

**BEYOND OPEN ACCESS:  
REUSE, IMPACT AND THE ETHOS OF OPENNESS  
IN DIGITAL SCHOLARLY EDITING**

Anna-Maria Sichani

DiXiT fellow - Huygens KNAW

DiXiT Convention II, "Digital Editions: Academia, Society, Cultural Heritage" Cologne, 16- 18 March 2016



**Ich**   
**open**  
**access**

# Overview

## **The virtuous cycle of open digital content**

- Open Access and the 'ethos of openness'
- Open and useable digital content: value through reuse
- Measuring impact, expanding reuse

## **Digital Scholarly Editions made (not-so) open to reuse**

- Reading the numbers
- Revealing the barriers
- 'Re-use' in the history of modern textual scholarship
- Models of reuse and/as aspects of value
- Why it really matters?

# OPEN ACCESS

OA removes *price barriers* (subscriptions, licensing fees, pay-per-view fees) and *permission barriers* (most copyright and licensing restrictions).

(Peter Suber, [Open Access](#), MIT Press, 2012)

Free immediate access to, and unrestricted reuse of scholarly research

(PLOS, <https://www.plos.org/open-access/>)

# The Open Definition



The **Open Definition** sets out principles that define “openness” in relation to **data and content**.

It makes **precise** the meaning of “open” in the terms “**open data**” and “**open content**” and thereby ensures **quality** and encourages **compatibility** between different pools of open material.

It can be summed up in the statement that:

*“Open means **anyone** can **freely access, use, modify, and share** for **any purpose** (subject, at most, to requirements that preserve provenance and openness).”*

Put most succinctly:

*“Open data and content can be **freely used, modified, and shared** by **anyone** for **any purpose**”*



open source<sup>®</sup>  
initiative



OPEN  
GLAM



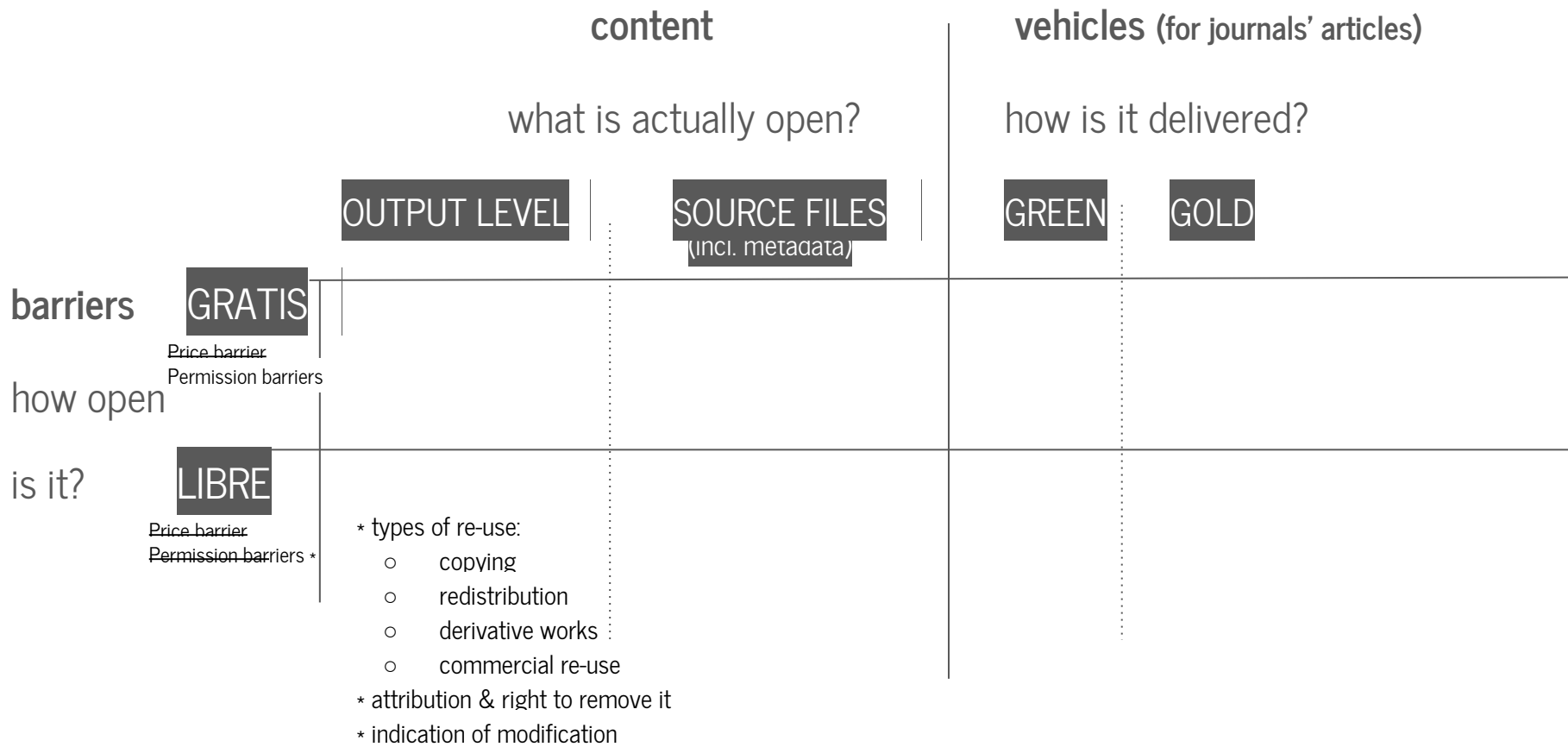


# PENNESS

principles:

1. Availability and Access
2. Re-use and Redistribution
3. Universal Participation

# Open Access Matrix



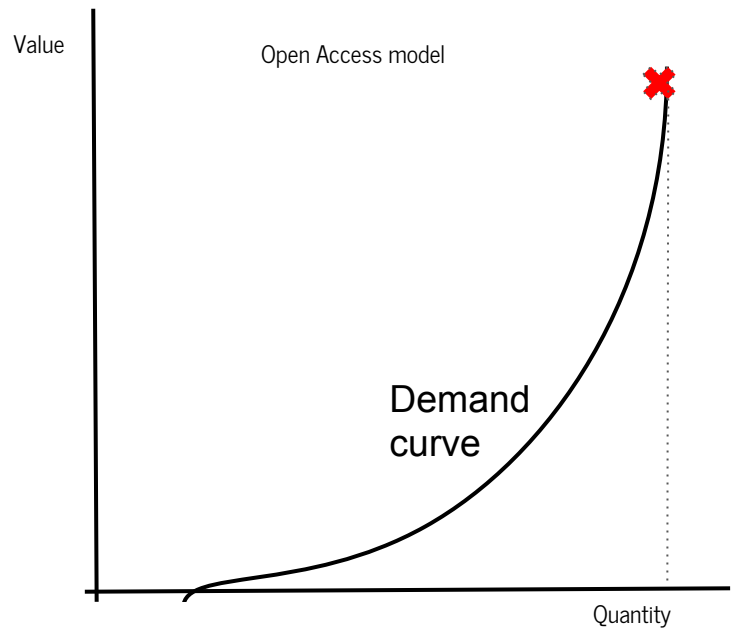
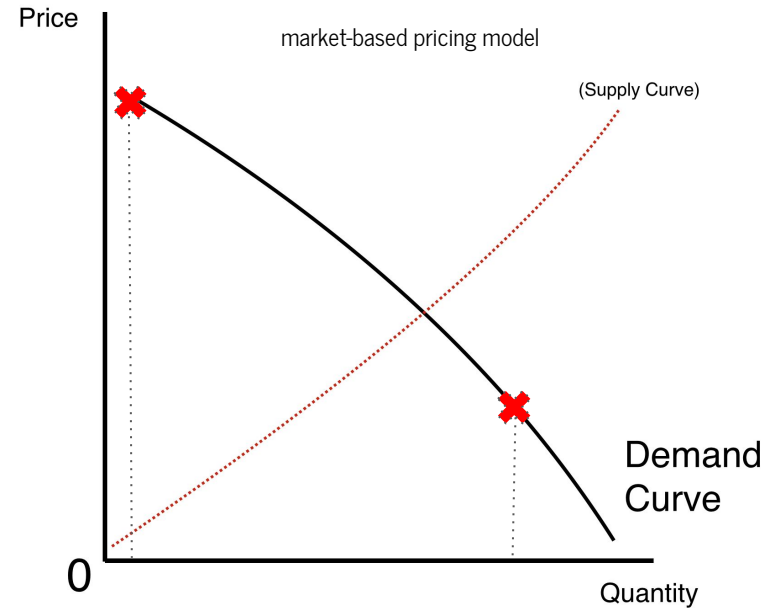


# OA and value through re-use

Instead of making content valuable by making it scarce,

Open Access makes knowledge valuable by making it widely available for re-use.

Law of demand / supply - Scarcity principle



# Impact & value assessment of open digital content

## Qualitative methods

stakeholder interviews  
resource surveys  
user feedback  
focus groups  
questionnaires

## Quantitative methods

webometrics / altmetrics  
log file analysis  
scientometric / bibliometric analysis  
content analysis

## Selective Bibliography

Warwick, C., et al. (2006), *The LAIRAH Project: Log Analysis of Digital Resources in the Arts and Humanities. Final Report to the Arts and Humanities Research Council*. 2006, School of Library, Archive and Information Studies, University College London: London  
<http://www.ucl.ac.uk/infostudies/claire-warwick/publications/LAIRAHreport.pdf>

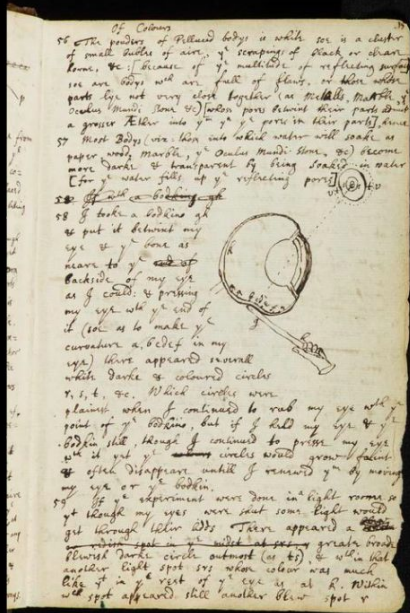
Meyer, E., K. Eccles, M. Thelwall and C. Madsen (2009). *Final Report to JISC on the Usage and Impact Study of JISC-funded Phase 1 Digitisation Projects & the Toolkit for the Impact of Digitised Resources (TIDSR)* [http://microsites.oii.ox.ac.uk/tidsr/system/files/TIDSR\\_FinalReport\\_20July2009.pdf](http://microsites.oii.ox.ac.uk/tidsr/system/files/TIDSR_FinalReport_20July2009.pdf)

*Toolkit for the Impact of Digitised Scholarly Resources (TIDSR)*, <http://microsites.oii.ox.ac.uk/tidsr/>

Simon Tanner (2012), *Measuring the Impact of Digital Resources: The Balanced Value Impact Model*. King's College London, [www.kdcs.kcl.ac.uk/innovation/impact.html](http://www.kdcs.kcl.ac.uk/innovation/impact.html)

*Evaluating and Measuring the Value, Use and Impact of Digital Collections* (2012), Lorna Hughes (ed), Facet Publishing.

Lorna M. Hughes, Paul S. Eil, Gareth A. G. Knight, Milena Dobrova (2013). "Assessing and measuring impact of a digital collection in the humanities: An analysis of the SPHERE (Stormont Parliamentary Hansards: Embedded in Research and Education) Project", *Digital Scholarship in the Humanities*, Volume 30, Issue 2, p. 183 - 198.



Page: 15

Home / Newton Papers / MS Add.3975

About Contents Thumbnails View more options ▾

## Download



Images made available for download are licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License (CC BY-NC 3.0)



Request reproduction rights to this image



Add this image / page to your personal book



Download original metadata for this document

## Share

If you want to share this page with o

<http://cucl.lib.cam.ac.uk/view/>

# Things are changing...

Search keyword(s)

Items &gt; Theatres -- Germany -- Cologne -- Opernhaus (1909 - 1944)



## Theatres -- Germany -- Cologne -- Opernhaus (1909 - 1944)

TYPE OF RESOURCE

still image

GENRE

Photographs

DIVISION

Billy Rose Theatre D...

[More Details](#)[Cite This Item](#)IMAGE ID  
TH-55826

PERMALINK

<http://digitalcollections.r...>

DOWNLOAD OPTIONS

300px

760px

ORDER

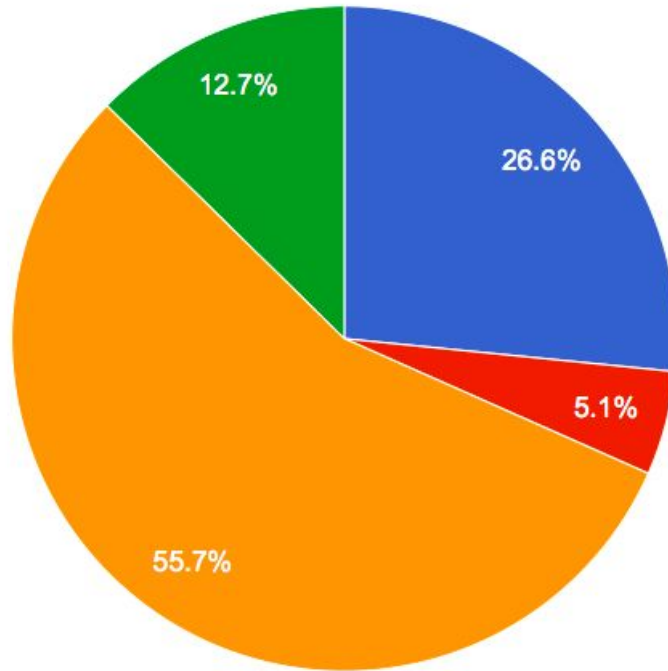
[Original Scan](#)[Art Print](#)[All download options ▾](#)

“free in every sense of the term, including the use, copying, distribution, and modification”  
(NEH 2012)

# Digital Scholarly Editions and Open Access

**Ich**   
**open**  
**access**

## Catalogue of Digital Editions and Open Access

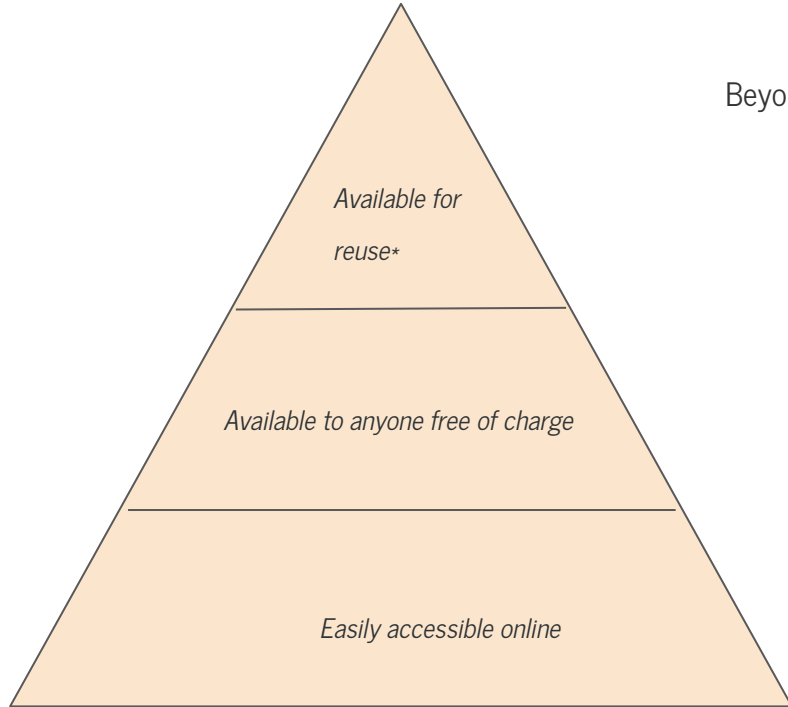


- 0: Proprietary, all material is copyrighted. The 'source' is closed and not reusable by other research projects. To access the material, users must pay a subscription
- 0.5: Same as above, but the subscription is free of charge
- 1: Open Access. The texts may be accessed through specific software, but the source is not accessible
- 1.5: Open Access and Open Source. All data underlying the digital edition is freely available for access, study, redistribution and improvement (reuse)

data from Greta Franzini's Catalogue of Digital Editions

[https://github.com/gfranzini/digEds\\_cat](https://github.com/gfranzini/digEds_cat)

# Digital Scholarly Editions and Open Access



Beyond-the-output level' : openly available for reuse and redistribution < Open Source >

*\*without restriction, except that attribution needs to be given to the source*

publicly funded == gratis OA by default < Open Content in output level >

making [ ... ]'s vast work for the first time freely and conveniently accessible to scholars, students and general readers

# Barriers to Open Access

**what ?** cross - domain content & different copyright \*  
content / data beyond the output

**who ?** “Who owns what?”

control over your data

**how ?** still “zoned to print”, browse & read a finished and closed product

**Look and enjoy...**



***BUT PLEASE* DON'T TOUCH!**





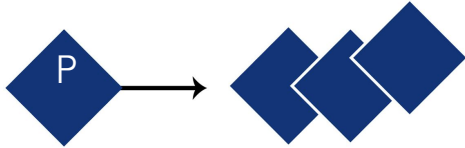


**We deserve more than a read-only world.**

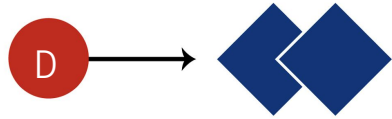
(courtesy SPARC)

\*for digital scholarly editing as well

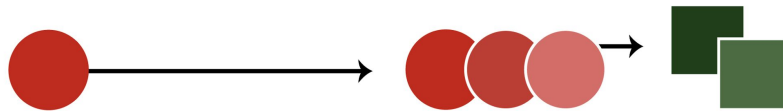
# 'Re-use' in the history of modern textual scholarship



icut: Greenwood Press, 1984), p. 393.  
4 'Continuous manuscript' is Hans Walter Gabler's term for the successive autograph notations that he uses as the copytext for his edition of *Ulysses*. See Gabler's Afterword to '*Ulysses*': *A Critical and Synoptic Edition* (New York: Garland, 1984), pp. 1894–6.  
5 Of the forty extant epiphanies, twenty-two (in Joyce's hand) are housed in



10. The Beckett Digital Manuscript Project, module 1 (BDMP1), edited by Dirk Van Hulle and Vincent Neyt (Brussels: University Press Antwerp), online at <<http://www.beckettarchive.org>>; Van Hulle (2011a) 102).



*[Electronic texts] are artefacts or mechanisms which, like the older technology of the book, are amenable to material and historical forms of investigation, challeng[ing] textual critics to respond to the new medium in terms of its own materiality, architecture and functioning, as distinct from those of print.*

# Re-use scenarios and their value

## 1. ***Open Source Critical Editions***

(Bodard, Gabriel, Juan Garcés (2009). "Open Source Critical Editions: A Rationale." In *Text Editing, Print, and the Digital World*, Kathryn Sutherland, Marilyn Deegan (eds), Ashgate, 83–98)

- ✓ Critical transparency: by revealing the code behind the output, the research behind the text, the decisions which are part of scholarly publication.

## 2. ***Agile editions (easily transformed, re-purposable)***

(Cummings, James (2009). "Converting Saint Paul: A New TEI P5 Edition of The Conversion of Saint Paul Using Stand-off Methodology." *Literary and Linguistic Computing* 24 (3): 307–17)

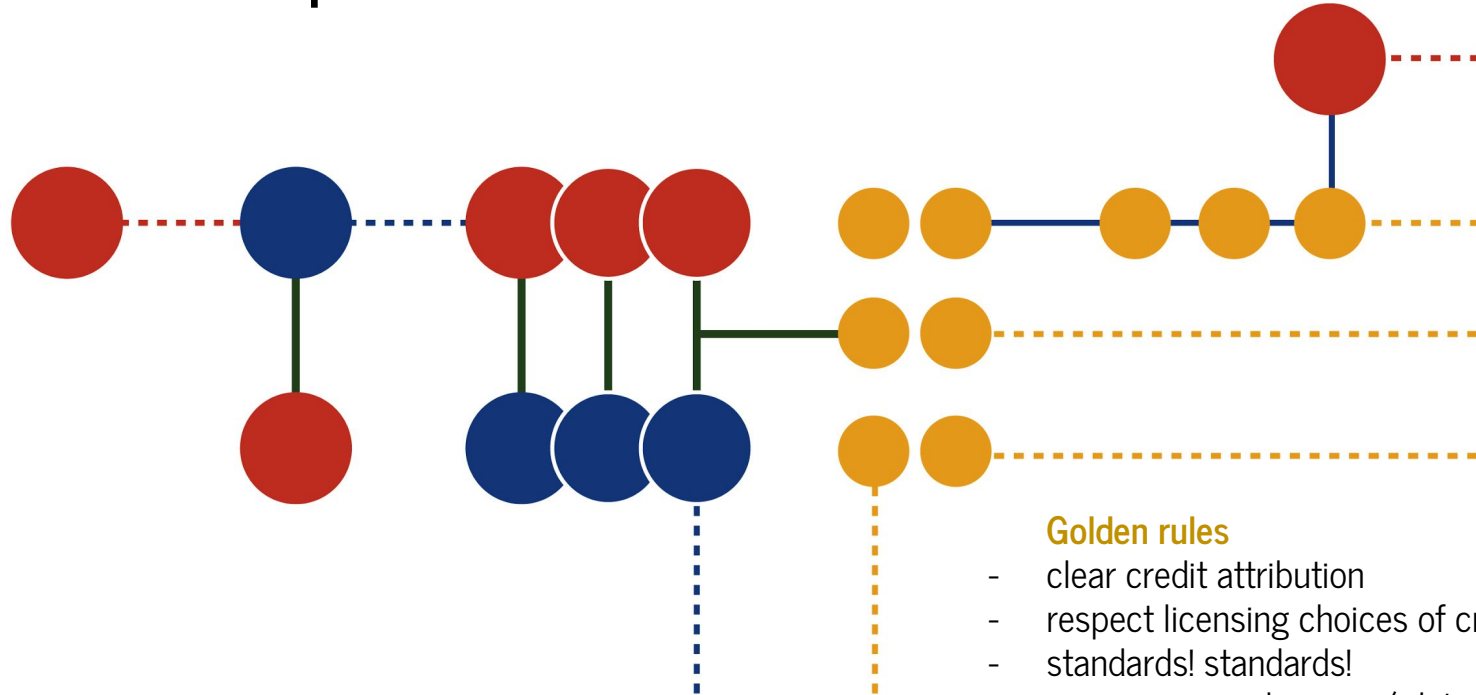
- ✓ opening & distribution of underlying XML files → discovery of our joint errors and misunderstandings of the TEI Guidelines

## 3. ***Digital scholarly edition 2.0***

(Boot, Peter, and Joris van Zundert (2011). "The Digital Edition 2.0 and The Digital Library: Services, Not Resources." In *Digitale Edition Und Forschungsbibliothek, Beitrage Der Fachtagung Im Philosophicum Der Universita't Mainz Am 13. Und 14. January 2011*, 141–152)

- ✓ a decentralised, diverse and distributed architecture of openly available sources, services and functionalities (VRE), initially enabled through cloud computing and cloud storage

# open - ended deep access model



## Golden rules

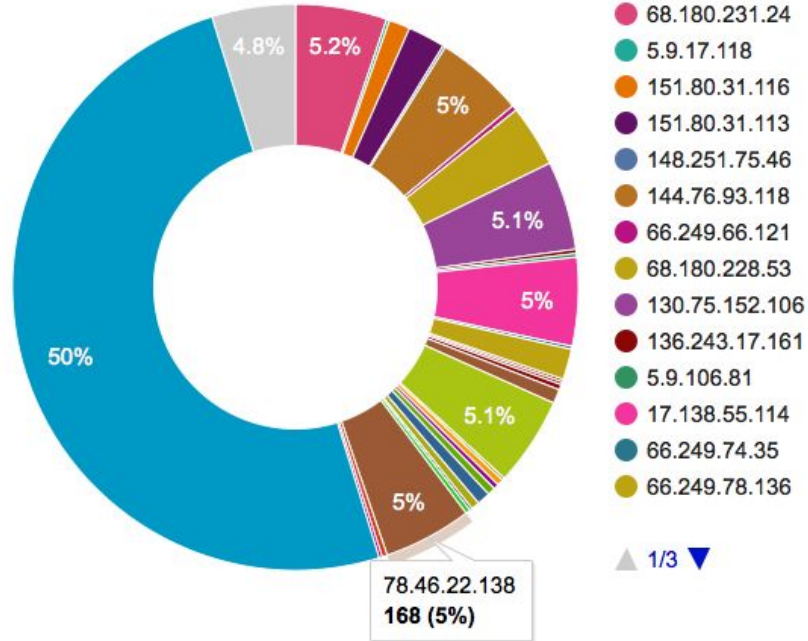
- clear credit attribution
- respect licensing choices of creators
- standards! standards!
- versioning mechanism / data management
- keeping track / communicate your reuse !

# open - ended **deep access** model

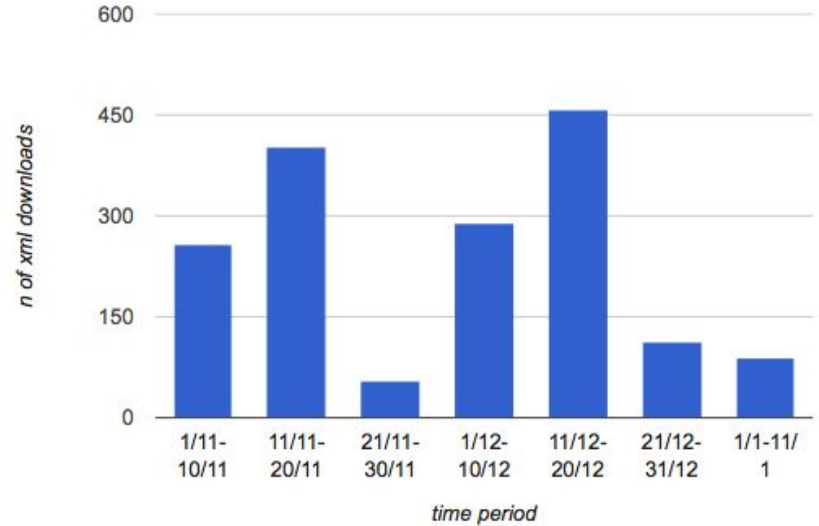
- source files
  - encoded XML
  - related schemas & ODDs
  - scripts for processing-visualisation
  - transformation scripts and query algorithms
  - entity relational models
  - metadata etc.
- **(even deeper)** parts of the project's documentation
  - (successful or not) funding proposals
  - feasibility studies
  - editorial conventions
  - budget plans etc.

**MAY THE \_\_\_\_\_**  
**SOURCE**  
**\_\_\_\_\_ BE WITH**  
**YOU**

how many xml have been downloaded per user?



Total downloads: 1665



Log Files Analysis [11/2015-1/2016]

Briefe und Texte aus dem intellektuellen Berlin um 1800 <<http://tei.ibi.hu-berlin.de/berliner-intellektuelle/>>

# open - ended deep access model

## ways of reuse - aspects of value :

1. base/raw material for new kind of research and scholarship
2. ideal test-bed for new computational/analytical approaches
3. easy way to produce derivatives and varying outputs (ePub, PDF) ⇒ add-ons / on demand services
4. reference material for teaching purposes
5. great potential for cost and time-avoidance
6. distributed archiving solution



# JONATHAN SWIFT ARCHIVE



## Jonathan Swift: Gulliver's Travels, Part IV Chapter vi

Orientation version: #235

Variant version: #189, #241

Home About

Orientation version Variant version

Choose an orientation version: 238

show

Chap. VI.

*A Continuation of the State of England, so well governed by a Queen as to need no first Minister. The Character of such an one in some European Courts.*

MY Master was yet wholly at a Loss to understand what Motives could incite this Race of Lawyers to perplex, disquiet, and weary themselves, and engage in a Confederacy of Injustice, merely for the Sake of injuring their Fellow-Animals; neither could he comprehend what I meant in saying they did it for Hire. Whereupon I was at much Pains to describe to him the Use of Money, the Materials it was made of, and the Value of the Metals; that when a Yahoo had got a great Store of this precious Substance, he was able to purchase whatever he had a mind to, the finest Clothing, the noblest Houses, the great Tracts of Land, the most costly Meats and Drinks, and have his Choice of the most beautiful Females; Therefore since Money alone, was able to perform all these Feats, our Yahoos thought, they could to spend or to save, as they found themselves inclined from their natural Bent, either to Profusion or Avarice. That the rich Man enjoyed the Fruit of the poor Man's Labour, were a thousand to one in proportion to the Former. That the Bulk of our People were forced to live miserably, by labouring every Day for small Wages to maintain me. I enlarged my self much on these

229 other Particulars to the same Purpose: But his Honour was still to seek: For he went upon a supposition that all Animals had a Taste, and especially enumerated those that were fit for Liquors to be distilled out of, and for Female Yahoos to be mated with. I assured him, that this whole Globe of Earth must be at least three times gone round, before one of our better Men could be wholly without Fresh Water, and the People put to the Necessity of sending over the Sea for Drink. I replied, that the Quantity of Food, more than its Inhabitants are able to consume, as well as Liquors extracted from Grain, or pressed out of the Fruit of certain Trees, which made excellent Drink, and the same Proportion in every other convenience of Life. But in order to feed the Luxury and Intemperance of the Males, and the Vanity of the Females, we sent away the greatest Part of our necessary Things to other Countries, from whence in return we brought the Materials of Diseases, Folly, and Vice, to spend among ourselves. Hence it follows of Necessity, that vast Numbers of our People are compelled to seek their Liveliness by Begging, Robbing, Stealing, Cheating, Pimping, Forswearing, Flattering, Suborning, Forging, Gaming, Lying, Fawning, Hectoring, Voting, Scribbling, Stargazing, Poisoning, Whoring, Canting, Libelling, Free thinking, and the like Occupations: Every one of which I was at much Pains to make him understand.

History of the Project

Abbreviations

Citation

Copyright

Staff

Download

app 66

- <<other</c><pr>229</pr> other Particulars (#235)
- other Particulars, (#189)
- other Particulars (#241)

app 28

- that (#235 #241)
- that, (#189)

app 57

- proportion (#235 #241)
- Proportion (#189)



## Ψηφιακές εκδόσεις χειρογράφων

### Copyright and Download

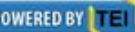
The copyright of the images belong to  
The XML and XSLT are available for do



Around a sequence and some notes of P  
Based on a work at [gallica.bnf.fr](http://gallica.bnf.fr).  
Permissions beyond the scope of this lic

### Downloads

The prototype comes **as it is** and we  
You can download the XML source file



You can also download the XSLT that

### iv. Χρονικότητα συγγραφής VS Χρονικότητα ανάγνωσης

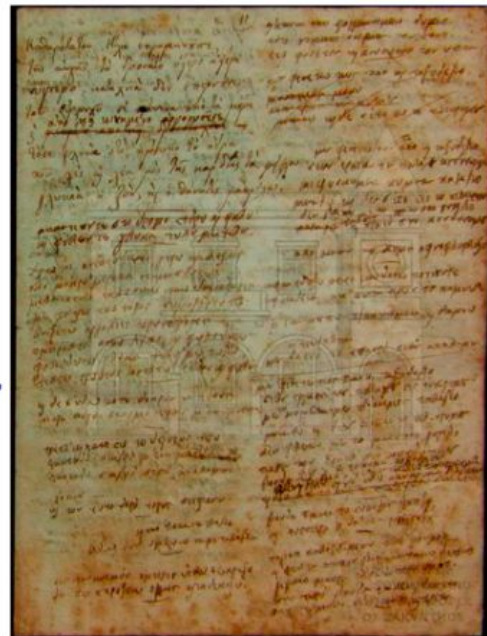
(zones appearing according to the order in which they were written VS in which they should be read)

π.χ. [http://research.cch.kcl.ac.uk/proust\\_prototype/index.html](http://research.cch.kcl.ac.uk/proust_prototype/index.html)

```
<zone n="1.2" ulx="26.335878" uly="380.75854" lrx="150.53915" lry="172.82443"  
rend="font-size:12 change="#ch06" xml:id="t2">
```

```
<zone rend="nobox change="#ch06-2" ulx="26.335878" uly="380.75854" lrx="150.53915"  
lry="172.82443">  
<metamark function="join" target="#x1">
```

→ Πώς αυτό μπορεί να εφαρμοστεί στην περίπτωση  
των σολωμικών αυτογράφων;



## Re - use of XML transcriptions

1. *Collected Works of Jeremy Bentham (70 vols)*

**Estimated savings:** £400,000 & up to 6 months of research staff time per vol.

2. *AHRC-funded project 'Convict Australia and Utilitarianism: Jeremy Bentham's Writings on Australia' (1/2016-12/2020)*

**Estimated savings:** > £80,000



# Transcribe Bentham

A Participatory Initiative

	Average no. of transcripts per year	Earliest date when all pages would be transcribed
Without <i>Transcribe Bentham</i>	1,076	2081
Overall <i>Transcribe Bentham</i> transcription rate	2,704	2036
1 Jan—30 Sept 2014 <i>Transcribe Bentham</i> transcription rate	5,564	2025

TAPAS Project



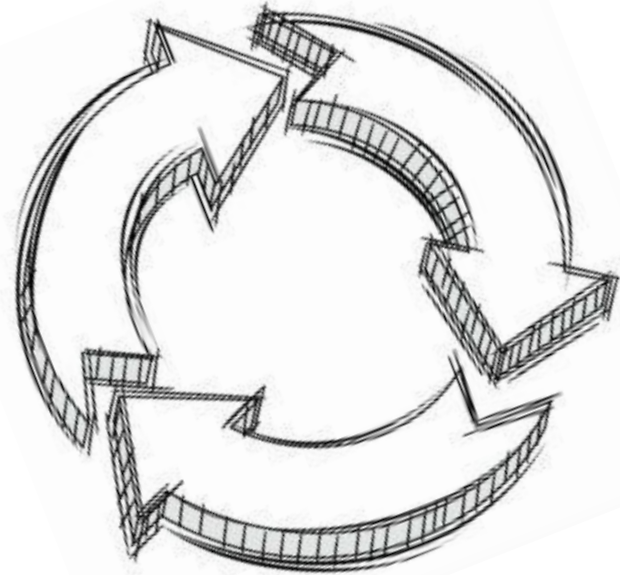
**REUSE  
IT  
OR  
LOSE  
IT**



We are here to make  
digital scholarly editions

**OPEN**

**U** **SABLE**  
**SED**  
**SEFUL**



@amsichani

anna-maria.sichani@huygens.knaw.nl