CIDOC CRM Educational Materials

Scope: determine aspects of semantic data education needed, relevant kinds of tools for use, and, consequently, educational materials required to support a learner in learning, appropriating and applying the CIDOC CRM.

Definition: a person shall be considered to have learned to use CIDOC CRM in a particular implementation scenario just in case they are able to answer novel questions in an objectively verifiable manner

Premise I: individuals learning the CIDOC CRM for application are concerned with appropriating a sufficient knowledge of the standard or its use in implementation environments in order to be able to do something with data in a semantic format (model it, map it, transform it, build a system to store / manage it, query it, visualize it)

Premise II: the appropriation of CIDOC CRM will come at different levels and involve different involvements in technical tasks. Part of learning CIDOC CRM for any actual use application is to

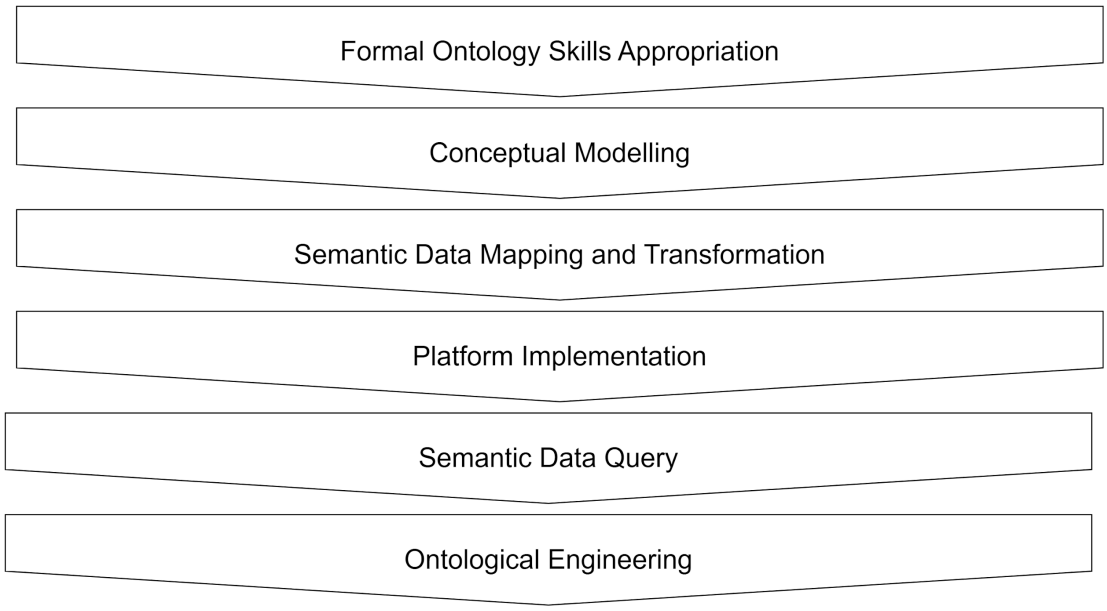
acquire a number of different skillsets which must be connected conceptually by the learner in order to arrive at a functional outcome.

Suggested Initial Approach:

Since using the CIDOC CRM for any practical purpose requires putting together a number of different skillsets, which may be mastered by one or more individual in a team and as part of an overall semantic data management workflow, a modular set of learning modules / materials would be best suited for supporting the overall learning of an individual. These modules / materials should be separate but have a consistent interrelation, allowing a learner to move from one module to another and apply skills and knowledge learned in one to their use in the other.

The fact that working with semantic data requires several distinct steps, methodologies and tools, suggests a useful way to segment the learning materials would be in relation to the different tasks that an individual needs to perform relative to appropriating and applying a knowledge of CIDOC CRM in a particular context.

An idealized representation of the path to knowledge looks somewhat like this:



Formal Ontology Skills Appropriation

Definition

At this level we are concerned with the question of understanding formal ontology itself, what it is, what it is for, of what it consists, what are its major concepts, what are the common ways it is implemented and with particular ontological models (i.e.: CIDOC CRM and extensions) and what they offer as a language for expressing data consistently and explicitly

Basics of Ontology

Definition

Understand what an ontology is, formal language, shared conceptualization, basics of classes, properties, instanation, IsA hierarchies etc., multiple IsA, multiple instantiation etc.

Existing Materials:

- Parts of Many Slide Decks

Relevant Tools/Software:

- Draw.io
- OntoMatch Game

Basics of RDF

Definition

Understand what RDF is, class and property, hierarchy declaration etc., application of an ontology into RDF

Existing Materials

Relevant Tools/Software

- Protege

Ontological Model Learning

Definition

Understand a particular model, model documentation, basic modelling strategy, higher level classes, patterns for different kinds of situations, limits of the model in question

Standards

- CRM Base
- CRM Extensions

Existing Materials

- CIDOC CRM Patterns List

Relevant Tools/Software

- OntoMe
- OntoMatch Game
- Draw.io + CHIN CRM data Model libraries
- CIDOC CRM Elements

Conceptual Modelling

Definition

Here we are concerned with the concrete application of a model towards the goal of creating consistent, integrated data following a standard. The goal is to appropriate good practice in applying an ontology in order to achieve a standard result with compatible data. Learners appropriate data modelling constructs (from supported standards) for situations and are able to apply them to relevant situations.

Data Modelling as Such (in conceptual world)

To add before the other modules, the bridge
The process of analysis etc.
Methodology of data modelling in the semantic world

c.e.s.ore@iln.uio.no

Basics of Modelling

Definition

Understand what a standard modelling approach is, ideas of patterns in general, definition of fields, collections, models

Standards

- Linked.Art
- SARI SRDMs

Existing Materials

- Linked.Art Patterns
- SARI Patterns

Relevant Tools/Software

- Zellij

Pattern Learning

Definition

Appropriation of selected modelling approach, learning specific models, collections and particular patterns

Standards

- Linked.Art
- SARI SRDMs

Existing Materials

- Linked.Art Patterns
- SARI Patterns

Relevant Tools/Software

- Zellij

Semantic Data Mapping and Transformation

Definition

Having understood an ontology and adopted a conceptual modelling framework, learners are interested in understanding how to convert existing and legacy data into the new representation. This is the process of semantic data mapping.

Basics of Mapping

Definition

Understand what aligning a dataset to an ontology (+ conceptual model) entails, strategies for approaching data alignment, particular problems, reconciliation

Existing Materials

Video by GB in Parthenos

Relevant Tools/Software

Principles of URI construction

Definition

Understand what a URI is, understand principles for building them

Existing Materials

Relevant Tools/Software

Mapping Tool Appropriation

Definition

Appropriation of use of a particular mapping tool to create and run a set of mapping instructions

Existing Materials

- 3M Manual
- GH Course on Karma
- GH Course on OntoRefine https://drive.google.com/open?id=1mlHc3Uhu8qFKhB2OyyUK4Y4bXaMSZFgo&authuser=gerald.hiebel%40uibk.ac.at&usp=drive_fs
-

Relevant Tools/Software

- 3M Software
- Karma
- OntoRefine
- Ontop (for relational DB)

Selecting and Implementing Semantic Information Systems

Definition

Having learned about ontology, an ontological model, a conceptual modelling framework and (potentially) how to transform data into semantic format, users want to manage and interact with semantic data using a practical data management platform.

I am leaving this section rather vague at this point since the variables may change significantly by platform and it is certainly not clear that there would be interest and ability to create a standardized documentation. Each platform has its own extant documentation. It may be sufficient to point to platforms and their documentation.

Classifying the platform types

- Triplestore / QuadStore / Graph Database
    - Apache jena
    - Virtuoso
- Semantic Data Management Platform
    - Arches
    - ResearchSpace
    - Metaphactory
    - WissKi
    - Geovistory
- Research Databases / Infrastructure System
    - Maintained database with set schema curating some sort of data [e.g. SeaLit]
- Aggregation Systems
    - Europeana
    - German Digital Library
    - Ariadne+

Platforms:

- Arches
- ResearchSpace
- Sparnatural

- Metaphactory
- WissKi
- Geovistory
- AQub

Installation

Definition

Model Setup

Thesaurus Setup

Data Loading

Querying

Definition

Learners having created semantic data want to query it a full semantic environment. We assume this to be RDF for the moment.

Basics of Data Query

Definition

Appropriate how queries are structured, different kinds of queries, principles of Sparql

Existing Materials

- Census Project Teaching Material [Takin]

Relevant Tools/Software

- Apache Jena

Query Pattern Learning

Definition

When following principles conceptual modelling principles, many queries are already available as canned queries to assess data quality etc. Learners can understand how to find and apply a query and understand its results.

Existing Materials

- Census Project Teaching Material [Takin]

Relevant Tools/Software

- Apache Jena
- Zellij

Ontological Engineering

Definition

Advanced users of CRM may want to construct extensions to the ontology or participate in the maintenance and development of CRMbase and its extension. This involves the knowledge of how to build ontological models using the principles adopted at CRM SIG, the manner of documentation, the rules of the overall management process.

Principles of Ontology Engineering

Definition

In order to create CIDOC CRM compatible models, or to move the CRM forward consistently, a learner should understand the methodology used in creating the ontology in the first place.

Existing Materials

- Principles document

Relevant Tools/Software

Principles of Ontology Documentation

Definition

In order to present an ontology that can be used in a manner compatible with CIDOC CRM, the ontology must follow basic documentation methods for describing and explaining the ontology. This module aids them to do this.

Existing Materials

- Scope Note Document
- Steve's Presenting an Extension Document

Relevant Tools/Software

Principles of CRM SIG Process Management

Definition

In order to make changes to the CIDOC CRM base model or an official extension, changes must be presented to and approved by the SIG. The SIG has rules of operation. This module presents and explains these principles, making the change management process clear to the learner.

Existing Materials

Relevant Tools/Software