

# LINKING TEXT AND IMAGE: TEI XML AND IIIF

Paolo Monella



*VeDPH Summer Camp 2020*

Thursday, 9 July 2020

[ve]dph

Venice Centre for  
Digital and Public  
Humanities



# WHAT IS IIIF?

IIIF is a protocol  
(a language shared between computers)  
to serve images with medatata.



# WHAT DOES IIIF STAND FOR?

International Image Interoperability Framework

# WHY IIIF?

Images stay on their server  
(e.g. the Ambrosiana Library)

and may be integrated  
(e.g. in "distributed" editions)

# CANVAS ← IMAGE(S)

E.g.: a MS page



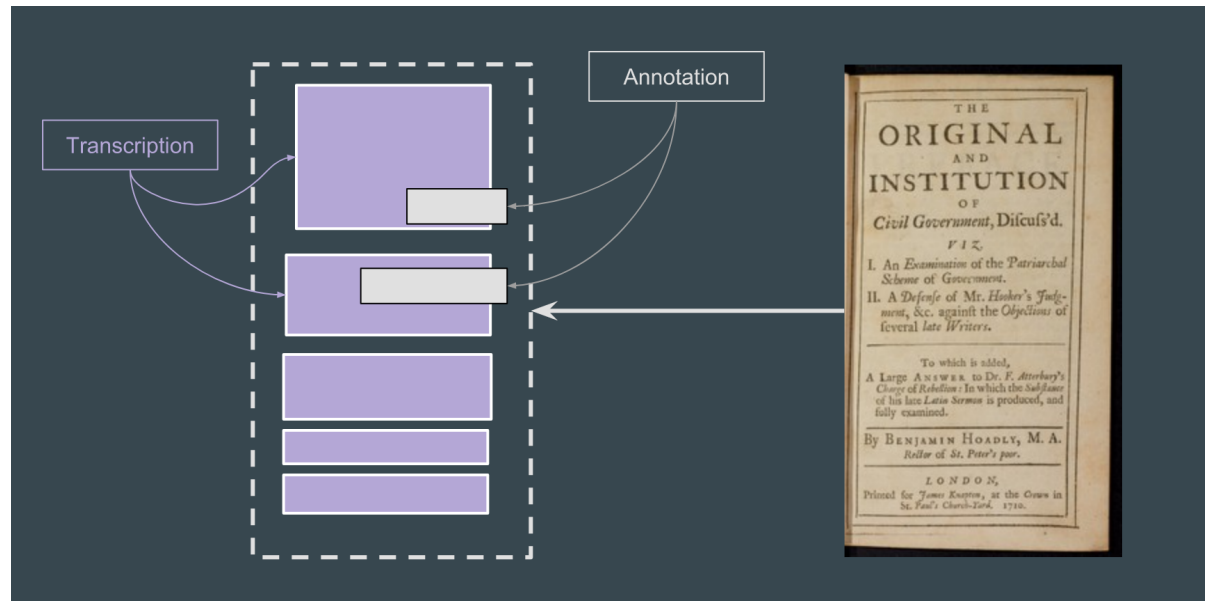
Source: Jason Ronallo, [Workshop: Shared Canvas](#)

# CANVAS ← IMAGE(S)

- One canvas may include many images, e.g.:
  - new photos of same MS
  - new photos taken after MS restoration
  - photos with radiant light, IR or UV

# CANVAS ← IMAGE(S)

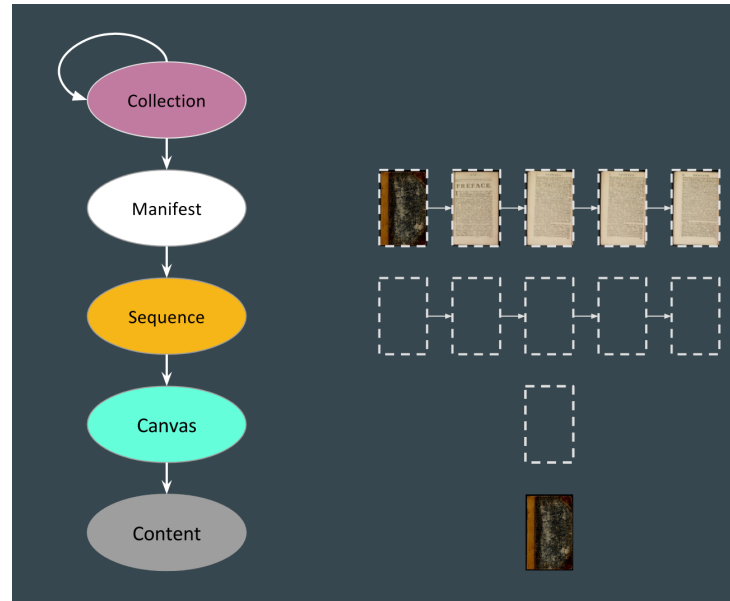
E.g.: a MS page



Source: Jason Ronallo, [Workshop: Shared Canvas](#)

# COLLECTION

E.g.: a whole MS



Source: Jason Ronallo, [Workshop: Shared Canvas](#)



# IIIF FOR PHILOLOGISTS

MS page

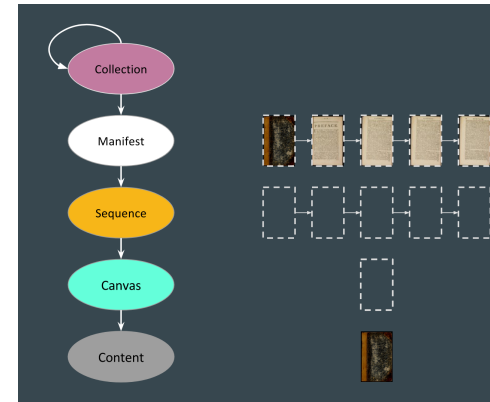
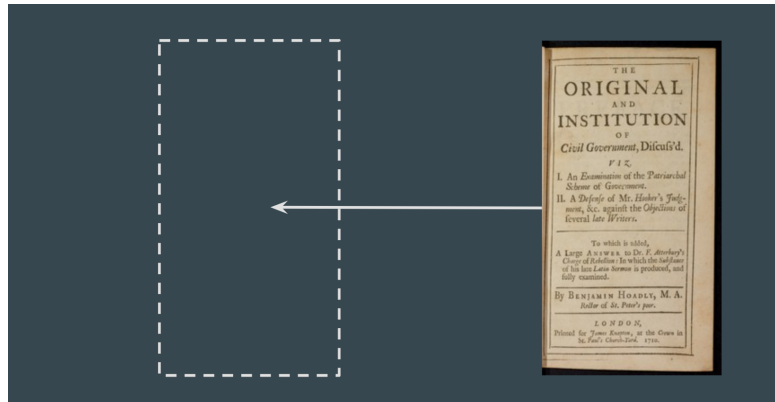
Manuscript

Canvas

Collection

Image API

Presentation API



# JSON METADATA FILES

# JSON METADATA FILES

## MANIFEST

- *Collection* of MS photos
- <https://iiif.bodleian.ox.ac.uk/iiif/manifest/ae9f6cca-ae5c-4149-8fe4-95e6eca1f73c.json>
  - `sequences/0/canvases/26/@id` ↓

# JSON METADATA FILES

## CANVAS JSON FILE

- *Canvas*: a MS page
- <https://iiif.bodleian.ox.ac.uk/iiif/canvas/699adfb3-a5da-4b09-9759-41eee3713bc3.json>
  - `images/0/resource/@id` ↓:

# JSON METADATA FILES

## IMAGE JSON FILE

- *Canvas*: a MS page
- <https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/info.json>

# JSON METADATA FILES

## METADATA → IMAGE FILE

Image metadata file	Actual image file
JSON	JPG
<a href="https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/info.json">https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/info.json</a>	<a href="https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/full/full/0/default">https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/full/full/0/default</a>
→ `/info.json`	→ `/full/full/0/default`

# JSON METADATA FILES

- Manifest: <https://iiif.bodleian.ox.ac.uk/iiif/manifest/ae9f6cca-ae5c-4149-8fe4-95e6eca1f73c.json>
- Canvas: <https://iiif.bodleian.ox.ac.uk/iiif/canvas/699adfb3-a5da-4b09-9759-41eee3713bc3.json>
- Image: <https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/info.json>
- Image **file**: <https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/full/full/0/default>

# IMAGE API

- Resizing, cropping etc. happens on the server
- I can ask for a specific size, region etc. of an image through an URL
- the server creates an *ad hoc* image and only returns that



# IMAGE API

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/full/full/0/default>



# IMAGE API

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/500,3550,3000,140/full/0/default>



# IMAGE API

## LET'S MANIPULATE AN IMAGE

- Full image
- resizing (size)
- cropping (region)
- quality (color, gray...)
- rotation

**FULL IMAGE**

# FULL IMAGE

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/full/full/0/default>

# FILE FORMATS

# FILE FORMATS

Full image with file extension at the end (**.jpg**):

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/full/full/0/default.jpg>

# FILE FORMATS

...with file extension at the end (**.jp2**):

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/full/full/0/default.jp2>



# FILE FORMATS

The server has no copy of this image in **.png** format, so this returns an error:

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/full/full/0/default.png>

# FILE FORMATS

...with file extension **before**:

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3.jp2/full/full/0/default>

# CROPPING

# CROPPING

Full image, no cropping:

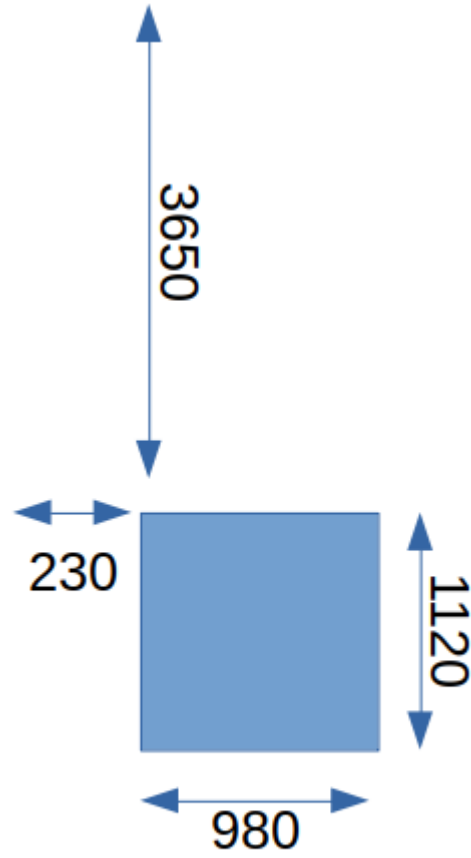
<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/full/full/0/default>

# CROPPING

Illuminated initial:

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/230,3650,980,1120/full/0/default>

# CROPPING



230,3650,980,1120

# CROPPING

Panoramic view of the city of Venice:

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/510,550,3580,3200/full/0/default>

# CROPPING

First line of the rubric (*Ci commence li livres du graunt  
Caam qui parole de la graunt Ermenie de Perffe*):

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/500,3550,3000,140/full/0/default>



# CROPPING

Second line of the rubric (*et def tartars et dÿnde. Et des  
granz merveille. qui p le monde font*):

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/530,3670,2400,140/full/0/default>

# CROPPING

See paragraph [Region](#) in the API documentation

# RESIZING

# RESIZING

Full size, no resizing:

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/full/full/0/default>

# RESIZING

Squeeze to 50 (width) x 600 (height):

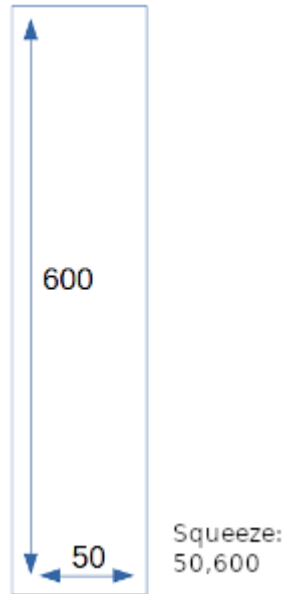
<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/full/50,600/0/default>



# RESIZING

Width 600 px:

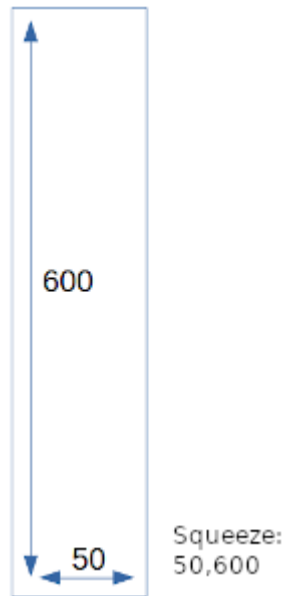
<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/full/50,/0/default>



# RESIZING

Height 600 px:

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/full/,600/0/default>



# RESIZING

See paragraph [Size](#) in the API documentation





# ROTATION

# ROTATION

No rotation (0°):

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/230,3650,980,1120/full/0/default>

# ROTATION

Rotate 90°:

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/230,3650,980,1120/full/90/default>

# ROTATION

Rotation is allowed only for multiples of 90° in this IIF server, so 45° returns an error:

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/230,3650,980,1120/full/45/default>

# ROTATION

Mirroring (with exclamation mark!):

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/230,3650,980,1120/full/!0/default>

# QUALITY

# QUALITY

Default:

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/230,3650,980,1120/full/0/default>

# QUALITY

Color (same as 'default'):

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/230,3650,980,1120/full/0/color>



# QUALITY

Gray:

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/230,3650,980,1120/full/0/gray>

# QUALITY

Bitonal (black and white):

<https://iiif.bodleian.ox.ac.uk/iiif/image/699adfb3-a5da-4b09-9759-41eee3713bc3/230,3650,980,1120/full/0/bitonal>

# QUALITY

See paragraph [Quality](#) in the API documentation

# LET'S LINK OUR TEI DIGITAL EDITION TO AN IMAGE

1. TEI → *local* image
2. TEI → *IIIF* image

# **TEI → LOCAL IMAGE ENCODING**

# TEI → LOCAL IMAGE ENCODING

1. Download image [218r.jpg](#) to your local folder
2. Write this in your TEI code:

```
<pb n="218r" facs="218r.jpg"></pb>
```

# TEI → LOCAL IMAGE VISUALIZATION

# TEI → LOCAL IMAGE

## VISUALIZATION

Let's use [OxGarage](#) to convert TEI XML to HTML.

Instructions:







< Text Encoding Initiative >

## **OxGarage Conversion**



**Convert from:** [?](#)



**Documents**



**Presentations**



**Spreadsheets**

Convert from: ?



### Documents

- Cocoa tagging
- Compiled TEI ODD
- DocBook Document
- Markdown tagging
- Microsoft Word (.doc)
- Microsoft Word (.docx)
- ODD Document
- OpenOffice 1.0 Text (.sxw)
- OpenOffice Text (.odt)
- Plain Text (.txt)
- Rich Text Format (.rtf)
- TCP XML Document
- TEI P4 XML Document
- TEI P5 XML Document



**Convert from:** ?



**Documents**

- Cocoa tagging
- Compiled TEI ODD
- DocBook Document
- Markdown tagging
- Microsoft Word (.doc)
- Microsoft Word (.docx)
- ODD Document
- OpenOffice 1.0 Text (.sxw)
- OpenOffice Text (.odt)
- Plain Text (.txt)
- Rich Text Format (.rtf)
- TCP XML Document
- TEI P4 XML Document
- TEI P5 XML Document
- TEI Tite XML Document
- WordPerfect (.wpd)

**Convert to:** ?

- Comma-Separated Values (.csv)
- DocBook Document
- ePub
- LaTeX
- Markdown tagging
- Microsoft Word (.doc)
- Microsoft Word (.docx)
- National Library of Medicine (NLM) DTD 3.0
- OpenOffice 1.0 Text (.sxw)
- OpenOffice Text (.odt)
- PDF
- Plain text
- RDF XML
- Rich Text Format (.rtf)
- TEI Simple XML Document
- VerbatimXML tagging
- xHTML



**Choose the file, upload images and press convert**

**Select file to convert:** [?](#)

Sfoggia...

Nessun file selezionato.

**+ Show advanced options** [?](#)

**Upload images:** [?](#)

You can upload image files and .zip files containing images

Sfoggia...

Nessun file selezionato. - Remove

+ Add more images

Convert

Reset





# TEI → IIIF IMAGE MANIFEST

Linking `<msDesc/>` (MS description, in the TEI header) to a whole IIIF JSON **manifest** (metadata about all images of a MS):

```
<sourcedesc>  
  <msdesc facs="https://iiif.bodleian.ox.ac.uk/iiif/manifest  
    <msidentifier>  
      <settlement>Oxford</settlement>  
      <repository>Bodleian Library</repository>  
      <idno>MS. Bodl. 264</idno>  
    </msidentifier>  
  [etc.]  
</msdesc>  
</sourcedesc>
```

# TEI → IIIF IMAGE PAGE

Linking <pb/> to a canvas JSON:

```
<pb n="218r" facs="https://iiif.bodleian.ox.ac.uk/iiif/canvas/
```

# TEI → IIIF IMAGE PAGE

Linking <pb/> to an image JSON:

```
<pb n="218r" facs="https://iiif.bodleian.ox.ac.uk/iiif/image/6
```

# TEI → IIIF IMAGE PAGE

Linking <pb/> to a **image file** (full image):

```
<pb n="218r" facs="https://iiif.bodleian.ox.ac.uk/iiif/image/6
```

# TEI → IIIF IMAGE LINE

Linking `<lb/>` to a **region** of an image file.

First line of the rubric (*Ci commence li livres du Grant  
Caam...*):

```
<lb n="218r.1" facs="https://iiif.bodleian.ox.ac.uk/iiif/image
```

# TEI → IIIF IMAGE LINE

...second line of the rubric (*et def tartars et dÿnde...*):

```
<lb n="218r.2" facs="https://iiif.bodleian.ox.ac.uk/iiif/image
```

# TEI → IIIF IMAGE GRAPHIC ELEMENT

```
<figure>  
  View of Venice  
  <figdesc>Illumination: view of the city of Venice</figdesc>  
  <graphic url="https://iiif.bodleian.ox.ac.uk/iiif/image/69  
</graphic></figure>
```



# LEARNING MATERIALS AND CONTACTS

1. DARTH, [Hacking Mirador 2016](#)
2. Paolo Monella, [Bologna 2019](#)
3. Tiziana Mancinelli, [Bologna 2020](#)
4. Paolo Monella and Roberto Rosselli Del Turco, [AIUCD 2020](#)
5. VeDPH Summer camp 2020, [main page](#)
6. Paolo Monella, [contacts](#)