CLARA

Creating and linking
Licensable educational resources

https://project.inria.fr/clara/

GDD
TALN
DRUID
DuKe
SemLIS

IRISA

Abstract

CLARA project aims to empower teachers to facilitate the creation of licensable educational resources (ER) based on existing ones. Our approach will suggest a relevant set of ERs that are coherent with a course sketch and have compatible licenses.

The main challenges we face are (1) how to enrich a network of ERs using AI algorithms, and (2) how to guarantee a minimal set of license-compatible ERs relevant to a given course goal with query relaxation techniques. We will exploit ERs provided by the French Ministry of Education and the X5-GON* project.

Goals

- **WP1**. To build a corpus of linked ER with a rich semantic representation (available metadata, licenses, similarities, dependencies, etc.).
- **WP2**. To enrich the semantic annotations with machine learning algorithms (concepts, prerequisites, temporality of concepts, etc.).
- **WP3**. To design a query engine that facilitates finding relevant license-compatible educational resources.
- **WP4**. To involve teachers to test our solutions all along the project.

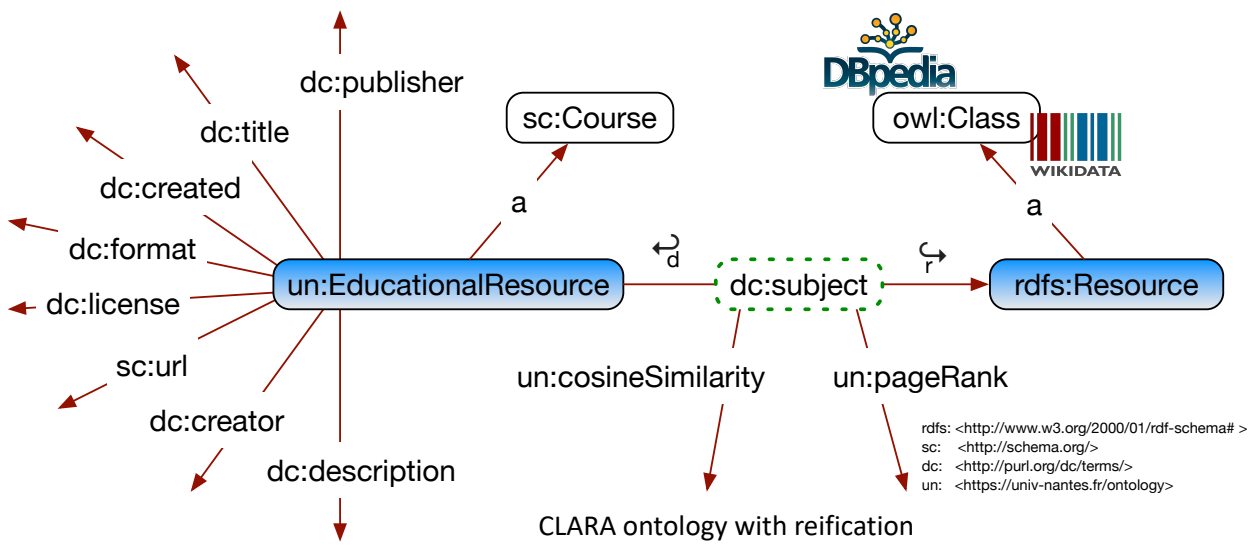
Ongoing planning

Workpackage	Partners	Year 1				Year 2				Year 3			
	LS2N IRISA DCSA GTNum	T3	T6	T9	T12	T15	T18	T21	T24	T27	T30	T33	T36
WP0	Project management												(coordination LS2N)
WP1	Building a collection of OER												(coordination LS2N)
Task 1.1 (OER corpus)													
Task 1.2 (RDF graph)													
Task 1.3 (License analysis)													
WP2	Enriching the graph of OER												(coordination IRISA)
Task 2.1 (State-of-the-art analysis)													
Task 2.2 (Representation learning)													
Task 2.3 (Identification of learning paths)													
WP3	OERs integration for a target cours												(coordination LS2N)
Task 3.1 (Sketch definition)													
Task 3.2 (Ranking results)													
Task 3.3 (Relaxing goals)													
WP4	Use case testing												(coordination LS2N)
Task 4.1 (Repository testing)													
Task 4.2 (Intermediate prototype testing)													
Task 4.3 (Final prototype testing)													

We are here

Building a knowledge graphe of ER (WP1)

- Current knowledge graph describes 43 619 ERs.
- ERs are lectures (pdf, video, audio, etc.)
- Each ER is related to a set of topics with a *Wikifier* that matches concepts with wikipedia pages.
- For each topic a relevance is calculated per ER : pageRang and cosineSimilarity.
- We use statement-level annotations (reification) to precise that an ER talks about a topic to a certain extent.
- Reified triples represent the topic's relevance by ER.
- Currently we have 135 814 topics (dc:subject).
- In average there are 188 topics by ER.



CLARA ontology with reification

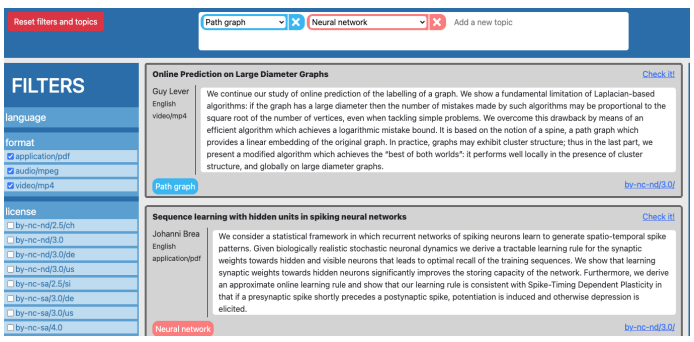
Identifying learning paths (WP2)

Many educational resources are available for public use. However, combining these resources in order to create **coherent sequences**, also called **learning paths**, is a challenging task. Furthermore, identifying other resources to complete or complement existing courses can be a useful task, yet it might require a considerable effort.

Goal and target approach (WP2)

The aim of the thesis is to assist and support educators in forming learning paths using existing resources and adding relevant resources to existing learning paths, thus designing curricula.

A preliminary interface (WP4)



Querying linked ERs (WP3)

- If the user query **Q** does not give enough interesting ERs, how to expand the scope of **Q** to guarantee a **k** number of ERs?
- **Q** can be expanded in a huge number of queries, which one could be the most interesting or close to the user query? How to take into account reification?
- Once a **k** number of ERs is found, how to rank them from the most to least interesting vis-a-vis of **Q**?

Goal and target approach (WP3)

Defining a ***query processing strategy*** that guarantees:

- A minimal relevant set of ERs.
- Compatibility of licenses of ERs.
- Quality of experience for the user.

Target approach:

- Mixing of **Ontology-based relaxation** and **topic's similarity** techniques to expand the scope of **Q**.
- **Similarity strategy** for ordering the potentially huge number of rewritten queries.
- **Ranking strategy** for the returned set of ERs.

Current and future work (ALL)

- State-of-the-art about query relaxation approaches in RDF.
- State-of-the-art about representation learning for knowledge graphs.
- User test of the preliminary interface.
- Collection of new educational resources.

*X5GON European project https://wp3.x5gon.org/