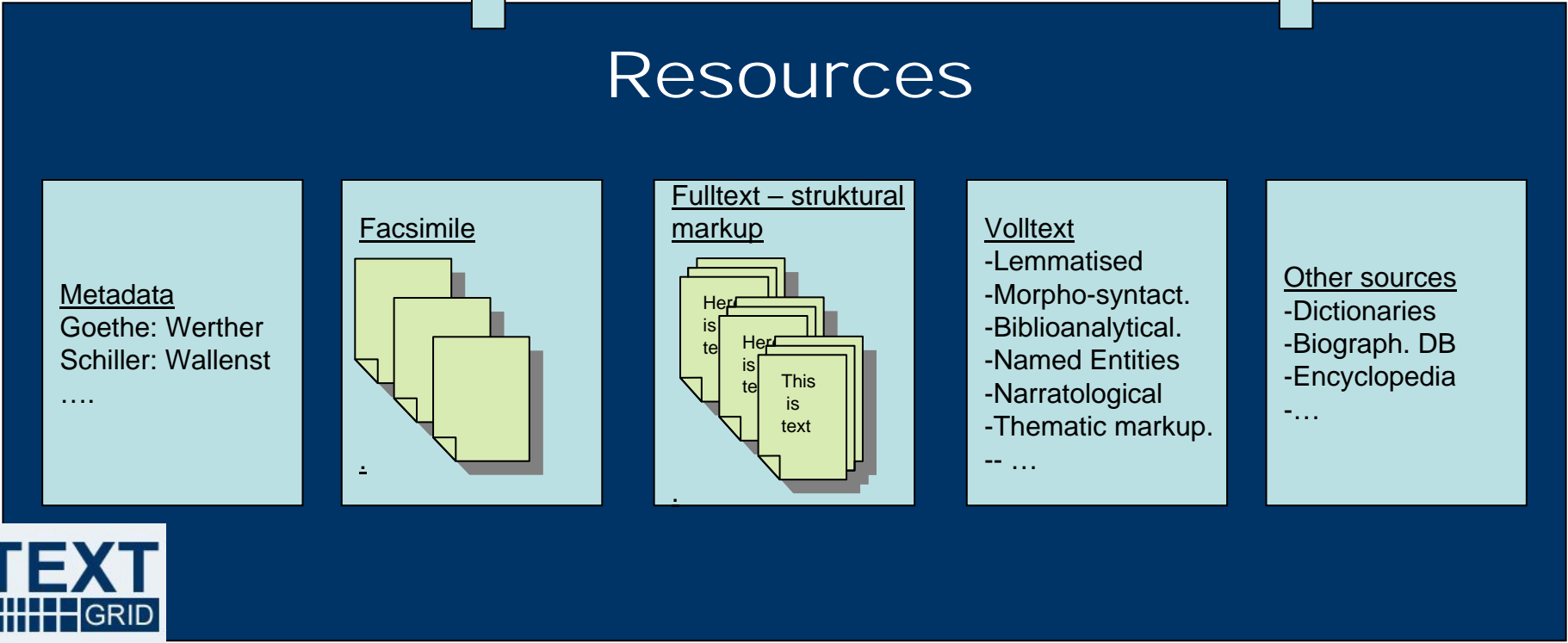
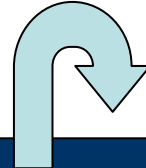
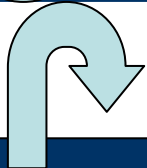
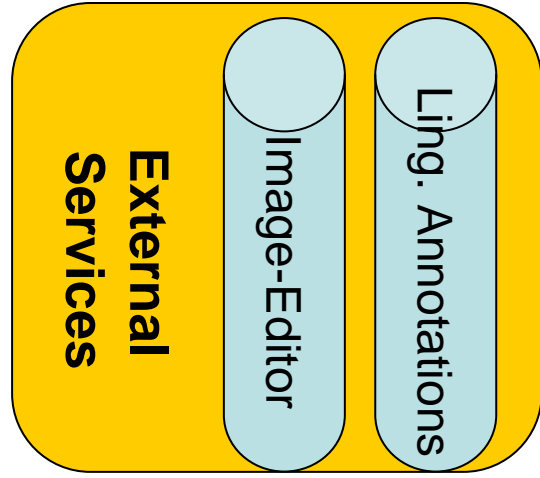
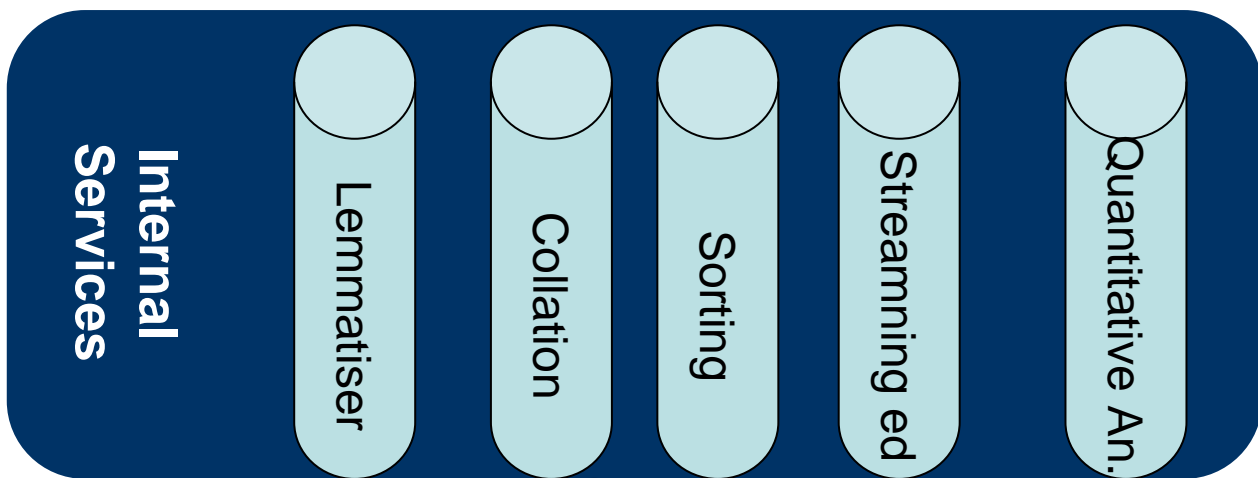
Text Publisher Web



Project Browser/
Navigator

Tokenizer

Sort Tool



TextGridLab

File Edit XML Search Tools Window Help

Analyze Names Project / User Administration Research

Brief an Herder.xml

```
<?xml version="1.0" encoding="UTF-8"?>
<TEI xmlns:xs="http://www.w3.org/2001/XMLSchema"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:tei="http://www.tei-c.org/
  xmlns:ll="http://www.linglit.tu-darmstadt.de" xmlns="http://www.tei-c.org/ns/1.0"
  xsi:schemaLocation="http://www.tei-c.org/ns/1.0 schemata/teilight-schema.xml">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title/>
      </titleStmt>
      <publicationStmt>
    </p/>
```

Names Open Source

- Johann Gottfried Herder
- Schäckespear
- C. F. Weiße
- Kästners
- Henrich Jung-Stilling
- Abraham Gotthelf Kästner
- Herder
- Salzmann
- Pegelow

Wikipedia

Wikipedia is sustained by people like you. Please [donate](#) today. [Log in / create account](#)

article discussion edit this page history

Johann Gottfried Herder

From Wikipedia, the free encyclopedia

Johann Gottfried von Herder (August 25, 1744 - December 18, 1803) was a German philosopher, poet, and literary critic. He is associated with the periods of Enlightenment, Sturm und Drang, and Weimar Classicism.

Johann Gottfried Herder

German philosophy
Enlightenment philosophy

WIKIPEDIA
The Free Encyclopedia

navigation

- Main page
- Contents
- Featured content



SEASR / MONK



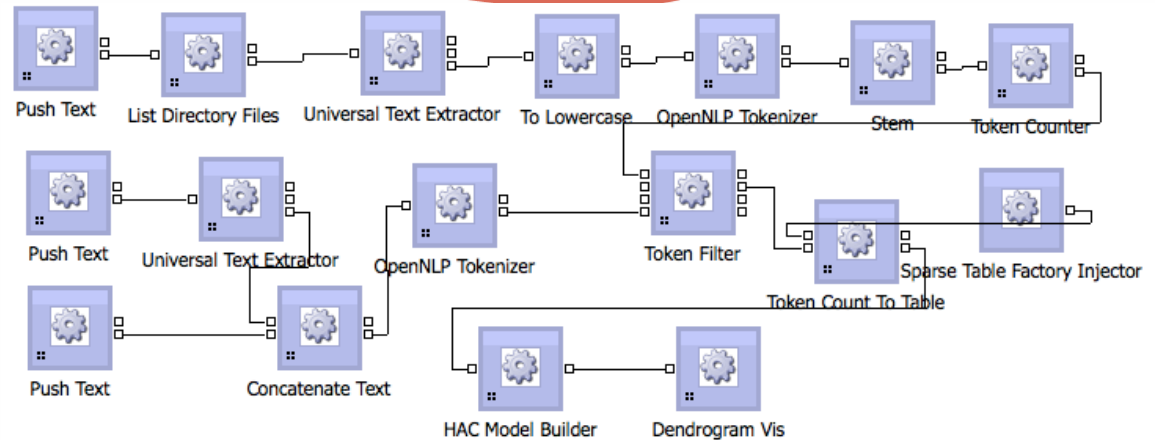
SEASR (Software Environment for the Advancement of Scholarly Research)

MONK (Metadata Offer New Knowledge)

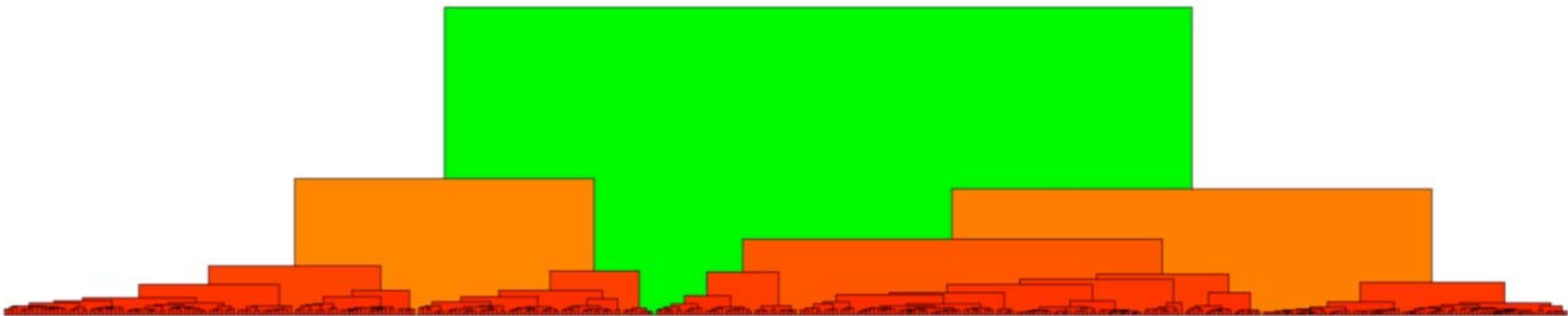
Andrew W. Mellon Foundation

Text Clustering

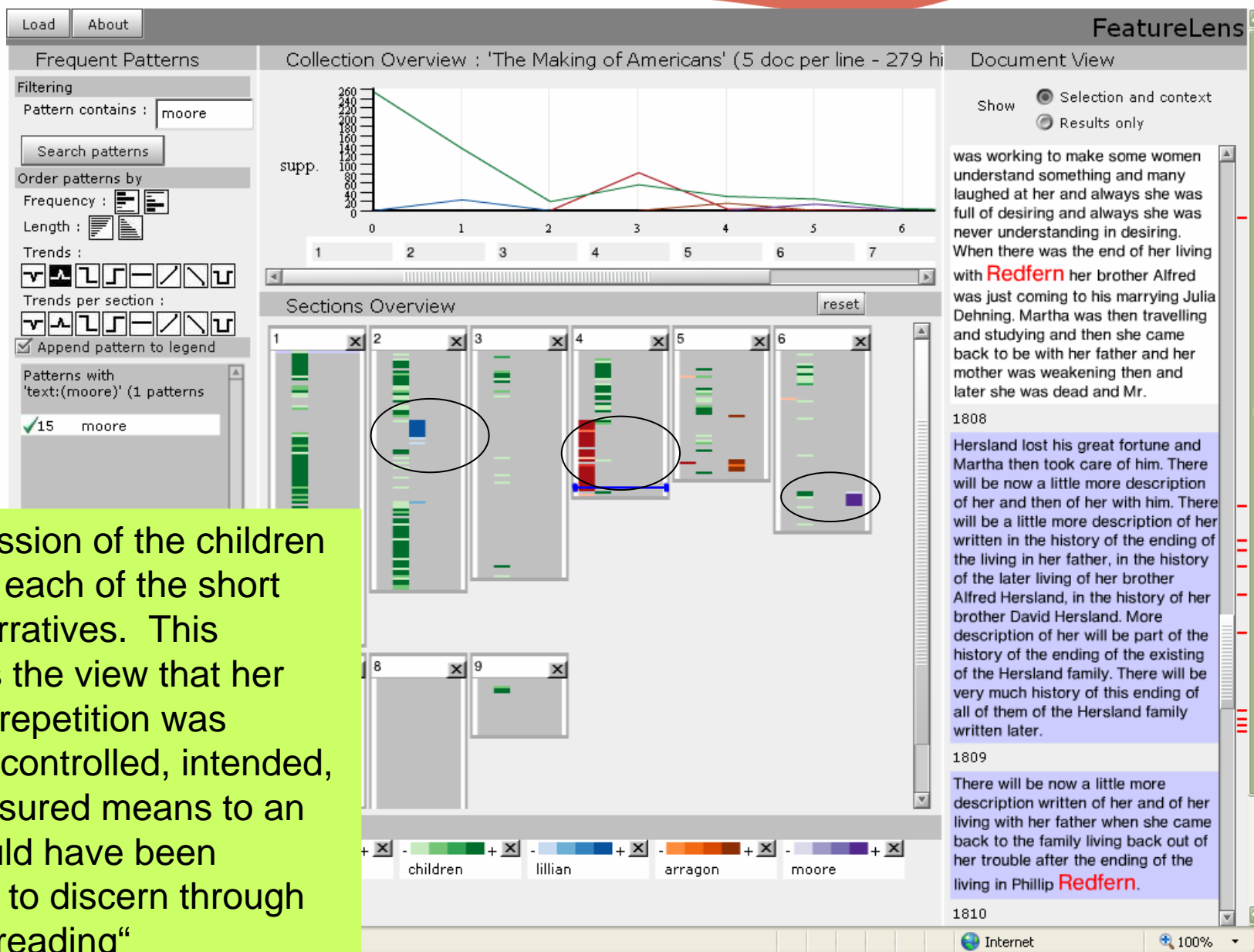
- Clustering of Text by token counts
- Various filtering options for stop words, Part of Speech
- Dendrogram Visualization



Done with the Jarko.com download



Feature Lens



“The discussion of the children introduces each of the short internal narratives. This champions the view that her method of repetition was patterned: controlled, intended, and a measured means to an end. It would have been impossible to discern through traditional reading“

Enables Scholar to Ask...

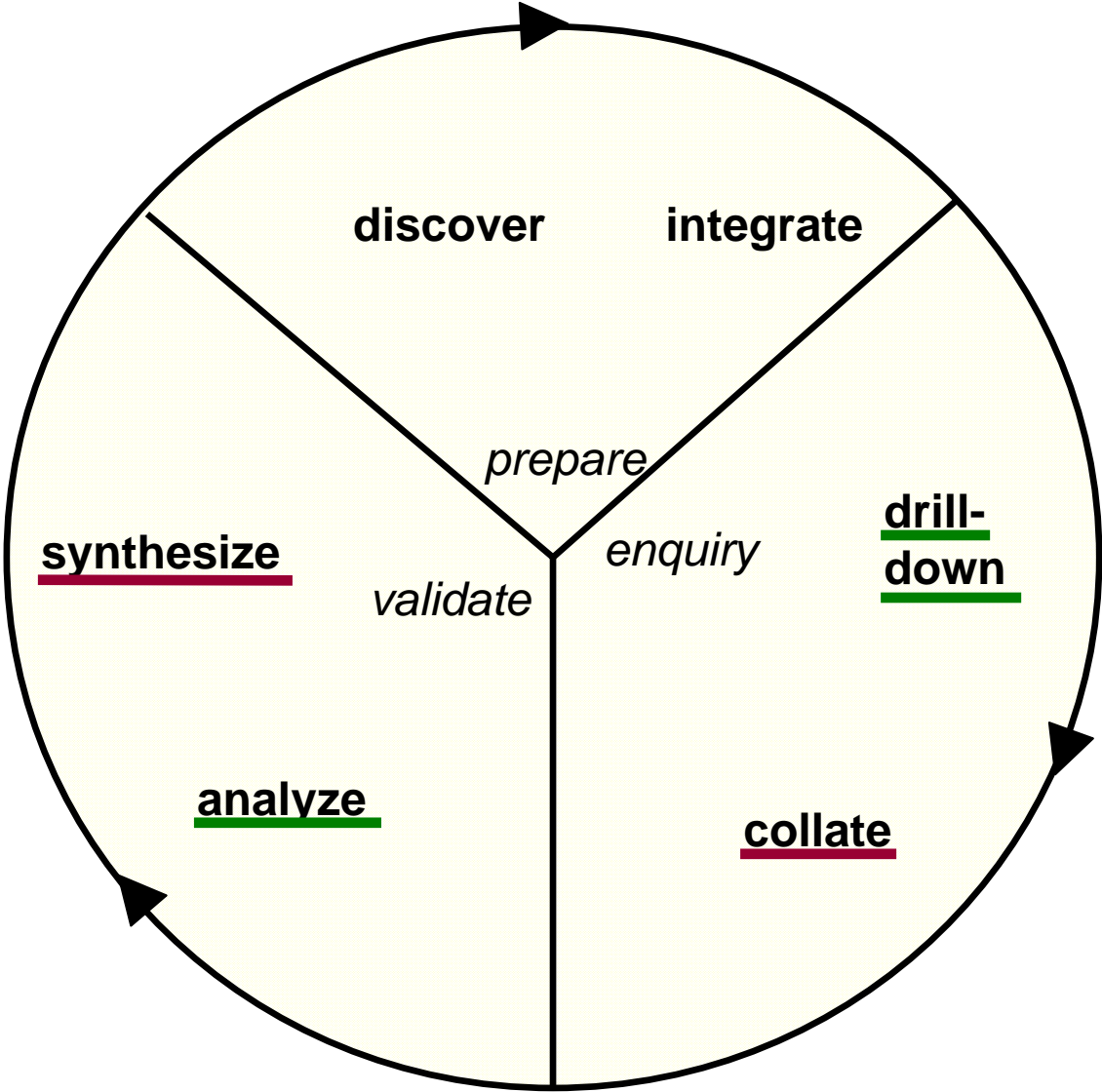
Pattern identification using automated learning

- Which patterns are characteristic of the English language?
- Which patterns are characteristic of a particular author, work, topic, or time?
- Which patterns based on words, phrases, sentences, etc. can be extracted from literary bodies?
- Which patterns are identified based on grammar or plot constructs?
- When are correlated patterns meaningful?
- Can they be categorized based on specific criteria?
- Can an author's intent be identified given an extracted pattern?

why link qualitative and quantitative? they always have been linked ...

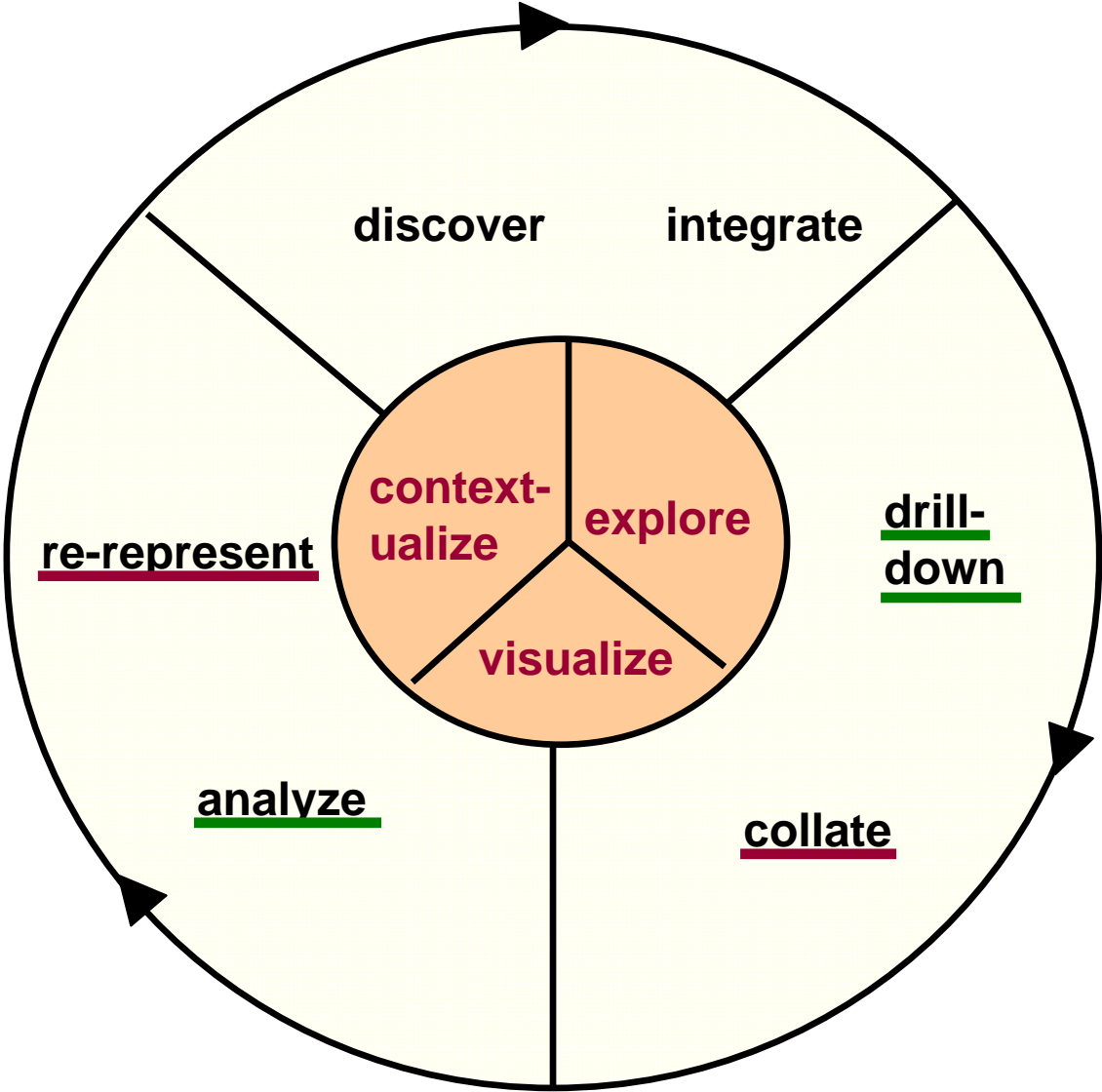
- create (one) - validate (many) research hypothesis (extrapolate)
- create (many) - validate (one) research hypothesis (replicate, show trends)
- explain / illustrate a trend (many) through individual examples (one)
- analyze an observation (one) through statistical analyses (many)

research lifecycle



inspired by <http://www.archimuse.com/papers/ukoln98paper/section6.html>

research lifecycle



inspired by <http://www.archimuse.com/papers/ukoln98paper/section6.html>

finally

- challenges:
 1. get the data (automatic harvest or manual selection/upload?)
 2. integrate/normalise the data (semi-automatic?)
 3. get the analysis/visualisation right, along which dimensions?
- cue for the architecture:

data will be redundant, to reuse existing systems and be open:
(a) active use, (b) various analysis frameworks, (c) preservation
- usability: hide complexity !
immediate results (automatic), and allow refinement (user)