



A banner with a geometric pattern of squares and triangles in shades of blue, purple, and orange. In the top left corner is the Onto Commons logo. The main text reads: "ONTOCOMMONS IN THE STANDARDISATION ENVIRONMENT" in large, bold, black letters. Below this is "An interview with Barry Smith" in a smaller font, followed by "SUNY Distinguished Professor of Philosophy at the University at Buffalo" in a smaller font. At the bottom left is the University at Buffalo logo. On the right side of the banner is a portrait of Barry Smith, an older man with white hair and glasses, wearing a dark suit and white shirt.

26 April 2022



"Many people have found a way to understand new kinds of data thanks to the successful use of ontologies. Ontologies make data more generally understandable and usable. They do this by providing controlled vocabularies, which are used to tag data organized using often meaningless codes using terms which make the data understandable and relevant."

Ontologies and standards are strongly interrelated and can be often be seen as two sides of the same coin. Indeed, standards reflect consensus on the semantics of terms, though different standards might employ different ways to explain the same or very similar concepts semantically. Given the crucial role played by both aspects in OntoCommons, we have interviewed **Barry Smith**, SUNY Distinguished Professor of Philosophy at the [University at Buffalo](#) and one of the [External Advisory Board \(EAB\)](#) members of the OntoCommons project, to better understand what are the challenges and priorities to focus on in this context.

What are the main priorities or challenges for standardisation in the ontologies field? How can projects, like OntoCommons, contribute to this?

Nowadays there are plenty of ontologies that have not been properly managed because the websites or repositories on which they are published have broken links or contain information that is out-of-date. This represents a problem to overcome: if we want to have a successful ontology market, we need to ensure that there is evidence of available resources that is made visible to potential users. Otherwise, the interested researchers cannot access updated data and research cannot proceed further. In this context, one of the main services that the OntoCommons



Contact

scenario highlights how crucial it is to have public repositories so that everybody can find the ontology that they should be using, with documentation as to how the ontology should be used, descriptions of success stories, and clear rules relating to how the ontology is being maintained.

OntoCommons can avoid the problem of ontology overproliferation in the materials and manufacturing fields by applying the best practices for ontology building and use that have been developed and tested over some 15 years by the OBO Foundry. This is already being done by initiatives like the Industrial Ontologies Foundry ([IOF](#)), and I believe that cooperation with the IOF Foundry will prove vital for OntoCommons, since IOF is developing a repository of ontologies for materials science and engineering disciplines. The IOF is focusing especially on applications relating to manufacturing, and their work therefore does not comprehend the entire domain. Hence, there is still room for a more comprehensive public repository, which would ideally represent the main work of OntoCommons in tandem with the work of the IOF in a single well-curated repository and associated ecosystem of data and applications. In this way we can avoid re-inventing the wheel. To achieve this, the two project teams would need to apply the same standardisation rules. Standards, indeed, help overcome the issue of having too many ontologies for every single domain or areas of expertise and the new standard ISO/IEC 21838 is attempting to achieve this for the principal top-level ontologies.

OntoCommons aims to build the OCES, the Ontology Commons Ecosystem, that provides specifications and guidelines for the methodological support to industry-targeted ontology development, maintenance and usage for documenting data. In your opinion, can it be feasible to get a standardisation recognition for the OCES?

I think that there would be an opening for standardisation work not only as regards the principles used in ontology building and maintenance but also concerning the principles for annotating or tagging data. Again, the OBO Foundry and subsequent efforts in its train have already done useful work in this regard. I think that it would be good to try to build a strategy for this, as I believe that standardising principles in multidisciplinary initiatives of this sort have great potential.

OntoCommons is going to implement Focus Areas on Cooperation on Infrastructures, Cooperation on Standardisation, Top Reference Ontology, Industrial Domain Ontologies and Demonstrators, which are key assets of the project. Do you think they are strategic for the OntoCommons' success and strategic for its community?

These are good, indeed indispensable, ideas as they facilitate the sort of exchange and reuse of information that we need to encourage in an era where data is expanding at such enormous rates. But I believe that something that is indispensable to make OntoCommons successful is the organisation of ontology training. It is important to have practical outreach efforts that provide opportunities for potential industrial users to learn about the products of the OntoCommons work and to see how they can be of practical value. One major reason for the success of the OBO Foundry was the degree to which ontology development work was accompanied by many different outreach efforts, directed to university audiences, to IT specialist communities, to conferences, and so forth, always with the intention to teaching the same set of principles to different audiences in order to maximize the possibility that a single ontology eco-system would develop in which everyone could be involved.

The OntoCommons Roadmap will include a number of recommendations for policy makers towards existing ontology and data documentation to ensure long-term relevance and implementation. In your opinion, how can this contribute to the [Digital Single Market agenda](#)?

In industrial engineering you can have horrible problems when sharing data as not all companies believe in open-source data and open-source tools. The key problem is mainly linked to intellectual property rights, so that what one company develops will in general not be available for use by other companies. This represents a paradox, as in the worst-case scenario no one would be willing to share information. In order to overcome this while building the OCES, it would be necessary to look at the Creative Commons licenses to ensure that ontologies are released in a way that everybody – and every company – is allowed to use the ontologies and data freely, without incurring legal problems. I have been working with the IOF to ensure that we have clear licensing processes to avoid the above-mentioned issues, and to ensure that ontologies are developed after the hub-and-spokes model, where all the ontologies near the hub are free to use by all, while some of the ontologies distant from the hub are subjected to IP restrictions imposed for example by the companies who develop them for internal use.

In all of this, both ontologies and standards play a crucial role in facilitating the exchange of data. Indeed, if data is organised using the same ontologies and standards, it is easier to use it in building bridges to external data. Companies – most prominently for example in the pharmaceutical area – are starting to recognise this and are becoming more willing to adopt both standards and ontologies built by experts for public use and based on adequate licensing policies. All of this, facilitate the needed exchange of information across multiple domains.
