

Extending the DSE: LOD support and TEI/IIIF integration in EVT

Outline

- Introduction
 - The birth and development of EVT
 - Expanding the DSE: The distributed edition
- LOD (Linked Open Data)
 - LOD support in EVT, current and forthcoming
 - TEI/LOD integration
- IIF (International Image Interoperability Framework)
 - IIF support in EVT, current and forthcoming
 - TEI/IIF integration
- Zenodo

The birth and development of EVT

- EVT ([Edition Visualization Technology](#)) was born as the edition viewer component of the Digital Vercelli Book project
- Developed to be a general purpose tool, many new features implemented thanks to collaboration with other projects
 - **EVT 1:** a diplomatic edition / single-witness critical edition viewer based on XSLT 2 and Web technologies [final version EVT 1.3 released December 2019]
 - **EVT 2:** a “reboot” of the project based on the MVC design pattern and the AngularJS framework [beta 2 due Spring 2020]
 - **EVT 2+:** porting to Angular 8 under way [don't ask]
- Modular design to improve flexibility and power
- UI/X designed (and tested) to maximize ease of use

Expanding the DSE

- A DSE is a **dynamic object**, a research tool which assists the scholar in data interpretation and analysis
- This also means that it can engage in dialogue and interaction with other Internet-based resources by
 - taking advantage of **semantic web** technologies and **Linked Open Data** to enrich the edition content
 - **text/image** linking, pointing to digital collections of images of manuscripts maintained by external repositories
 - modelling **intertextual relationships** through canonical text services
 - becoming itself a **source** of shareable, reusable data

The distributed edition

- This is an EVT goal dating back quite some time
- It is only recently that all the necessary technologies seems to be falling into place
 - LOD repositories
 - IIF framework for images
 - CTS and DTS protocols for text
 - Zenodo as a safe, long term general purpose repository
 - GitHub as a server of live pages
- To use those effectively we need to work both on the general methodology and to experiment with the existing tools
- Great opportunities, but also some risks and complications

AIUCD 2020 - Milano,
15-17 gennaio 2020

< EVT >
Edition Visualization Technology

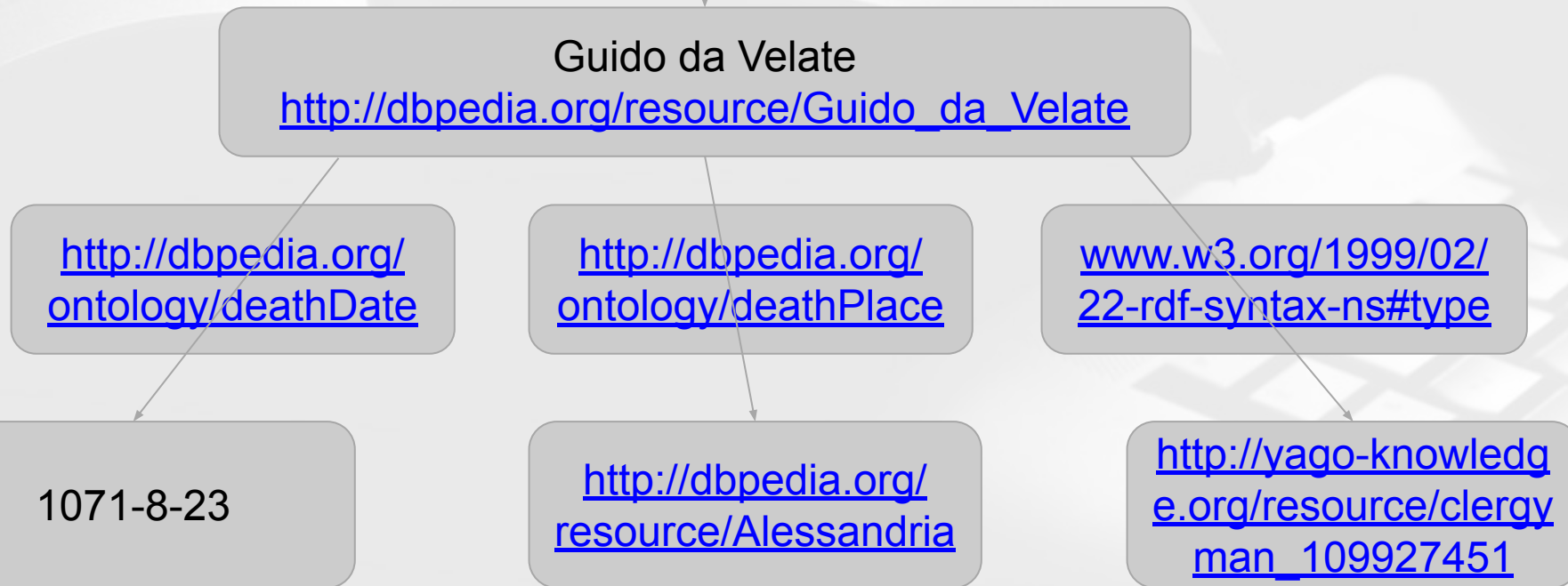


LOD Linked Open Data

Using LOD / Creating LOD

- **Using** (referencing) external resources

```
<persName  
ref="http://dbpedia.org/resource/Guido\_da\_Velate">  
Guidonem</persName>
```



Using LOD / Creating LOD

- **Creating** new LOD resources

```
<persName  
ref="http://dbpedia.org/resource/Guido\_da\_Velate">  
Guidonem</persName>
```

The work I published
<http://myproject.it/myeditionuri>

refers to

www.ancientwisdoms.ac.uk/media/ontology/sawsOntology.owl#refersTo

Guido da Velate

http://dbpedia.org/resource/Guido_da_Velate

Using LOD / Creating LOD

- **Creating** new LOD resources

```
<persName  
ref="http://dbpedia.org/resource/Guido\_da\_Velate">  
Guidonem</persName>
```

“Guidonem”

is an attested name for

<http://xmlns.com/foaf/0.1/name>

Guido da Velate

http://dbpedia.org/resource/Guido_da_Velate

Using LOD: EVT support 1 (via local list)

<teiHeader>

```
<listPerson>
  <person xml:id="Rothari">
    <persName>Rotari</persName>
    <sex>M</sex>
    <birth>606</birth>
    <death>652</death> [...]
  </person>
</listPerson>
```

<body>

```
<persName ref="#Rothari">
  <w>rothari</w></persName>
```

Using LOD: EVT support 1 (via local list)

<teiHeader>

<listPerson>

<person>

<pe

<se

<bi

<de

</perso

</listPerson

<body>

<persName r

<w>rotho

1 IN NOMINE **Dāi** INCIPIT
2 EDICTVM QVEM RE
3 NOVAUIT **DOŃ** ROTHAR
4 RI UIR EXCELLENTISSIMO
5 REX GENTI LANGOBAR
6 DORVM CVM PRIMA
7 TOS IUDICES SVOS

8 **INCIĀ** PROLOGVS

9 Ego in **dī** nomine **rothari**

Rotari

More Info	Occurrences
	MI
	606 652
	636-652
	Rex Langobardorum et Italiae 636-652
	Duca longobardo di Brescia, nato agli inizi del VII secolo, nel 636 venne eletto re dei Longobardi dopo la morte del predecessore Arioaldo. V. anche la voce corrispondente su DBpedia .

vir **exell**, rex

10 **Ē** septimo decimū rex gentis langobar
11 **dorū** annuente **dō** propiante regni **me**

Using LOD: EVT support 1 (via local list)

<teiHeader>

```
<listPerson>
  <person xml:id="Rothari">
    <persName>Rotari</persName>
    <sex>M</sex>
    <birth>606</birth>
    <death>652</death>
    <note>V. anche la voce corrispondente su
      <ref target="http://dbpedia.org/resource/Rothari">
        DBpedia</ref>.
    </note>
  </person>
</listPerson>
```

<body>

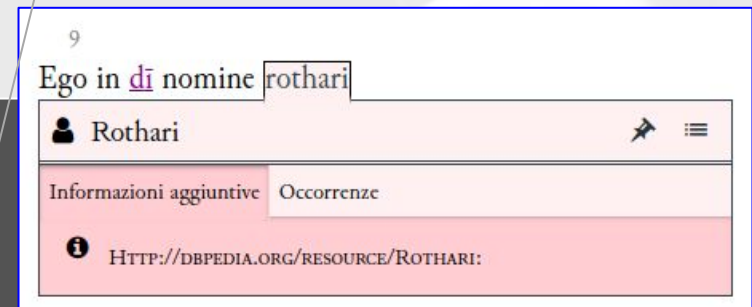
```
<persName ref="#Rothari">
  <w>rothari</w></persName>
```

Using LOD: EVT support 1 (via local list)




<teiHeader>

```
<listPerson>
  <person xml:id="Rothari">
    <persName>Rotari</persName>
    <sex>M</sex>
    <birth>606</birth>
    <death>652</death>
    <note>V. anche la voce corrispondente su
      <ref target="http://dbpedia.org/resource/Rothari">
        DBpedia</ref>.
    </note>
  </person>
</listPerson>
```


Partial EVT support
(showing URL)



9
Ego in di nomine rothari

 Rothari  

Informazioni aggiuntive Occorrenze

 HTTP://DBPEDIA.ORG/RESOURCE/ROTHARI:

<body>




```
<persName ref="#Rothari">
  <w>rothari</w></persName>
```

Using LOD: EVT support 2 (direct link)

EVT support
forthcoming
(currently showing
empty info box)

9

Ego in dī nomine rothari

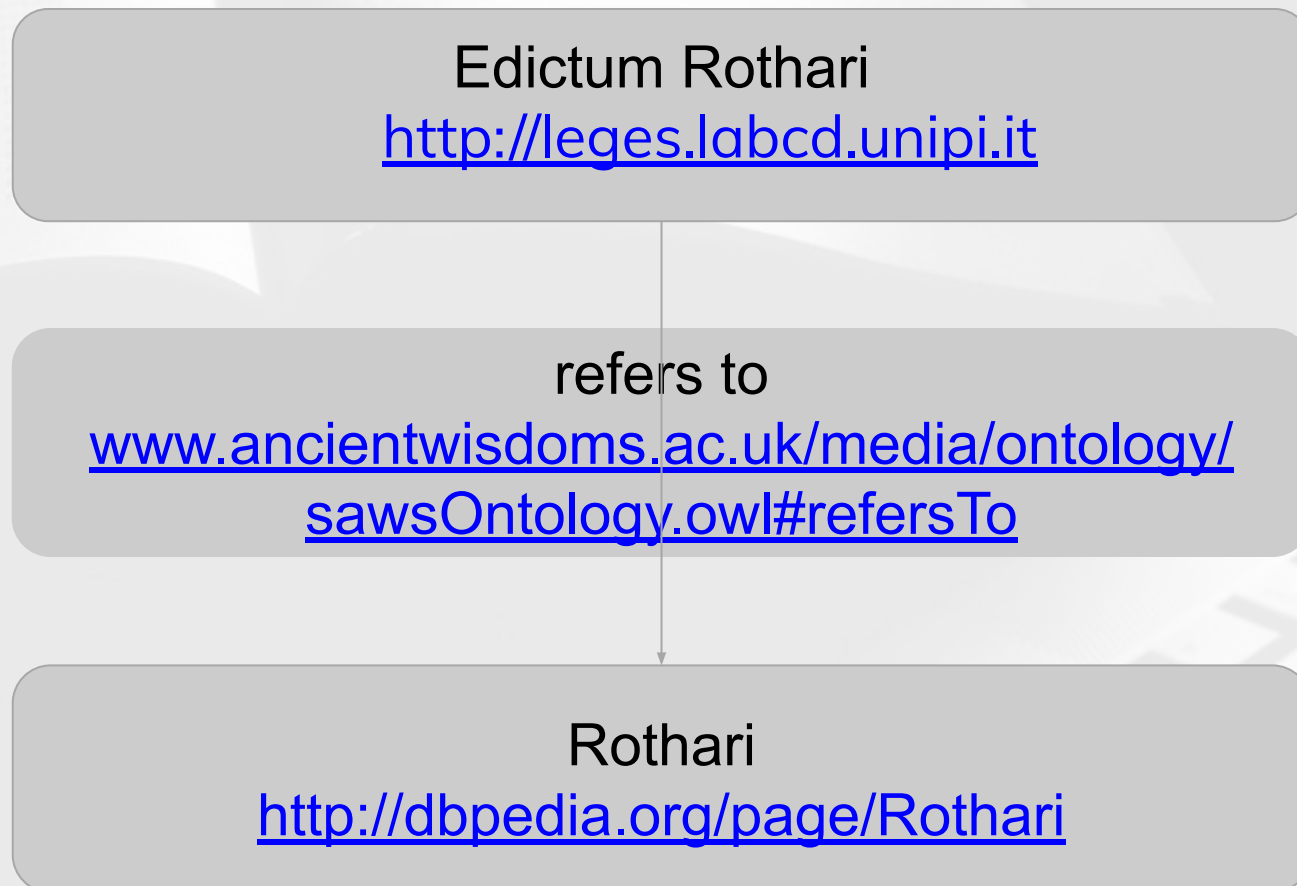
		
Informazioni aggiuntive	Occorrenze	

<body>

<persName ref="<http://dbpedia.org/resource/Rothari>">
<w>rothari</w></persName>

Creating LOD

- **Creating** new LOD resources (triples) → External SW needed



LOD support in EVT

- **Current**
 - Encoder creates **local list** (e.g. <listPerson>)
 - Encoder **links** local element to **LOD**
(<person><ref="<http://dbpedia.org/resource/Rothari>">)
 - EVT visualizes **info box** from <person>
 - EVT shows **clickable link** to LOD resource
- Project's **own software can extract** RDF triples from XML
- **Forthcoming**
 - Support **direct link**: text → LOD resource (no local list)
 - **Import info from LOD** resource and visualize in info box
 - **Extract** RDF triples from XML (?)

TEI/LOD Integration

- **Current TEI/LOD integration**
 - Linking from TEI to LOD resources
 - **Via local data structure** (<listPerson>, <listPlace>, etc.): minimal overhead
 - Or **directly from the text** (e.g. <persName>): closer to LOD philosophy, but burdensome if the same link occurs multiple times
- **Issues in TEI/LOD integration**
 - TEI is flexible, but it currently is document-oriented more than LOD-oriented
 - Re-orienting the TEI **Guidelines and Schema** ([Chiarcos-Ionov 2019](#))
 - E.g.: allowing for (LOD) **external URIs** in some attributes
 - Tools needed:
 - **Editing**: Oxygen and other authoring tools could facilitate LOD linking
 - **Visualization**: EVT could retrieve and show information retrieved directly from DBpedia displaying it (in part or fully) in a dedicated container
 - Tools to **extract** and expose semantic information (LOD triples) from TEI

AIUCD 2020 - Milano,
15-17 gennaio 2020

< EVT >
EditionVisualizationTechnology



IIF

International Image Interoperability Framework

IIIF: what and why

- IIIF: International Image Interoperability Framework
 - Exchange and integrate image-based resources in Web-based systems
- Use cases:
 - IIIF image server ([e-codices](#), [Biblioteca Vaticana](#), [Biblioteca Ambrosiana](#))
 - IIIF collection (MS images: [Sang. 730](#), Edictum Rothari)
 - IIIF canvas (each image / MS page: [folio 20](#))
 - TEI linking to the URI of the external IIIF image (e.g. `<pb facs="IIIF_URI_here">`)

IIF metadata

- Whole collection / MS ([Sang. 730](#))
 - <https://www.e-codices.unifr.ch/metadata/iiif/csg-0730/manifest.json>
 - [Presentation API](#)
- Individual image / page ([folio 20](#))
 - https://www.e-codices.unifr.ch/metadata/iiif/csg-0730/canvas/csg-0730_020.json
 - [Image API](#)
- JSON metadata files reference actual image files:

[manifest.json](#)

```
▼ service:  
  @context: "http://iiif.io/api/image/2/context.json"  
  ▼ @id: "https://www.e-codices.unifr.ch/loris/csg/csg-0730a/csg-0730a_020.jp2"  
  profile: "http://iiif.io/api/image/2/level2.json"
```

IIF Image API by examples

- https://www.e-codices.unifr.ch/loris/csg/csg-0730/csg-0730_020.jp2/full/full/0/default/jpg

whole image
(no **region**
cropping)

full **size**
(no resizing)

0°
rotation

default
quality
(color,
gray...)

jpg
format
(tif,
png...)

IIF Image API by examples

- https://www.e-codices.unifr.ch/loris/csg/csg-0730/csg-0730_020.jp2/full/full/0/default/jpg

whole image
(no region
cropping)



IIF Image API by examples

- https://www.e-codices.unifr.ch/loris/csg/csg-0730/csg-0730_020.jp2/full/full/0/default/jpg
- https://www.e-codices.unifr.ch/loris/csg/csg-0730/csg-0730_020.jp2/1400,880,2850,600/full/0/default/jpg



region
cropping

Two directions for TEI/IIIF integration

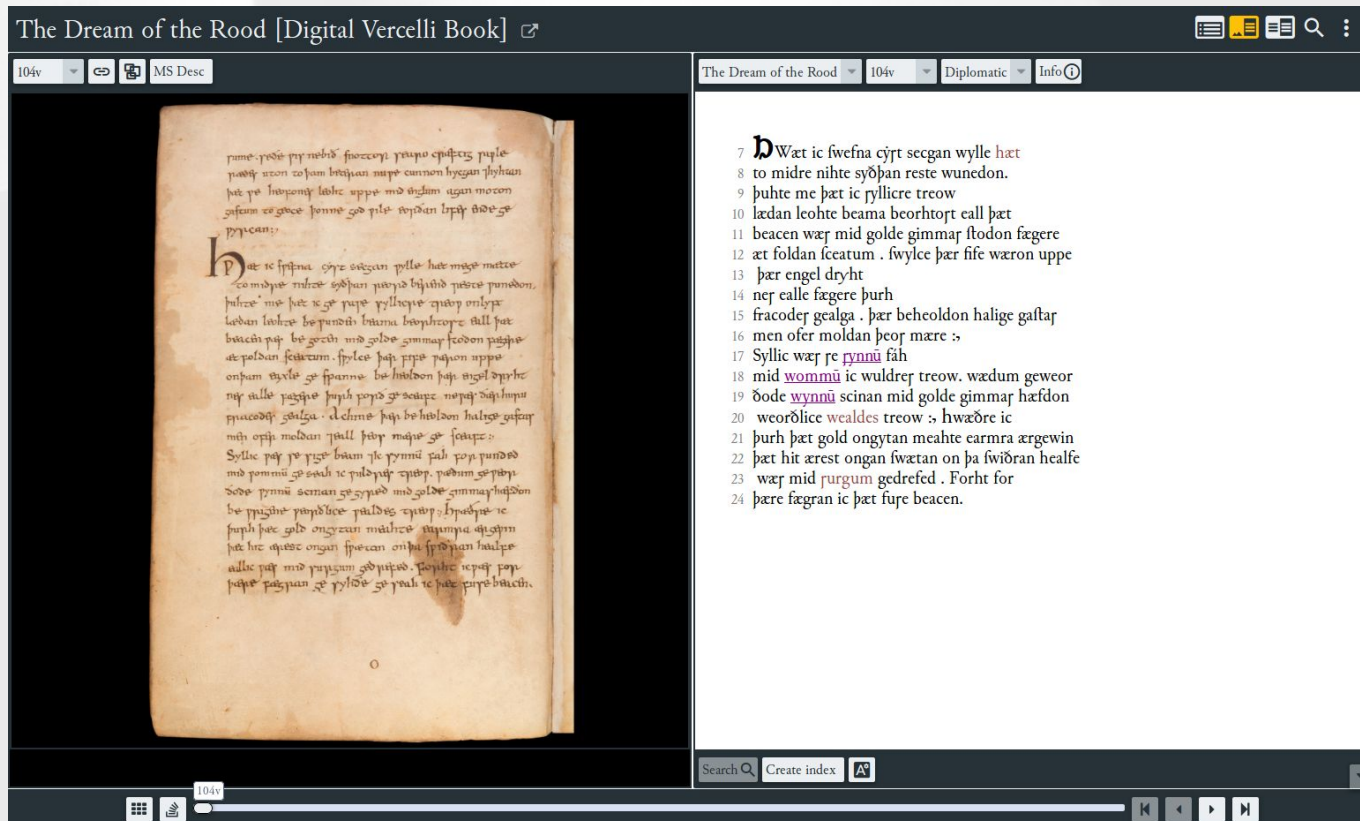
1. **IIIF → TEI**: adding a transcription to an IIIF-based digital facsimile
[IIIF manifest.json](#) → [TEI XML transcription file](#) ([details](#))
2. **TEI XML → IIIF URI**: embedding IIIF URIs in a TEI-based diplomatic transcription

```
<pb  
facs="https://www.e-codices.unifr.ch/loris/csg/csg-0730/csg-0730\_020.jp2/full/full/0/default/jpg"/>
```


IIIF support in EVT



Page → Local image


```
<pb facs="data/images/single/VB_fol_104v.jpg"/>
```





The screenshot displays the Digital Vercelli Book interface. The left pane shows a high-resolution image of a manuscript page (104v) with Old English text. The right pane shows the corresponding diplomatic transcription of the text, with line numbers 7 through 24. The transcription includes a large initial 'H' and various words in Old English, some of which are highlighted in red (rubrication). The interface includes a search bar, a 'Create index' button, and navigation controls at the bottom.

The Dream of the Rood [Digital Vercelli Book] [↗](#)

104v   MS Desc

The Dream of the Rood 104v Diplomatic Info 

7 **H**Wæt ic swefna cȳrt secgan wylle hæc
8 to midre nihte syððan reste wunodon.
9 þuhte me þæt ic ƿyllice treow
10 lædan leothe beama beorhtost eall þæt
11 beacen wæs mid golde gimmar stodon fægere
12 at foldan fceatum . swylce þær life wæron uppe
13 þær engel dryht
14 neƿ ealle fægere þurh
15 fracodeƿ gealga . þær beheoldon halige gastaƿ
16 men ofer moldan þeow mare :
17 Syllic wæs se **ƿynnū** fāh
18 mid **wommū** ic wuldref treow. wædum geweor
19 ðode **wynnū** scinan mid golde gimmar hæfdon
20 weorðlice wealdes treow : hwæðre ic
21 þurh þæt gold ongytan meakte earmra ærgewin
22 þæt hit arest ongan swætan on þa fwiðran healf
23 wæs mid furgum gedrefed . Forht for
24 þære fægran ic þæt fuƿe beacen.

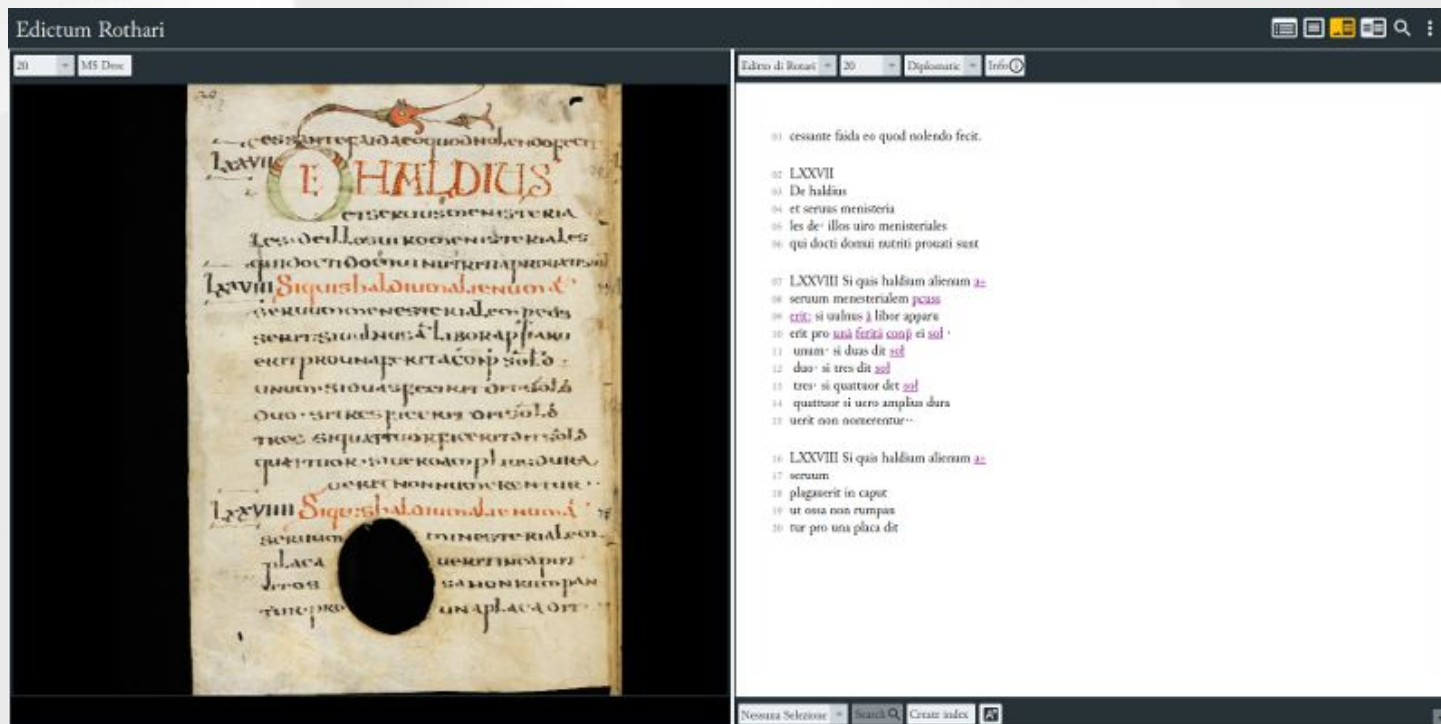
Search  Create index 

IIIF support in EVT

Page → IIIF full image



```
<pb facs="https://www.e-codices.unifr.ch/loris/csg/csg-0730/  
csg-0730_020.jp2/full/full/0/default/jpg"/>
```

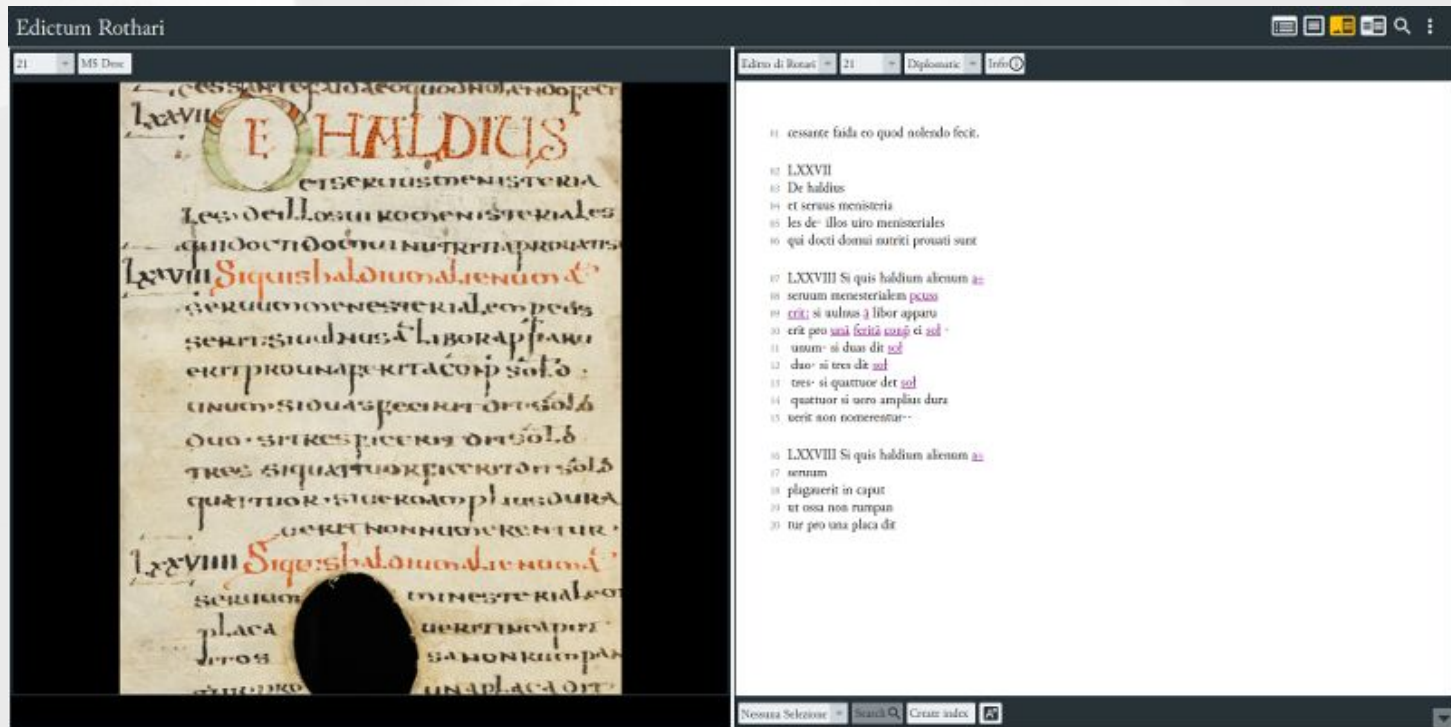


IIIF support in EVT

Page → IIIF image region



```
<pb facs="https://www.e-codices.unifr.ch/loris/csg/csg-0730/  
csg-0730_020.jp2/800,600,3800,5000/full/0/default/jpg"/>
```



IIIF support in EVT

Region (<lb>, <p>, <div>...) → <zone> (region) → IIIF full image

```
<surface>  
  <zone lrx="1200" lry="423" ulx="360" uly="180"  
    xml:id="SG730_hs_20_01"/>  
</surface>
```

```
<pb facs="https://www.e-codices.unifr.ch/loris/csg/csg-0730/  
csg-0730\_020.jp2/full/full/0/default/jpg"/>
```

```
<div facs="#SG730_hs_20_01">
```

IIIF support in EVT

Region (<lb>, <p>, <div>...) → <zone> (region) → IIIF full image

```
<surface>  
  <zone lrx="1200" lry="423" ulx="360" uly="180"  
    xml:id="SG730_hs_20_01"/>  
</surface>
```

Issue: Tool needed to draw rectangles and get coordinates

- Local image: Ch. Sparks, TEI zoner (see [1](#), [2](#), [3](#))
- Local and IIIF image: Oxygen [Image Maps in TEI](#)
- IIIF image (under development): T. Mancinelli, [Barberino generator](#)

IIIF/EVT future development

Region → IIIF image region



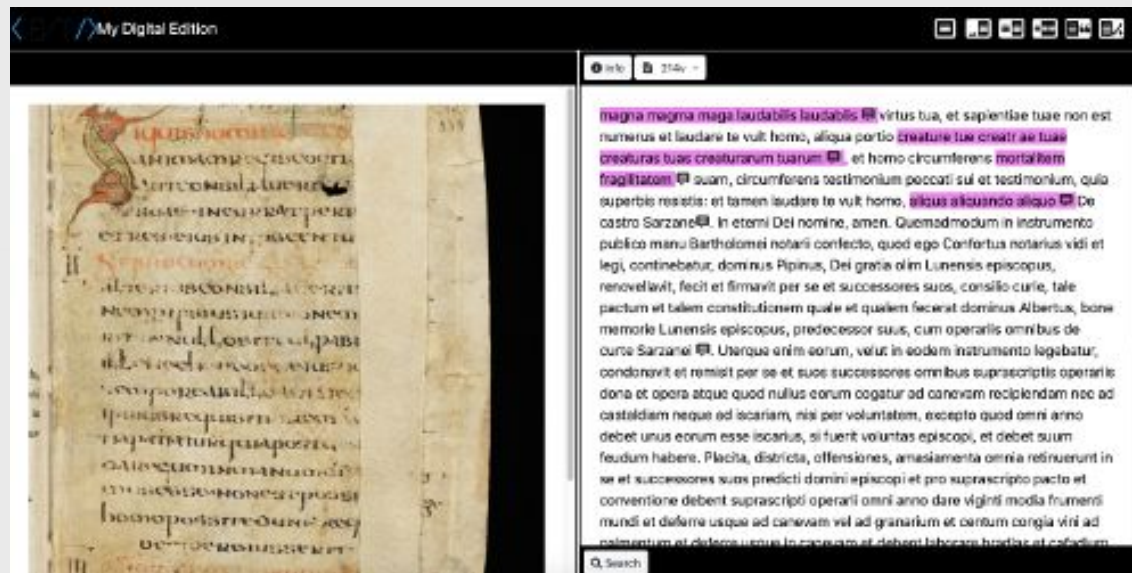
```
<lb facs="https://www.e-codices.unifr.ch/loris/csg/csg-0730/csg-0730\_020.jp2/1400,880,2850,600/full/0/default/jpg">
```

- EVT can already import image regions through *IIIF Image API* URIs like this
- However, visualization strategies must be implemented to show them (hotspot, pop-up window within text, text-embedded image, etc.)

IIIF/EVT future development

- EVT2+ (Angular 8 version, in development)
- Presentation API support
- EVT `config.json` configuration file

```
{ "title": "My Digital Edition",  
  "manifestURL": "https://www.e-codices.unifr.ch/metadata/iiif/csg-0730/manifest.json" }
```



TEI/IIIF integration

- Tool needed to define area coordinates in IIIF-served images
 - A variant of [TEI Zoner](#) (T. Mancinelli's [Barberino generator](#))
 - Oxygen in Author mode ([Image Map Editor](#))
- Concurrent mechanisms to define image portions
 - TEI `<zone lrx="1200" lry="423" ulx="360" uly="180"/>`
 - IIIF Image API URI https://www.e-codices.unifr.ch/loris/csg/csg-0730/csg-0730_020.jp2/1400,880,2850,600/full/0/default/jpg
- IIIF calculates coords differently than TEI
 - TEI: both rectangles (coords for upper left and lower right, ulx, uly, lrx, lry attributes) and polygons (points attribute)
 - IIIF: only rectangles (coords for upper left vertex, then rectangle dimensions)
- Tool needed to convert from IIIF format into TEI

The distributed edition

The distributed edition

- A concept dating back to the end of the '90s
 - Ore, Espen S. “Monkey Business—or What Is an Edition?” *Literary and Linguistic Computing* 19, no. 1 (2004): 35–44.
- A view of DSEs not only taking advantage of LOD, but becoming resources on their own
- To do this several requisites must be met
 - suitable repositories
 - long term sustainability
 - data sharing methodology
 - clear editorial responsibility
- We now have suitable frameworks ([Zenodo](#), [OpenAIRE](#), [GitHub](#)) and sensible dissemination criteria ([FAIR Principles](#))

“Good things come in small packages”

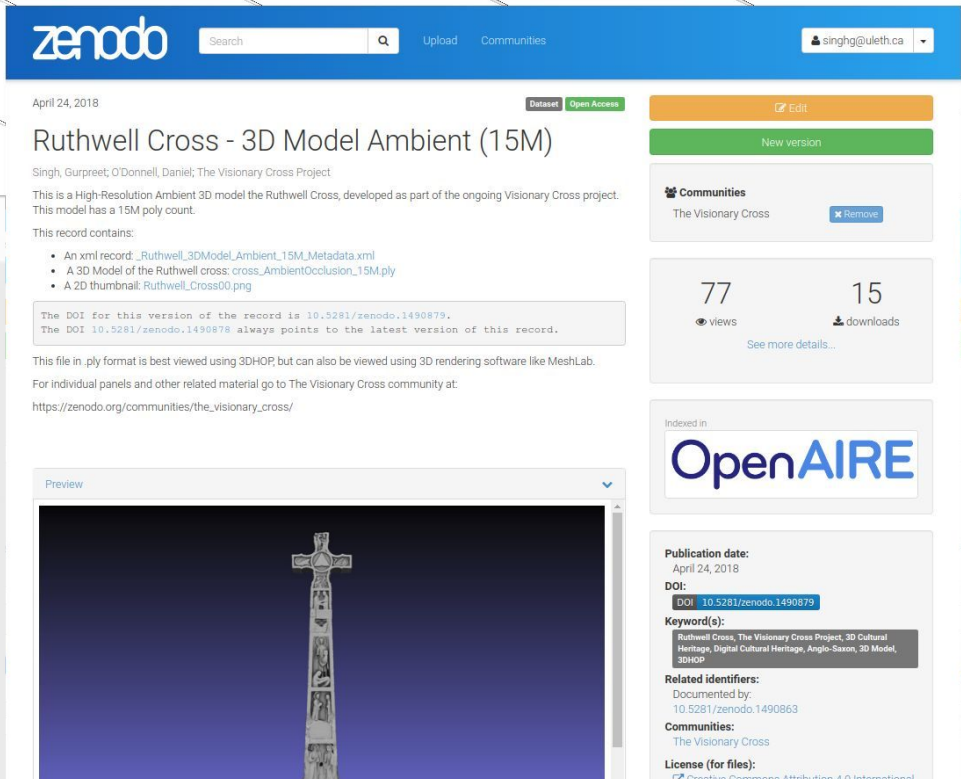
- A spin-off of the Visionary Cross project and a consequence of its “data-first” approach
- General plan: making all the edition data available for other researchers to use as they see fit (no intermediation)
- Not a LOD approach, strictly speaking, but one that fits well with the open, distributed edition concept
- It will require work and experimentation (also to make EVT suitable for the task)

Project approach

- Flexible
 - direct access to data (also thanks to descriptive metadata)
 - access single / group of objects
- Extensible
 - add, rearrange, or reuse material without negotiation
- Authoritative
 - preserve credit/responsibility for all contributions
- Durable
 - permanently discoverable and available
 - low/no maintenance

Zenodo record

- Metadata record + file(s)
- RDF (internal and external relations)
- Two kinds of DOIs:
 - “Conceptual” (latest)
 - “Version” (current)
- RESTful files URLs
 - No link rot



The screenshot shows a Zenodo record page for a 3D model of the Ruthwell Cross. The page includes a search bar, user profile, and navigation links. The record title is "Ruthwell Cross - 3D Model Ambient (15M)" and it was created on April 24, 2018. The record contains three files: an XML metadata file, a 3D model file, and a 2D thumbnail. The page also displays the number of views (77) and downloads (15), and is indexed in OpenAIRE. The publication date is April 24, 2018, and the DOI is 10.5281/zenodo.1490879. The record is part of the "The Visionary Cross" community and is licensed under Creative Commons Attribution 4.0 International.

Conclusion

Conclusion

- Our approach
 - having a look at standards, frameworks, protocols currently available (LOD, IIF) and see how they are supposed to work
 - direct experimentation with available tools (TEI, EVT)
- The good
 - the networked / distributed DSE is nearly there
 - current frameworks are sound and work well
- The bad
 - got to be careful in choosing the best methodological approach
- The ugly
 - when TEI and IIF/LOD speak different languages a “translator” is needed (i.e. new tools must be implemented)

AIUCD 2020 - Milano,
15-17 gennaio 2020

< EVT >
EditionVisualizationTechnology



Thank you for your attention!