Biblissima

CAMPUS CONDORCET
Paris–Aubervilliers

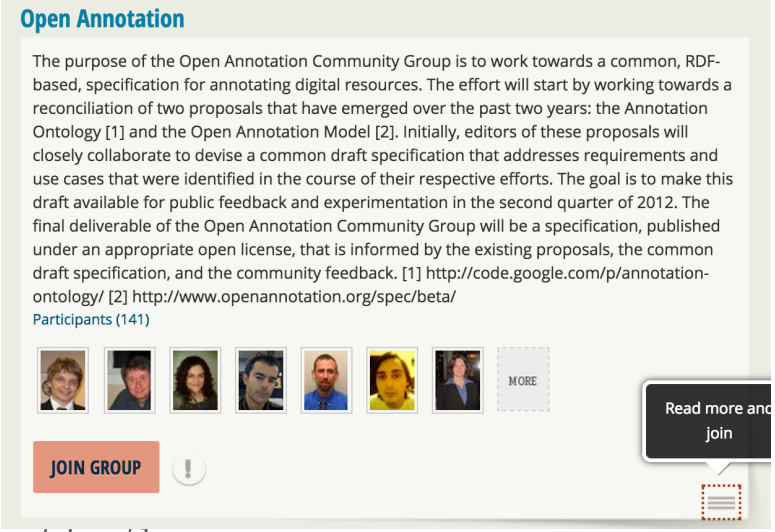
OA – Shared Canvas – TEI

And how they are used in the Biblissima project

Stefanie Gehrke, coord. Metadata+Prototype Biblissima - Workshop TEI and neighbouring standards - 15/09/15 Huygens, Den Haag

What is Open Annotation ?

Open Annotation Collaboration and
Annotation Ontology initiative
jointly fonded
W3C Open Annotation Community Group



Purpose: towards a common **RDF-based** specification for

annotating digital resources

Reconciliation of two proposals:

- Annotation Ontology [http://code.google.com/p/annotation-ontology/]
- Open Annotation Model [http://www.openannotation.org/spec/beta/]

=> **Open Annotation Data Model (2013)**: http://www.openannotation.org/spec/core/

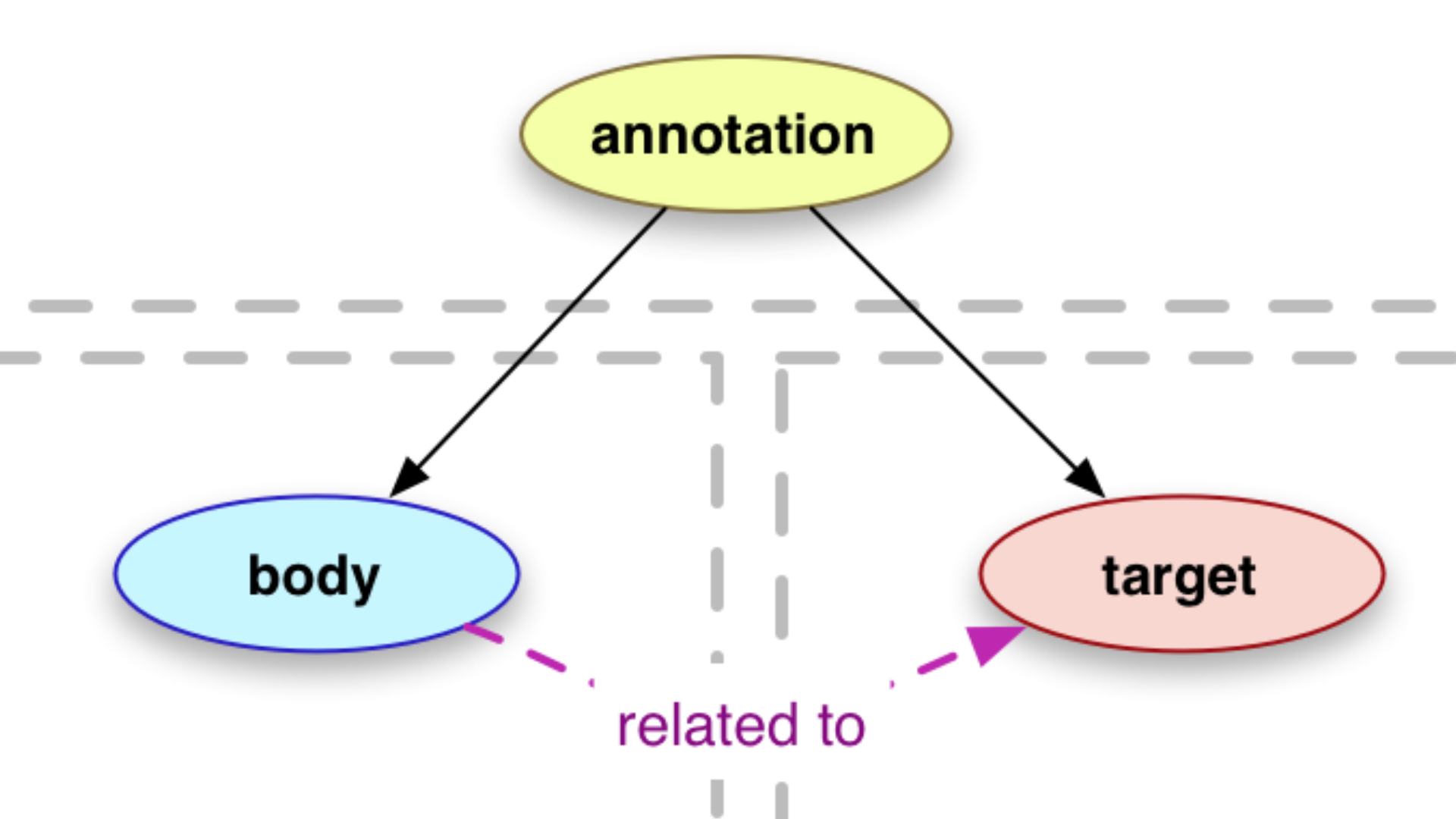
OA: Core object-properties and classes

Namespace oa: http://www.w3.org/ns/oa#

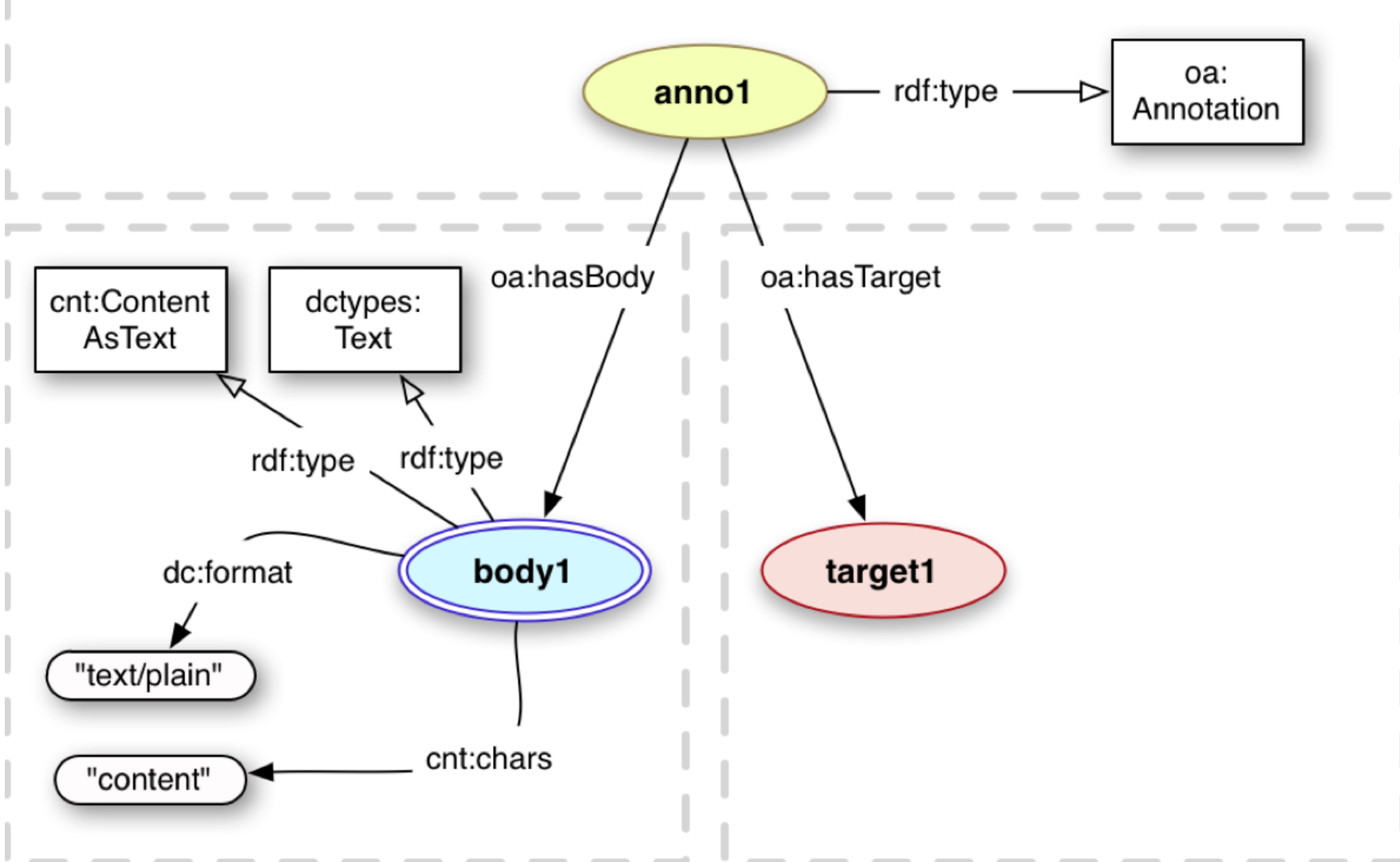
oa:hasBody	oa:Annotation
oa:hasTarget	oa:Motivation
rdf:type	

- creating associations between resources, annotations
- the most common use case: attaching a piece of text to a single web resource
- publishing via the web as Linked Open Data in RDF

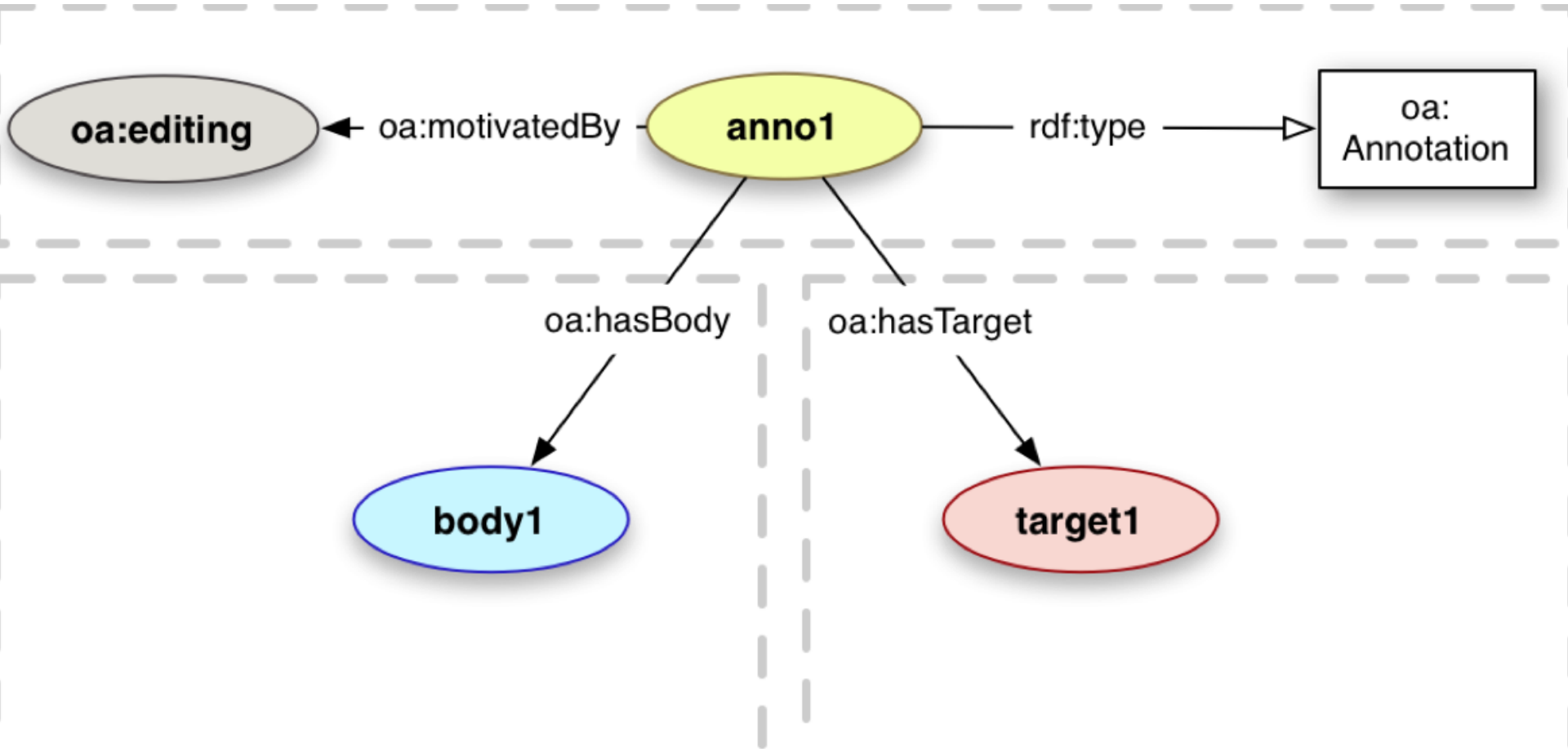
Annotation in OA


annotation
body
target
related to

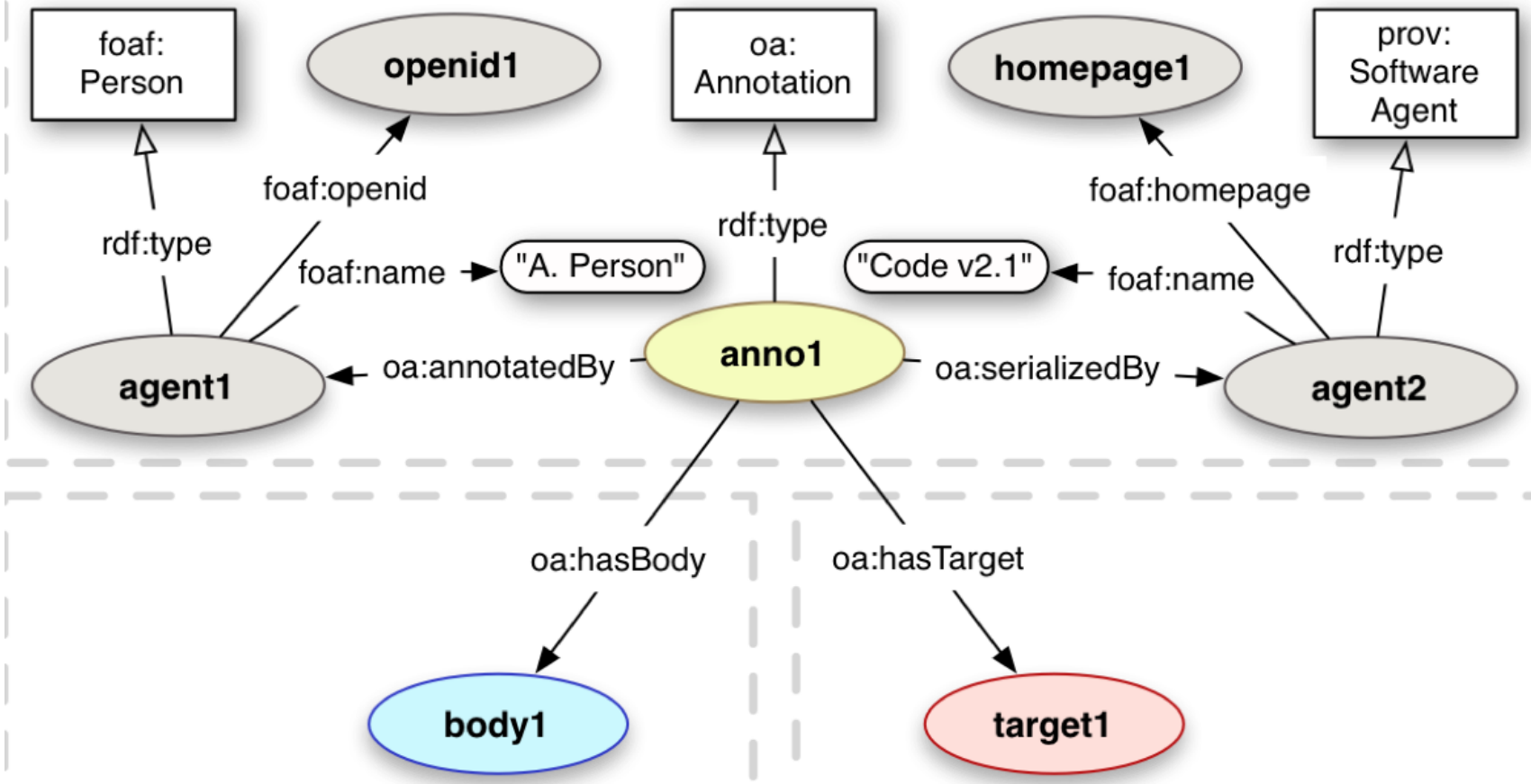
Textual Annotation in OA


anno1
rdf:type
oa:
Annotation
oa:hasBody
oa:hasTarget
cnt:Content
AsText
dctypes:
Text
rdf:type
rdf:type
dc:format
body1
target1
"text/plain"
cnt:chars
"content"

The Motivation of an Annotation in OA

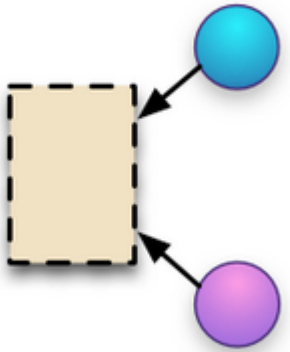

oa:editing
oa:motivatedBy
anno1
rdf:type
oa:
Annotation
oa:hasBody
oa:hasTarget
body1
target1

The Actors involved in the Annotation Process in OA


foaf:
Person
openid1
oa:
Annotation
homepage1
prov:
Software
Agent
foaf:openid
rdf:type
rdf:type
foaf:homepage
rdf:type
foaf:name
"A. Person"
"Code v2.1"
foaf:name
agent1
oa:annotatedBy
anno1
oa:serializedBy
agent2
oa:hasBody
oa:hasTarget
body1
target1

Relation between OA and Shared Canvas / IIIF

- Shared Canvas Data Model is based on OA
- Namespace sc: http://www.shared-canvas.org/ns/



- Editors: Robert Sanderson, Los Alamos National Laboratory Benjamin Albritton, Stanford University – Robert Sanderson, today Stanford University, is also editor of the OA Data Model

Relation between OA and Shared Canvas / IIIF

- IIIF APIs are compliant to the Shared Canvas Data Model
  (API: Application Programming Interface)

- Presentation API (2.0): http://iiif.io/api/presentation/2.0/

  Manifest in JSON-LD

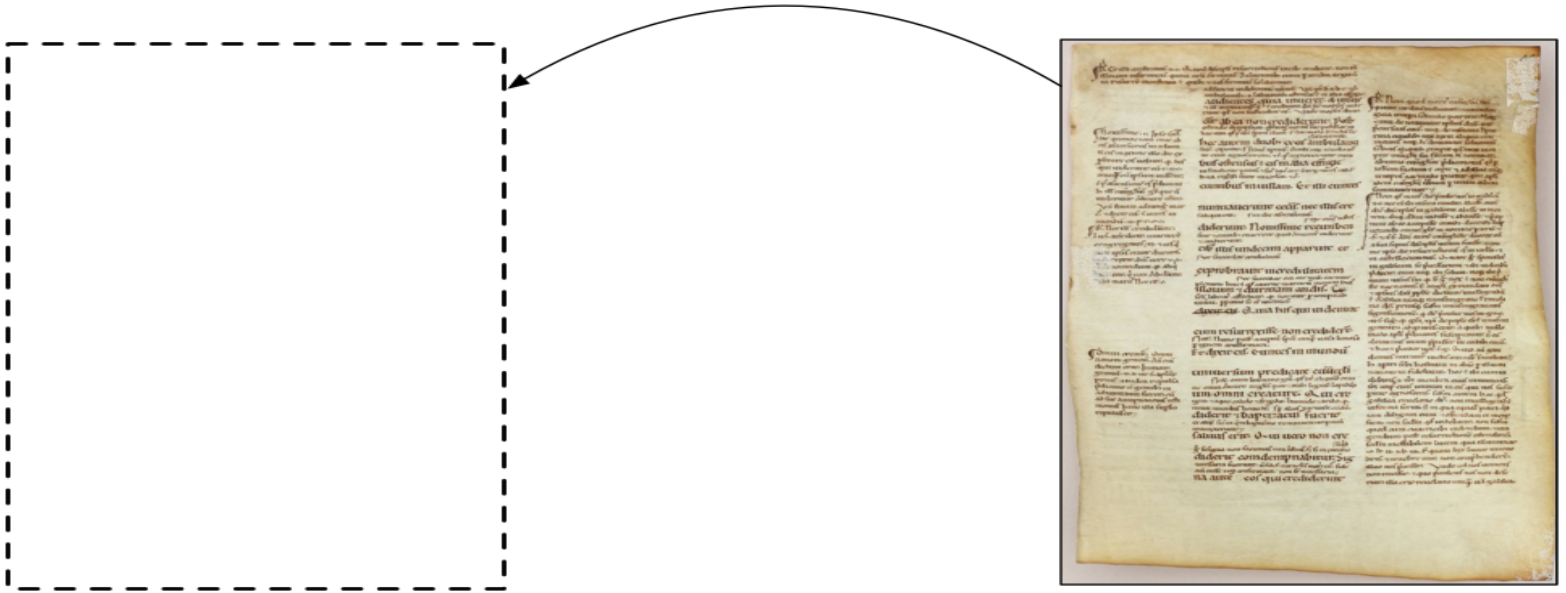
- Image API (2.0): http://iiif.io/api/image/2.0/

Differences between OA and Shared Canvas / IIIF

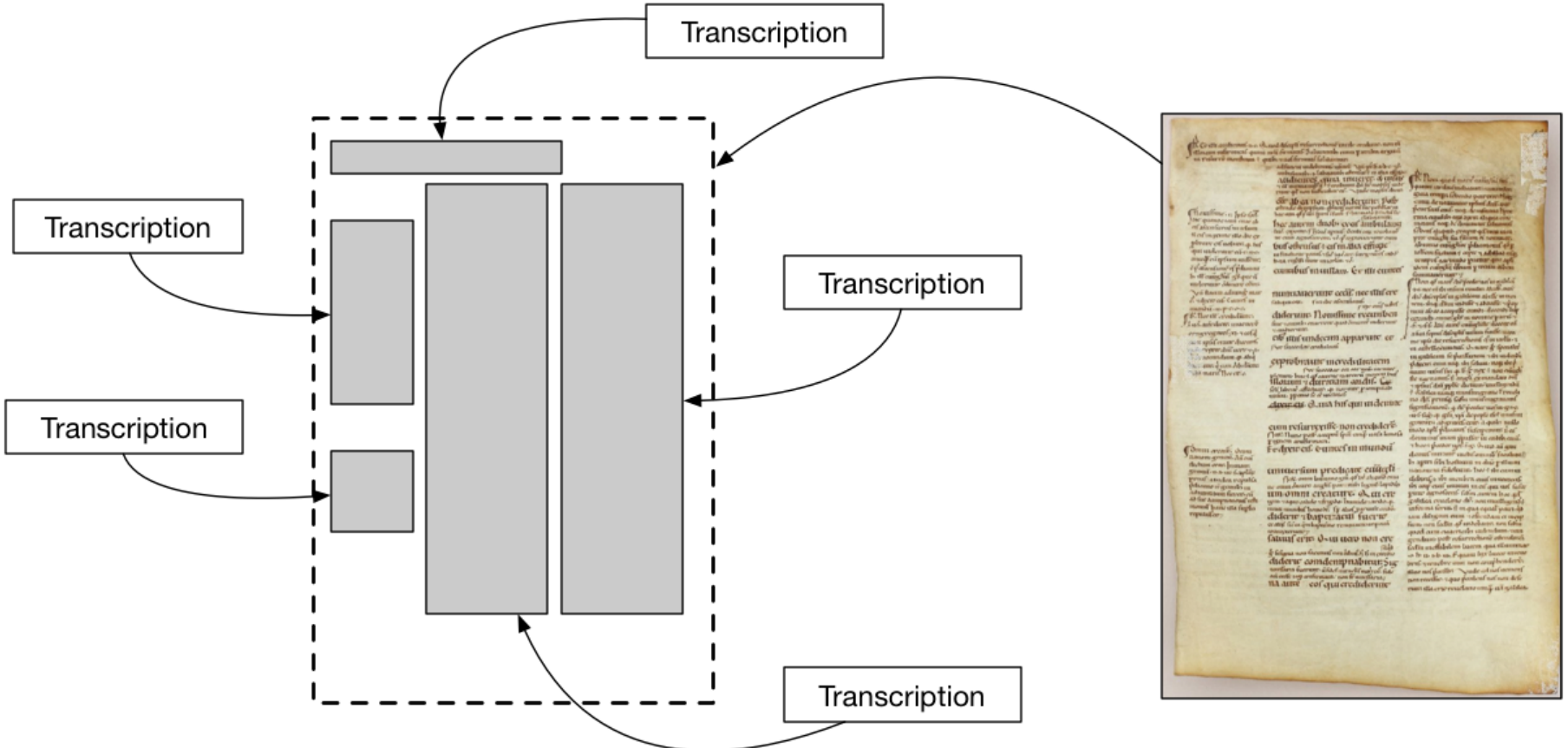
- Creation of a new class called **sc:Canvas** (abstract space used for building a view of an object, target of the annotation)

- Creation of a new instance of **oa:Motivation** called **sc:painting**
- Creation of **sc:Manifest** (ore:Aggregation), **sc:Sequence**, **sc:Range**, **sc:AnnotationList**
- Creation of **sc:Collection**

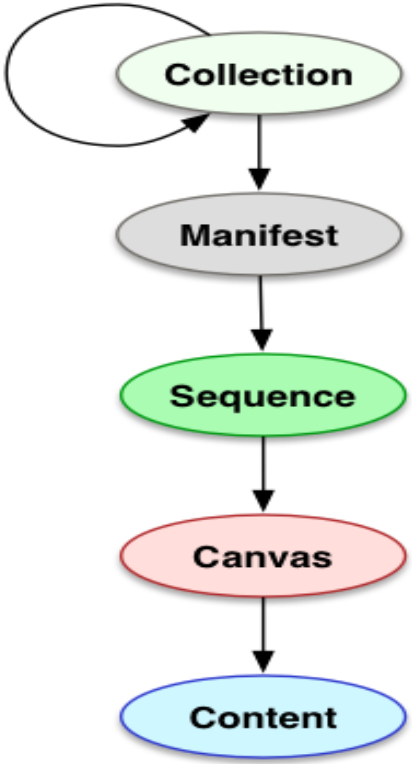
Shared Canvas



Shared Canvas



Shared Canvas - Structure



Why is it important (for the TEI-community) ?

- transcending silos => Linked Open Data Model for collaborative access to digital surrogates
- interoperability via the principles of the semantic web
- images and metadata still kept at repositories
- IIIF Consortium (IIIF-C) formed in May 2015: Bodleian Library, British Library, Stanford University Library, ARTstor, Bayerische Staatsbibliothek, Cornell University, Bibliothèque nationale de France, Nasjonalbiblioteket (Norway), Princeton University Library, Wellcome Trust, Yale University
- developed within the context of the study of medieval manuscripts

Shared Canvas/IIIF and TEI

- not embedded in TEI
- a sc:Manifest can be generated from a TEI-file
- probably not an alternative to TEI (who will encode a transcription in a JSON-LD file ?)
- supports the display of transcriptions deriving from TEI-files in an image-viewer
- still missing : an object-property that relates a manifest to the physical object for which it provides a view

Biblissima and IIIF / Shared Canvas

IIIF / Shared Canvas is used in order to establish a single access-point to three virtual libraries Gallica (BnF), Bibliothèque virtuelle des manuscrits médiévaux (IRHT) and Bibliothèques virtuelles humanistes (CESR).

Manifests (work in progress):

https://graph.global/static/data/universes/iiif/biblissima.json

http://bvmm.irht.cnrs.fr/manifest/

Sample Prototype:

https://demo.logilab.fr/biblissima/id/Illumination/Mandragore/23726

France, Paris, Bibliothèque nationale de France, Français 150 f. 25

texte enluminé : vaulx (j. de), premières oeuvres Jacques de Vaulx (cartographe, 15..?-15..)

lieu précis d'origine : Le Havre (Seine-Maritime, France)

pays d'origine : France

descripteurs : angleterre, atlantique, bateau, belgique, belle.île, carte, compas, cordouan, espagne, france, irlande, man, manche, mer.du.nord, oléron, ouessant, pays.bas, poids.mesures, ré, scilly, sein, wight, yeu, île, îles.anglo.normandes, îles.cies, îles.frisonnes

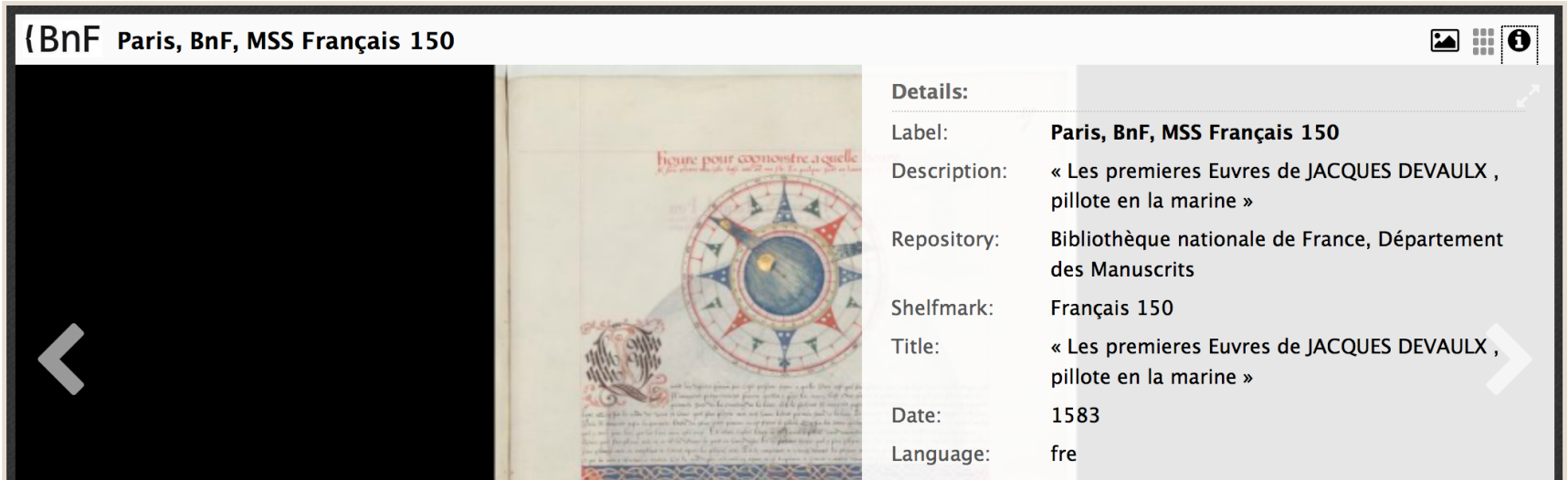
descripteurs géo : Angleterre, Archipel Frison, Atlantique (océan), Belgique, Belle-Île (Morbihan), English Channel, Espagne, France, Irlande, Isles of Scilly, L'Île-d'Yeu (Vendée), Mer du Nord (Atlantique Nord), Oléron, Île d' (Charente-Maritime), Ouessant (Finistère), Pays-Bas, Ré, Île de (Charente-Maritime), cordouan, Île de Man (Royaume-Uni), Île de Wight (GB), Île-de-Sein (Finistère, France), Îles Anglo-Normandes (Royaume-Uni), Îles Cíes (Galice, Espagne)

légende : poids et mesures: distances

url de la notice : http://mandragore.bnf.fr/jsp/feuilleterNoticesImageArk.jsp?id=23726

images : http://gallica.bnf.fr/ark:/12148/btv1b550024840/f74.highres

manuscrit : France, Paris, Bibliothèque nationale de France, Français 150



Generating a IIIF manifest from a TEI-file

Metadata about the physical object (descriptive metadata):

msDesc/… => metadata: « label » - « value » pairs

Structural metadata:

msItem/locus @from, @to, @on => sc:Range

msPart => sc:Range

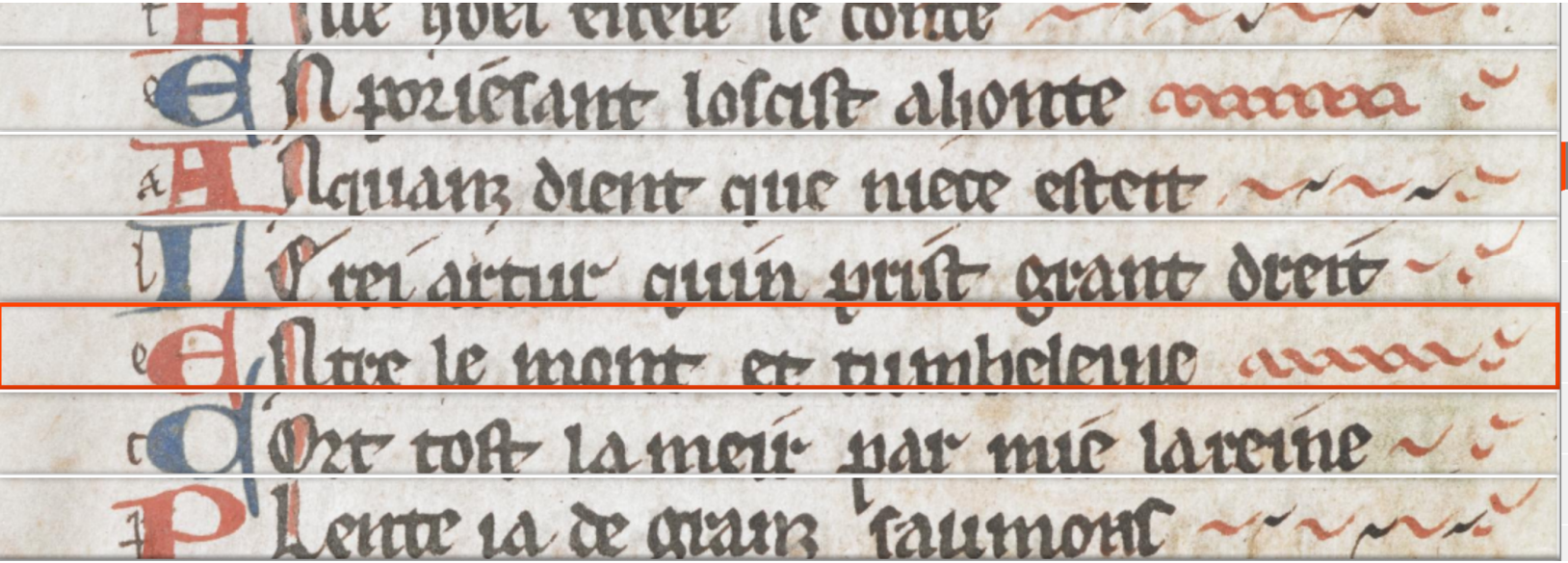
Transcription:

text/body/div/pb => sc:AnnotationList per sc:Canvas

facsimile/surface + graphic + zone => sc:Canvas + image + chars on URI.json#xywh=X,Y,W,H

Biblissima Demo of the Roman du Mont Saint-Michel

http://demos.biblissima-condorcet.fr/roman-bl-caen/m1/

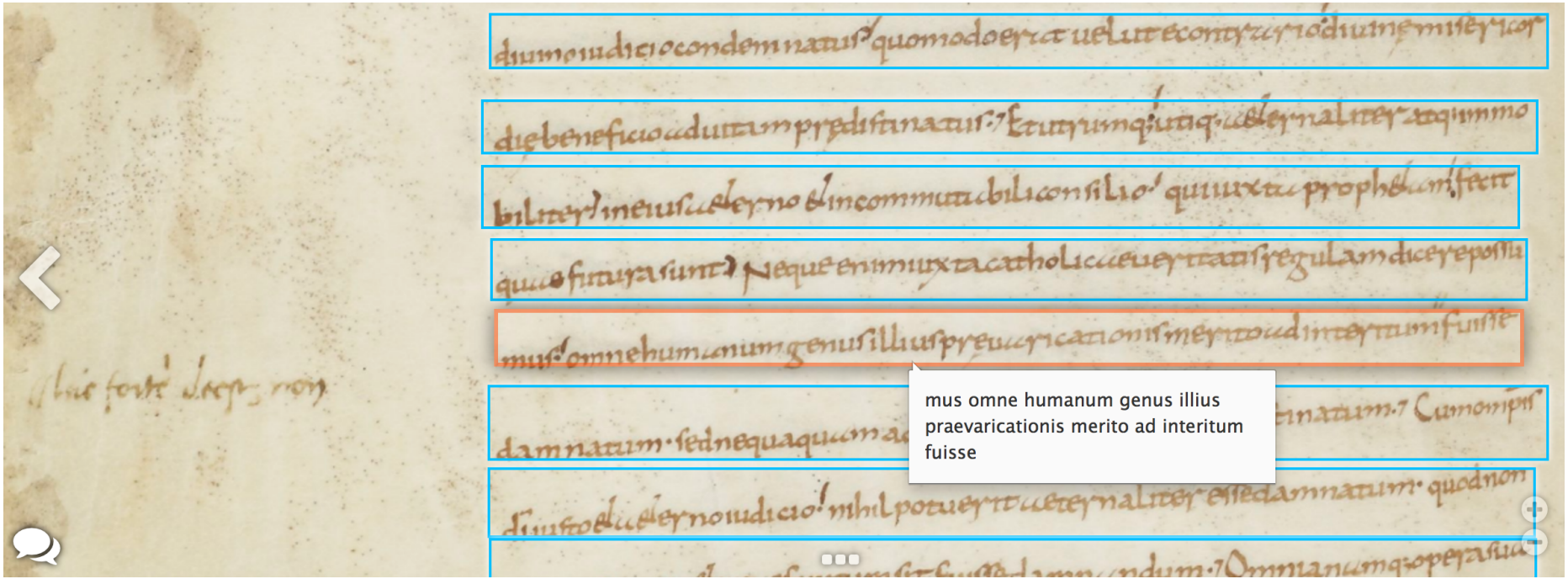


XSLT: https://github.com/stefaniegehrke/TEI-2-SC

JSON-LD: http://demos.biblissima-condorcet.fr/iiif/metadata/BL_Add_10289/list/transscript_8v.json

Biblissima Demo of Florus de Lyon

http://demos.biblissima-condorcet.fr/florus/absolutio/mirador/#652d0c4c-d1a5-4422-a39b-ce258f4ff4ee



http://demos.biblissima-condorcet.fr/florus/#absolutio

Biblissima and IIIF/Shared Canvas: Planning

- display of transcriptions (inventories / sales catalogues and literary texts)

=> work on automatic line-detection

- provide a workspace for researchers

Other Projects

Shelley-Godwin-Archive:

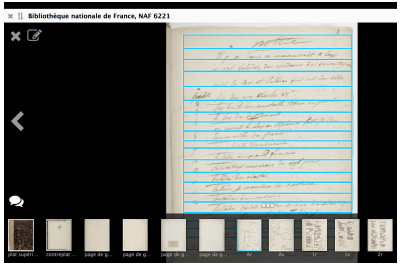
- http://shelleygodwinarchive.org/about
- https://github.com/umd-mith/unbind

The Sentences Commentary Text Archive (Jeffrey Witt):

- http://lombardpress.org/category/imagesmanifestsiiif/
- http://images.scta.info/
- http://mirador.scta.info/#ebfdb9a3-455c-4c5d-94de-d72b91f744e7

Sample files IIIF

http://iiif.vischina.freizo.org/m2/#uuid-98fa42ae-c7e9-4c83-badb-3c57e91bf49b
(BnF NAF 6221 - for display of line by line transcription click - Recueil de 154 lais, ballades, rondeaux et serventois, dont 71 au moins d'Eustache Deschamps - resp. Benjamin Albritton, Stanford University)



JSON-LD samples:

- http://dms-data.stanford.edu/data/manifests/BnF/jr903ng8662/manifest.json
- http://dms-data.stanford.edu/data/manifests/Stanford/ege1/manifest.json
- http://iiif.io/api/presentation/2.0/example/fixtures/
  (number 43-48 and 61 about transcription)