Boost 4.0* work on ontologies

Javad Chamanara (L3S) <chamanara@l3s.de>
Dave Raggett (W3C) <dsr@w3.org>

MMO Workshop, Brussel, 6 June 2019

* see: https://boost40.eu/

Boost 4.0 Project

- Large scale EU funded H2020 *lighthouse* project on big data in manufacturing
  - 20 M€ funding, 100 M€ private investment, 50 partners across 16 countries
  - 10 pilots across major European manufacturers, plus 3 evaluation projects across automotive, manufacturing equipment and household appliances
- Objectives
  - **Global Standards:** Contribution to the international standardization of European Industrial Data Space data models and open interfaces aligned with the European Reference Architectural Model Industry 4.0 (RAMI 4.0).
  - **Secure Digital Infrastructures:** Adaptation and extension of cloud and edge digital infra- structures to ensure high performance operation of the European Industrial Data Space; i,e, support of high-speed processing and analysis of huge and very heterogeneous industrial data sources.
  - **Trusted Big Data Middleware:** Integration of the four main open source European initiatives (Industrial Data Space, FIWARE, Hyperledger, Big Data Europe) to support the development of open connectors and big data middleware with native blockchain support in the European Industrial Data Space.
  - **Digital Manufacturing Platforms:** Open interfaces for the development of big data pipelines for advanced analysis services and data visualization supported by the main digital engineering, simulation, operations and industrial quality control platforms.
  - **Certification:** European certification program of equipment, infrastructures, platforms and big data services for their operation in the European Industrial Data Space.

Boost Core Ontology

- Bringing together ideas from several existing ontologies
  - The Reference Architecture Model for Industry 4.0 - RAMI vocabulary
    - With sets of classes, object properties, data properties, annotation properties and named individuals for the reference architecture and administration shell
  - The Standards Ontology (STO)
    - The Standard Ontology, STO, represents standards, standardization organizations and standardization frameworks for the Industry 4.0 area, see paper: Industry 4.0 Standards Landscape from a Semantic Integration Perspective
  - The Semantic Sensor Network (SSN) ontology and its descendent: Sensor Observation, Sample and Actuator SOSA ontology
    - The SSN ontology is designed for describing sensors and their observations, the involved procedures, the studied features of interest, the samples used to do so, and the observed properties, as well as actuators. SSN follows a horizontal and vertical modularization architecture by including a lightweight but self-contained core ontology called SOSA for its elementary classes and properties
  - Supply Chain Operation Reference Vocabulary (SCORVoc)
    - Cross-industry approach to lay the groundwork for more efficient and effective information exchange in supply networks. SCORVoc is an OWL vocabulary which fully formalizes the latest SCOR standard, while overcoming identified limitations of existing formalizations
- Concepts include device, sensor, machine, data, process and platform
- To be used in conjunction with IOF ontologies
- What other ontologies should we be looking at?

Ontologies Registry

- Boost 4.0 project aims to provide a registry of ontologies relating to big data in smart factories
  - Central registry web page
  - GitHub repository
  - VoCol/WebVOWL for public access
- Graphical browsing with WebVOWL
  - Interactive force-directed graphical layout (demo)
  - Text based browser (demo)
- We still need to work on richer documentation
  - Use cases and requirements
  - Introductions, working assumptions and application tutorials
  - Tooling and community support

Boost Core Ontology

- What are the concepts, with definitions, in the upper level the ontology?
  - Agent: something that bears some form of responsibility for an activity taking place, for the existence of an entity, or for another agent's activity.
  - Activity: something that occurs over a period of time and acts upon or with entities; it may include consuming, processing, transforming, modifying, relocating, using, or generating entities.
  - Device: a physical piece of technology - a system in a box. Devices may, of course, be built of smaller devices and software components (i.e. systems have components)
  - Event: An Event in the real world (e.g. SensorInput, Stimulus)
  - Entity: Any object that is relevant from a user or application perspective. (IoT-A Definition)
  - Process: A process is a unique activity performed to meet predefined outcomes
- What are the industrial use cases demonstrating the value of the ontology?
  - IDS: Industrial Data Space
  - Supply chain
  - Smart factories and big data processing/analysis

Overlaps with other ontologies

- RAMI4.0:
  - Asset
  - AdminShell
  - Communication
  - Machine
  - Platform
- AutomationML:
  - Attribute
  - External Interface/Reference
  - Project
  - CAEXFile
- STO:
  - Standard
  - StandardOrganization
  - Publication
  - TechnicalCommittee
  - Domain
- SSN:
  - Accuracy
  - Condition
  - Deployment
  - Device
  - FeatureOfInterest

Boost Core Ontology

- What are the (main) relations in your ontology?
  - is-a
  - same-as
  - has
- What is the knowledge your specific ontology represents?
  - Knowledge necessary for a pragmatic description of current practices
    - BCO represents entities involved and interacting in the domain of Industry 4.0 including machines, devices, sensors, processes, data, and artefacts
  - Explanation of the world according to one of the philosophical views called realism / conceptualism / nominalism
- How does your ontology represent the relations between different granularity views on the same object?
  - ...
- How does your ontology represent materials?
  - components, physical, chemical, mechanical, etc. properties, history of actions and effects on material
- What type of processes do you address? How does your ontology represent these processes?
  - Manufacturing, supply chain, delivery,
  - In general as a flow of activities. Currently, only a linear flow is considered.
- How does your ontology represent manufacturing?
  - As a set of interrelated processes that include platforms, machines, operations, material, product
- How does your ontology address the circular connection between physical properties, materials models (see definition in RoMM Review of Materials Modelling VI) and measurement?
  - in this draft: as a thing that has measurable properties, and those properties can be measured by sensing devices capable of measuring X, Y, Z.
- What is the representation language and implementation (logics)?
  - OWL, RDF

Purpose of Materials Ontologies

To be effective, and to provide a return on investment, ontologies need to be designed with a clear understanding of the use cases that the ontologies will be put to, e.g.

- Using a catalogue to select materials based upon their properties
- Searching for manufacturing processes appropriate to a given material
- Understanding the tolerances practical using different manufacturing processes
- Modelling macro physical characteristics for assemblies constructed from different materials, e.g. deformation under load, and failure modes
- Modelling micro physical characteristics, e.g. fracturing due to repeated stress
- Modelling chemical characteristics, e.g. in relation to corrosion in different environments
- Modelling electronic characteristics, e.g. semiconductors and nanotechnologies
- Predicting how materials and assemblies will degrade over time and use
- Selecting suppliers for raw and worked materials
- Cost of the raw material and the reliability of its supply chain
- Cost for working the material into useful components
- Processes for disposing or recycling materials as production waste or product end of life

Outstanding Questions for Ontologies

- What are the main business cases for adopting the use of ontologies?
  - What are the major business schools saying on this?
- What the alternatives and their pros and cons vis-à-vis ontologies?
  - What are the major business schools saying on this?
- Ontologies are complex to develop and to use – how can this be made easy enough for the average developer?
  - W3C Easier RDF initiative and the Graph Standardisation Business Group
- How flexible are ontologies in respect to continually evolving requirements?
  - Different use cases will emphasise different requirements and vary in complexity
  - A single ontology is likely to be too inflexible and act as a brake on innovation
- Simple rules and operational semantics vs logic problem solvers
- Failure to support real world situations
  - Uncertainty, incompleteness, inconsistency, likelihood of errors
  - Reasoning: deductive, inductive, abductive, causal, counterfactual, temporal, spatial
  - Solution is to combine symbolic and statistical approaches, but new work is needed!

Lightweight vs Formal Ontologies

- Lightweight ontologies describe sets of concepts and their relationships
  - Easy to create where there is agreement on the meanings
  - Reasoning through operation of explicit rules
- Formal ontologies express axioms and inferences
  - Hard to create, but general in coverage
  - Reasoning with a Logic engine
- Mapping data between different ontologies
  - Peer to peer mappings
    - Rely on explicit rules
  - Back to first principles mappings
    - Rely on mapping to a shared upper ontology (hard)
  - Statistical models by analogy to machine translation of human languages
- For more details see:
  - A Lightweight Ontology Approach to Scalable Interoperability by Hongwei Zhu and Stuart Madnick, MIT Sloan Working Paper 4621-06, June 2006

Future ontologies will be generated by machine learning with human supervision

- Manual design of ontologies is hard and requires deep understanding and technical expertise in knowledge representation
  - This acts as a brake on widespread adoption
- The future lies with machine learning based upon curated sets of examples and counter-examples
  - Much easier for the average developer
  - Use of controlled subset of natural language
  - Bootstrap from corpora maintained by industry groups
- Continuous learning as new cases appear
- Lends itself to techniques exploiting computational statistics

Thanks you