CDM-Core: A Manufacturing Domain Ontology in OWL2 for Production and Maintenance

Luca Mazzola*, Patrick Kapahnke, Marko Vujic and Matthias Klusch

I2S - Intelligent Information Systems Research Team, ASR- Agents and Simulated Reality Department, German Research Center for Artificial Intelligence (DFKI) Stuhlsatzenhausweg 3, D-66123 Saarbruecken, Germany

Keywords: CDM-Core, Applied Ontology, Semantic Annotation, Knowledge Engineering, CREMA H2020 RIA Project, Ontology Quality Measurement.

Abstract: Ontology engineering is known to be a complex, time-consuming, and costly process, in particular, if an ontology has to be developed from scratch, and respective domain knowledge has to be formally encoded. This paper presents the largest publicly available manufacturing ontology CDM-Core in the standard formal ontology language OWL2. The CDM-Core ontology has been developed within the European research project CREMA in close collaboration with the user partners in order to sufficiently cover the CREMA use case domains of metal press maintenance and automotive exhaust production. CDM-Core makes use of many relevant standard vocabularies and ontologies, with only about one fifth of its size being CREMA use case specific. The practical applicability of CDM-Core for semantic annotation of domain-related process models, sensor data and services has been approved by the user partners, and its quality according to selected common criteria of verification and validation was successfully evaluated. From the public release of the CDM-Core, we expect to cover the lack of a base common ontology for the manufacturing domain, thanks to feedbacks from industrial reuse and improvements from the community.

1 INTRODUCTION

Developing ontologies is considered nowadays a standard activity in research project dealing with semantics. Unfortunately, this is not a common result of applied projects, where the effort and knowledge required to develop an ontology from scratch is considered not sustainable, in respect of the expected benefits. For this reason, in the context of the EU-founded Horizon2020 CREMA project, we decided to develop the "CREMA Data Model, Core module" (CDM-Core), as a manufacturing ontology taking into account both the general manufacturing domain applicability and the specific project use cases coverage. As a result, this is the first publicly available applied manufacturing ontology (Mazzola et al., 2016), composed by three different parts: a general manufacturing-related flat layer, a set of domain-specific or standard-based vertical slices such as "Conditional Monitoring" or "Semantic Sensor Net-

*Corresponding author, mazzola.luca@gmail.com
1Available at: http://sourceforge.net/projects/cdm-core/
2CREMA is "Cloud-based Rapid Elastic Manufacturing" and its website is http://www.crema-project.eu/

work", and some use cases specialized segments (automotive exhaust production and metallic press maintenance), that can be a guidance for developing other specific applications.

The rest of the paper is organized as follows: In section 2 the CDM-Core requirements are presented, together with the obtained results from the distributed process for the ontology creation; some of the CDM-Core usages for semantic annotation of process models, services and data streams are showcased in Section 3; whether Section 4 closes the paper with an analysis of the quality measures for the developed ontology.

2 THE CDM-Core ONTOLOGY

From the use cases description and the user partner inputs, a set of high level requirements was designed, and subsequently validated.

The elicited requirements include multifaceted aspects, such as the CDM-Core capability to represent domain knowledge for both use cases, in order to allow annotation of process model, services and sensors
data; its expected (logical) consistency; the adoption of selected relevant freely available standards; and the usage of a standard W3C modelling language. Based on these specifications, the development started with the engineering phase, as briefly described in the next section.

2.1 Ontology Engineering

The CDM-Core is modeled in the standard W3C ontology language OWL2-DL (Consortium et al., 2012). The CDM-Core has been developed by the task partners according to the distributed ontology engineering (OE) methodology DILIGENT (Tempich et al., 2005). Coherently with the literature (Sure et al., 2009), (Simperl et al., 2010) there are some common steps, that were followed also in this case.

Initially, the Domain analysis concerns the elicitation of requirements from use cases, the identification of relevant information sources and the ranking and adoption of relevant semantic data models which are available for public reuse. Then there is the Conceptualisation of knowledge, where these inputs are tranformed into of a semi-formal conceptual model of objects and concepts with taxonomic relationships. Following there is the Formalisation of the conceptual model where the conceptual model is translated into a knowledge representation language with formal semantics. Eventually, the Evaluation of the formal ontology phase analyzes the sufficient coverage and description of the domains by users and domain experts and the syntactic correctness, consistency and normalisation of the ontology by the ontology engineer. At the end of each phase, an iteration loop can happen, if some of the involved professionals ask for a refinement based on the reached results. In our particular case, we observed multiple iteration, in particular triggered by the user partners.

As result, the requirement that CDM-Core represents knowledge of the CREMA use cases was eventually achieved to a satisfactory degree. In fact, the semantics of the given process models and sensor data can be basically described using the ontology.

The distributed engineering of the shared CDM-Core ontology has been performed by task partners according to the DILIGENT methodology (Pinto et al., 2004).

The role of ontology engineer is manifold and includes (a) the support of domain analysis and conceptualisation of knowledge, (b) the formalisation of the conceptual model in OWL2, and (c) the technological evaluation of the CDM-Core. A set of selected partners plays the role respectively of domain experts and users of the CDM-Core. All task partners are members of the control board for ontology analysis, revision and evaluation.

The CDM-Core ontology engineering process is cyclic. It is based on four main steps with controlled iterations: the first step is Build where the ontology engineering team builds a very small and basic consensual version of the CDM-Core ontology. These initial activities are carried out by the domain experts intensively supported by the ontology engineer. The second operation is a Local refinement where each domain expert performs an in-depth refinement of the shared CDM-Core version at the local site, towards a refinement of the conceptual model per use case. These activities are carried out concurrently and at geographically dispersed sites. Every local ontologies is evaluated by domain experts and ontology engineer, and then formalised. As third step, the Analysis and revision requires that the control board analyses the locally refined ontologies and revises the shared CDM-Core ontology accordingly, by means of identification of similarities and their respective alignment. Eventually, after a new release of the ontology, a Local update is performed by domain experts before initiating further local refinements (i.e. 2nd operation).

The user partners informed that for the considered processes and sensor data in the use cases currently no standard data models were used at their sites.

The result of the initial search and assessment of relevant non-semantic standard data models carried out by the task partners is shown in Table 1. In particular, the main domains of listed data models according to their public description are shown in the column Domain Coverage; in column Public Reuse the availability of these data models for their translation (e.g. to OWL2) and inclusion into the publicly available semantic data model CDM-Core is shown. Instead, in Table 2 are presented the initial set of considered semantic data models and extensions in the OWL2 language.

2.2 Result

The distributed development based on the presented requirements guided the creation of the ontology. In Table 3 the actually reused public ontologies in CDM-Core are stated, together with their role and their characteristics (numbers of classes, properties and axioms). The main result, beside the ontology itself, is its usage for covering the identified requirements, examples of which are presented in the next section.
### Table 1: Selected non-semantic standard data models for CREMA use cases.

<table>
<thead>
<tr>
<th>Data Model</th>
<th>Modeling Language</th>
<th>Domain Coverage</th>
<th>Public Reuse</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISO 13372:2012</td>
<td>UML Txt/Tab</td>
<td>Condition monitoring and diagnostics of machines, predictive maintenance</td>
<td>(Y)*</td>
</tr>
<tr>
<td>ISO 10303 (STEP) APs</td>
<td>EXPRESS EXPRES-G</td>
<td>AP214 in AP242:2104 - Core data for automotive mechanical design processes</td>
<td>N</td>
</tr>
<tr>
<td>ISO STEP PDM Schema V1.2</td>
<td>Graphical notation, Txt/Tab</td>
<td>Product data management (common subset extracted from STEP APs 214, 203, 212, 232)</td>
<td>(Y)</td>
</tr>
<tr>
<td>UN/CEFACT CCL, UNTDED-ISO 7372</td>
<td>Graphical notation, Txt/Tab, XML(S)</td>
<td>Supply chain and cross-border trading transaction messages for buy, ship and pay business processes</td>
<td>Y</td>
</tr>
<tr>
<td>ASD S-2000M V6.0</td>
<td>UML Txt/Tab</td>
<td>Material management incl. spare parts, focus on aerospace industry</td>
<td>Y</td>
</tr>
<tr>
<td>ISA-88</td>
<td>UML B2MML</td>
<td>Batch control configuration and communication between components in batch manufacturing plants</td>
<td>N</td>
</tr>
<tr>
<td>ISA-95</td>
<td>UML B2MML</td>
<td>Business logistics and manufacturing control incl. production scheduling, maintenance management - at the level of enterprise, site, area</td>
<td>N</td>
</tr>
<tr>
<td>ISO 3166, ISO 4217</td>
<td>Txt/Tab, XML</td>
<td>Country and currency codes</td>
<td>(Y)</td>
</tr>
</tbody>
</table>

* only the Informative sections

### Table 2: Selected semantic data models and extensions in OWL2 language for CREMA use cases.

<table>
<thead>
<tr>
<th>Ontology</th>
<th>Type</th>
<th>Relevant Domain Coverage</th>
<th>Standard</th>
<th>Public Reuse</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSN (extends DUL)</td>
<td>Domain</td>
<td>Sensory, Sensor Networks</td>
<td>W3C</td>
<td>Y</td>
</tr>
<tr>
<td>MASON</td>
<td>Upper</td>
<td>Manufacturing</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>DUL (DOLCE+DnS Ultralite)</td>
<td>Upper</td>
<td>Concepts of physical and social context, temporal and spatial relations</td>
<td>W3C</td>
<td>Y</td>
</tr>
<tr>
<td>GeoNames</td>
<td>Domain</td>
<td>Geolocation</td>
<td>W3C</td>
<td>Y</td>
</tr>
<tr>
<td>ONTO-PDM [PDC12]</td>
<td>Upper</td>
<td>Manufacturing product data</td>
<td>IEC 62264, ISO 10303</td>
<td>N</td>
</tr>
<tr>
<td>SCORvoc</td>
<td>Domain</td>
<td>Supply chain operations reference (SCOR)</td>
<td>APICS (Y)**</td>
<td></td>
</tr>
<tr>
<td>CM (extends SSN)</td>
<td>Domain</td>
<td>Condition Monitoring, Machinery Maintenance</td>
<td>ISO/IEC 13372</td>
<td>Y</td>
</tr>
</tbody>
</table>

** only the original N answer is overridden by the written permission received by APICS

### Table 3: Publicly available ontologies (or non-semantic standards translated) used inside CDM-Core release.

<table>
<thead>
<tr>
<th>Ontology</th>
<th>Type</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>MASON (Lemaignan et al., 2006)</td>
<td>Upper Ontology</td>
<td>Classes: 224, Properties: 40, Axioms: 370</td>
</tr>
<tr>
<td>SSN (Compton et al., 2012)</td>
<td>Domain</td>
<td>Classes: 52, Properties: 55, Axioms: 127</td>
</tr>
<tr>
<td>ConditionMonitoring (Günel et al., 2013)</td>
<td>Domain</td>
<td>Classes: 182, Properties: 41, Axioms: 363</td>
</tr>
<tr>
<td>Org (Reynolds, 2014)</td>
<td>Domain</td>
<td>Classes: 15, Properties: 37, Axioms: 662</td>
</tr>
<tr>
<td>DUL (Gangemi, 2012)</td>
<td>Upper Ontology</td>
<td>Classes: 76, Properties: 109, Axioms: 517</td>
</tr>
<tr>
<td>SCORvoc (Petersen et al., a) (Petersen et al., b)</td>
<td>Domain</td>
<td>Classes: 279, Properties: 297, Axioms: 7657</td>
</tr>
</tbody>
</table>
3 USAGES

This section presents examples of semantic annotations of process models, services, and sensor data of the CREMA use cases based on the CDM-Core. This is complemented with examples of how these semantic annotations can be used by other CREMA components in the use cases, for example for the automatic implementation and optimization of process models with services (SOA paradigm).

3.1 Process Model Annotation

In CREMA, the process models are described in the standard BPMN (Business Process Model Notation). The formal semantics of a process model in BPMN can be described by means of its annotation with elements of formal ontologies. Such annotation allows applications to semantically reason on them in general, and assist process managers in their semantic service-based implementation in particular. Services can be more precisely selected and automatically composed for implementing a given process model based on its semantic annotation. This is in line with the semantic SOA paradigm and reference model (MacKenzie et al., 2006).

In CREMA, a component will perform an optimal service-based implementation of given process models with advanced means of semantic service selection and planning. That requires the semantic annotation of both process models and available services. The semantic annotation of process models based on the CDM-Core is manually done by the process manager with the help of some specialized interfaces.

There is no standard format for the semantic annotation of process models in BPMN yet (Boissel-Dallier et al., 2015). On an abstract level, the semantics of processes and services can be basically described in terms of their input (I), output (O), preconditions (P) and effect (E) of their execution. In particular, the task partners first informally described the basic IOPE semantics of all process tasks of given process models in the CREMA use cases. These informal semantic description consists of text, including relevant main concepts which were either identified in or newly added to the CDM-Core for this purpose. Next, these semi-formal descriptions were manually transformed into the formal semantic IOPE-based annotations with CDM-Core.

Figure 1 shows an example of a semantic IOPE-based annotation of one process task of one process model. In particular, this process task is concerned with the resource allocation of a suitable robot. Given a production schedule containing a list of orders to be fulfilled, this task identifies a robot cell and corresponding robot, and leases it for the schedule. A robot must be able to perform welding operations and has to be equipped with particular clamps to hold a certain type of exhaust as specified in the tasks of the schedule. The requirements described above are specified through different parameters, including the production schedule input, the robot cell and robot outputs, and a range of internal variables. These semantic annotations (such as the ones in Fig. 1) were made relying on domain-specific and upper ontologies together with the use case specific part of CDM-Core.

The following Listing provides an example of the semantic annotation of the process task in part A of the Figure 1 as extension in the BPMN XML source code. Additional attributes named inputs, outputs, preconditions and effects are used to attach the semantic annotations to the standard BPMN definitions.

```xml
<xml version="1.0" encoding="UTF-8">
<bpmn:definitions ...>
  <xmlns:tco=http://<URI1>/wp8/tco.owl#>
  </xmlns:tco>
</bpmn:definitions>
</xml>
```

3 Semantic SOA is considered key to the development of semantics-empowered intelligent applications for the future Internet in various domains including manufacturing 4.0, and supported by an increasing number of industrial stakeholders such as Software AG, SAP, IBM, Siemens.

3.2 Service Annotation

Semantic services are services whose functional and non-functional semantics are described with formal ontology-based annotations. On an abstract level, a semantic service description includes a semantic IOPE-based profile and a semantic process model that
describe what this service does and how it actually works (Klusch, 2008a), (Klusch, 2008b).

According to Semantic SOA, process models can be implemented with executable services which semant-ics are described with a shared formal ontology. The implementation of each step or task of a process model with a relevant single or composite service by a process designer can be supported by means of automated high-precision semantic selection and planning of annotated services, either in fully automatically or in semi-automatically (with user interaction).

Since the CDM-Core is defined in OWL2, one natural choice of the semantic service description format would be OWL-S (Web Ontology Language for Web Services (Martin et al., 2004)) which allows a grounding of semantic services in WSDL or REST services (Lathem et al., 2007). As mentioned above, semantic service descriptions will be used to determine a functionally optimal assignment of services to given annotated process models.

3.3 Data Stream Annotation

In CREMA, CDM-Core can also be used to semantically annotate sensor data of the use cases. The user partners provided general information about the metal press system, the robots, their components and the attached sensors, as well as the relevant sensor data schema. The given data schema for the stream of time-stamped and sequentially ordered data buckets can be in different format, such as CSV, TSV (Tab Separated Values) or JSON (Gray et al., 2011).

In particular, each sensor observes or measures one property such as pressure or temperature. The semantic annotation of streamed sensor data will be automatically done using the CDM-Core concepts. That requires the mapping of the sensor measurement label to the concept in the ontology, which defines the formal semantics of this label in XML-RDF encoded OWL2. As a result, the data item is described by a set of RDF triples as an instance of the corresponding concept in the ontology. This mapping table and respective naming of sensor classes and measurements is given by the partner generating the stream itself.

Figure 2 presents an example of semantic annotation of multi-variate sensor data from multiple sensors attached to the hydraulic drive system of a metal press. The semantic annotation of sensor data with the CDM-Core allows for a qualitative, that is domain knowledge-based, rather than quantitative statistics-based data interpretation.

Figure 1: Usage of the CDM-Core ontology for (A) Process Task inside the BPMN semantic annotation, (B) Service semantic Annotation, and (C) a matching using the plug-in approach (Paolucci et al., 2002) between the semantic annotation on top and middle of this example.
4 QUALITY MEASURES

Ontology evaluation is the task of measuring the quality of an ontology according to given criteria. It is "a technical judgment of the content of the ontology with respect to a frame of reference during every phase and between phases of their lifecycle" and can be classified into ontology verification and validation (Gómez-Pérez et al., 2003). There are several approaches, methods and tools for ontology evaluation available, but no best practices and guideline for the selection of measures for ontology quality criteria.

The result has been evaluated according to a selected subset of ontology quality criteria and measures defined in (Vrandecic, 2010). In particular, the selected criteria subsume the given requirements for the CDM-Core. Since there is no publicly available ontology for the manufacturing domain, our CDM-Core could not be evaluated against some gold standard as a reference. For this reason it was performed by all task partners with user partners as stakeholders.

In summary, all requirements for the CDM-Core are satisfied. The individual results are described in the following in the context of the selected criteria.

Verification was concerned with evaluating if CDM-Core specification is formally correct and meaningful in terms of syntactic validity, and logical consistency. Syntactic validity refers to the syntactically correct encoding of the ontology specification, which can be tested with appropriate validation tools such as the OWL validator4, SWOOP5, CityPulse Ontology Validator6, Eyeball7, OBO-Edit8, and OOPS!9.

Logical Consistency requires that the ontology specification does not include or allow for any logical contradiction. In other words, an ontology is inconsistent10 if it does not allow any formal model to satisfy its axioms. Checking of consistency can be done with a classical ontological reasoner such as Pellet (Sirin et al., 2007). The above mentioned validation tools support consistency checking, but they differ in the extent they check for common problems or pitfalls.

Results: The CDM-Core is syntactically valid and consistent. The XML-RDF encoded specification of the CDM-Core ontology in OWL2 was successfully tested with the tool OOPS! (also, advanced evaluation option): it does not contain critical problems such as circular class hierarchies, redundant axioms, incompleteness, or other consistency problems.
sistent naming schemes, or other logically inconsistent definitions of concepts and relations.

**Validation** was concerned with evaluating if the ontology is practically useful for the targeted stakeholders in terms of its accuracy and completeness of both covering the use case domains and supporting the tasks for which the ontology has been designed, its computational efficiency, and its adaptability to manufacturing domains and tasks of other stakeholders (additional criterion that the task partners identified later).

**Accuracy** checks if the specification complies with the knowledge of the stakeholders. **Correctness** in this case means compliance to the gold standards” of use case descriptions and respective conceptual models. **Completeness** verify that the use case domains are not only accurately but fully covered by the ontology, and that semantic annotations of given use case process models, services, and sensor data are sufficiently supported. It also covers the structural quality of the ontology for which measurement the following measures were selected:

- **(TD)** maximal class Taxonomy Depth (Lozano-Tello and Gómez-Pérez, 2004), i.e. the maximum subsumption path length of the CDM-Core. It covers the intuition that a high TD reflects a more detailed concept knowledge represented by the ontology.
- **(RR)** Relationship Richness (Vrandecic, 2010), i.e. the ratio between the number of property names and the number of class names and property names of the CDM-Core. It reflects the diversity of relations in the ontology and cover the intuition that detailing of existing classes would increase the relational richness of the ontology.
- **(AR)** Attribute Richness (Tartir et al., 2005), i.e. the average number of properties (attributes) per class. It suggests the intuition that high attribute richness indicates more information about each class on average.

**Computational efficiency** is the reasoning complexity (RC) of CDM-Core, i.e. the complexity that applies to the common reasoning tasks for the OWL2 fragment that is actually being used in the CDM-Core. **Adaptability** represents an indicator for the effort expected for effectively reusing the developed ontology in different cases inside the same domain.

**Results:** CDM-Core is accurate: as a consequence of the approval of its sufficient matching with the underlying conceptual model by the stakeholder, definitions and descriptions in CDM-Core are correct. It is also complete according to the stakeholders. As a result of its joint engineering, the CDM-Core was eventually approved by the stakeholders to represent all relevant instances, concepts and relations of the conceptual model. Moreover, it allowed the annotation of each of the given process models with concepts and properties; all sensor measurements of the given data stream schema were semantically mapped to corresponding elements in it; and every used sensors, robot cells and metal press is represented by an appropriately designed individual in CDM-Core.

The structural quality factors of the CDM-Core are: $TD = 7$, $RR = 0.3993$, $AR = 0.8156$. Our interpretation of these values is that the developed CDM-Core features a very high number of domain-specific classes with very high attribute richness (AR), which indicates a high amount of detailed information about each class on average. Its class hierarchy is of the same moderate maximal depth (TD) as, for example, the generic manufacturing ontology MASON and the standard W3C SSN ontology, it partly builds upon.

The worst case reasoning complexity (RC) is computed as the intersection of the complexity of the different OWL2 fragments in the CDM-Core, and is equivalent to SROIQ(D). Anyway no definition of CDM-Core covers jointly all the operators of this complexity, which indicates that the reasoning complexity in practice might be of some magnitude lower. **Adaptability** is limited due to its focus on covering the use case domains described (CREMA_consortium, 2016a), (CREMA_consortium, 2016b) and allowing the tasks of annotating the given process models, services and sensor data. However, the CDM-Core in particular builds on and includes generic and standard-based ontologies. These generic parts can serve other stakeholders to model knowledge of different manufacturing domains and tasks. The normalized proportion of the generic to the CREMA use case domain-specific parts is 21.09%.

The CDM-Core is specified in OWL2-DL and can be in principle extended and specialized monotonically, i.e. without the need to remove axioms.

### 5 CONCLUSIONS

In this paper we presented the first publicly available ontology for the manufacturing domain, together with the process for its development. It is one intermediate result of a shared effort of different organizations in the context of a collaborative project. The validation showed its capability for covering the requirements elicited as prerequisite for the ontology engineering phase. Additionally, some measures for the structural quality of the produced CDM-Core ontology are presented, together with its applications for process model, services, and data stream annotations. The resulting ontology is released for public reuse and we expect that industries can reuse it, provide feedbacks
and ask for improvement to the community.

This work was partially supported by the Commission of the European Union within the CREMA H2020-RIA project (Grant agreement no. 637066).

REFERENCES


