The Missing Link Between Digital Humanities and Open Science
2

https://books.google.fr/books?id=tb4UAAAAQAAJ&hl=fr&pg=PA632#v=onepage&q&f=false
Scimus ex dominico praecepto, quod caro infirma (626A) sit, spiritus promptus. Non ergo nobis blandiamur, quia Dominus concepit carнем infirmam esse. Proprietae enim praedictm spiritum promptum, ut ostenderet, quid cui debet esse subjectum, salutem, ut caro servavit spiritu, infirmor fortiori, ut ab eo etiam ipsam fortitudinem assumat. Collocatibur spiritus cum carne de communi salutе, nec iam de inconmodis carcerum, sed ipso ago et praeclor cogitamus. Timentis forsan caro gladium gravem, et crucem exsangue, et rabie bestiarum, et summam ignum poenam, et omne carnifici ingenium in tormento. Sed spiritus contra ponat sibi et carri, acerba licet ista, a multi tamen aequo animo excepta, imo et ultra appetita, fameae et gloriae causa; nec a visis tantum, sed etiam a foeminis, ut vos quoque, beneficae, (626B) sexui vestro respondate. Longum est, si numeremus singulos, qui se gladio confecerint, animo suo duci. De foeminis ad manum est Lucretia, quae visi stuprati passa, cultum sibi sedest in conspectu propinquirum, ut gloriae castissimi sui pararet, Mutius, manum suam deexteram in aram cremavit, ut hoc factum eius fama haberet. Minus fecerunt philosophi: Heracleitus, qui se bubulo stercore obtulerat exeuntiam; item Empedocles, qui in ignes Aetnae mons dissipavit; et Periphragmus, qui non olim se rogo immisit: cum foeminae quoque contemptis vere: Didius, ne post virum dilectissimum nubere cogeretur; item quum Asa Rumblasis uxor, quae iam ardentie Carmagnanum suum supplicium Scipioni vidisset, cum falls suis in incendium patae devolavit. Regulus, dux Romanorum, (626C) captus a Carmagnenibus, cum se unum pro multis captivis Carmagnenibus compensari nolebisset, maluit hostibus reddi et in anicis genus etipatus undique extrinsecus davis transfluxerat, tot cruces sensit. Historia foeminae ibenis appetit, et utique aspidae, serpentae tauro vel urso horrores, quasi Cleopatra immisit sibi, non in manus immune perveniens. Sed mortis metus non tantum est, quantus tormentorum. Itaque, cessat carnis filia meretricin (626A) athenensis? quae conscius coniuratio, cum propitiae tonzeuger et tyranno, et non prodit coniurator, et novissimis linguam suam comestam in faciem tyranni exsputat, ut nihil agere se scret tormenta, etsi ultra perseveraret. Nam quod hodie apud Lacedaemones solemnes maxima esse, &quosfortius, id est, flagellato, non laetet. In quo sacro, ante aram nobis quique adolecentes, flagello effigentur, ascendentibus parentibus et propriis, et usi perseverent adhortantibus. Ornamentum enim et gloria deputatur maior quidem titulo, si anima potius cessaret plagas, quam corpus. Igitur, si tantum terraeae gloriae licet de corpos et animi vigore, ut gladium, ignem, crucem bestias, tormenta contemnent, sub praemia laudis humanae, possim dicere, modicse sunt istae passiones (626B) ad consecutiem gloriae cœlestis et divinae mercedis. Si tantum vitrum? quanti margantem? Quis ergo non libentissime tantum pro vero habeat erogare, quantum ali pro fallo?
A Song

Tu strange, this heart within my breast,
  Reason opeing, and her powers,
Cannot one gentle moment rest,
  Unlesse it knows what's done in yours.

In vain, I ask it of your eyes,
  Which subtlest was my fears controul,
For art, has taught them to disguise
  Which nature made, t'explaine the Soul.

In vain, that sound, your voyce affords
  Flatters sometimes, my easy mind.
But of too vast extent are words,
  In them, the Jewell truth to find.

Then let my fond enquiry cease,
  And so let all my troubles end.
For sure, that heart shall ne'er know peace,
  Which on another's, does depend.

Hear this poem

ABOUT THE POEM

The psychology of courtship on display in this song prompted Charles H. Hinnant to observe that, like many of Finch's songs and poems, the work reflects an "obsession with love's uncertainties" (The Poetry of Anne Finch: An Essay in Interpretation [Newark: Univ. of Delaware Press, 1994], pp. 52-53). Finch's song resembles those incorporated into plays, which reflect courtship dynamics between...
'Tis strange, this heart within my breast text: Anne Finch Digital ...
... knows what's done in yours. In vain, I ask it of your eyes, 8 Which subtly you'd my fears controle. For art, has taught them to disguise Which nature made, to explain the Soul. In vain, that sound ...
library.uncg.edu/dp/annefinch/item.aspx?pid=1828&uuid=853
18k – Cached

Anne Finch Transcript
... knows what's done in Yours. In vain I ask it of your Eyes, Which subtly you'll my Fears controle. For Art has taught them to disguise, Which Nature made I explain the Soul. In vain that Sound, your ...
library.uncg.edu/dp/annefinch/transcriptOnly.aspx?pid=1828&uuid=853&sc=853
8k – Cached
St. Paul's Cathedral

St. Paul's Cathedral was—and remains—an important church in London. In 962, while London was occupied by the Danes, St. Paul's monastery was burnt and raised anew. The Church survived the Norman conquest of 1066, but in 1007 it was burnt again. An ambitious Bishop named Maurice took the opportunity to build a new St. Paul's, even petitioning the king to offer a piece of land belonging to one of his castles (Times 115). The building Maurice initiated would become the cathedral of St. Paul's which survived.
A STRANGE SIGHTED TRAVELLER

AN honest Country foole being gentle bred,
Wras by an odd conceited humor led,
To trauell and some English falshions see,
With such frange fights as heere at London be.
Stuffing his purse with a good golden some,
This wandring knight did to the Cittie come,
And there a seruingham he entertaines,
An hosteller in Newgate not remains.
He shew'd his Maufer fights to him most strange,
Great tall Pauls Steeple and the royall-Exchange:
The Boile at Billings-gate and London stone,
And at White-Hall the mondious great Whales bone,
Brought him to the banck-side where Beares do dwell
And into Shor-ditch where the whores keepe hell,
Shew'd him the Lyons, Gyants in Guild-Hall,
King Lud at Lud-gate the Baboons and all,
At length his man, on all he had din pray,
Shew'd him a theueuilh trick and ran away,
The Traveller turnd home exceeding chauill,
And swore in London he had seene the Deuil.
A Strange Sighted Traveller

AN honest Country foole being gentle bred,
Was by an oddle conceited humor led,
To traveill and some English fallacies see,
With such strange fights as heere at London be.
Stuffing his purse with a good golden some,
This wandering Knight did to the Cittie come,
And there a Seruairman he entertaines,
An homelie in Newgate not remains.
He shewed his Manyer fights to him most strange,
Great tall Pauls Steeple and the royall Exchange;
The Boile at Billings gate and London stone,
And at White Hall the monstrous great Whales bone,
Brought him to the bank side where Beares do dwell
And unto Shor-ditch where the whores keepe hell,
Shew'd him the Lyons, Gyants in Guild-Hall,
King Lud at Lud gate the Babeunes and all,
At length his man, on all he had did pray,
Shew'd him a deceitfull trick and ran away,
The Traveller turnd home exceeding glad,
And swore in London he had seen the Peuell.
pure humanitus adinventam iudex idoneus ferret pro fide sententiam


[Conclusio]

2 Et inprimis protestatur quod fides non subicitur humano iudicio, et haec est una conclusio. Patet quia fides est donum Dei suparnaturale et est de illis de quibus written in the manuscript: "proloegum Sententiarum" and "humanitas adinventam iudex idoneus ferret pro fide sententiam".
### Plaoul Commentary

#### Table of Questions

<table>
<thead>
<tr>
<th>filter list</th>
<th>Modus</th>
<th>Status</th>
<th>Build</th>
<th>GitHub Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principium I</td>
<td>In Progress</td>
<td>build passing</td>
<td>View on GitHub</td>
<td></td>
</tr>
<tr>
<td>Principium II</td>
<td>In Progress</td>
<td>build passing</td>
<td>View on GitHub</td>
<td></td>
</tr>
<tr>
<td>Principium III</td>
<td>In Progress</td>
<td>build passing</td>
<td>View on GitHub</td>
<td></td>
</tr>
<tr>
<td>Principium IV</td>
<td>In Progress</td>
<td>build passing</td>
<td>View on GitHub</td>
<td></td>
</tr>
<tr>
<td>Lectio 1, De fide</td>
<td>draft</td>
<td>build passing</td>
<td>View on GitHub</td>
<td></td>
</tr>
<tr>
<td>Lectio 2, De fide</td>
<td>In Progress</td>
<td>build passing</td>
<td>View on GitHub</td>
<td></td>
</tr>
<tr>
<td>Lectio 3, De fide</td>
<td>In Progress</td>
<td>build passing</td>
<td>View on GitHub</td>
<td></td>
</tr>
<tr>
<td>Lectio 4, De fide</td>
<td>In Progress</td>
<td>build passing</td>
<td>View on GitHub</td>
<td></td>
</tr>
<tr>
<td>Lectio 5, De fide</td>
<td>In Progress</td>
<td>build passing</td>
<td>View on GitHub</td>
<td></td>
</tr>
<tr>
<td>Lectio 6, De fide</td>
<td>In Progress</td>
<td>build passing</td>
<td>View on GitHub</td>
<td></td>
</tr>
<tr>
<td>Lectio 7, De fide</td>
<td>In Progress</td>
<td>build passing</td>
<td>View on GitHub</td>
<td></td>
</tr>
<tr>
<td>Lectio 8, De fide</td>
<td>In Progress</td>
<td>build passing</td>
<td>View on GitHub</td>
<td></td>
</tr>
<tr>
<td>Lectio 9, De fide</td>
<td>In Progress</td>
<td>build passing</td>
<td>View on GitHub</td>
<td></td>
</tr>
<tr>
<td>Lectio 10, De fide</td>
<td>In Progress</td>
<td>build passing</td>
<td>View on GitHub</td>
<td></td>
</tr>
<tr>
<td>File Name</td>
<td>Size</td>
<td>Date</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>-------</td>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>app_lectio1.xml</td>
<td>60.5 KB</td>
<td>2014-08-14</td>
<td>app changes</td>
<td></td>
</tr>
<tr>
<td>erlang_lectio1.xml</td>
<td>18.7 KB</td>
<td>2014-02-22</td>
<td>/Users/JCWitt/WebPages/petrusplaoui-mirror/projectfiles/Textfiles</td>
<td></td>
</tr>
<tr>
<td>lectio1.xml</td>
<td>79.6 KB</td>
<td>2016-07-06</td>
<td>started adjusting app types to reflect to schema guidelines; also shifted ref-&gt;seg to cit-&gt;ref</td>
<td></td>
</tr>
<tr>
<td>notes_lectio1.xml</td>
<td>8.5 KB</td>
<td>2014-07-06</td>
<td>auto</td>
<td></td>
</tr>
<tr>
<td>pal_lectio1.xml</td>
<td>27.6 KB</td>
<td>2014-02-22</td>
<td>/Users/JCWitt/WebPages/petrusplaoui-mirror/projectfiles/Textfiles</td>
<td></td>
</tr>
<tr>
<td>reims_lectio1.xml</td>
<td>42.7 KB</td>
<td>2015-03-09</td>
<td>Auto generated facsimile, surface, and zone coordinates into all diplomatic transcriptions</td>
<td></td>
</tr>
<tr>
<td>sorb_lectio1.xml</td>
<td>42.8 KB</td>
<td>2015-03-09</td>
<td>Auto generated facsimile, surface, and zone coordinates into all diplomatic transcriptions</td>
<td></td>
</tr>
<tr>
<td>evict_lectio1.xml</td>
<td>44.5 KB</td>
<td>2015-03-09</td>
<td>Auto generated facsimile, surface, and zone coordinates into all diplomatic transcriptions</td>
<td></td>
</tr>
<tr>
<td>vat_lectio1.xml</td>
<td>36.7 KB</td>
<td>2015-03-09</td>
<td>Auto generated facsimile, surface, and zone coordinates into all diplomatic transcriptions</td>
<td></td>
</tr>
<tr>
<td>versionlog_lectio1.xml</td>
<td>1.8 KB</td>
<td>2013-08-12</td>
<td>test of function commit message</td>
<td></td>
</tr>
</tbody>
</table>
# Information for: Lectio 1, De fide

(select links below for more information)

**Publication Information** (pubInfo datastream | metadata count: 4)

<table>
<thead>
<tr>
<th>dc11:title</th>
<th>Lectio 1, De fide</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://scta.info/property/status">http://scta.info/property/status</a></td>
<td>draft</td>
</tr>
<tr>
<td>rdf:type</td>
<td><a href="http://scta.info/resource/expression">http://scta.info/resource/expression</a></td>
</tr>
</tbody>
</table>

**Content Information** (contentInfo datastream | metadata count: 0)

**Linking Information** (linkingInfo datastream | metadata count: 6)

<table>
<thead>
<tr>
<th>dc:isPartOf</th>
<th><a href="http://scta.info/resource/pp-deFide">http://scta.info/resource/pp-deFide</a></th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://scta.info/property/hasManifestation">http://scta.info/property/hasManifestation</a></td>
<td><a href="http://scta.info/resource/lectio1/critical">http://scta.info/resource/lectio1/critical</a></td>
</tr>
<tr>
<td><a href="http://scta.info/property/hasManifestation">http://scta.info/property/hasManifestation</a></td>
<td><a href="http://scta.info/resource/lectio1/reims">http://scta.info/resource/lectio1/reims</a></td>
</tr>
<tr>
<td><a href="http://scta.info/property/hasManifestation">http://scta.info/property/hasManifestation</a></td>
<td><a href="http://scta.info/resource/lectio1/sorb">http://scta.info/resource/lectio1/sorb</a></td>
</tr>
<tr>
<td><a href="http://scta.info/property/hasManifestation">http://scta.info/property/hasManifestation</a></td>
<td><a href="http://scta.info/resource/lectio1/svict">http://scta.info/resource/lectio1/svict</a></td>
</tr>
<tr>
<td><a href="http://scta.info/property/hasManifestation">http://scta.info/property/hasManifestation</a></td>
<td><a href="http://scta.info/resource/lectio1/vat">http://scta.info/resource/lectio1/vat</a></td>
</tr>
</tbody>
</table>

**Uncategorized** (miscInfo datastream | metadata count: 55)

<table>
<thead>
<tr>
<th><a href="http://scta.info/property/editedBy">http://scta.info/property/editedBy</a></th>
<th>Jeffrey C. Witt</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://scta.info/property/expressionType">http://scta.info/property/expressionType</a></td>
<td><a href="http://scta.info/resource/lectio">http://scta.info/resource/lectio</a></td>
</tr>
<tr>
<td><a href="http://scta.info/property/gitRepository">http://scta.info/property/gitRepository</a></td>
<td><a href="https://bitbucket.org/jeffreycwitt/lectio1">https://bitbucket.org/jeffreycwitt/lectio1</a></td>
</tr>
</tbody>
</table>
Utrum trinitas vere praedicitur de essentia
Responsio sic propter omni modam ydemptitatem
identitatem essentiae ad totam trinitatem Idem

Utrum generet eum natura
Responsio sic quia est natura filius ut dicit AUGUSTINUS in littera quia generare est actus naturae

Utrum generet eum
secundum rationem
exemplaris vel ydeae ideae
Responsio aliquid procedere
Frequency of Use of Biblical Quotations

A Graph that shows the frequency of use of biblical quotations in medieval Sentences commentaries

Back to Graph List
Digital Humanities and Critical Edition: The Text Encoding Initiative
Bibliographic search

Digital libraries of manuscripts and rare books

Identifying texts in manuscripts

Computer assisted collation

Computer-assisted stemmatics

Identifying sources in a text

Structuring the edited text

Analysing the edited text
WHY?

- To share a common modelling
- To privilege sense over sensibility aspect
- To be free from proprietary formats
The Text Encoding Initiative (TEI)

- Rules for modelling (the Guidelines)
- A way to implement those rules (XML schema)
- A community of users (the TEI Consortium)
XML?

- eXtensible Markup Language
- Uses „tags“, which may have „attributes“
- TEI XML = full-text + database
ReferentIn: Thema:

Archemand femur cromes de juba et. N. omnesse in alii. omnium pependit pagum unam, salvi in omen Salvator. Quem laburo

est hominem. uerum audemtur sancti et ut etque sunt securi res, non rebus. Hominum in pecore, frons autem

in suis in linguam bona quae mandata. melita seu maurena de habita. seu enim reddidit qui de pecore linguas possum, quae

mille eccle. gessit. ad honorem de omperna et gesset gremios seue, beatus maurez maurey socer et in moderno eccle guesencen

vestri omni fluidi socios socios, quos municipis dominii. Et. hacta perussione abbas sigur de ianuas ibidem finem erit

et vuas. pecore, gessid. labcrum rabsolutum suum possidens, ac iniquus, remussiones nunc, quae erit omni annosse cum

biceps pecore. gessid, absolutum suum possidens, ac iniquus, remussiones nunc, quae erit omni annosse cum

inhiber nunc gessid, suum possessio, quos municipis dominii. ac iniquus, remussiones, nunc, ac iniquus, remussiones, nunc.
<p>
  <seg type="suscription"><lb/>Arthemandus senior comes de Quibor et
  <choice>
    <abbr>M</abbr>
    <ex><ex>argarita</ex></ex>
  </choice>
  <choice>
    commissa uxor illius</choice>
  <seg type="salut">omnibus presentem paginam intuentibus
  salutem in omnium salvatore</seg>. <seg type="préambule">Quoniam labilis</seg> est
  hominum memoria veterum auctoritate sanctum est, ut ea que fiunt scripture testimonio
  roborentur.</p>
Representing the Text: XML

- Describes the data in a document with tags, optionally qualified with attributes
- Tree structure
<?xml version="1.0" encoding="UTF-8"?>
<text>
  <p n="1">I am reading a book by <persName>Jack London</persName></p>
  <p n="2">I live in <placeName>London</placeName></p>
</text>
Controlling the Text: TEI Schema

- provide rules for structuring the document beyond the rules of XML
- Implements the TEI Guidelines
Displaying the Text: CSS and XSLT

- CSS: simple, comes from the HTML world
- XSL: much more developed and powerful

Usage:
- Transforming XML into HTML, for display
- Converting from one XML format to another
Querying the Text: XQuery and XPath

- Query languages, defined by W3C recommendations
- XQuery: more complex and powerful, but also more difficult to use
- XPath: selects parts of XML documents; much used within XSLT
Exercise
EXAMPLE: A POEM

Helen Mort, „North of Everywhere“, in Blackbox Manifold, issue 7, Oct. 2011

http://www.manifold.group.shef.ac.uk/issue7/HelenMort7.html
Encoding the metadata: the header

- `<fileDesc>` contains all the basic information:
  - `<titleStmt>` grouping together `<title>` and `<author>`
  - `<publicationStmt>`: which licence for this edition, etc.
  - `<sourceDesc>`: full-text description of the source from which this edition derives
Encoding the structure:

- `<lg>`: "(line group) contains one or more verse lines functioning as a formal unit, e.g. a stanza, refrain, verse paragraph, etc."
  
- `<head>`: "(heading) contains any type of heading, for example the title of a section, or the heading of a list, glossary, manuscript description, etc."
  
- `<l>`: "(verse line) contains a single, possibly incomplete, line of verse."
Encoding the data (1a): mentions place names

- <placeName>: „contains an absolute or relative place name.”

Example:

... past <placeName>Sheffield</placeName>’s border lands
Encoding the data (1b.1): identify the places

- In the header, add section <profileDesc> (<text-profile description> provides a detailed description of non-bibliographic aspects of a text, specifically the languages and sublanguages used, the situation in which it was produced, the participants and their setting “)

- <settingDesc>: „(setting description) describes the setting or settings within which a language interaction takes place, or other places otherwise referred to in a text, edition, or metadata. “

- <listPlaces>: „(list of places) contains a list of places, optionally followed by a list of relationships (other than containment) defined amongst them.“

- <place>: „contains data about a geographic location”
  - <placeName>
  - <location>
    - <geo>
  - <note>
Example:

<place xml:id="she">
    <placeName>Sheffield</placeName>
    <location>
        <geo>53.3787 -1.4753333</geo>
    </location>
    <note>Sheffield is a city and metropolitan borough in South Yorkshire, England.</note>
</place>
Encoding the data (1b.2) : link place names to places

Add @ref attribute to <placeName>, pointing to the @xml:id attribute of a <place>

*Example:*

... past <placeName ref="#she">Sheffield</placeName>’s border lands
Let’s visualise our encoding:

http://ciham-digital.huma-num.fr/datacite/
WHAT HAPPENED?

- XSLT script => transformation of the XML into a web page using *Leaflet*, a javascript script
- Leaflet: uses [OpenStreetMap](https://www.openstreetmap.org) data layers to add the places you have encoded on a map
Encoding the data (2) : mentions of animals

<term>: „contains a single-word, multi-word, or symbolic designation which is regarded as a technical term.”

@ref on <term>: “(reference) provides an explicit means of locating a full definition or identity for the entity being named by means of one or more URIs.”

Example:

... a resting <term ref="https://en.wikipedia.org/wiki/Gull">gull</term>
Let’s visualise our encoding (again):

http://ciham-digital.huma-num.fr/datacite/
Thanks
Digital Humanities and Critical Edition: Opening up TEI data with LOD: an introduction
Digital Humanities and Critical Edition: Opening up TEI data with LOD: an introduction
Outline

- Open Science?
- Principles of Semantic web and Linked Open Data
- TEI to LOD
- SPARQL demo
- Exercice
From OA to OS

- **Open Access**: free and persistent access to research data and publications

- **Open Data**: files made publicly available by official organisms for re-use

- **Open Process**: right to openly observe the underlying data and workflows of research projects

- **Open Science**: free and persistent access to research data with the right to observe openly these data with digital tools.

- **Open Science = Open Access + Open Process**
The information pyramid

- **Data**: Numbers, facts, measurements…
- **Information**: Interpreted data
- **Knowledge**: Broader interpretation
- **Wisdom**: Strategic and long term decisions
The semantic web

• A parallel web, that differs from the original web by the kind of knowledge presented and accessed.

• The knowledge found on the SW is *formal* knowledge with:
  ‣ a machine readable notation ;
  ‣ a formal syntax ;
  ‣ a formal semantics with inference mechanisms.

• The SW started as a vision by Tim Berners-Lee. The vision became true via *Linked Data*.
LOD

Open Data + Linked Data

= 

Linked Open Data
Linked Data

- Design principles for sharing machine-readable interlinked data on the Web:
  - Name resources with unique identifiers (URIs);
  - Use the architecture of the web to get some information about these resources (Http);
  - Use a standard model to give info about these resources (RDF).
In practice

http://has-winter-school/missing-link/exercice/poem

Poem « North of Everywhere »

Metadata
Content-type: application/xhtml+xml

Data
<!DOCTYPE html SYSTEM "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <title>North of Everywhere</title>
  </head>
  <body> (...)
The semantic web
« standards stack »

- User interface and applications
- Trust
- Proof
- Unifying Logic
  - Querying: SPARQL
  - Ontologies: OWL
  - Rules: RIF/SWRL
  - Taxonomies: RDFS
  - Data interchange: RDF
  - Syntax: XML
- Identifiers: URI
- Character Set: UNICODE
- Cryptography
The Linking Open Data Cloud
• RDF = « Resource Description Framework »

• RDF is a way of expressing some information about identified things (resources) with very simple sentences.
RDF sentences = « triples »

- A triple is the minimal building block of RDF. It is composed of:
  - 1. a **subject**, identifying the resource the statement is about;
  - 2. a **predicate**, that identifies a property of the subject;
  - 3. an **object**, identifying the resource linked to the subject by the property.
Helen Mort is the author of « North of everywhere ».
Helen Mort is the author of « North of everywhere ».
was born in « Sheffield »

The object of the triple can also be a « literal »
a graph model

- Characteristics:
  - relations are part of the data;
  - each triple is autonomous, complete, persistant;
  - a distributed model.
Helen Mort is the author of the text represented in the TEI document. The title of the text is the string « North of Everywhere ».
Last night, my body was a compass needle

« Hermaness » is the English name of a place.
This place is identified by the URI « http://dbpedia.org/page/Hermaness ».
The magic

Some implicit simple sentences in the TEI structure that might interest a given audience or be useful for some automatic treatment are selected and «repackaged» in RDF with a transformation language…

Database
TEI files

XSLT / XQuery

RDF graph

Publish

Reuse
Representation requires 2 steps

• Step 1: Conceptualisation
  ‣ a point of view on the reality represented oriented by our goals
  ‣ here we have 2 individuals: a woman and a book
resource 3 - relation: is a woman

resource 4 - relation: reads
Representation requires 2 steps

• Step 2: Language selection
  ‣ the terms (URI and literals)
  ‣ a syntax

*ex:* abbreviates our namespace « http://www.example.org»
resource 3 - relation: is a woman

resource 4 - relation: reads
• Pairs of resources are connected by the binary relation they belong in:
  ‣ `ex:helen ex:reads ex:book`

• Unitary relations are connected to a class:
  ‣ `ex:helen rdf:type ex:woman`
a set of RDF triples is a graph

ex:helen ex:reads ex:book
ex:helen rdf:type ex:woman
Literals to associate a natural language fragment to a resource

```
ex:book

ex:reads

ex:helen

ex:woman

foaf:name
```

« helen »
Linked Open Vocabularies

• Definition

• Common vocabularies: foaf, dc:terms

• Ontology: one example (LAWD)
all you need is LOV :)

https://lov.okfn.org/dataset/lov/
# In LOV at a glance

## About LOV
- Vocabularies describe and link Data on the Web
- Vocabularies are also data
- Quality vocabularies are in LOV
- LOV provides sustainable resources
- Contribute, join the community

## LOV Features
- Vocabulary Documentation
- Data Access
- Vocabulary Search Engine

## Application Ecosystem
- Applications using LOV
- Useful Related Applications

## Acknowledgements

## Historical Creators

- Pierre-Yves Yandenibussche
- Bernard Vatant
Helen Mort is a poet born in Sheffield. She won the Gregory Award from The Society of Authors in 2007 and the Betty Travers Society of Authors’ Poetry Prize in 2010, the youngest ever poet to win both awards. In 2013, she became the youngest ever Poet in Residence at the Wordsworth Trust, winning the Wordsworth Trust Poetry Award. In 2019, she was shortlisted for the Picador Prize and for the Rathbones Folio Prize.
LAWD

• Linking Ancient World Data

• LAWD is designed to be a *minimal* ontology for connecting vocabularies useful in describing data concerning the ancient world.
An ontology for Linked Ancient World Data

LAWD

An ontology for Linked Ancient World Data

The goal of LAWD is to fill in the cracks between the data used and published by projects with a focus on the Ancient World and the data that is available through them. These breakdowns range from geographic, in the form of a lack of digital data, to structural, in the form of datasets that don’t exist. LAWD aims to bridge these gaps and provide a complete view of the ancient world.
namespace: http://lawd.info/ontology/

prefixes:

- dc: http://purl.org/dc/elements/1.1 (Dublin Core Elements)
- dct: http://purl.org/dc/terms/ (Dublin Core Terms)
- geo: http://geovocab.org/spatial# (GeoVocab.org)
- prov: http://www.w3.org/ns/prov# (PROV-O, the Provenance ontology)
<owl:ObjectProperty rdf:about="http://lawd.info/ontology/hasCitation">
   <rdfs:comment xml:lang="en">Indicates a Citation that contains the source for a lawd:Attestation</rdfs:comment>
   <rdfs:range rdf:resource="http://lawd.info/ontology/Citation"/>
   <rdfs:domain rdf:resource="http://lawd.info/ontology/EvidentiaryItem"/>
</owl:ObjectProperty>
# SRP identifies a CanonicalWork with "work/book-of-the-dove"
# of which SRP's person/239 (Bar Hebraeus) is the author
# and of which the English title is "Book of the Dove"
<srp:work/book-of-the-dove> a lawd:ConceptualWork;
   dc:creator <srp:person/239/>;

# SRP identifies a Manuscript with "manuscript/paris-bnf-syr-313"
# which supports (i.e., carries, physically contains) something of interest at a
# particular range (1-56)
<srp:manuscript/paris-bnf-syr-313> a lawd:WrittenWork ;
   dct:hasPart <srp:manuscript/paris-bnf-syr-313/1-56> .

# SRP says the thing of interest at that range is a Text
# which embodies Bar Hebraeus' work book of the dove, as defined above
<srp:manuscript/paris-bnf-syr-313/1-56> a lawd:WrittenWork ;

# OCLC/Worldcat identifies something that contains an edition of that Text
<oclc:7485581/519-599> a lawd:Edition ;
   dct:source <srp:manuscript/paris-bnf-syr-313/1-56> .

# OCLC/Worldcat identifies something that is a translation of that Text
<oclc:652251598> a lawd:Translation ;
   dc:source <srp:manuscript/paris-bnf-syr-313/1-56> .

# Do we need to distinguish Modern translation and ancient version? Can we show the language of the translation?
SCTA Search

Search for a resource of the selected type whose title contains the search term(s)

* For example to search a quotation used in a commentary or text that contains the word "fides", select "quotation" and the search for "fides"

- Quotation
- Item
- Question Title
- Commentary
- Name
- Work
- Book
- Distinction

Type search text here

Validator

Or perform custom SPARQL Query

```
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

SELECT ?s ?p ?o
WHERE {
  ?s ?p ?o
}
LIMIT 1000

Validator
```
Recap

• LOD enhances data created in digital humanities projects and scholarly editions. You can publish and consume LOD at the same time.

• You can build your ontology (knowledge base of the relations between resources) just by building on existing ontologies and filling the gaps.

• The syntax and the way you express triple is a technical issue. What counts for you as designers of « information artifacts » is the data model. It is fully an editorial task.

• Understanding how your digital humanities publications can connect with other databases in the LOD world can influence the way you define your choice of tags…

• From the same TEI collection of files, you can extract several different datasets. There might be a task for continuous re-purposing of exposed RDF data…

• Nonetheless, to meet the founders expectations concerning re-use, the data must be well documented, which means documented in with information useful for research uses. That’s where TEI and its rich expression capabilities becomes a key asset.
The Semantic Web Technology Stack
(not a piece of cake...)

Most apps use only a subset of the stack
Querying allows fine-grained data access
Standardized information exchange is key
Formats are necessary, but not too important
The Semantic Web is based on the Web
Linked Data uses a small selection of technologies
http://fr.dbpedia.org/sparql

demo
Thank you

@emma_morlock
Exercice

• Look at your texts with LOD in mind:
  
  ‣ Work in small groups
  ‣ Choose a corpus (your project / postcards from Letters 1916)
  ‣ Identify one dataset that you would like to:
    ✓ publish as LOD
    ✓ and/or
    ✓ re-use as LOD
  ‣ Use the MVP canvas to test the idea from an « information designer » point of view
• The Letters of 1916 project is the first public humanities project in Ireland. It is creating a crowd-sourced digital collection of letters written around the time of the Easter Rising (1 November 1915 – 31 October 1916).

• Proposed selection of postcards:
  ‣  http://letters1916.maynoothuniversity.ie/explore/letters/43
  ‣  http://letters1916.maynoothuniversity.ie/explore/letters/73
  ‣  http://letters1916.maynoothuniversity.ie/explore/letters/84
  ‣  http://letters1916.maynoothuniversity.ie/explore/letters/95
  ‣  http://letters1916.maynoothuniversity.ie/explore/letters/2362
Minimum Viable Product

• @lissijean: « The smallest work you can do to test your product »

• Philosophy:
  ‣ validate before building
  ‣ focus on problems, not features
The Business Model Canvas

- **Key Partners**: Key partners
- **Key Activities**: Key activities
- **Value Propositions**: Value proposition
- **Customer Relationships**: User relationships
- **Customer Segments**: Target audience
- **Key Resources**: Key resources
- **Channels**: Channels
- **Cost Structure**: Costs structure
- **Revenue Streams**: Revenue/retribution
Useful links

• Dbpedia example: http://dbpedia.org/resource/Helen_Mort

• https://lov.okfn.org/dataset/lov/

• http://www.foaf-project.org/

• http://dublincore.org/documents/2012/06/14/dcni-terms/?v=terms
DBPEDIA DEMO
Basic exploration of a RDF graph with simple SPARQL queries
Introduction

Example representation of RDF triples
2 assertions

<Helen> reads a book

<Helen> is a woman
2 RDF triples

PREFIX ex: <http://www.example.org>

ex:helen ex:reads ex:book
ex:helen rdf:type ex:woman
Graph representation

ex:helen

ex:reads

ex:book

foaf:name

« helen »

rdf:type

ex:woman
Another RDF triple
Another RDF triple
ex:book

ex:reads

ex:helen

foaf:name

ex:chapter

ex:hasPart

ex:thing

ex:chapter

owl:sameAs

test:section

ex:woman

rdf:type

rdf:type
Demo

http://dbpedia.org/sparql
SPARQL

Simple Protocol

And

Query Language
Query structure

- SELECT distinct * == select all resources
- WHERE { } == the query
- LIMIT, GROUP BY, ORDER BY…
Variables
as in a Fill-in-the-Blank exercise -1

_________ is a woman
Variables
as in a Fill-in-the-Blank exercise - 2

WHERE { ?s rdf:type ex:woman }
Virtuoso SPARQL Query Editor

Default Data Set Name (Graph IRI)
http://dbpedia.org

Query Text
select distinct ?Concept where {{} a ?Concept} LIMIT 100

Results Format: HTML
Execution timeout: 30000 milliseconds (values less than 1000 are ignored)
Options:
- Strict checking of void variables
- Log debug info at the end of output (has no effect on some queries and output formats)

(The result can only be sent back to browser, not saved on the server, see details)

Run Query  Reset
Simple Queries 1

• Find resource with the english label « Prague »

• Find all the properties of this resource

• Find the types of this resource

• Choose a type

• Find the resources with the type
• select distinct * where {?s rdfs:label "Prague"@en} LIMIT 100

• select distinct * where {<http://dbpedia.org/resource/Hanlin,_Burma> ?p ?o} LIMIT 100

• select distinct * where {<http://dbpedia.org/resource/Prague> rdf:type ?o} LIMIT 100

• select distinct * where {?s rdf:type <http://dbpedia.org/ontology/PopulatedPlace>} LIMIT 1000
• Simple inference in action with Corese, a Semantic Web Factory (triple store & SPARQL endpoint) implementing RDF, RDFS, SPARQL 1.1 Query & Update, developed by INRIA

• http://wimmics.inria.fr/corese
  ‣ tutorial: http://wimmics.inria.fr/node/34
  ‣ Linked Data Navigator using Corese and SPARQL Template TransformationLanguage: https://corese.inria.fr/
Further reading