

**D**E BVS  
DOMINI  
CIS SER  
MO BEATI  
AVGVSTINI  
EPISCOPI  
DE SCA TRI  
NITATE  
EX LIBRO  
IPSIVS AD

PETRVM  
DIACO  
NV M DE  
FIDE



## Manuscript-AI

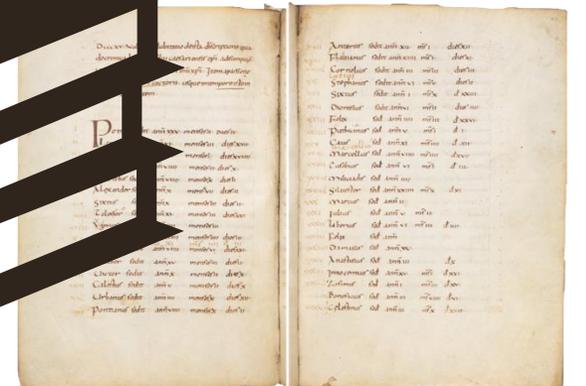
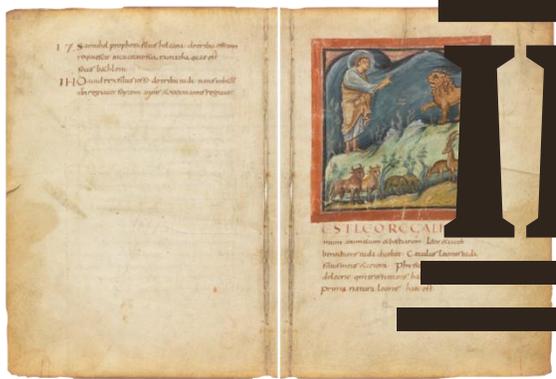
A dedicated ontology and AI-tool to facilitate integration and accessibility of heterogeneous catalogue data on medieval manuscripts

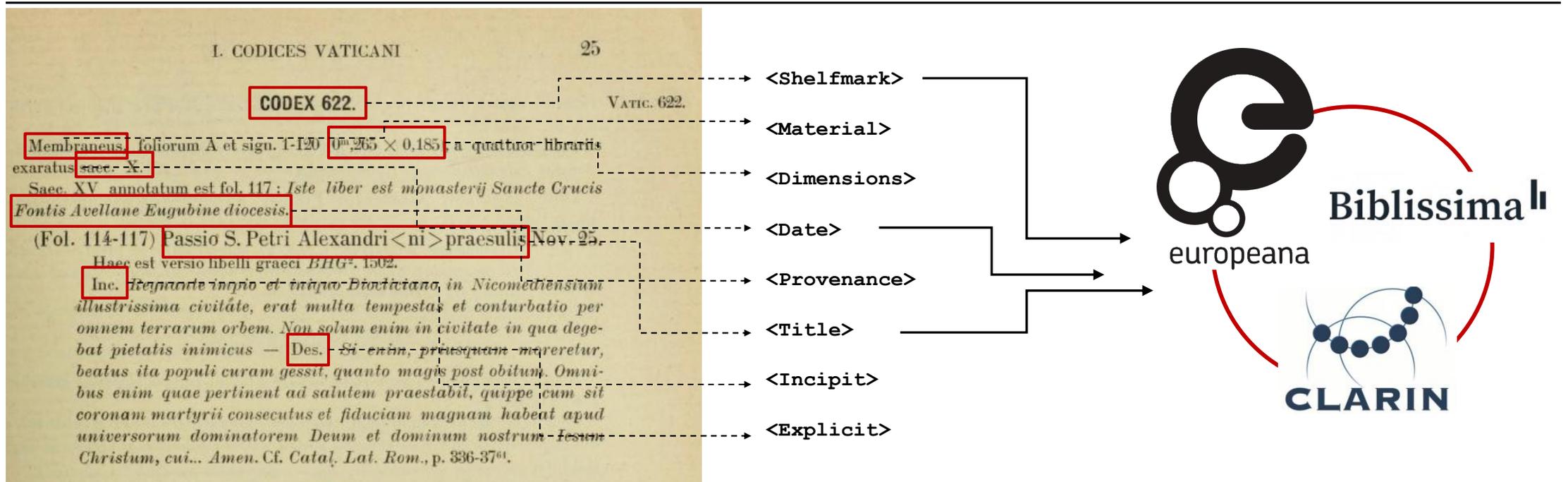








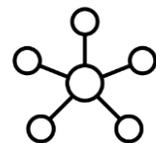


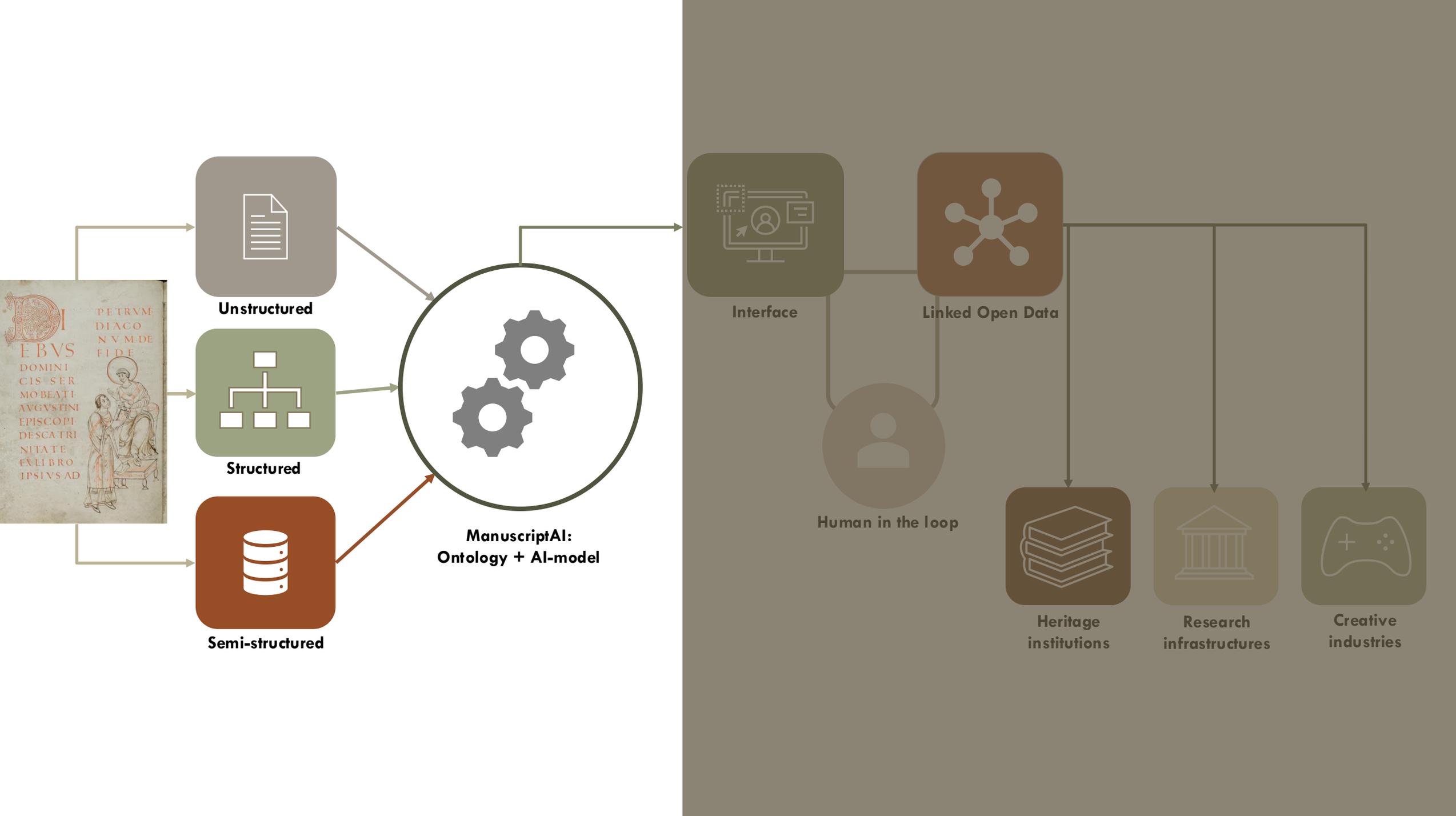


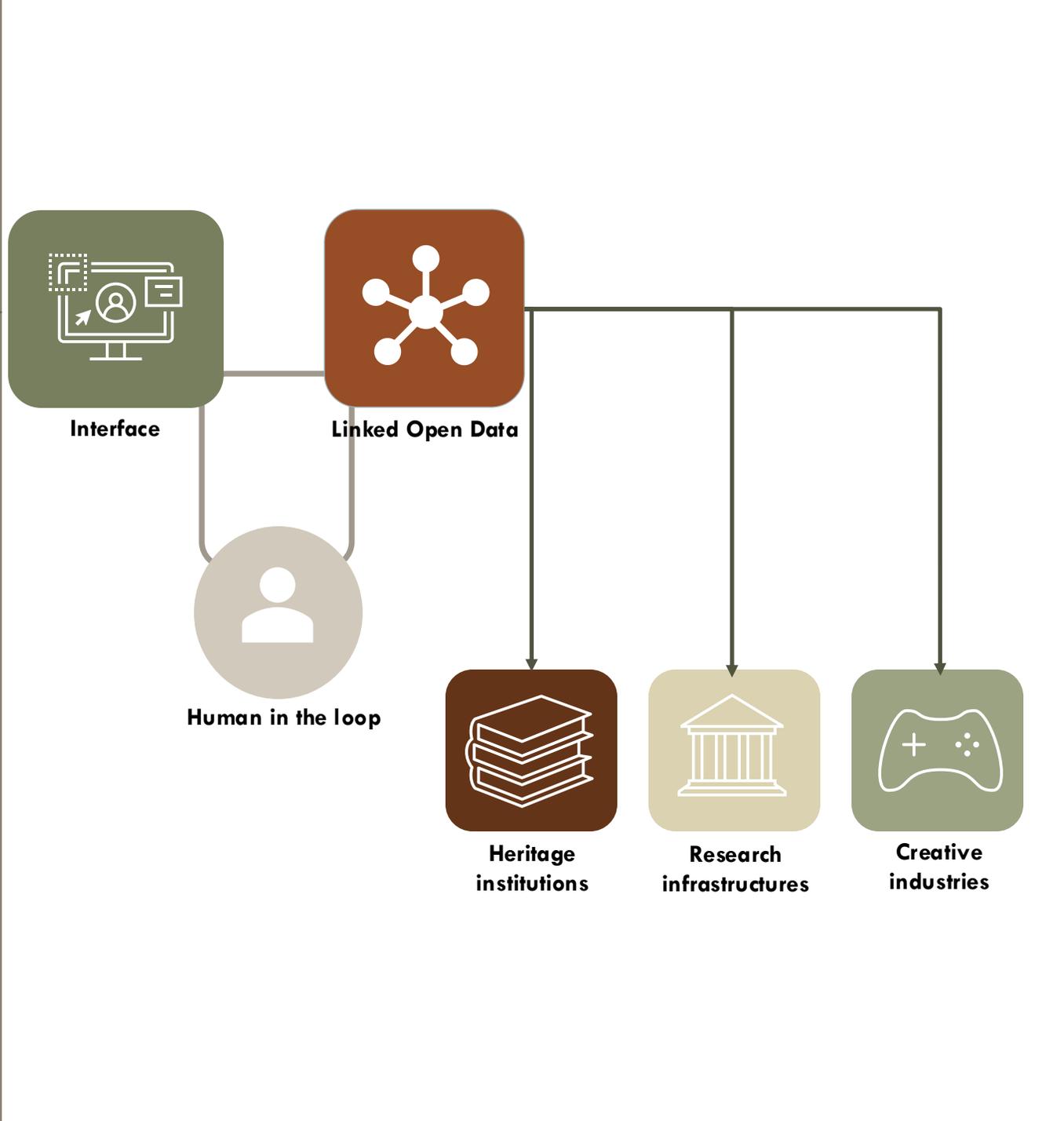
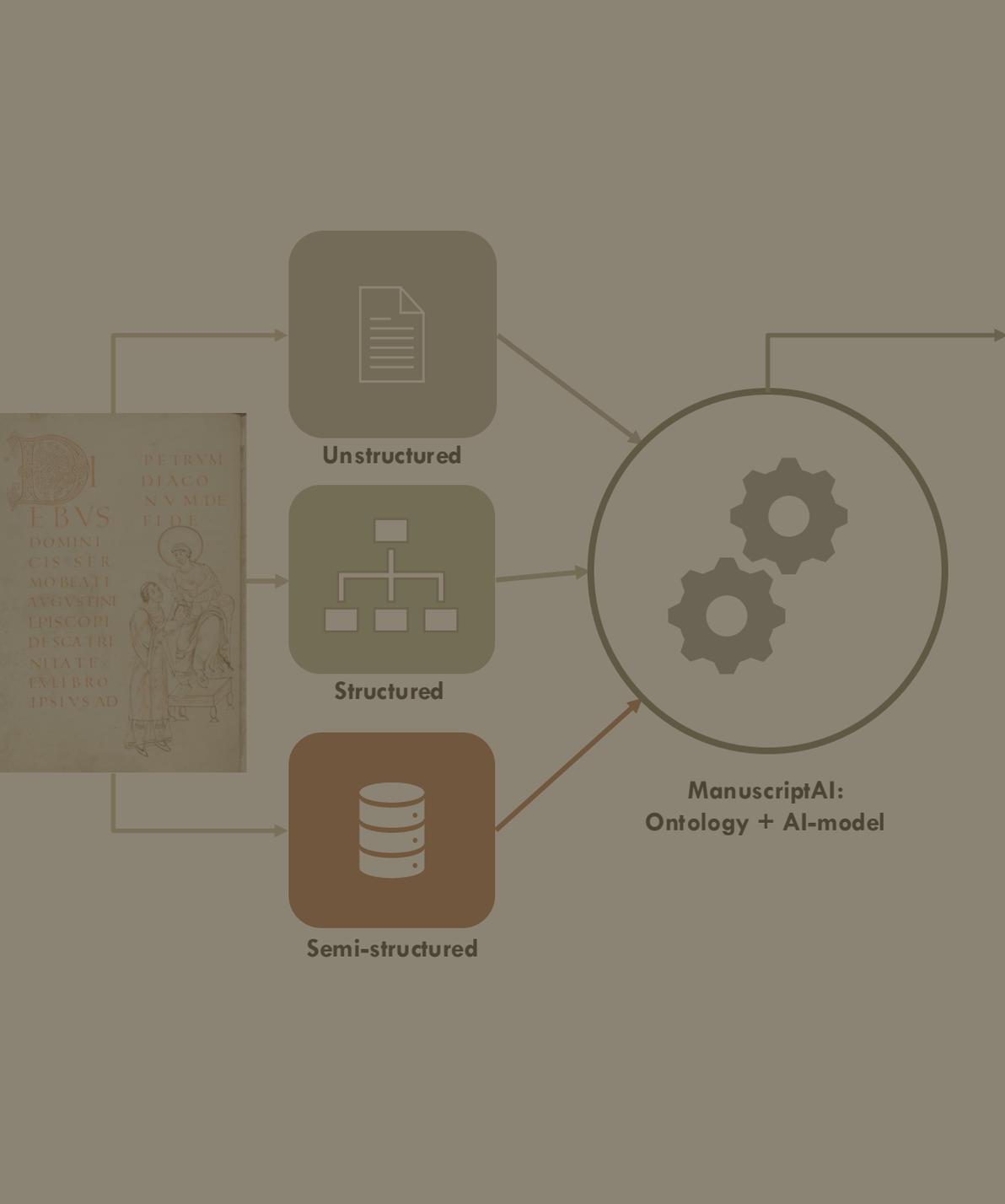
Customised mapping to  
semi-structured data



Manual creation of  
Linked Open Data

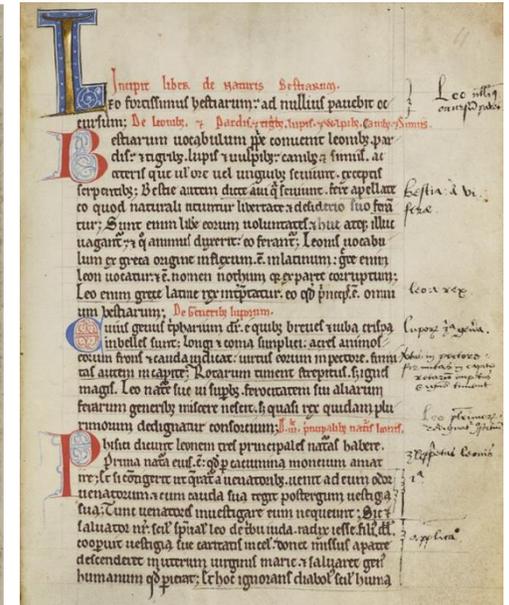
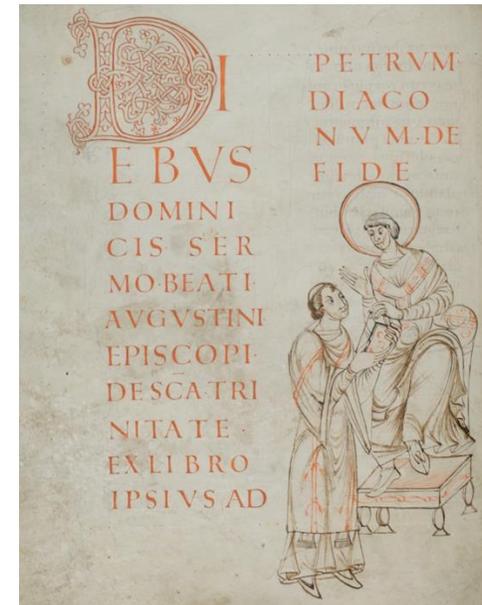




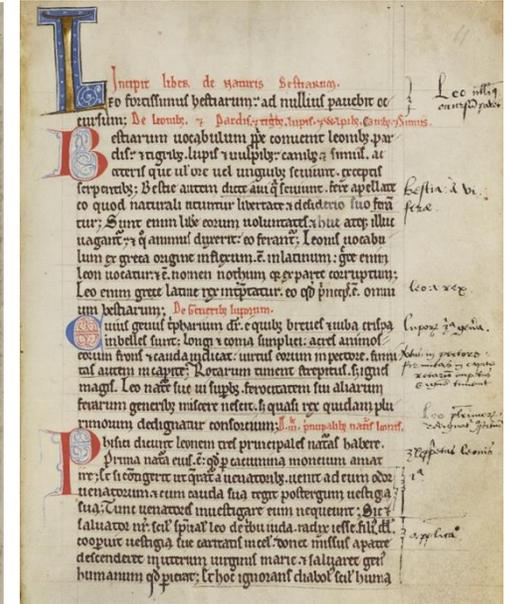
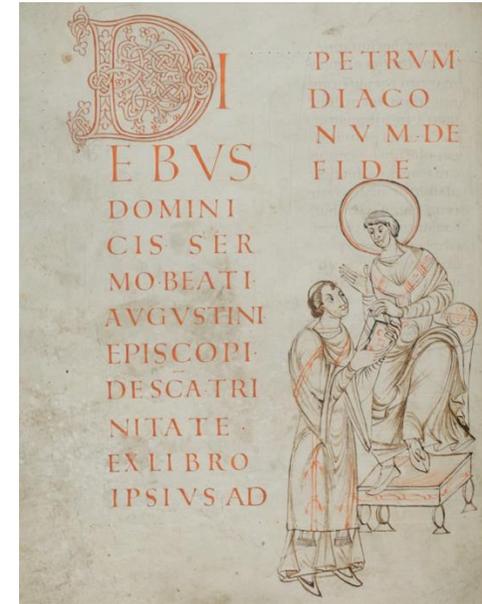


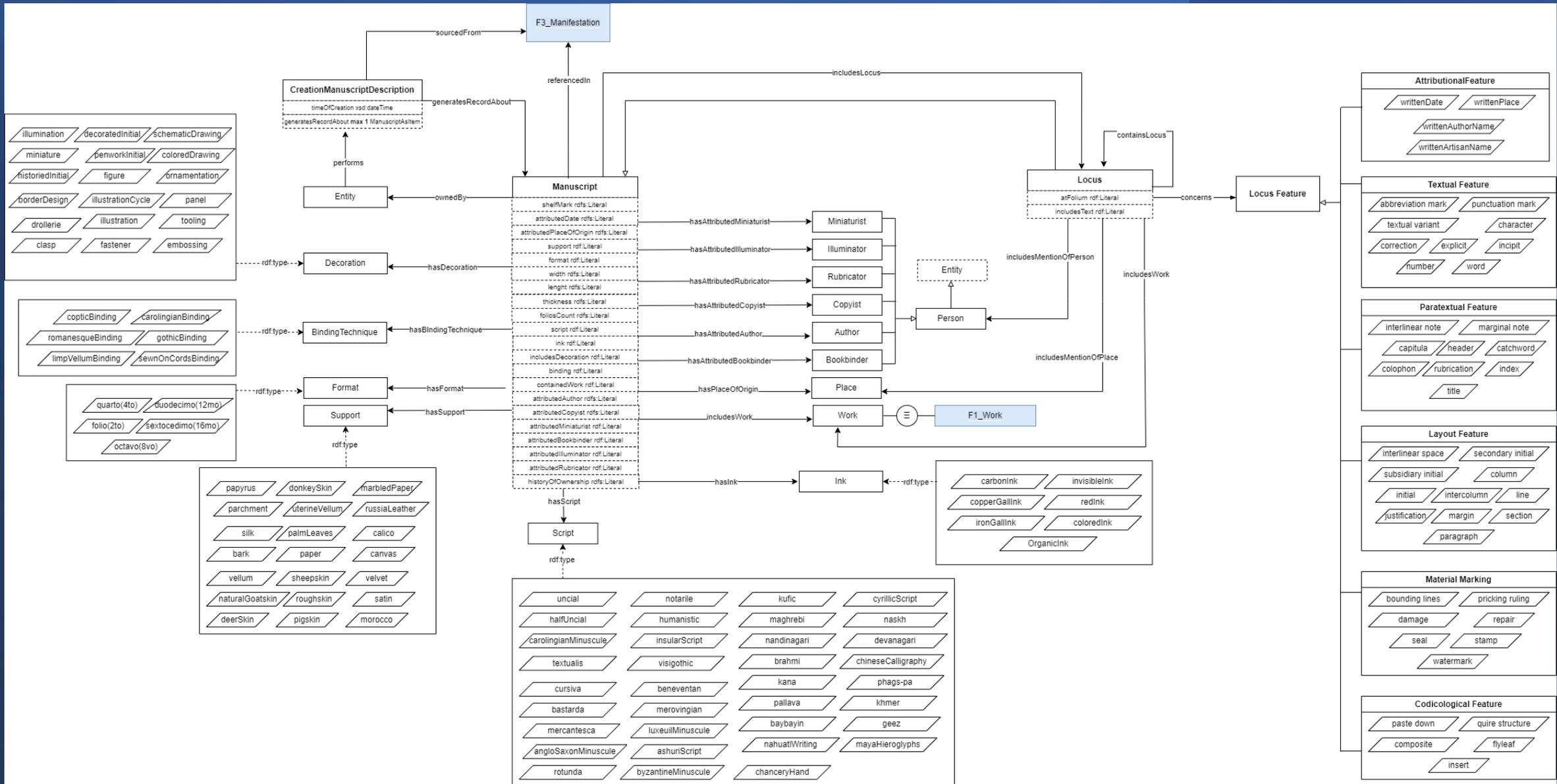


- Manuscripts are not just books, nor are they just objects
- No ontology has critical mass and sophistication varies
- Existing ontologies and catalogue data neglect the contents
- Digital manuscript studies needs good metadata to flourish



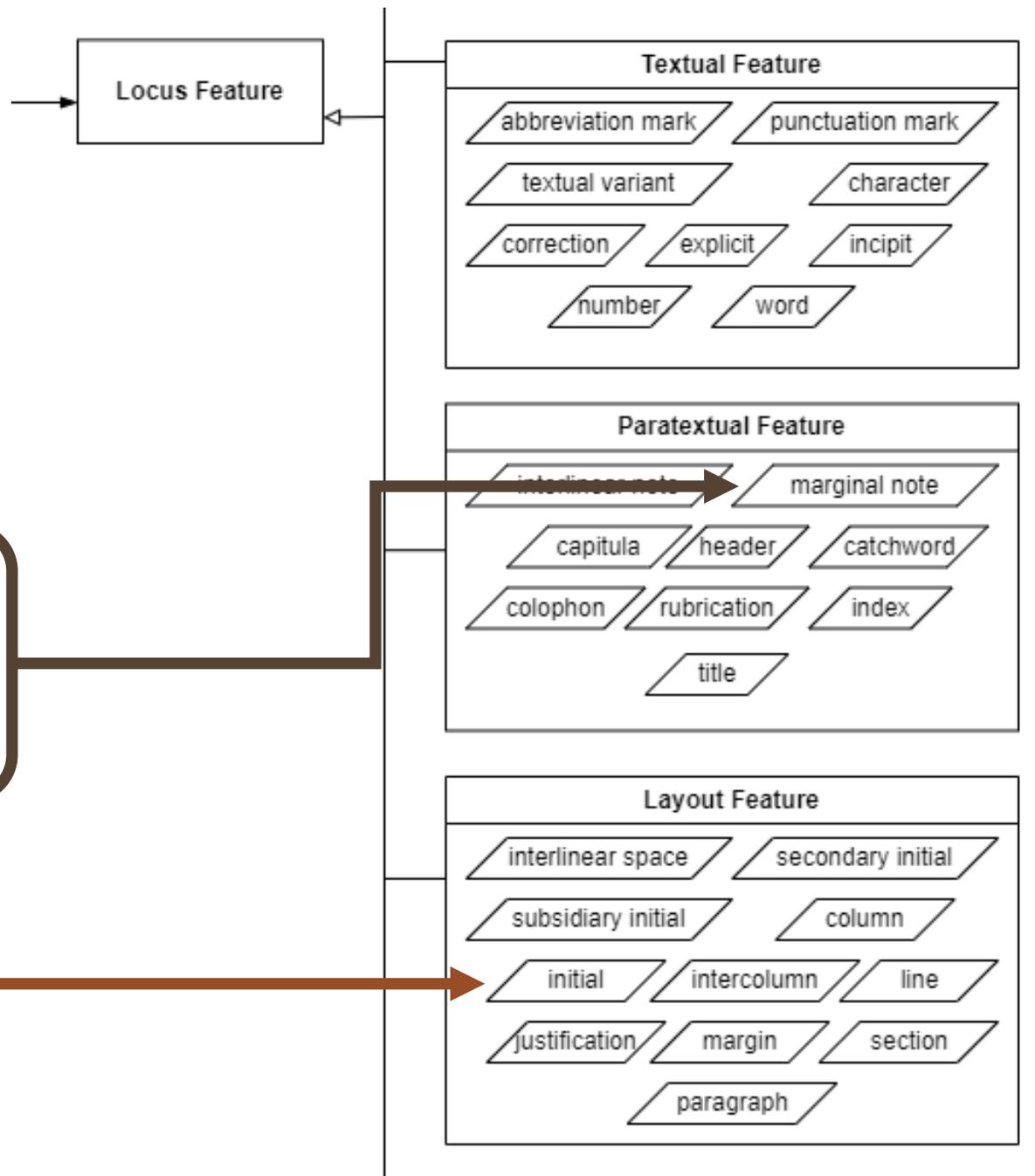
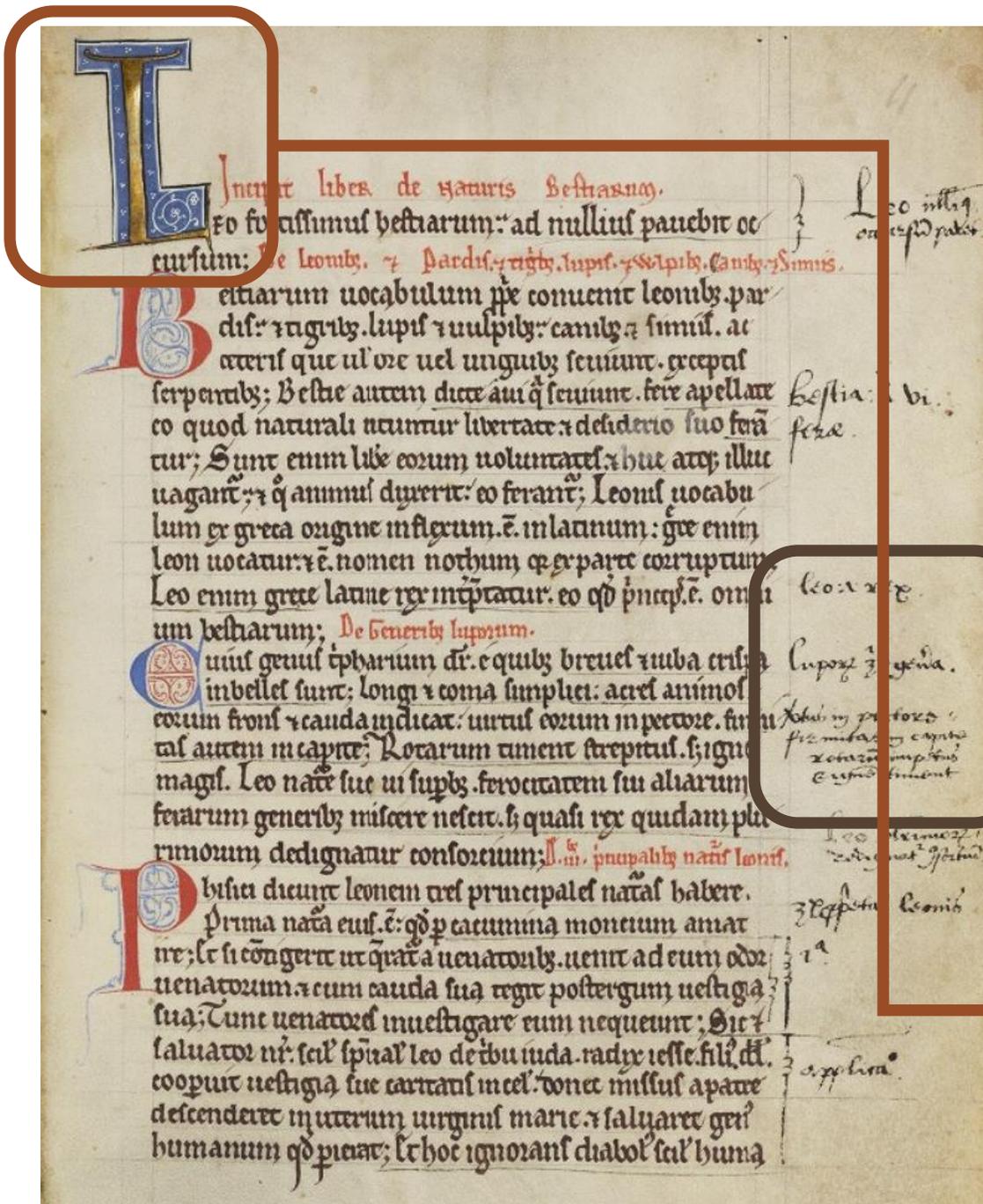
- The ManuscriptAI project
- The need for an ontology for manuscripts
- **Presentation of the ontology**
- Some specific features
- The ontology as part of our AI-framework
- The sustainability challenge



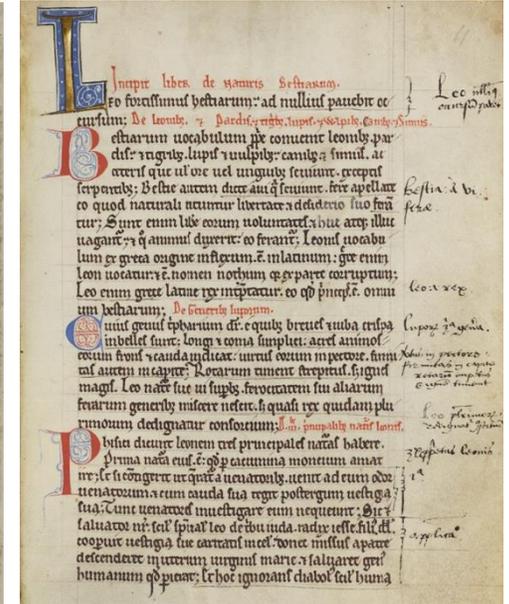
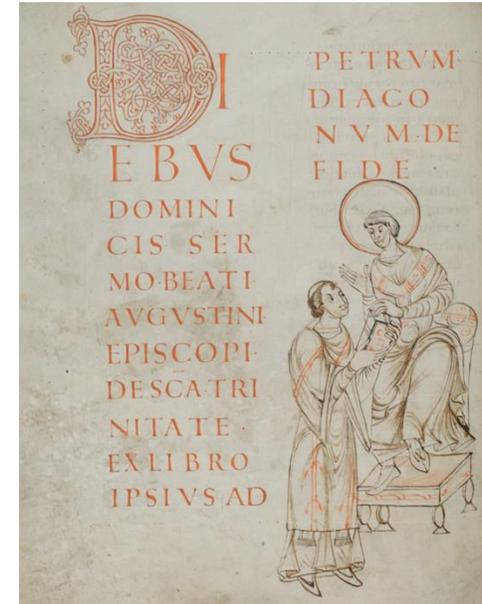








- The ManuscriptAI project
- The need for an ontology for manuscripts
- Presentation of the ontology
- Some specific features
- The ontology as part of our AI-framework
- The sustainability challenge





determine what works best for your needs. Each option is designed to help you structure your data in a machine-understandable format, making it ready to be seamlessly linked with other collections of data.

### Create LOD

Use this path if you want to structure data using Manuscript-AI Ontology from scratch (or explore the full expressiveness of the ontology) and see your knowledge graph grow.

Level of Automation:



START

### Manuscript Annotation Desk

Use this path to identify the right information in your data that need to be sent to the agent framework to convert your data into LOD using Manuscript-AI Ontology.

Level of Automation:



START

### Select and Convert

Use this method if you wish to automate the extraction of key values from specific data segments and convert them into LOD using the Manuscript-AI ontology.

Level of Automation:



START

### Drop & Classify

Use this method if you wish to just drop your data and let agents take care of classifying them under the Manuscript-AI ontology.

Level of Automation:



START

## MANUSCRIPT INFORMATION

To uniquely identify each manuscript, please provide the **shelfmark** of the manuscript in the field below without any spaces or punctuation marks. For example, enter "L1234" instead of "L 1234" or "L-1234". This identifier will be used to create a unique URI for the manuscript in the knowledge graph, just as the user ID is used to identify users.

Shelfmark\*

NEXT

Please fill in the information below. If something is unknown or not applicable, you may leave the field(s) empty.

Attributed Date of the Manuscript

Attributed Place of Origin of the Manuscript

Includes Work(s)

## TURTLE RDF SYNTAX

Prefixes used:

```
@prefix ex: <http://www.example.org/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix MS4AI: <http://ontology.tno.nl/manuscriptAI/> .
```

Knowledge Graph:

```
ex:6789 rdf:type MS4AI:Entity ;
  MS4AI:performs ex:idl4ixdyhjo ;
  rdfs:label "Giulia"^^xsd:string .

ex:idl4ixdyhjo rdf:type MS4AI:CreationManuscriptDescription ;
  MS4AI:timeOfCreation "2025-03-28T12:37:21.230Z"^^xsd:dateTime ;
  MS4AI:generatesRecordAbout ex:RIC2345 .

ex:RIC2345 rdf:type MS4AI:Manuscript ;
  MS4AI:shelfmark "RIC2345"^^xsd:string ;
  MS4AI:containedWork "Divina commedia"^^xsd:string ;
  MS4AI:attributedAuthor "Dante"^^xsd:string ;
  MS4AI:attributedPlaceOfOrigin "Florence"^^xsd:string ;
  MS4AI:attributedDate "1345"^^xsd:string .
```

## EXPAND CONNECTIONS

Here, you can link your data with other collections by providing URI/IRI identifiers for the work(s), author(s), copyist(s) and place of origin(s) of the manuscript.

SEARCH ON WIKIDATA

SEARCH ON EUROPEANA

### Work

+

X

### Author

+

X

### Copyist

+

### Rubricator

+

### Miniaturist

+

## TURTLE RDF SYNTAX

### Prefixes used:

```
@prefix ex: <http://www.example.org/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix MS4AI: <http://ontology.tno.nl/manuscriptAI/> .
```

### Knowledge Graph:

```
ex:6789 rdf:type MS4AI:Entity ;
  MS4AI:performs ex:idl4ixdyhjo ;
  rdfs:label "Giulia"^^xsd:string .

ex:idl4ixdyhjo rdf:type MS4AI:CreationManuscriptDescription ;
  MS4AI:timeOfCreation "2025-03-28T12:37:21.230Z"^^xsd:dateTime ;
  MS4AI:generatesRecordAbout ex:RIC2345 .

ex:RIC2345 rdf:type MS4AI:Manuscript ;
  MS4AI:shelfmark "RIC2345"^^xsd:string ;
  MS4AI:containedWork "Divina commedia"^^xsd:string ;
  MS4AI:attributedAuthor "Dante"^^xsd:string ;
  MS4AI:attributedPlaceOfOrigin "Florence"^^xsd:string ;
  MS4AI:attributedDate "1345"^^xsd:string ;
  MS4AI:includesWork <https://www.wikidata.org/wiki/Q40185> ;
  MS4AI:hasAttributedAuthor <https://www.wikidata.org/wiki/Q1067> .
```

## Well done! Now, it is time to download

The knowledge graph you've created leverages machine-understandable, semantic-based vocabulary terms. To enhance interoperability, select additional ontologies compatible with Manuscript-AI (which extends these ontologies). Once selected, you can download the complete knowledge graph along with the associated ontology.

- Use **Manuscript-AI** only
- Includes alignment with **CIDOC-CRM**
- Includes alignment with **LRM (New FRBR Version)**

Download



determine what works best for your needs. Each option is designed to help you structure your data in a machine-understandable format, making it ready to be seamlessly linked with other collections of data.

### Create LOD

Use this path if you want to structure data using Manuscript-AI Ontology from scratch (or explore the full expressiveness of the ontology) and see your knowledge graph grow.

Level of Automation:



START

### Manuscript Annotation Desk

Use this path to identify the right information in your data that need to be sent to the agent framework to convert your data into LOD using Manuscript-AI Ontology.

Level of Automation:



START

### Select and Convert

Use this method if you wish to automate the extraction of key values from specific data segments and convert them into LOD using the Manuscript-AI ontology.

Level of Automation:



START

### Drop & Classify

Use this method if you wish to just drop your data and let agents take care of classifying them under the Manuscript-AI ontology.

Level of Automation:



START

# Manuscript-AI Annotation Desk

Please fill the fields for the description of the manuscript. You can either copy and paste snippets or directly type into the form fields. You can describe up to 5 manuscripts at a time. **Note:** If you fill any fields for a given manuscript, remember to provide a shelfmark as *ID* (or simply an ID).

## Manuscript Data

Sample Catalog Entry (Manuscript Description)

Shelfmark/Call Number:

BnF Ms. lat.

12345

Title: Book of Hours (Use of Paris) Date: c. 1450-1460 Material (Support): Parchment Extent: 120 folios (ff. 1r-120v) Dimensions: Approx. 150 x 100 mm Script (Handwriting Form): Gothic textualis, well-formed with neat ruling Decoration:

Illuminated initials throughout

4 full-page miniatures with floral borders

Numerous smaller decorative initials in gold and ultramarine

Binding: Later (17th-century) calf over wooden boards, with simple tooling and remnants of clasps

Contents & Notable Features:

Text includes the Hours of the Virgin, Office of the Dead, and several devotional prayers typical of a Paris liturgical use.

Manuscript

1

Manuscript

2

Manuscript

3

Manuscript

4

Manuscript

5

Select a Field:

Manuscript ID

ADD SELECTED TEXT

Manuscript ID:

BnF Ms. lat. 12345

Century of creation:

1456

Support type:

Parchment

Handwriting form:

No data added yet

Decorations:

Illuminated initials throughout 4 full-page miniatures with floral borders

### Manuscript-AI Annotation Desk

Please fill the fields for the description of the manuscript. You can either copy and paste snippets or directly type into the form fields. You can describe up to 5 manuscripts at a time. **Note:** If you fill any fields for a given manuscript, remember to provide a shelfmark as ID (or simply an ID).

#### Manuscript Data

Sample Catalog Entry (Manuscript Description)

Shelfmark/Call Number:  
BnF Ms. lat. 12345

Title: Book of Hours (Use of Paris)Date: c. 1450-1460Material (Support): ParchmentExtent: 120 folios (ff. 1r-120v)Dimensions: Approx. 150 × 100 mmScript (Handwriting Form): Gothic textualis, well-formed with neat rulingDecoration:  
Illuminated initials throughout  
4 full-page miniatures with floral borders  
Numerous smaller decorative initials in gold and ultramarine  
Binding: Later (17th-century) calf over wooden boards, with simple tooling and remnants of clasps  
Contents & Notable Features:  
Text includes the Hours of the Virgin, Office of the Dead, and several devotional prayers typical of a Paris liturgical use.

Select a Field: Manuscript ID **ADD SELECTED TEXT**

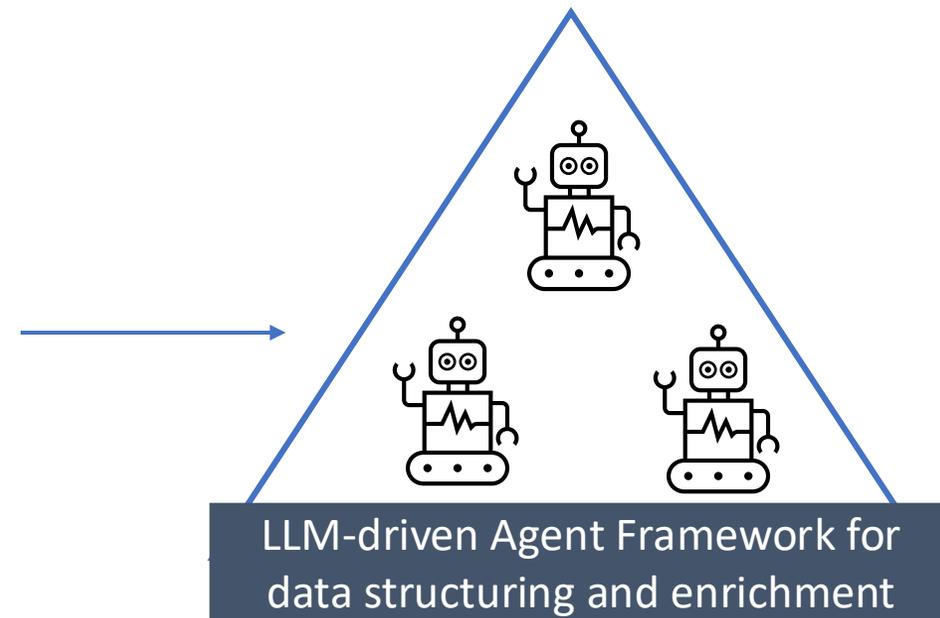
Manuscript ID:  
BnF Ms. lat. 12345

Century of creation:  
1456

Support type:  
Parchment

Handwriting form:  
No data added yet

Decorations:  
Illuminated initials throughout 4 full-page miniatures with floral borders



The data is sent to an **LLM-driven agent framework** that uses system messages and custom functions to **structure and model information under the Manuscript-AI ontology**. Additionally, the framework suggests links and URIs from Wikidata, enabling users to link their data to external knowledge bases.



12345

Prefixes

```

ex: http://example.org/
rdf: http://www.w3.org/1999/02/22-rdf-syntax-ns#
rdfs: http://www.w3.org/2000/01/rdf-schema#
xsd: http://www.w3.org/2001/XMLSchema#
ms4ai: http://ontology.tno.nl/manuscriptAI/

```

Properties

ms4ai:attributedDate

1456

[ms4ai:borderDesign](#)

ms4ai:hasDecoration

[ms4ai:illumination](#)

[ms4ai:miniature](#)

ms4ai:hasSupport

[ms4ai:parchment](#)

ms4ai:includesDecoration

Illuminated initials throughout 4 full-page miniatures with floral borders

ms4ai:shelfmark

BnF Ms. lat. 12345

ms4ai:support

Parchment

MARK AS REVIEWED

DOWNLOAD TURTLE

Downloads



manuscripts (54).ttl



[Open file](#)

knowledge\_graph (75).ttl

[Open file](#)

ns.rdf

[Open file](#)

See more





determine what works best for your needs. Each option is designed to help you structure your data in a machine-understandable format, making it ready to be seamlessly linked with other collections of data.

### Create LOD

Use this path if you want to structure data using Manuscript-AI Ontology from scratch (or explore the full expressiveness of the ontology) and see your knowledge graph grow.

Level of Automation:



START

### Manuscript Annotation Desk

Use this path to identify the right information in your data that need to be sent to the agent framework to convert your data into LOD using Manuscript-AI Ontology.

Level of Automation:



START

### Select and Convert

Use this method if you wish to automate the extraction of key values from specific data segments and convert them into LOD using the Manuscript-AI ontology.

Level of Automation:



START

### Drop & Classify

Use this method if you wish to just drop your data and let agents take care of classifying them under the Manuscript-AI ontology.

Level of Automation:



START

✓ Upload files

🔍 Inspect files

3 Review files

4 Map Concepts

```
1 v [
2 v {
3   "ms_id": "BNF-FR-001",
4   "doc_code": "BNF-FR-001",
5   "identifiant": "BNF-FR-001",
6   "date": "XIIIe siècle",
7   "era": "Moyen Âge",
8   "medium": "parchemin",
9   "dim_data": "25 x 35 x 78",
10  "works_list": "La Chanson de Roland; Les Lais de
11  "beginning_text": "Au commencement, le verbe fut
12  "ending_text": "Fin du récit.",
13  "handwriting_form": "Textualis gothique",
14  "script style": "Textualis gothique",
15  "handwriting_notes": "Écrit avec des initiales o
16  "script_notes": "Écrit avec des initiales ornées
17  "art_deco_kind": "miniatures; initiales décorées
18  "art_deco_description": "Décorations riches en m
19  "covering": "Cuir sur planches de bois",
20  "folio_range": "1r-150v",
21  "ink_type": "encre de galate",
22  "ink_used": "encre de galate",
23  "authors": "Auteur inconnu",
24  "title": "La Chanson de Roland"
```

0

## Manuscripts

Manuscript1

Manuscript2

## Manuscript Content

```
"art_deco_description": "Décorations riches en motifs
floraux",
"covering": "Cuir sur planches de bois",
"folio_range": "1r-150v",
"ink_type": "encre de galate",
"ink_used": "encre de galate"
```

STRUCTURE CONTENT

1 Upload files 2 Inspect files 3 Review files 4 Map Concepts

```
1 v [
2 v {
3   "ms_id": "BNF-FR-001",
4   "doc_code": "BNF-FR-001",
5   "identifiant": "BNF-FR-001",
6   "date": "XIIIe siècle",
7   "era": "Moyen Âge",
8   "medium": "parchemin",
9   "dim_data": "25 x 35 x 78",
10  "works_list": "La Chanson de Roland; Les Lais de
11  "beginning_text": "Au commencement, le verbe fut
12  "ending_text": "Fin du récit.",
13  "handwriting_form": "Textualis gothique",
14  "script style": "Textualis gothique",
15  "handwriting_notes": "Écrit avec des initiales o
16  "script_notes": "Écrit avec des initiales ornées
17  "art_deco_kind": "miniatures; initiales décorées
18  "art_deco_description": "Décorations riches en m
19  "covering": "Cuir sur planches de bois",
20  "folio_range": "1r-150v",
21  "ink_type": "encre de galate",
22  "ink_used": "encre de galate",
23  "authors": "Auteur inconnu",
24  "..."
25  }
```

0

### Manuscripts

Manuscript1 Manuscript2

### Manuscript Content

"art\_deco\_description": "Décorations riches en motifs floraux",

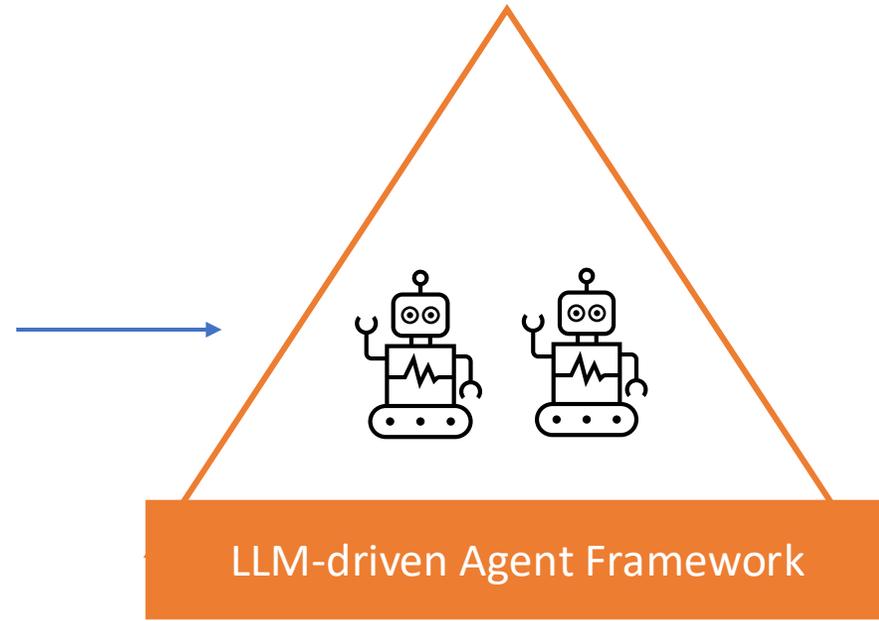
"covering": "Cuir sur planches de bois",

"folio\_range": "1r-150v",

"ink\_type": "encre de galate",

"ink\_used": "encre de galate"

STRUCTURE CONTENT



The user-chunked data is sent to specialized agents in the LLM-driven agent framework, which recognize the relevant concepts of interest

✓ Upload files

✓ Inspect files

✎ Review files

4 Map Concepts



File: input1.csv  
Manuscripts: undefined

#### EXTRACTED MANUSCRIPTS



BNF-FR-001



Manuscript11

CONVERT TO TURTLE

Actions:

MARK AS REVIEWED

COMPARE

#### Extracted properties

Manuscript id:

BNF-FR-001

EDIT

Century of creation:

XIIIe siècle

EDIT

Support type:

parchemin

EDIT

Dimensions of the manuscript width:

25 cm

EDIT

Upload files   Inspect files   Review files   Map Concepts

File: input1.csv  
Manuscripts: undefined

EXTRACTED MANUSCRIPTS

- BNF-FR-001
- Manuscript11

CONVERT TO TURTLE

Actions: MARK AS REVIEWED   COMPARE

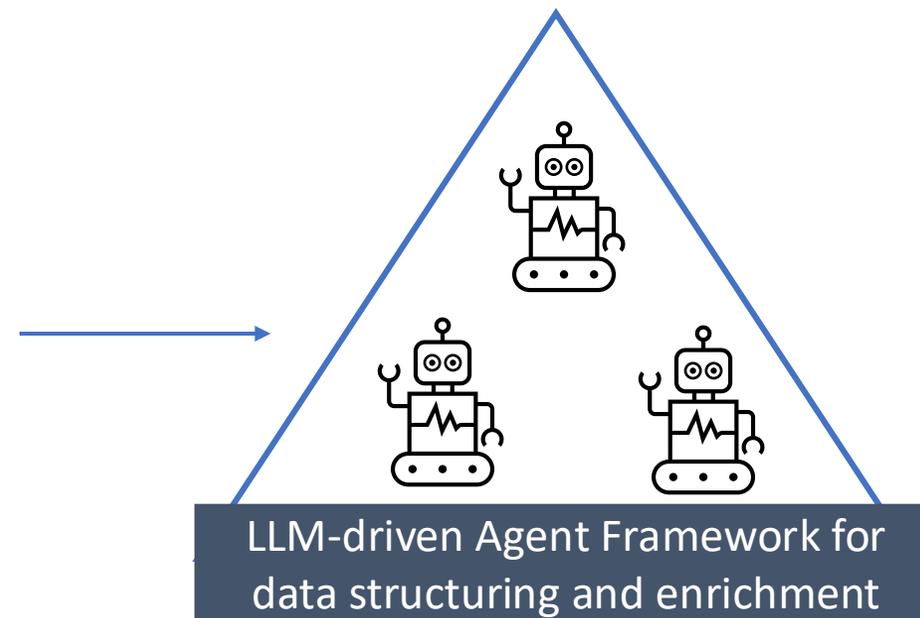
Extracted properties

Manuscript id:  
BNF-FR-001   EDIT

Century of creation:  
XIIIe siècle   EDIT

Support type:  
parchemin   EDIT

Dimensions of the manuscript width:  
25 cm   EDIT



The corrected data is sent to the LLM-driven agent framework, which structures it according to the manuscript-ai ontology and further enriches it by suggesting additional links to the Wikidata knowledge base.



<code>ms4ai:ink</code>	encre de galate
<code>ms4ai:script</code>	Textualis gothique
<code>ms4ai:shelfmark</code>	BNF-FR-001
<code>ms4ai:support</code>	parchment
<code>rdfs:comment</code>	Bon état de conservation.

Incipit/Explicit Loci

```
ex:LOC20250403_0009178469 rdf:type ms4ai:Locus ;  
ms4ai:concernsFeature ms4ai:incipit ;  
ms4ai:includesText
```

Au commencement, le verbe fut...

```
ex:LOC20250403_0009171769 rdf:type ms4ai:Locus ;  
ms4ai:concernsFeature ms4ai:explicit ;  
ms4ai:includesText
```

Fin du récit.



determine what works best for your needs. Each option is designed to help you structure your data in a machine-understandable format, making it ready to be seamlessly linked with other collections of data.

### Create LOD

Use this path if you want to structure data using Manuscript-AI Ontology from scratch (or explore the full expressiveness of the ontology) and see your knowledge graph grow.

Level of Automation:



START

### Manuscript Annotation Desk

Use this path to identify the right information in your data that need to be sent to the agent framework to convert your data into LOD using Manuscript-AI Ontology.

Level of Automation:



START

### Select and Convert

Use this method if you wish to automate the extraction of key values from specific data segments and convert them into LOD using the Manuscript-AI ontology.

Level of Automation:



START

### Drop & Classify

Use this method if you wish to just drop your data and let agents take care of classifying them under the Manuscript-AI ontology.

Level of Automation:



START

## Manuscripts

BNF-FR-001

CONFIRM REVIEW

### Manuscript ID:

BNF-FR-001

#### Raw Data (data\_analyzed)

```
ms_id,doc_code,identifrier,date,era,medium,dim_data,works_list,beginning_text,ending_text,handwriting_form,  
BNF-FR-001,BNF-FR-001,BNF-FR-001,XIIIe siècle,Moyen Âge,parchemin,25 x 35 x 78,La Chanson de Roland; Les L  
floraux.,Données analysées par l'outil.,Données analysées par l'outil.,Aucune donnée non identifiée,Aucune
```

#### Century of creation:

13th century

#### Support type:

parchment

#### Handwriting form:

Textualis gothique



Manuscripts

BNF-FR-001

## RDF Triples

### Prefixes

```
ex: http://example.org/  
rdf: http://www.w3.org/1999/02/22-rdf-syntax-ns#  
rdfs: http://www.w3.org/2000/01/rdf-schema#  
xsd: http://www.w3.org/2001/XMLSchema#  
ms4ai: http://ontology.tno.nl/manuscriptAI/
```

### Properties

<code>ms4ai:attributedAuthor</code>	Auteur inconnu
<code>ms4ai:attributedCopyist</code>	Scribe anonyme
<code>ms4ai:attributedDate</code>	13th century
<code>ms4ai:attributedMiniaturist</code>	Artiste inconnu
<code>ms4ai:binding</code>	Cuir sur planches de bois
<code>ms4ai:conservationIntervention</code>	Restauration en 1995

### **Create LOD path: Data Insertion & Real-Time Growth:**

- The user manually inputs initial manuscript data.
- As each new value is added, the knowledge graph (KG) immediately expands, allowing the user to see real-time changes.

### **Manuscript-AI Annotation Desk path**

- The user works within the Manuscript-AI annotation desk to assign specific values (e.g., attributed author, century).
- The agent framework converts this annotated data into RDF, aligned with the Manuscript-AI ontology, and enriches it with links from Wikidata.

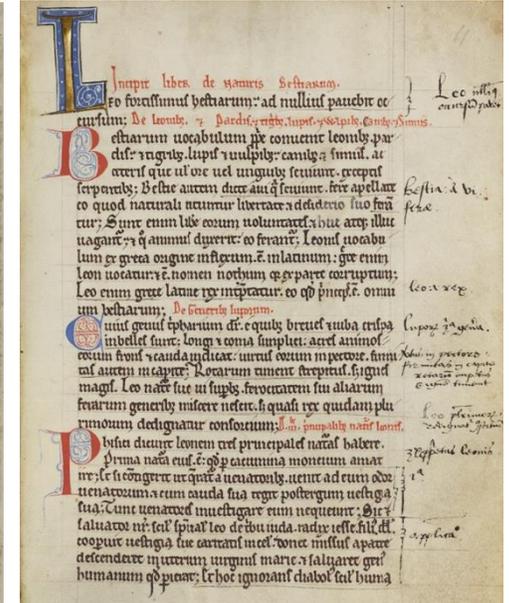
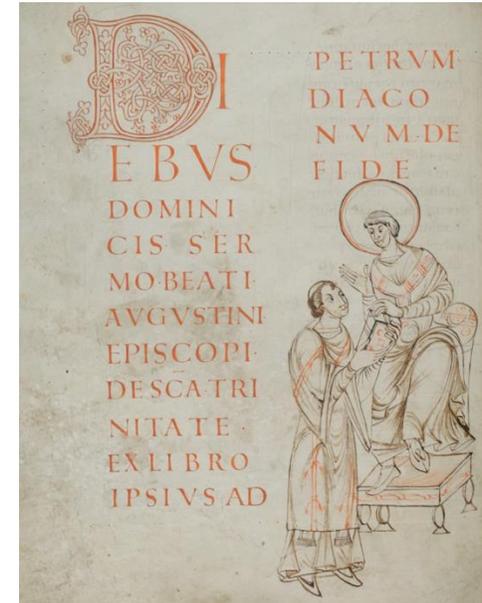
### **Select and Convert Path**

- The user selects only a portion of the data (chunk) for a specific manuscript.
- The agent framework automatically identifies and proposes relevant values for key categories.
- The user reviews and corrects these suggestions if needed.
- The refined data is then converted into RDF (Manuscript-AI ontology) and enriched with external references.

### **Drop and Classify Path: Automated Chunking & Enrichment**

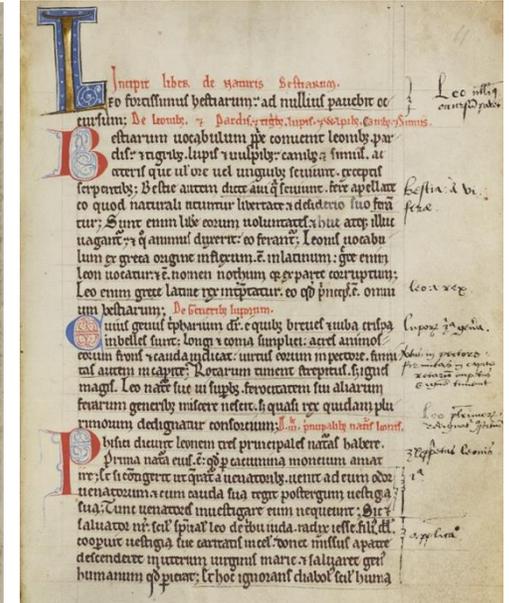
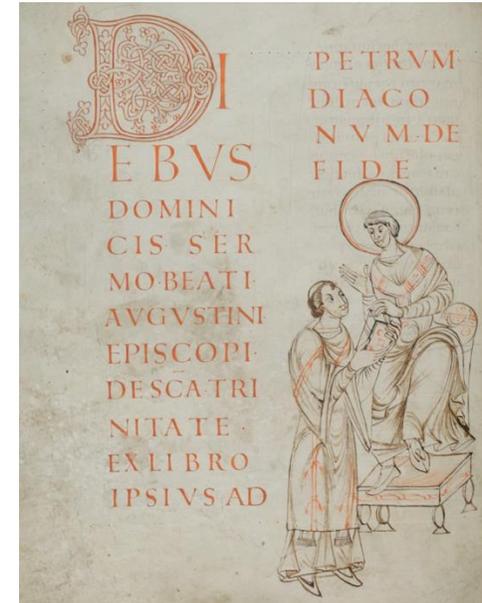
- The chunking of data happen automatically.
- The user's role is primarily to validate the final results
- The agent framework structures the data, applies the Manuscript-AI ontology, and provides enrichment links, with only final confirmation or adjustments required from the user.

- The ManuscriptAI project
- The need for an ontology for manuscripts
- Presentation of the ontology
- Some specific features
- The ontology as part of our AI-framework
- The sustainability challenge



# Challenges and the way forward:

- Flexibility
- Community
- Implementation
- Authority





- Radboud Institute for culture and history ([shari.boodts@ru.nl](mailto:shari.boodts@ru.nl))
- Netherlands Organisation for Applied Scientific Research ([giulia.biagioni@tno.nl](mailto:giulia.biagioni@tno.nl))

**TNO** innovation  
for life

KB } nationale  
bibliotheek



ai·ch

Biblissima



Radboud Universiteit



UNIMORE  
UNIVERSITÀ DEGLI STUDI DI  
MODENA E REGGIO EMILIA



