

Métadonnées et ontologies pour l'édition critique numérique de textes littéraires

Colloque Nouvelles perspectives sur les éditions critiques

Montréal

30 septembre – 1^{er} octobre 2024

Ioana GALLERON
Sorbonne Nouvelle – Consortium ARIANE

L'urgence des métadonnées

- Augmentation du nombre d'éditions disponibles en ligne et volatilité des sites web
- « Jungle » des métadonnées (standards: EAD, DC, TEI, OLAC e. a.), irrégularité des pratiques (v. Idmhand et Galleron, Montréal, 2022)
- Non-engagement des bibliothèques dans un travail de recensement et mise à disposition
- Absence de véritables éditeurs numériques
- Publication dans des entrepôts de données (Zenodo, Nakala) aux moteurs de recherche plus ou moins performants

Les 1001 formes du teiHeader

	fileDesc	encodingDesc	profileDesc	revision Desc
Projet 1	<p>titleStmt: titre « transcription de », auteur</p> <p>respStmt: 1 responsable</p> <p>editionStmt: outil employé</p> <p>publicationStmt: labo</p> <p>sourceDesc: bibl (auteur+titre)</p>	–	–	–
Projet 2	<p>titleStmt: titre ouvrage</p> <p>respStmt: 5 responsables</p> <p>editionStmt: date+MàJ+resp</p> <p>publicationStmt: labo, licence, note sur l'édition</p> <p>sourceDesc: msDesc, listBibl</p>	<p>langUsage</p> <p>projectDesc</p> <p>editorialDecl</p>	<p>keywords: littérature</p> <p>catRef: description par rapport à un schéma interne</p>	oui
Projet 3	<p>titleStmt: nom fichier .jpg</p> <p>respStmt: 1 responsable</p> <p>editionStmt: [vide]</p> <p>publicationStmt: [vide]</p> <p>sourceDesc: titre, date, lieu de rédaction, aspects matériels</p>	–	<p>Domaine: « journaux et papiers »</p> <p>Corpus: « journaux »</p> <p>Nature: « pages de journal »</p>	–

Les 1001 formes du teiHeader

	fileDesc	encodingDesc	profileDesc	revision Desc
Projet 4	titleStmt: titre + auteur editor (pas respStmt) editionStmt: ingénieur recherche publicationStmt: labo + licence sourceDesc: [vide]	projectDesc	correspDesc keywords: correspondance, littérature	oui
Projet 5	titleStmt: titre+auteur respStmt: 1 responsable editionStmt: [vide] publicationStmt: [vide] sourceDesc: bibl	–	keywords: actualités politiques, antiquités, botaniques, livres + liste entités nommées (personnes)	–
Projet 6	titleStmt: title, author, editor, funder editionStmt: resp publicationStmt: institution, licence, comment citer sourceDesc: biblFull	projectDesc editorialDecl tagsDecl	creation textDesc textClass > keywords « domaine », « genre », « forme », « relation »	oui



AMIS

Advanced Metadata Intelligent System

- Cible: communauté de la science des textes (lettres, histoire, linguistique...)
- Développement d'une application web avec trois objectifs principaux
 1. faciliter la création de métadonnées (assistance à la conception)
 2. offrir des recommandations personnalisées
 3. améliorer la découvrabilité



Fonctionnement d'AMIS

Entrée

Analyse

Sortie

Fichier de métadonnées à
structure minimale

OU

Fichier comportant des
données et des métadonnées
à structure minimale

OU

Document texte sans
métadonnées

- « Lecture » du fichier
- Identification de caractéristiques (traitement TAL + base de connaissances)
- Recherche de textes/ documents similaires dans différents entrepôts de données

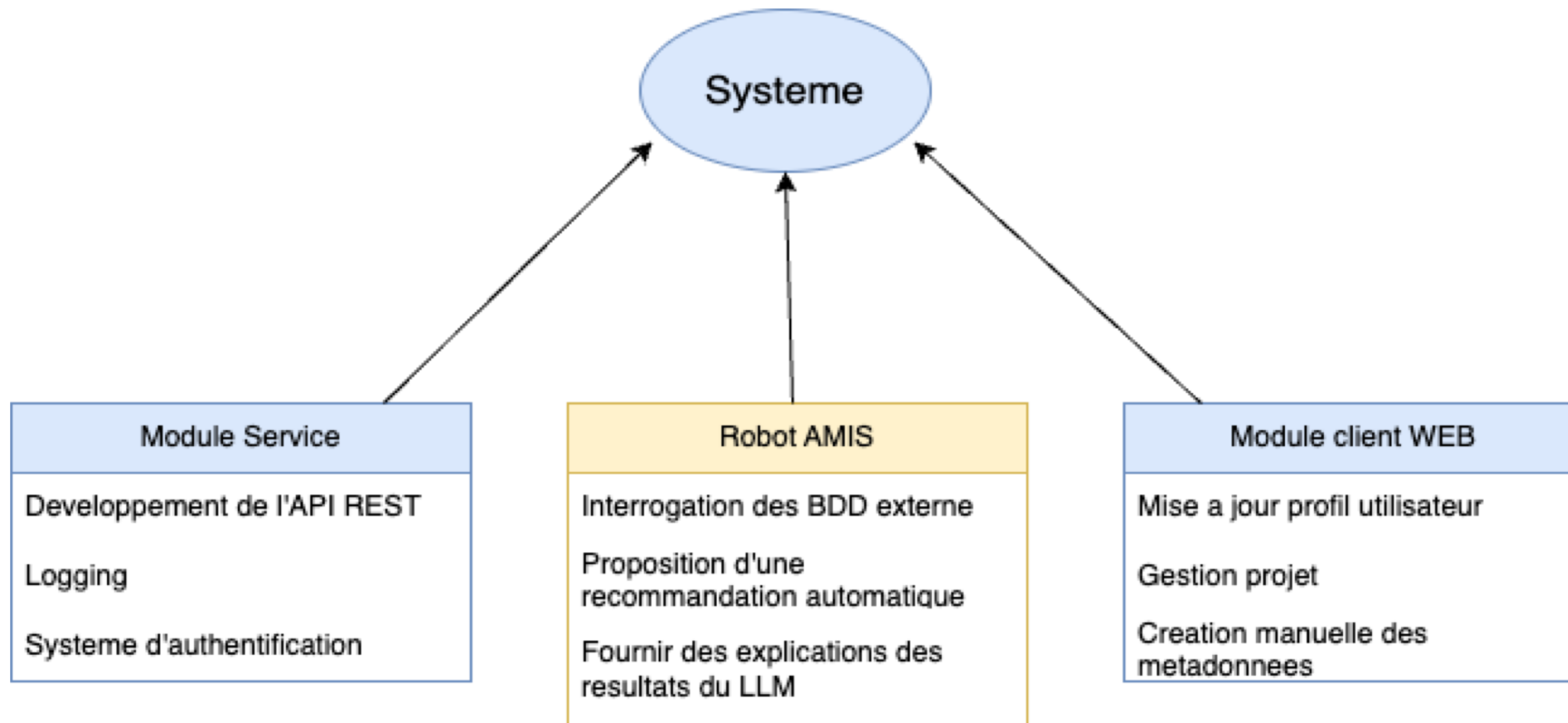
- Suggestion de métadonnées
- Suggestion de vocabulaires contrôles existants
- Choix/ rejet par l'utilisateur
- Production de:
 - ✓ un fichier xml/TEI ou csv avec les métadonnées enrichies
 - ✓ un rapport de travail



OSCAR

Open Science Clusters' Action
for Research & Society

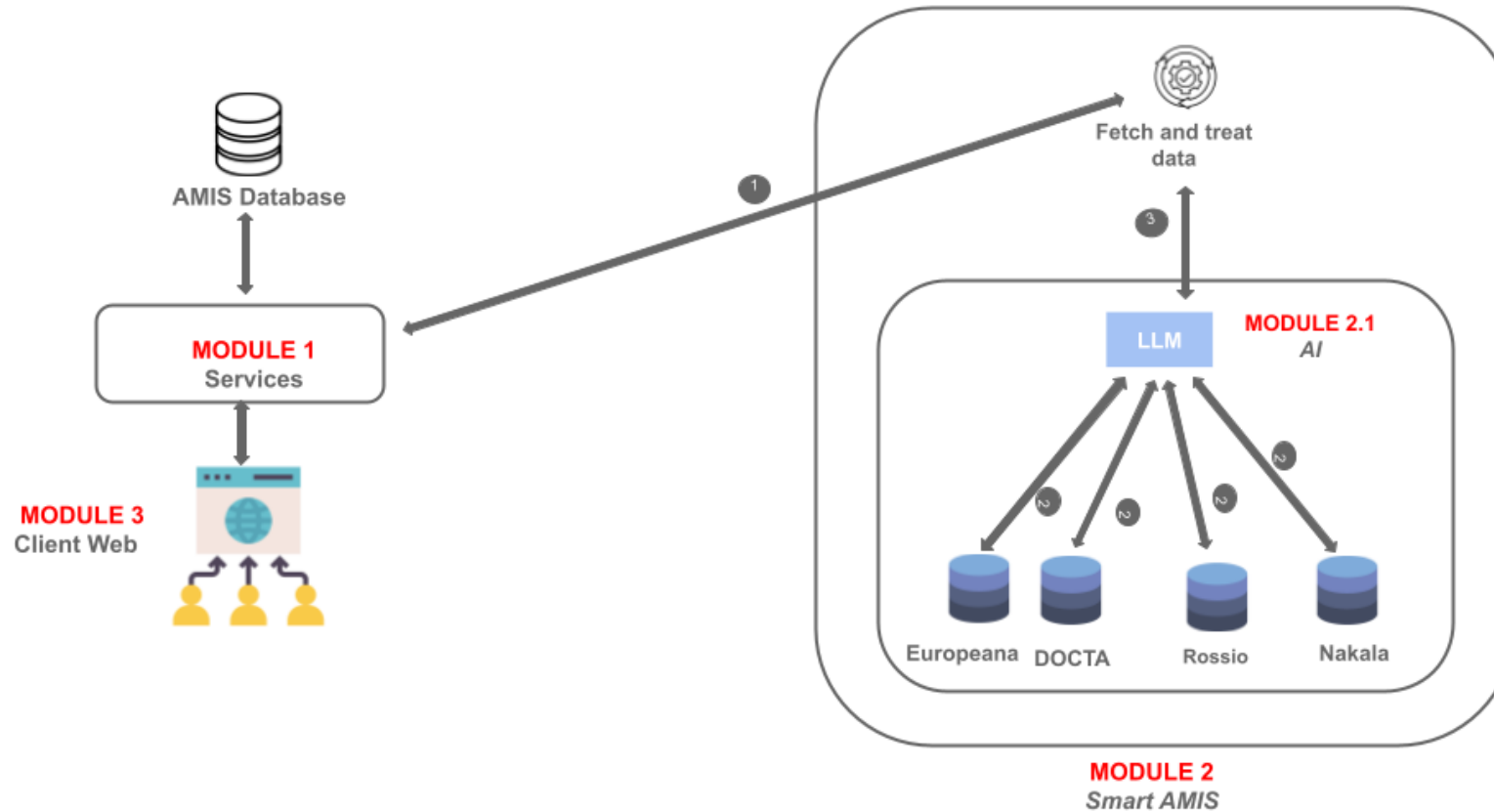
Modules AMIS





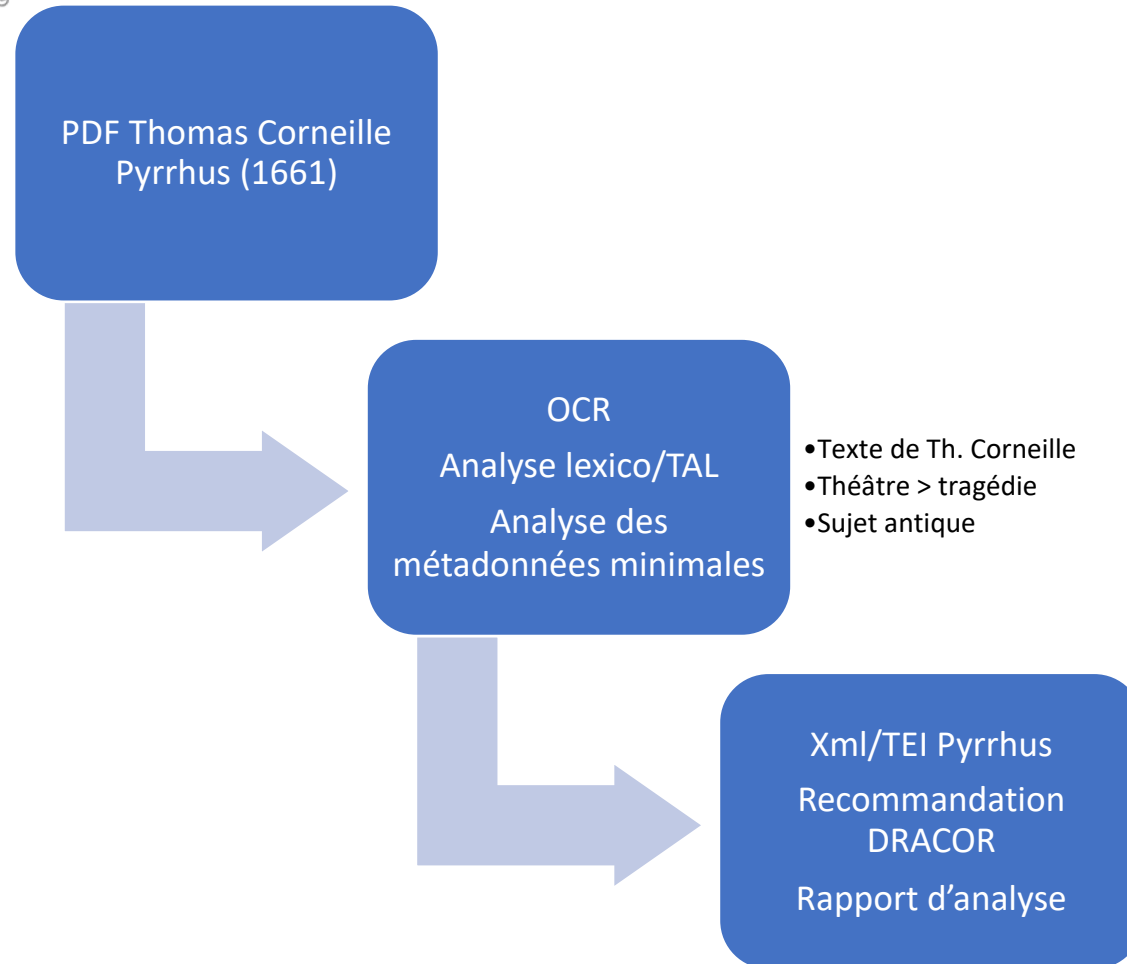
OSCAR
Open Science Clusters' Action
for Research & Society

Architecture générale d'AMIS





Un cas d'usage d'AMIS



Exemples de recommandations

fileDesc	encodingDesc	profileDesc	revisionDesc
<p>titleStmt: titre « transcription de », author respStmt: 1 responsable editionStmt: outil employé publicationStmt: labo sourceDesc: bibl (auteur+titre)</p>	–	–	–
Post-analyse AMIS			
<p>titleStmt: « XXX. Édition numérique » Author: OK respStmt: détailler les rôles</p> <ul style="list-style-type: none"> - Transcription - Correction de la transcription - Encodage - Correction de l'encodage <p>Etc. publicationStmt: suggestion de correction sourceDesc: exemples d'usage de biblFull</p>	Exemples de description des choix d'encodage	Renvoi à « Typologies textuelles » Textes du même genre Textes à proximité lexicale/ thématique Textes à proximité « sentiments »	Exemples de déclarations de révision

Ontologie(s) pour AMIS?



Chowlk (Polytechnica Madrid)
EcoPortal
GLODAP
IVOA Validator
LCT
LOV
OntoPortal
PaN Ontologies
Etc.

Resources

21 resources found.

Search...

Scientific domain

- Agricultural Sciences
- Engineering & Technology
- Humanities
- Medical & Health Sciences
- Natural Sciences
- Social Sciences
- Domain agnostic

AgroPortal

Metadata & ontologies Agricultural Sciences

A semantic artefact catalogue for agri-food

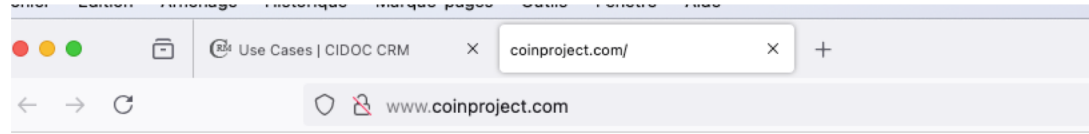


Organisations:

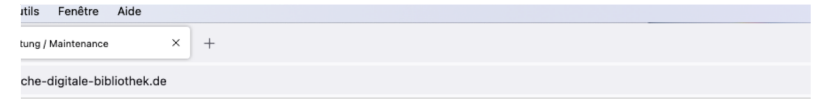
Agroportal

AgroPortal is an ontology repository (more generally a semantic artefact catalogue) for the agri-food domain which features ontology hosting, search, versioning, visualization, comment, and recommendation; enables semantic annotation; stores and exploits ontology alignments; and enables interoperation with the semantic web.

Ontologies pour AMIS?

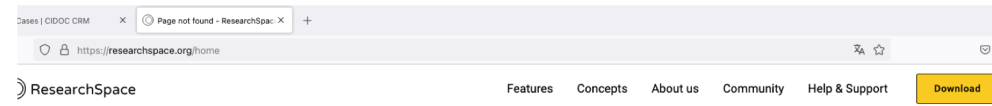


Warning: mysql_pconnect(): Too many connections in /home/coin/public_html/includes/config.php on line 14
Could not connect MySQL Database



**Das System steht wegen Wartungsarbeiten vorübergehend nicht zur Verfügung.
Wir bitten um Entschuldigung für die Unannehmlichkeiten.**

**Temporarily, the system is not available because of maintenance work.
We are sorry for the inconvenience.**

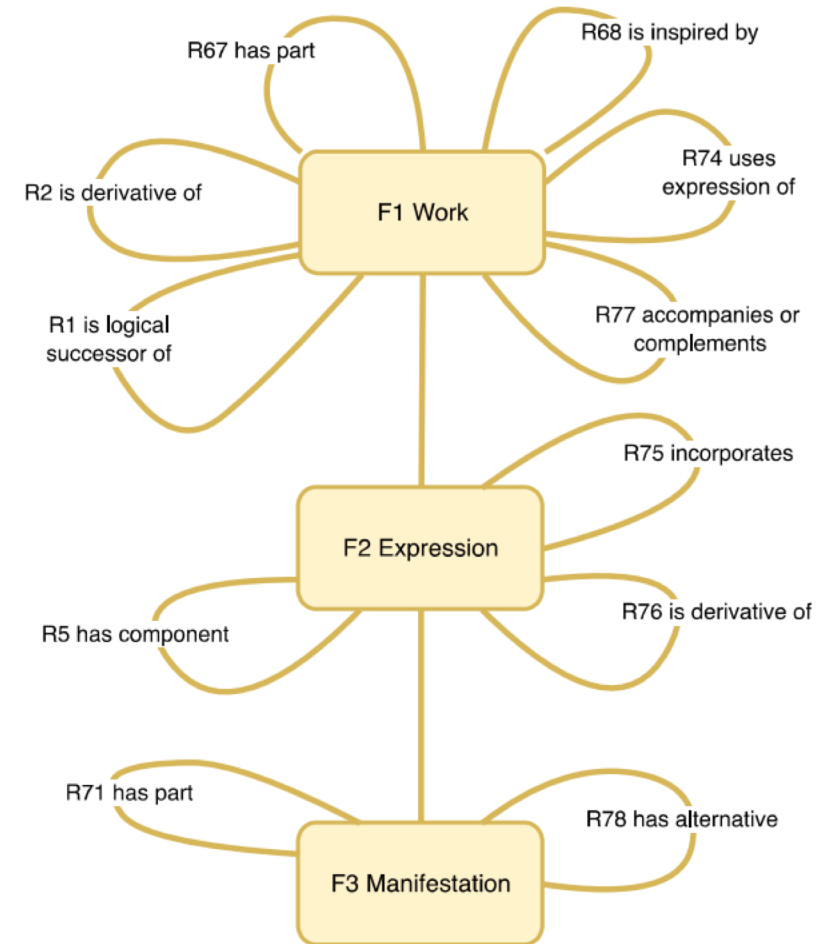


Ops, looks like you are lost...

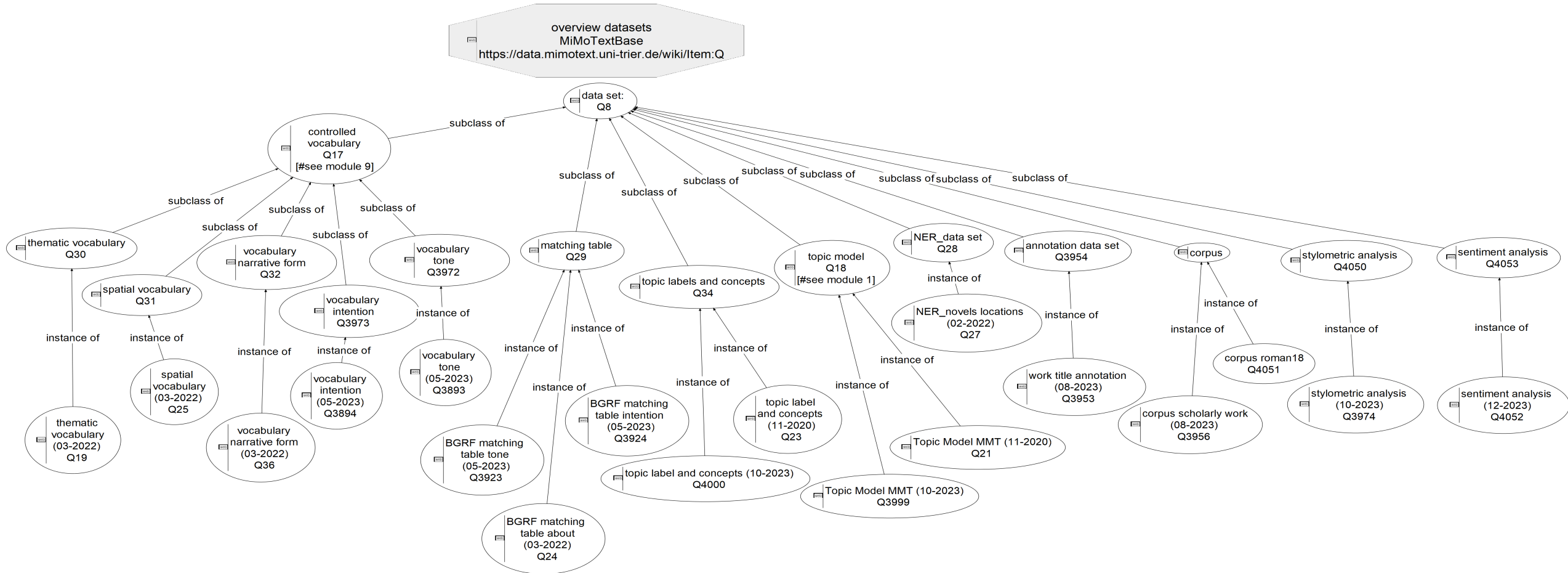
[Back to Homepage ↩](#)

CIDOC-CRM pour AMIS?

- Ce n'est pas la première extension souhaitée de CIDOC CRM, alignée avec les études littéraires: voir FRBRoo > LRMoo (dernière version 0.9 2023, mais erreur 500 sur le site de l'IFLA)
- Où sont « Citation » and « Citation_Scheme » proposés par Romanello et Pasin en 2010?
- Pas de structure de coordination des études littéraires similaire au Comité international pour la documentation (CIDOC) du Conseil international des musées (ICON)



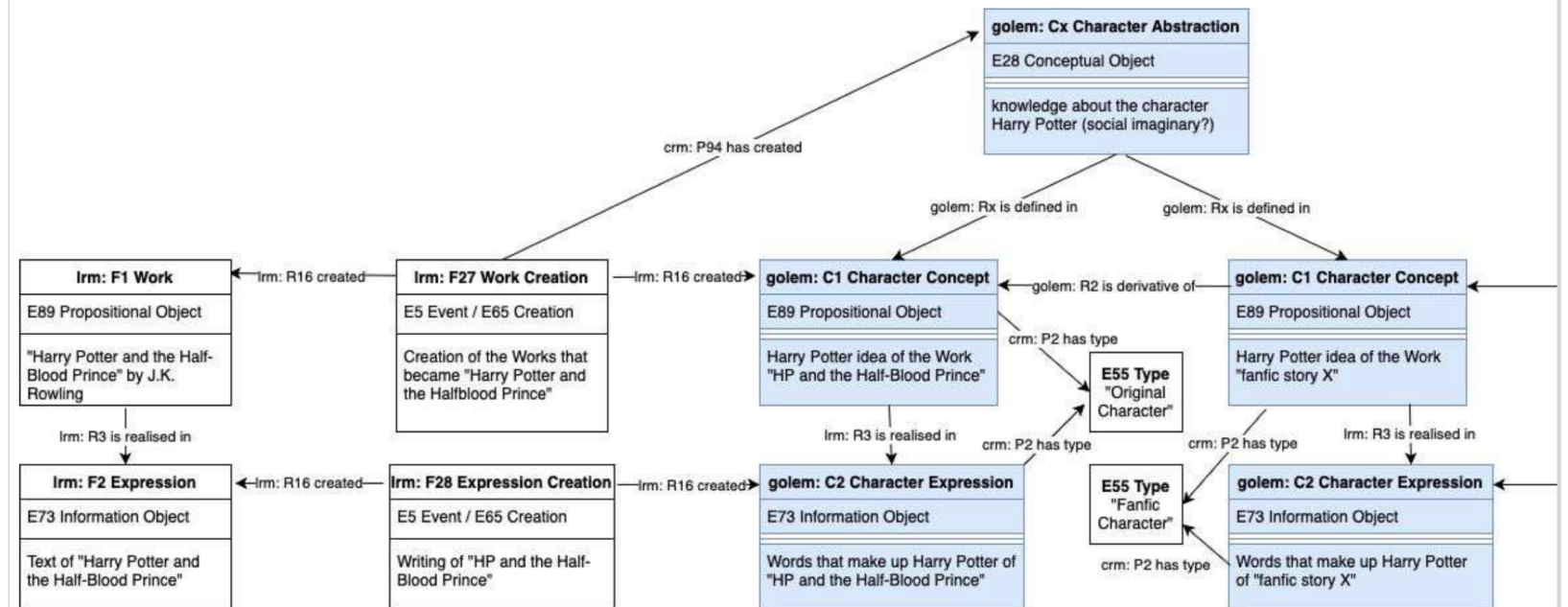
MiMoTex pour AMIS?



GOLEM pour AMIS?

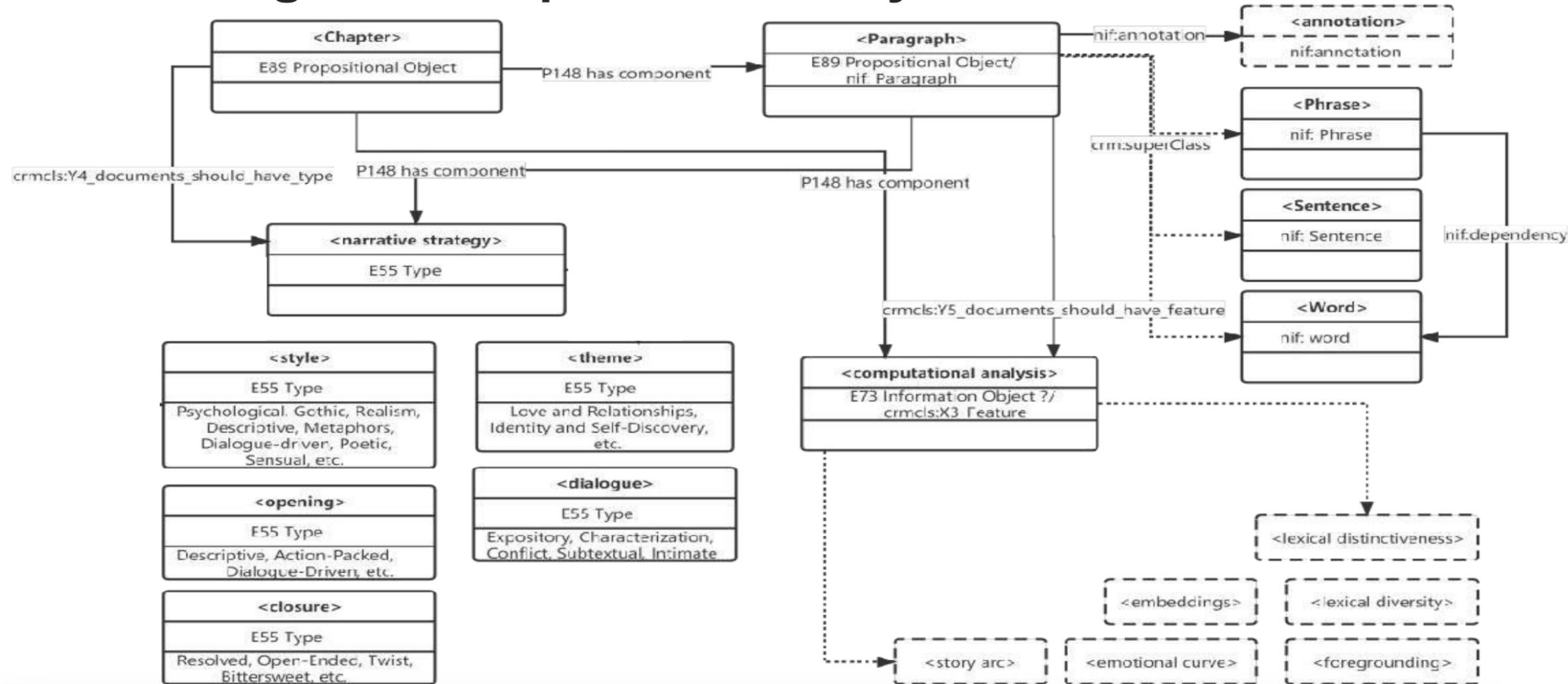
Pianzola, Federico et
Xiaoyan Yang
« First steps of the GOLEM
Ontology for narrative and fiction »,
*Workshop Ontologies for Narrative
and Fiction*, 3-4 July 2023

Fictional Characters & Fanfiction

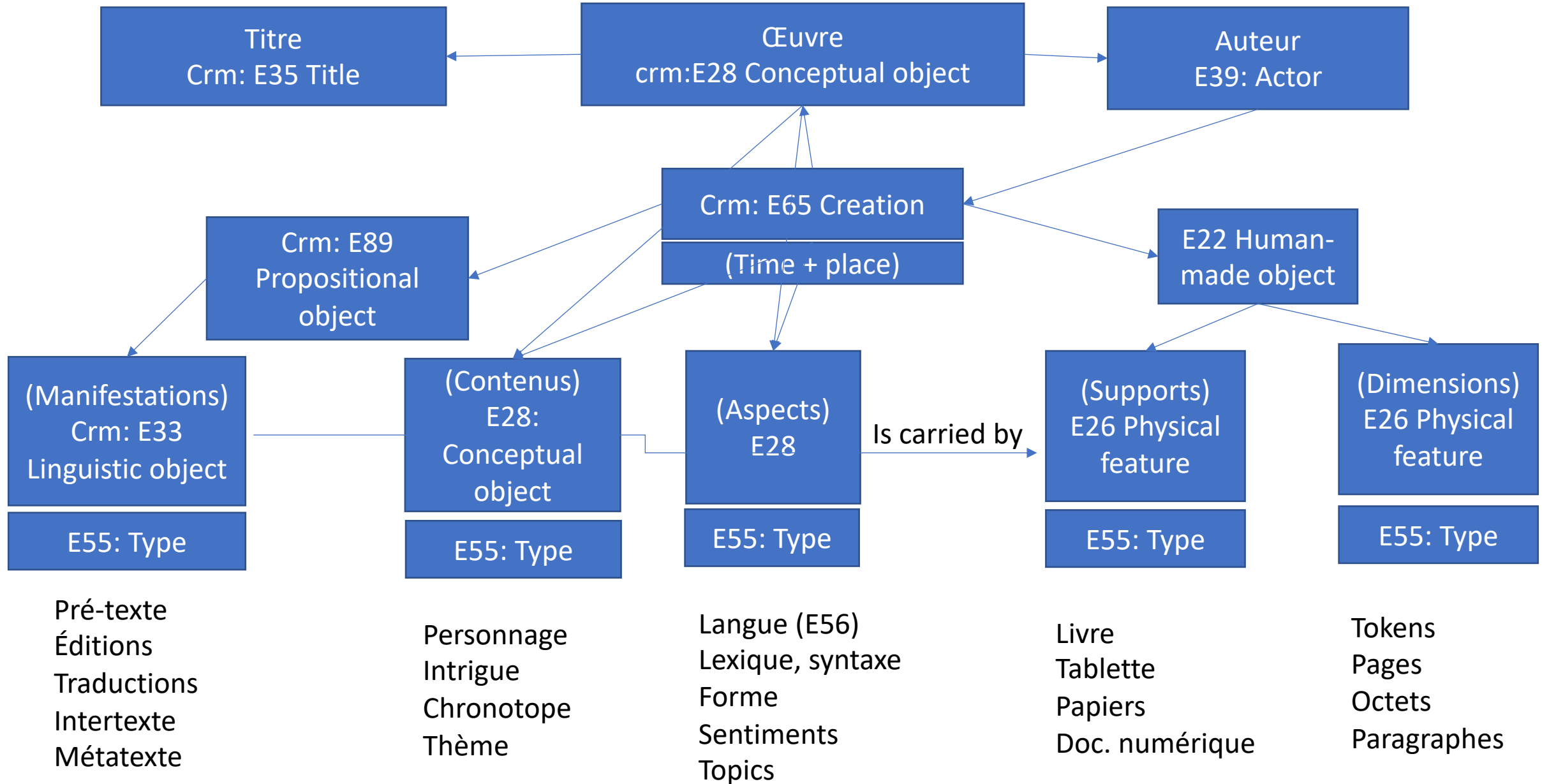


GOLEM pour AMIS

Narrative Strategies & Computational Analysis



Une proposition de schéma



MERCI POUR VOTRE ATTENTION!