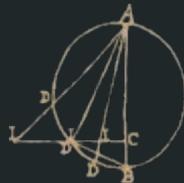


# Wir machen dann mal Digital Humanities Software - ein Erfahrungsbericht

Julia Damerow, Arizona State University, A Place Called Up Consulting  
Dirk Wintergrün, Max-Planck-Institut für Wissenschaftsgeschichte

FORGE 2016, Hamburg



# Voraussetzungen

Wissenschaftsgeschichte/ Historische Epistemologie

- kulturvergleichend
- sprachübergreifend
- epochenübergreifend
- disziplinenübergreifend

MPIWG gegründet 1994

> 150 Mitarbeiter, Stipendiaten und Gäste + 20-30 SHK

# Voraussetzungen

- Quellengattungen
  - Texte, Bilder, Materialdaten und Artefakte
- Sprachen
  - von Sumerisch über Latein und Arabisch zu Chinesisch
- Schriftsysteme
  - Vom Keilschrift bis zur lateinischen Schrift
- Epochen
  - von 3000 v. bis 2016 n.



# Praktische Konsequenzen

- Quellen können auf herkömmliche Art nicht beschafft werden
- Digitalisierung von Quellen
- Quellen liegen bereits als „digital born“ vor
  - Z.b. Datenbanken aus den Naturwissenschaften, Detektormaterial
- Frühzeitige Fokussierung des Institutes auf digitale Medien und digitale Publikationsformen
- Digitale Medien
  - Text
  - Bild
  - Datenbanken
- Anwachsen der Datenbestände über 4 Mio Digitalisate

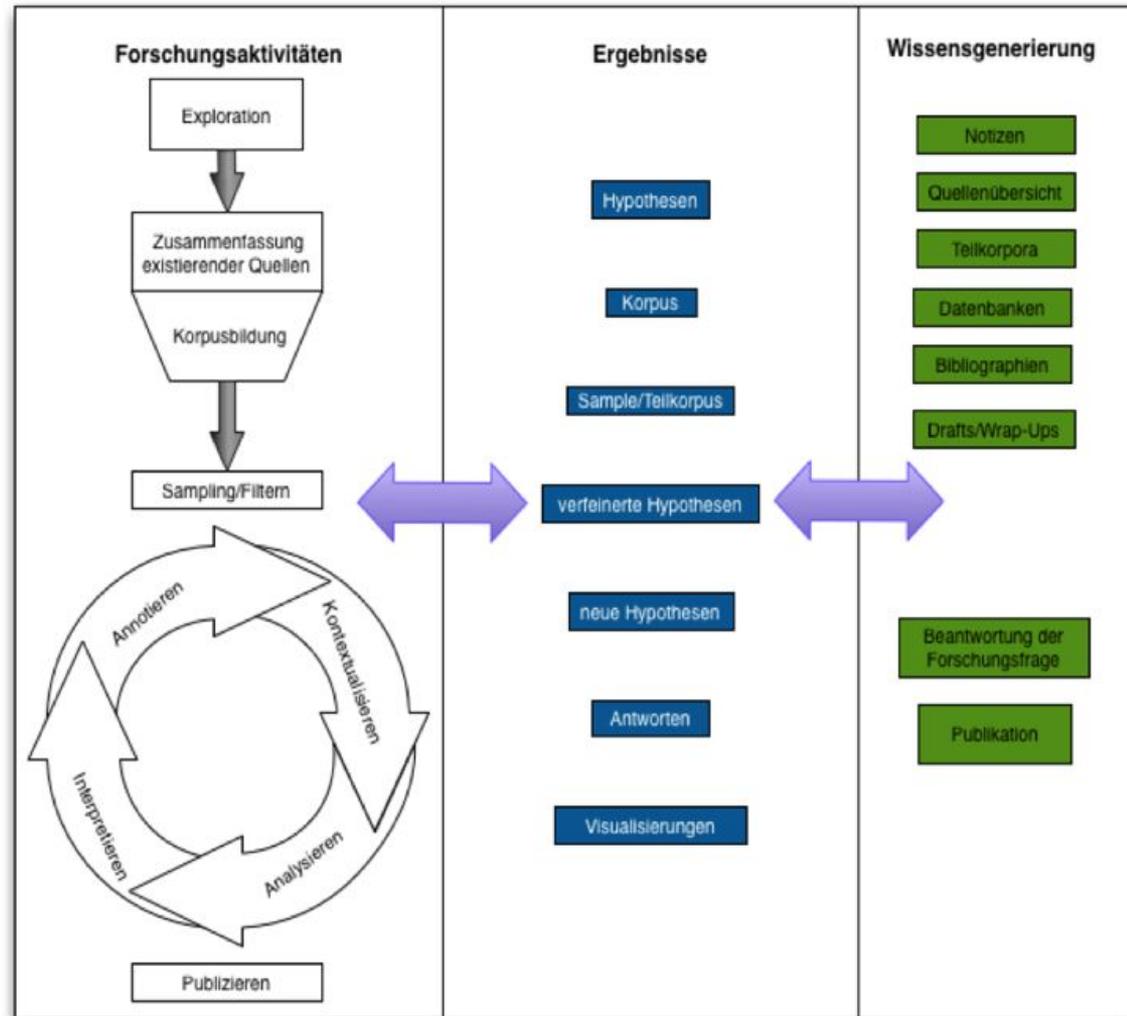
# Methoden

- Netzwerkanalysen
  - Topic Modelling
  - Text- und Datamining
  - OCR
  - Annotationen und Kommentare
  - ...
- Alle Methoden erzeugen digitalen Output als geisteswissenschaftliche Forschungsdaten

# Forschungskreislauf

## Pragmatischer Ansatz zur Verortung von Tools und Technologien

- Was muss wie bereit gehalten werden?
- Was muss aktualisiert und gewartet werden?
- Wo ist Generalisierung notwendig oder wünschenswert?



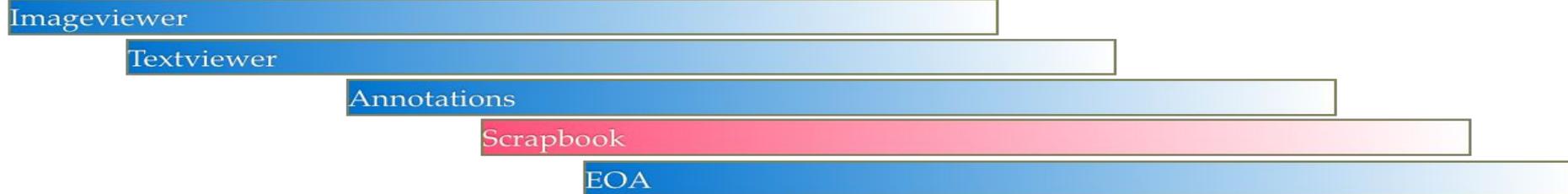
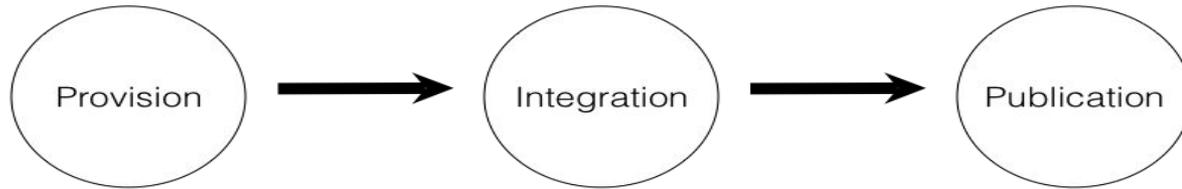
# Trennung Datenerzeugung und -bereitstellung

- Erzeugung von (organisationsübergreifender) Wissensbasis
- Wiederverwendung von Daten, Anwendungen, und Infrastrukturen

# ASU Projekte

- **Quadriga/VogonWeb**
  - <http://diging.github.io/quadriga/>
  - Annotation von Quellen mit Konzeptnetzwerken
  - Trennung von Annotationserstellung (VogonWeb) und -speicherung (Quadriga)
  - Einfacher Austausch von Annotationssoftware
  
- **Tethne**
  - <http://diging.github.io/tethne/>
  - Python package zum Parsen und Analysieren von bibliographischen Metadaten
  - Relativ aktive Anwenderbasis (unter anderem als Unterrichtstool)
  - Auf gutem Weg zu Community-supported Open-Source Tool

# Beispiel Projekte am MPIWG



ENTRY POINT

www.mpiwg-berlin.mpg.de/Galileo\_Prototype/MAIN.HTM

Galileo Galilei's Notes on Motion

Joint Project of

Biblioteca Nazionale Centrale, Florence

Galileo Galilei Ms. Gal. 72

List of folios Drawing Working level High resolution

Working Level of Folio 51 r Indices

Click on textblock, calculation, or drawing to see more!



Galileo's notes on motion and 1 pages. They contain numerous calculated numbers, diagrams, on the intellectual route follow

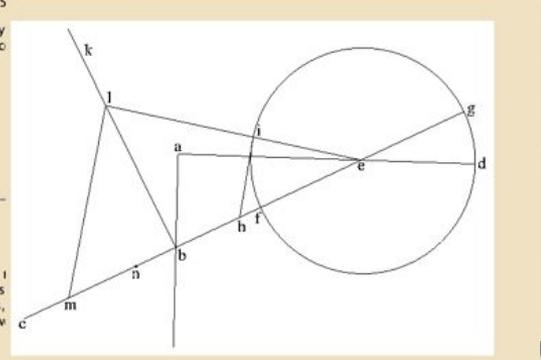
Drawing D01A of Folio Page 51 r

www.mpiwg-berlin.mpg.de/Galileo\_Prototype/HTML/F051\_R/D01A.HTM

Galileo Galilei Ms. Gal. 72

List of folios Drawing Working level High resolution

Working Level of Folio 51 r Indices



Galileo Galilei Ms. Gal. 72

List of folios Drawing Working level High resolution

Working Level of Folio 51 r Indices

Click on textblock, calculation, or drawing to see more!

Galileo Galilei Ms. Gal. 72

List of folios Drawing Working level High resolution

Working Level of Folio 51 r Indices

WELTBILD UND ERKENNTNIS

WELTBILD UND ERKENNTNIS x Musawarat Graffiti Archive ... x

einstein-virtuell.mplwg-berlin.mpg.de/home

Deaktivieren Cookies Keine CSS-Fehler Formulare Grafiken Informationen Verschiedenes Hervorheben Größe Extras Quelltext

Start Panoramas Hide Texts back Impressum Copyright



P212341: CDLJ 2002/1, no. 1 - Mozilla

Back Forward Reload Stop

http://www.cdli.ucla.edu/fmi/xsl/res

Search Print

cdli

CDLJ 2002/1, no. 1

Record 1 of 1 records found

Access System Information Back to Search Page Copyright

Primary Publication: CDLJ 2002/1, no. 1, Englund, Robert K., 2002  
 Collection: California Museum of Ancient Art, Los Angeles, California, USA  
 Museum no.: CMAA 002-C0005  
 CDLI no.: P212341  
 Provenience: Dreheim?  
 Genre: Administrative  
 Period: Ur III  
 Measurements (mm): 30x27x13

Download Full CDLI Files

View full catalogue entry

line art

Obv.

- 4(disz) gu4 4(disz) ab2
- e2 muhaldim
- u4 2(u) 5(disz)-kam
- zi-ga
- ur<sup>#</sup>-(d)en-il2-la2

Rev.

n lines broken  
 blank

- mu us2-sa ki-masz(ki) ba-hul

Please report any errors or other problems, and suggestions for improvement to: [cdli@ucla.edu](mailto:cdli@ucla.edu)

Virtual Laboratory

wp.mpgw-berlin.mpg.de/essays/index.htm

About: mplwg... Network Workbu... Java API Tutorial... uTrack - online... ECHOclass/DocuV... Presentation - D... Lesazeichen

Deaktivieren Cookies Keine CSS-Fehler Formulare Grafiken Informationen Verschiedenes Hervorheben Größe Extras Quelltext

Essays Experiments Technology Objects Sites People Concepts Library

About this Section

List of Essays

How to contribute

Picture of the Month

Picture Archive

Activities

Cooperations

Related Links

Press Reports

Search

News

All six editions of *Wissenschaften* (1832-1920) renowned *Grundriss der physiologischen Psychologie* from 1874 to 1911 are now online in the library (English Translation of the 5th edition by *Fechner: Principles of Physiological Psychology*, 1904).

Also new in the *journal section* of the library: the complete *Zeitschrift für Psychologie und Physiologie der Sinnesorgane* (1890-1904).

The virtual Laboratory was initiated by the research project *The Experimentation of Life*. It is work in progress. We look forward to your *comments, critiques and distributions*.

An American Physiologist Abroad: Francis Gano Benedict's European Tours

Elizabeth Newkirk [11/2010]

Listening to the Body Electric. Electrophysiology and the Telephone in the Late 19th Century.

Axel Volmar [11/2010]

Scientific and Medical Films in the 1920s-1930s

Andreas Killen [07/2009]

[more articles]

Home About Sitemap Contact Search Tools Survey myLab Intranet

Max Planck Institute for the History of Science, Berlin

# Virtual Spaces

- Desktopanwendung zur Erstellung virtueller Ausstellungen
- Trennung von Inhalten und Präsentation
  - Speicherung von Inhalten und semantischen Zusammenhängen in XML Modellen
  - Externe Speicherung von Quellen in Repositorien
  - Generierung von statischen HTML Seiten zur Präsentation
- Erweiterung der Software für Anpassung an technologischen Fortschritt
- Versionierbarkeit von Inhalten

Aber:

- Enge Kopplung von XML Modellen und Anwendung
- Fehlende Generalisierung für Anwendung über virtuelle Ausstellungen hinaus

# Projektumsetzung und Entwicklung am MPIWG

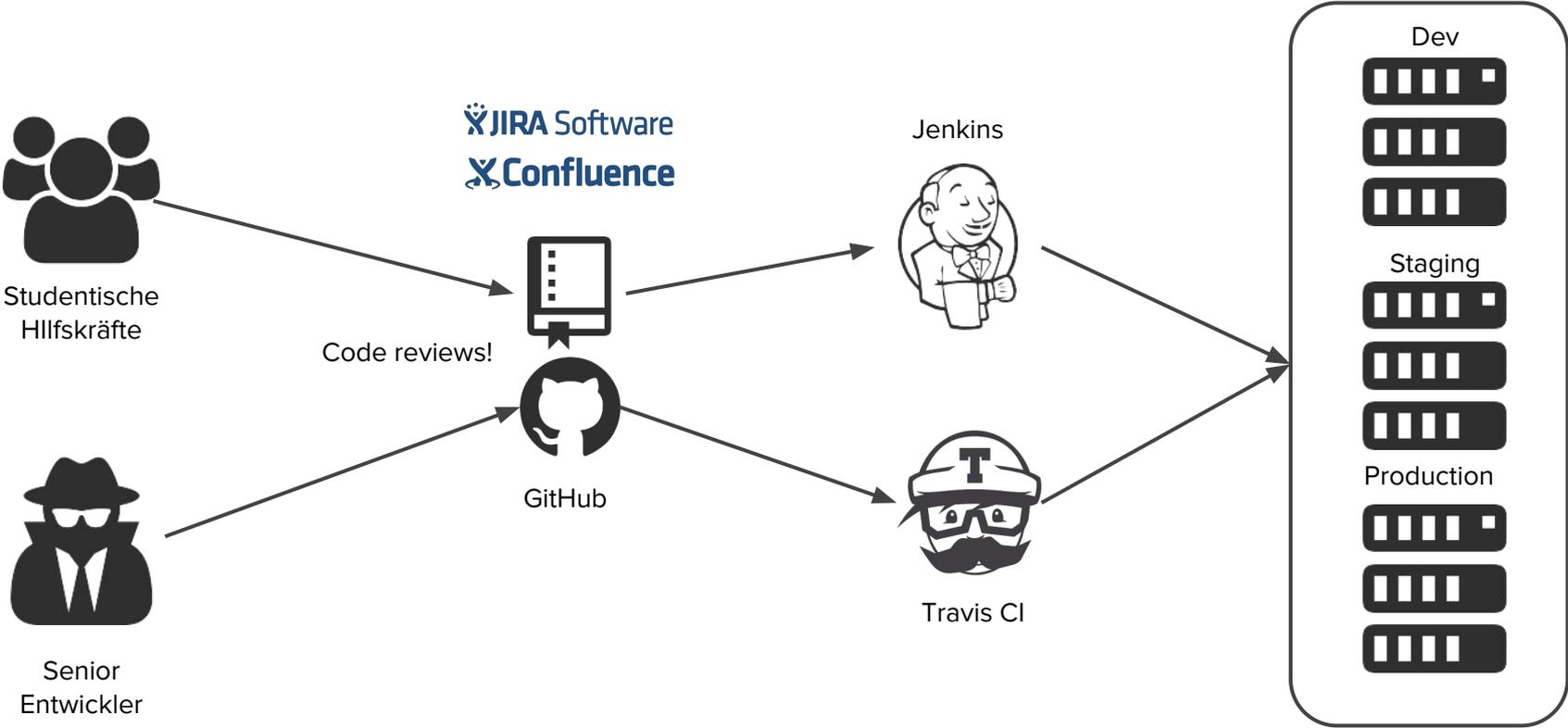
- Schaffung einer wissenschaftlichen Gruppe zur Arbeit mit digitalen Methoden 2000 mit 3 wiss. Mitarbeitern
- Eingebettet in den Forschungsalltag
- Kein „Kunden-Angebotsverhältnis“.
- Enge Einbindung der Bibliothek

# Projektumsetzung und Entwicklung in DigInG

- Digital Innovation Group (DigInG) – <http://diging.asu.edu>
- Studenten der Informatik und Geisteswissenschaften
- In die Lehre eingebettete Softwareentwicklung
- Agile Methoden mit GitHub, Jira, und CI Tools (Jenkins/Travis)
- Open Source
- Code Reviews!



# Entwicklungs-Workflow



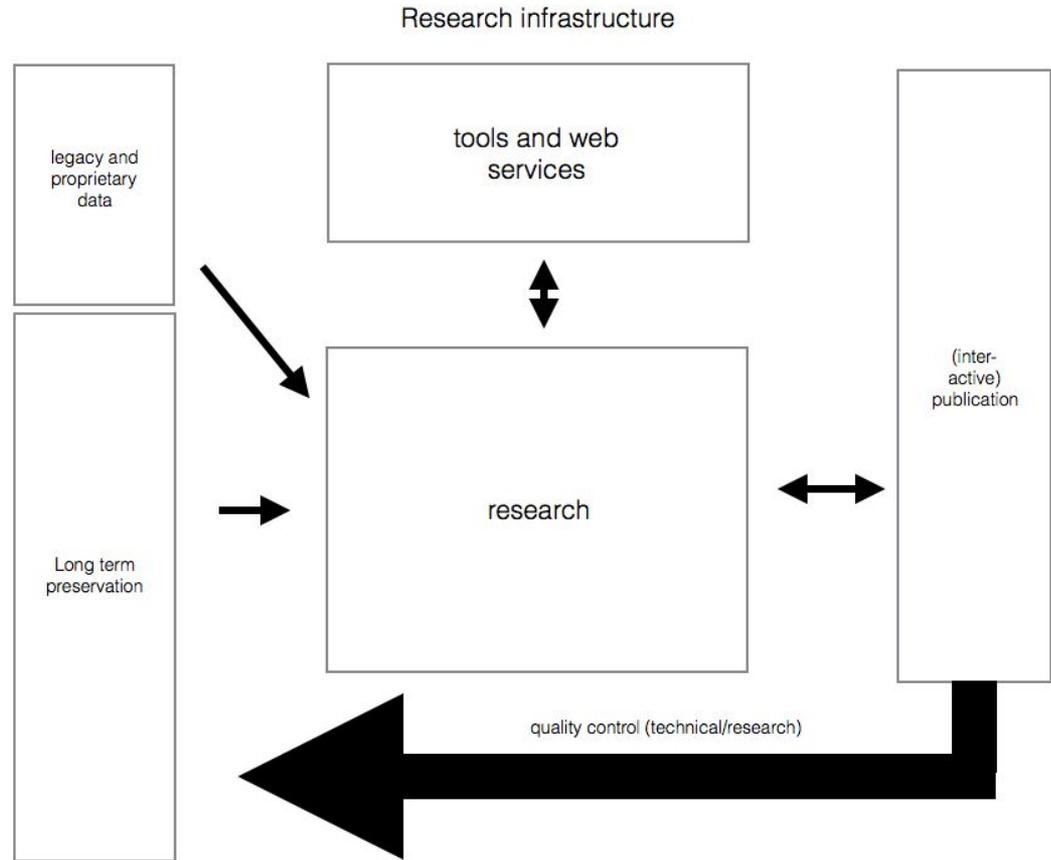
# Vor-/Nachteile

- Leichter finanzierbar
- Weniger Bündelung von Wissen in nur einem Programmierer
- Einbindung von Studenten ermöglicht andere Finanzierungsquellen/-argumente
- Relativ leicht skalierbar je nach Entwicklungsbedarf

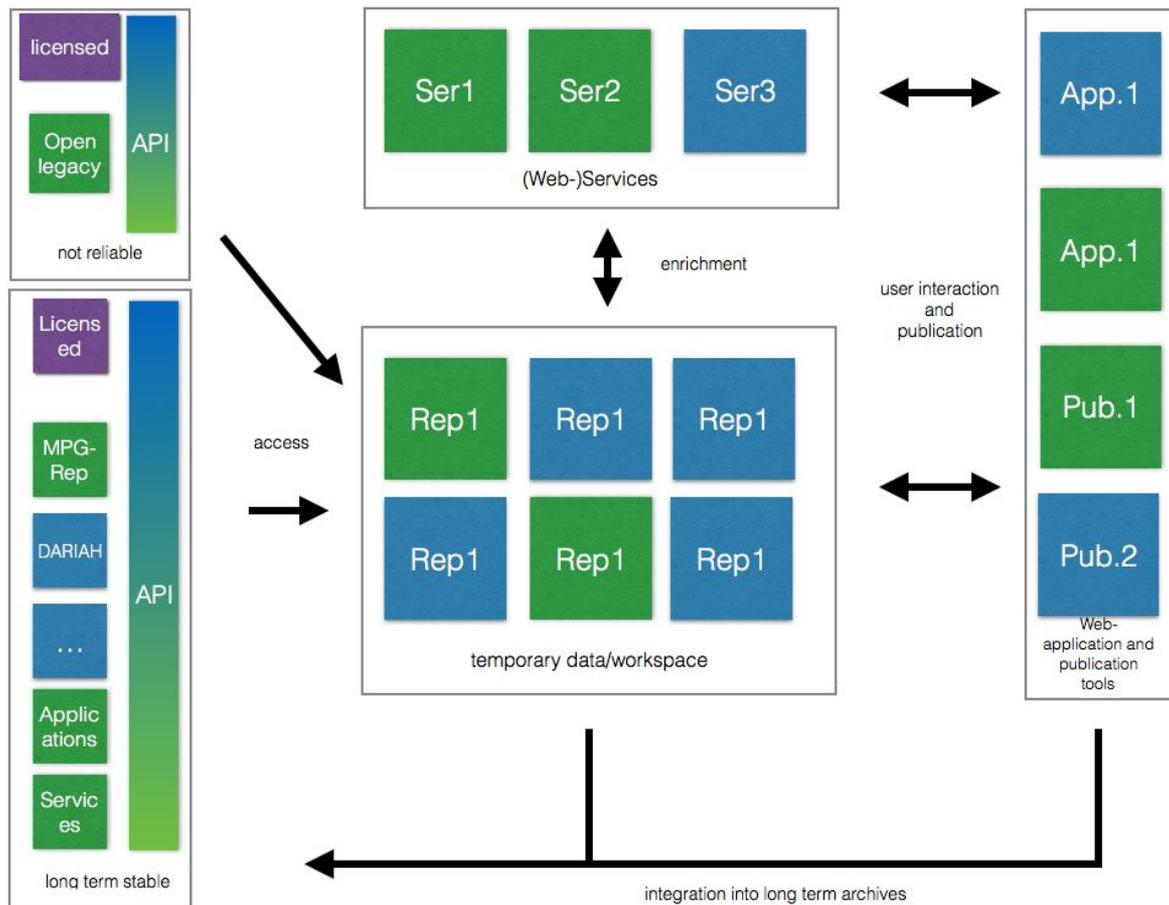
Aber:

- Langsam
- Viel “hand holding” bei der Entwicklung
- Wissensbündelung bei Senior Entwickler

# Lessons learned



# Lessons learned



# Lessons Learned

- Generalisierung ist schwierig trotz heterogener Anwendungsszenarien
- Entwicklung und Anwendung laufen parallel
  - “Agile development” im Prinzip adäquates Format
- (professionelle) Softwareentwicklung unterpriorisiert
  - Endet in der Regel mit Projektabschluss
- Kooperation mit anderen Projekten ist häufig schwierig, trotz Offenheit und guten Willens
  - Unterschiedliche Zeitpläne, disziplinierte Ansätze,...

**Und nu?**