

# 2: Anatomy of DH Projects

Matt Price

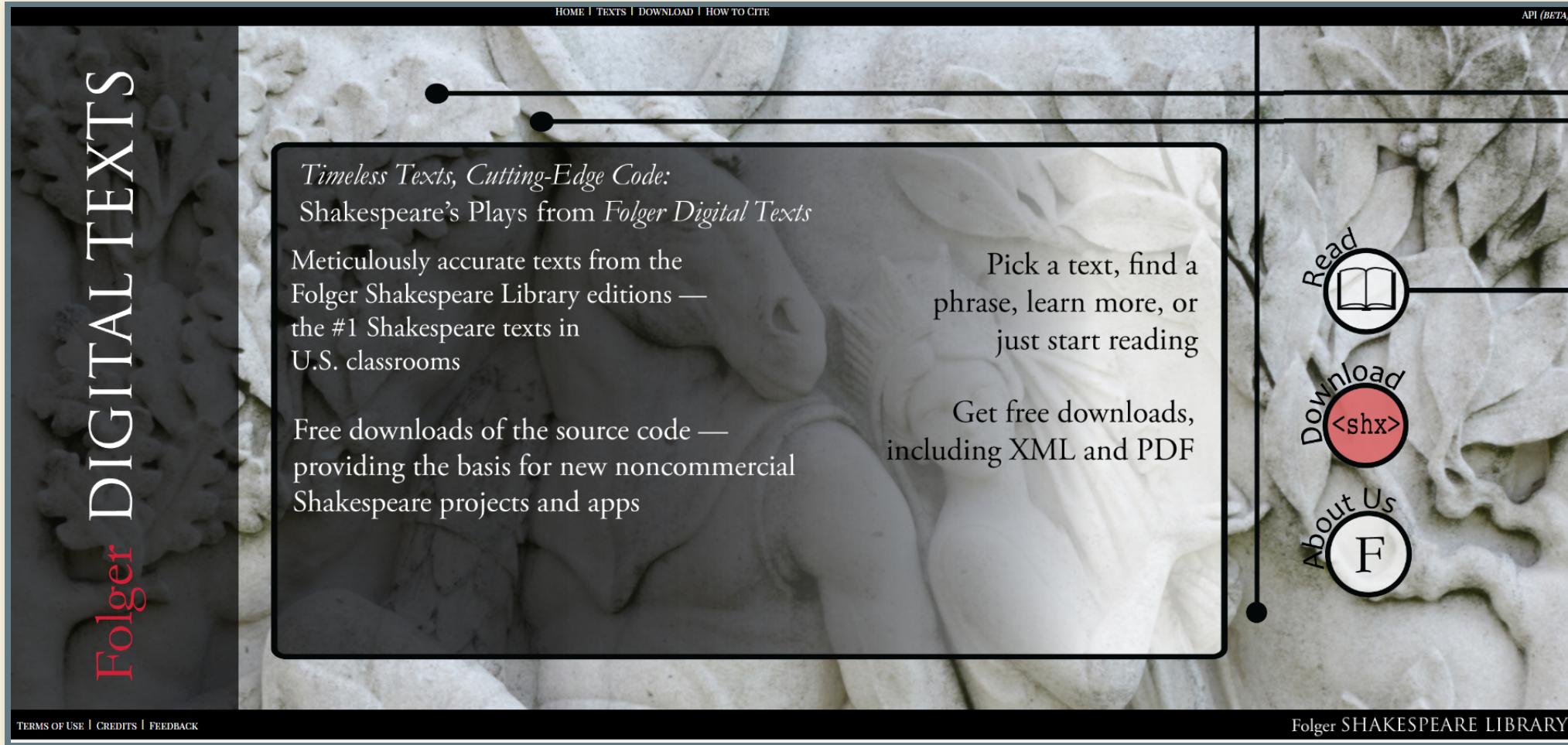
# Today

- First, some DH examples How do we Think about Technology?
- Very Basic Design Parameters
- Data and Data Models
- Assignment 1

# Project Examples

# Digital Editions, Archives, Narratives

# Folger: Shakespeare Library



HOME | TEXTS | DOWNLOAD | HOW TO CITE      API (BETA)

Timeless Texts, Cutting-Edge Code:  
Shakespeare's Plays from *Folger Digital Texts*

Meticulously accurate texts from the  
Folger Shakespeare Library editions —  
the #1 Shakespeare texts in  
U.S. classrooms

Free downloads of the source code —  
providing the basis for new noncommercial  
Shakespeare projects and apps

Pick a text, find a  
phrase, learn more, or  
just start reading

Get free downloads,  
including XML and PDF

Read

Download <shx>

About Us F

Folger SHAKESPEARE LIBRARY

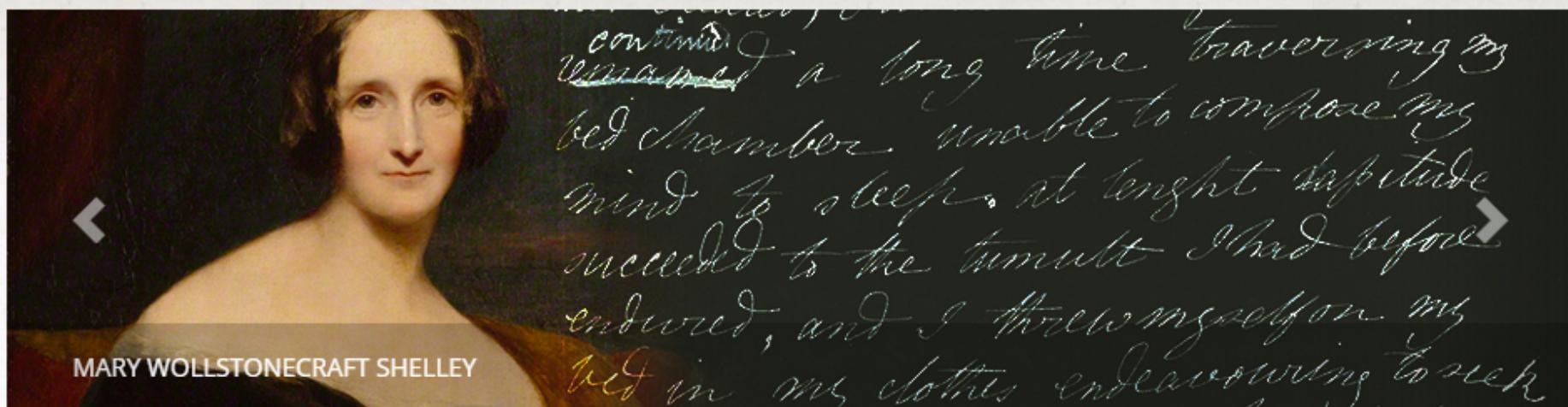
Folger DIGITAL TEXTS

TERMS OF USE | CREDITS | FEEDBACK

# Shelley-Godwin Archive

## THE Shelley-Godwin ARCHIVE

[HOME](#)   [ABOUT](#)   [EXPLORE THE ARCHIVE](#)   [SEARCH](#)   [USING THE ARCHIVE](#)



*General Editors:* Neil Fraistat, Elizabeth Denlinger, and Raffaele Viglianti

About the Archive

Featured Works

## The Archimedes Palimpsest: Multispectral Imaging



# Infinite Ulysses: interactive reader's edition by Amanda Visconti

 **Infinite Ulysses** Read Now About Credits Contact Sign Up! Log In

**Come read *Ulysses* with us!**  
We welcome readers of every background.



Highlight parts of the book and add your comments & questions.

Read interpretations & help left by other readers.

Customize the annotations you see to just the ones you need!

# Maps, Visualizations, Interpretations

## Mapping Medieval Chester

- Goal: making a digital map of medieval Chester by combining post-medieval maps with archaeological and historical evidence
- Digital advantage: interactive, layered, & transparent

# Mapping Homer's Catalogue of Ships

Courtney Evans and Ben Jasnow, with Jenny Strauss Clay and the UVA Scholars' Lab (2013)

Mapping the towns in Homer's Catalogue of Ships to analyze underlying organizing principles for the narration

Mapping the Catalogue of Ships

Home Credits Creative Commons

## The Iliad, Book 2 [\(View Fullscreen\)](#)

Hyria [x](#)

[Pleiades entry](#)

ἡρχον  
Αρκεσίλαος τε Προθοήνωρ τε Κλονίος  
τε, (495)  
οῖ θ' Υριην ἐνέμοντο καὶ Αὐλίδα  
πετρήεσσαν

11

## Project DECIMA

University of Toronto, History Dept.

Onto a 16th century map of Florence, DECIMA maps census data about Florence's inhabitants, their occupations, wealth, and daily life.

# ToposText: Ancient Texts + Maps

# Mapping Ararat

## MAPPING ARARAT

AN IMAGINARY JEWISH HOMELANDS PROJECT

[Home](#)

[News](#)

[The Project](#)

[Historical Research](#)

[About the Team](#)

[Contact](#)



Welcome to Mapping Ararat: An Imaginary Jewish Homelands Project. Using augmented reality, this project animates Major Mordecai Noah's 1825 unrealized plan to transform Grand Island, New York into Ararat, a "city of refuge for the Jews." Explore. Take a Tour. Learn. Imagine.

## About that Land Acknowledgment

- History of Life on this shore
- Dish with One Spoon
- Continuity and Disjuncture

# Communication

- Scholarly Writings
- Public Engagement
- Social Media
- Teaching & Collaboration

# Scholarly Books: The Networked Monograph

In addition to printed books, can scholarly monographs expand their forms to “take advantage [...] of the interactive, annotative, and computational affordances of the web”? Donald J. Waters, “Monograph Publishing in the Digital Age” (2016)

Digitally augmented scholarly monographs (e.g. Mellon-funded

 Fully interactive and searchable online with primary sources and other works	 High quality as judged by peers
 Portable across reader applications	 Able to support a user’s annotations independently of any particular reader technology
 Capable of supporting metrics of use that respect user privacy	 Reviewed and eligible for disciplinary prizes and awards
 Maintained and preserved in its digital form	 Expertly marketed, widely accessible, and owned (not rented) by the reader
 Economically sustainable	

## VR/3D Printing for Cultural Heritage Collections: Small Wonders (AGO)

- Miniature boxwood carvings from the early 1500s
- Studied and exhibited through digital imaging (micro CAT scanning), VR, and 3D printing



# Research Materials: Medieval Manuscripts and Social Media



# Teaching

## This Class

Heather Eason: five artists' representations of threshold crossings in Dante's *Inferno*. Omeka/Neatline undergraduate student project.

**Liminal Representation in the Inferno**

[View Fullscreen](#)



21

## Others

Website for [Khangchendzonga Conservation Committee](#), written and developed by students from HIS455, *Hacking History*, 2018.

# Collaboration & Infrastructures

The screenshot shows a GitHub repository page for the project `edgi-govdata-archiving / web-monitoring`. The repository is described as "Documentation and project-wide issues for the Website Monitoring project (a.k.a. "Scanner")". It features 108 commits, 5 branches, 0 releases, and 11 contributors. The page includes tabs for Code, Issues (25), Pull requests (0), Projects (0), Wiki, Insights, and Settings. Topics listed include `web-monitoring`, `documentation`, `project-management`, and `gsoc-2017`. A merge pull request from `lightandluck` is visible at the bottom.

Search or jump to... / Pull requests Issues Marketplace Explore

edgi-govdata-archiving / web-monitoring Unwatch 20

Code Issues 25 Pull requests 0 Projects 0 Wiki Insights Settings

Documentation and project-wide issues for the Website Monitoring project (a.k.a. "Scanner") <https://enviromon...>

web-monitoring documentation project-management gsoc-2017 Manage topics

108 commits 5 branches 0 releases 11 contributors

Branch: master ▾ New pull request Create new file Upload files

lightandluck Merge pull request #134 from edgi-govdata-archiving/lightandluck-patch-1 ...

# Why “Anatomy of DH”?

- Design & evaluation of DH projects specifically
- Design & evaluation of projects (digital and otherwise) more generally
- Understanding of knowledge production in digital environments

# DH pipelines

# Tools & Technologies



# Technology in DH

# Choosing a Tool

# Affordances

*Affordances of a technology = properties of a technology that enable certain tasks; tasks that users can perform with a technology (e.g. query, search, analysis, visualization)*

- OFD. Accessed

*Digital documents..have their own affordances. They can be easily searched, shared, stored, accessed remotely... But they lack the affordances that really matter to a group of people working*

# Users, Readers, Communities

# Technology-in-practice (Wanda Orlikowski)

*The notion of “technology-in-practice” suggests that the technology is not just the machinery, digital or analog, but on the one hand **the organizational or institutional culture around it**, and on the other hand, **the needs and practices of its user community**. Organizational culture and user community create a “**behavioural and interpretive template**” for the use of a technology.*

# Technology-in-practice (DH)

DH: technology-in-practice is software and data as used by scholarly communities of practice, within institutional and disciplinary concerns and constraints, in conversation with audiences.

# Technology-in-practice

# Users and Intended Audiences

# What groups of users/readers are invited in? What groups of users/readers are shut out?

Consider:

- expertise & interest; scholars & laypersons
- use of e.g. assistive technologies like screen readers, speech recognition, close captioning, etc.
- level of technical training (or lack of it)
- limited access to (powerful) computers
- limited internet bandwidth
- Tradeoffs & compromises

# User Experience, Interfaces, & Display

- UX (user experience) & usability testing
- User stories: descriptions, from a user's perspective, of a tool: who they are, what they want the tool to do, in what order...

# Data

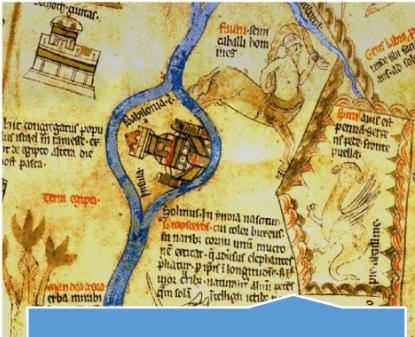
# Uses of Humanities Data

Ailleurs, où l'industrie du barbier était fort en honneur, les plaidoiries de ce travail étaient rares, et si Solon avait négligé ce décret, la prospérité avec laquelle il réussit à s'en occuper.

D'ailleurs, il y avait des cas où l'observation de la loi est étroitement en faute ou en droit. Si l'accusait d'un enfant au horceau, d'un absent ou d'une femme, il était nécessaire que quelqu'un pût témoigner de la vérité des paroles, qu'il n'y avait pas de moyen pour l'accusé de se défendre. Si l'Etat était directement intéressé à une affaire, si par exemple il s'agissait d'un décret ou d'un autre acte de l'autorité publique, alors aussi grave la défense de l'intérêt public fut abandonnée à l'initiative privée. La république désignait donc des orateurs, chose qui devait être faite avec le plus de soin possible. Dans tous ces cas, il avait bien fallu déroger à la règle, et admettre que le plaigneur se pouvait faire représenter par autres.

Il y avait également des cas où l'accusé, soit un citoyen, d'un matelot ou d'un soldat, dont la langue indocile se montrait récalcitrante à la parole, il devenait bien difficile d'appliquer la loi. C'est pourquoi il était nécessaire que l'accusé puisse moquer d'eux que de leur produire un plaidoyer absolument incroyable d'exposer le premier mot de son faire, disait-il, ce qui n'allait pas sans évidemment entraîner de graves pertes aux soldats blancs sous le harnais, qui dans maints combats n'avaient connu que d'énormes gênes et de lourdes pertes, exposées sans défense aux attaques d'un vil sycophante, et se voulant rassuré par une condamnation. L'argent qui devait servir à l'entretien de l'ordre public devait être dépensé pour que ce plaidoyer inexpérimenté pût appeler à son aide quelques ami bienveillant et discret, dont la langue officieuse voulait bien porter secours à sa défense. Et comme les amis habiles et dévoués étaient assez rares à Athènes, il

Text



Images & Maps



Sounds & Music



Films



Live or Digital Performances

ābēcēdē ☰ ●

→ abecedaria

Noun, f; 1x m. or n.

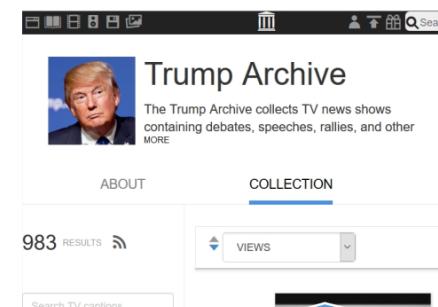
Att. sp.: F: abecede.

M. or N.: abecedes.

3 occ. (mainly in ByrM)

a. an ABC, the  
'in alphabet'

Words, Languages,  
Grammars  
*se abecedes*



Social Media  
Archives

# Classifying Hum Data

Unstructured

A corpus of literary texts

Semistructured

TEI-encoded

```
<msDesc>
  <msIdentifier>
    <settlement>Oxford
    </settlement>
    <repository>Bodleian Library
    </repository>
    <idno>MS. Add. A. 61
    </idno>
  </msIdentifier>
  <msContents>
```

The Travels  
of  
**Sir John Mandeville**

The version of the Cotton Manuscript

A	B	C	Dublin Core:Title
1 Dublin Core:Coverage	Dublin Core:Identifier		De gestis Regum Anglorum
2 London, British Library	MS Add. 23147	Add. 23147	Collection of English Poems by John
3 London, British Library	MS Add. 29729	Add. 29729	Collection of English poems, chiefly
4 London, British Library	MS Add. 34360	Add. 34360	Precedent book, etc.; copies of royal
5 London, British Library	MS Add. 48031A (formerly Yelverton Add. 48031A (formerly Yelverton 35)	Cotton Cleop. C.iii	Historical collections of Francis Thysius
6 London, British Library	MS Cotton Cleop. C.iii	Cotton Cleop. C.iv	Miscellaneous historiographical works
7 London, British Library	MS Cotton Cleop. C.iv	Cotton Faust. E.ii	Miscellaneous historiographical works
8 London, British Library	MS Cotton Faust. B.ii	Cotton Faust. E.v	Papers of members of the Society of
9 London, British Library	MS Cotton Faust. E.v	Cotton Galba E.vii	Chronicle of English history, AD 1-1500
10 London, British Library	MS Cotton Galba E.vii	Cotton Julius B.i	Chronicle of the city of London, AD 1-1500
11 London, British Library	MS Cotton Julius B.i	Cotton Julius B.ii	Chronicle of the city of London, AD 1-1500
12 London, British Library	MS Cotton Nero C.xi	Cotton Nero C.xi	New Chronicles of England and France
13 London, British Library	MS Cotton Nero C.xi	Cotton Vesp. A.xiv	Miscellaneous theological works
14 London, British Library	MS Cotton Vesp. A.xiv	Cotton Vesp. A.xxv	Versus vii psalmorum penitentiae; Liber fundationis ecclesie sancti Barnabae
15 London, British Library	MS Cotton Vesp. A.xxv	Cotton Vesp. B.ii	Cartulary of Westwood Priory; Utal
16 London, British Library	MS Cotton Vesp. B.ii	Cotton Vesp. E.ii	Chronicle of London to AD 1500; In
17 London, British Library	MS Cotton Vesp. E.ii	Cotton Vesp. E.vii	
18 London, British Library	MS Cotton Vitall. A.vii	Cotton Vitall. A.vii	

# But then, Thinking Hard about Data

- What “counts” cannot necessarily be counted
- Data representation = interpretation:
  - The process of modelling and collecting our data is an interpretive process that is shaped by our choices re. what aspects of the data we model; by our research question, argument, perspective, discipline, social context, institutional context, tools available etc.

# Humanities Data: Posner

It would be possible to enumerate all of the filmic conventions that recall the conventions of melodrama. Is there a villain? Is there a heroine? Are good and evil depicted in stark, black-and-white terms? You could even build a dataset like this and use it to show how film changed over time.

Title	Year	Virtuous Heroine	Cruel Villain	Terrible evil	Heroine in peril	Broad gestures
The Lonedale Operator	1911	x	x	x	x	x

# Data Universes

*Humanities data has depth in small universes. Our material has the depth of small universes.*



Sergiev Posad Museum of Toys,  
Russia, Public Domain,  
[https://commons.wikimedia.org/  
curid=5051554](https://commons.wikimedia.org/cu...)

# Data Models

- Data model: abstract representation of things or processes
- Data model: choice of virtual entities, relationships, properties/aspects → “toy universe”
- Data model: top row of your spreadsheet

# Metadata

- Metadata is structured data describing data. Think Author, Title, Date, Subject in your online library catalogue
- Standard metadata schemas include Dublin Core, MARC, and MODS (each of these schemas has a different set of information it collects).

We want your feedback on our new item page, use our feedback button to leave your comments.



### Peonies and Butterfly

Hokusai presents a fleeting moment from nature. The elegant blooms of the peonies sway in the breeze, and a colourful butterfly flutters down over them. The gentle breeze that captures the blossoming flowers is hinted by the sensitive graphic delineation of the flower forms.

Inscription: Japanese inscription on the side

Repository/Location: National Museum of Fine Arts (Valletta, Malta)

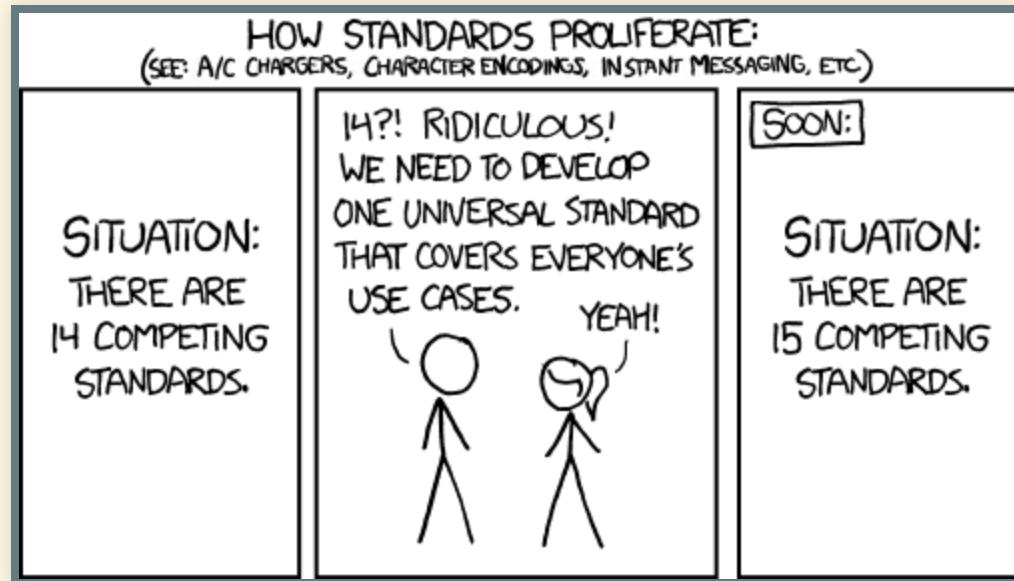
NEXT ITEM



Met gesloten ogen: De Adonis plant:

# Oops! Custom Metadata

# Metadata Standards

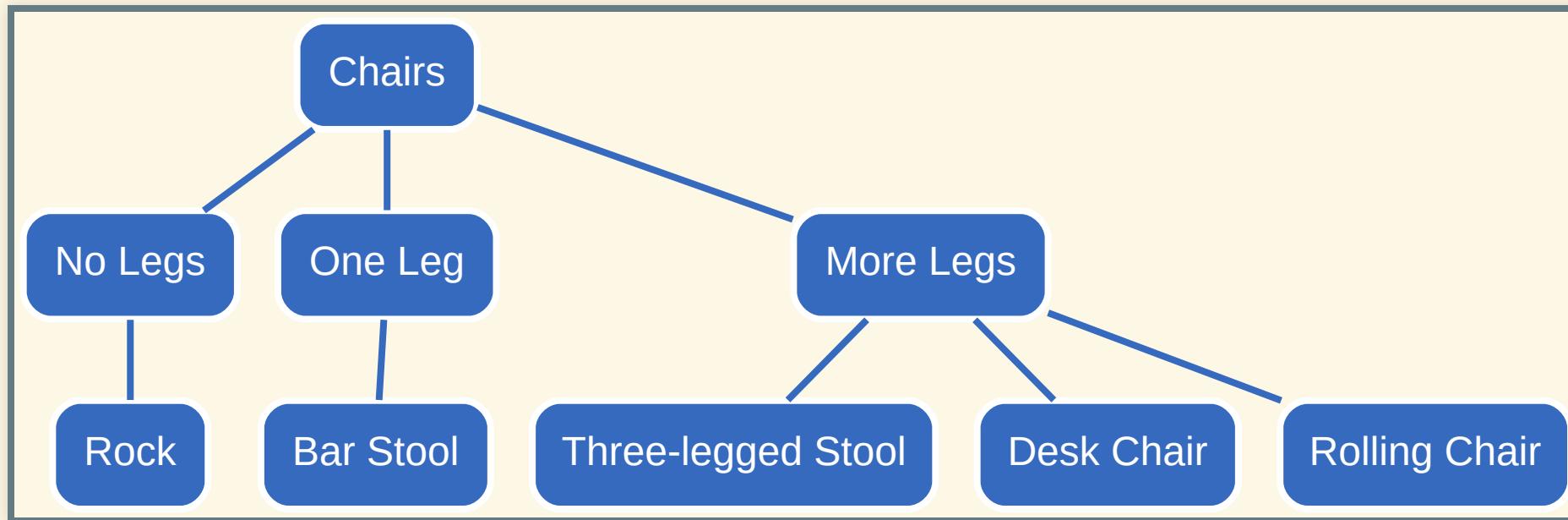


[\(\[MunroeRandallStandards\]\)](#)

# Vocabulary and Ontology

- **Controlled vocabulary:** set of permitted literal values to describe a given entity (e.g. chairs: desk chairs, rolling chairs, step stools, kitchen table chairs); may include categorization
- **Ontology:** making explicit the relationships between values in controlled vocabularies

# Vocabulary and Ontology

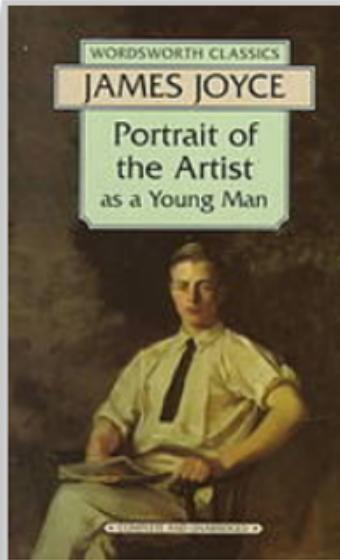


# Controlled Vocabulary

# Authority Control

- The practice of systematizing controlled vocabularies so descriptors of entities are unique, non-overlapping, and consistently used—e.g. having an accepted set of “correct” subject headings

# James Joyce



Joyce, James, 1882-1941.



Joyce, James (James Augustine Aloysius), 1882-1941



Joyce, James



James Joyce Irish novelist and poet



1882-1941 ג'יוס, ג'יימס



Джойс, Джеймс, 1882-1941



1941-1882 جویس، جیمز



Joyce, James (Irish author, 1882-1941)



Joyce, James Augustine Aloysius, 1882-1941.



# Gazetteers:

- place name dictionaries + maps + atlases + URIs  
(<https://pleiades.stoa.org/places/570491>)
- for time: (<https://perio.do/en/>)

# Research Data: Life Cycle



- “Data is always a partial representation of the object of study”  
([SchochBigSmartClean201])
- Failure-, incompleteness-, and pessimism-based approaches to data collection, modelling, analysis, and preservation

# Preservation-Ready Data

Data: accessible, usable, readable, preservable

- Human-readable and human-editable
- Separable from technical platforms
- Described via metadata standards used in your discipline
- Contextualized in clear documentation
- Housed in non-proprietary, open source standards and technologies
- Saved in, or reducible to, simple formats: .txt, TEI P5, .csv, JSON, .pdf, .jpg, tiff
- Embedded in your disciplinary community

# DH pipelines

# Reflection 1: DH Project Profile

[Quercus Link](#)

# Bibliography

- [DruckerHumanitiesApproachesGraphical2011] Drucker, Humanities Approaches to Graphical Display, *Digital Humanities Quarterly*, 005(1), 2011.
- [GaleyHowprototypeargues2010a] Galey & Ruecker, How a prototype argues, *Literary and Linguistic Computing*, 25(4), 405-424 (2010). doi:10.1093/llc/fcq021
- [Gladwellsociallifepaper2002] @miscGladwellsociallifepaper2002 The social life of paper: looking for method in the mess, shorttitle = The Social Life of Paper, author = Gladwell, Malcolm, year = 2002, month = may, url = /home/matt/Zotero/storage/BNH58GA9/retrieve.html, howpublished = http://link.galegroup.com/apps/doc/A84156646/LitRC?sid=googlesearch\_a, journal = The New Yorker, language = English
- [HarpringIntroductionControlledVocabularies2010] Harpring, Introduction to Controlled Vocabularies, 2010.