

Submission 4112

The submission information is shown below.

- [Update information](#)
- [Update authors](#)
- [Add file](#)
- [Delete the submission](#)
- [Email to authors](#)
- [Email to reviewers](#)
- [Show reviews](#)
- [Add review](#)
- [Request review](#)
- [Assignment](#)
- [Edit note](#)
- [Attach document](#)

Submission 4112	
Title	Collaborative Development of a Process Chemistry Domain Ontology, PROCO
Author keywords	Process Chemistry Ontology Knowledge Management Data Standards
Abstract	<p>Process chemists embracing data mining, machine learning and artificial intelligence rapidly discover that the lack of structured data hampers their efforts. Although raw and processed instrument data have been addressed with public ontologies such as Allotrope and somewhat by commercial enterprise content management solutions, the chemical context of that data has been largely neglected. Of particular interest was the domain of process chemistry that studies the development and optimization of production processes for chemical compounds. Recognizing this foundational gap and the impact it would have on developing new and integrated scientific data capture systems like electronic laboratory notebooks, Merck chemists and IT reached out to peers in other companies and academia to address this need. The result was PROCO, a process chemistry domain ontology. The scope was set using specific use cases and is being rounded out by modeling public databases such as ORD.</p> <p>Development of PROCO was based on “up-scaling” the chemist’s semantic skill sets and the domain knowledge of the ontologists. Simplified public ontology tools such as WebProtege provided a collaborative online space for ontology developers and subject matter experts with features like commenting, suggesting, and approving changes. To increase internal adoption and applicability of PROCO, the ontology was loaded into Merck’s master ontology for discovery, pre-clinical, and early development space (MDO) via its CENTree ontology management system. Specific applications use MDO as the master source to develop their ‘application ontologies.’ As an example, ELN application ontology covers experimental metadata and feeds it into the Perkin-Elmer Signals notebook to define values of drop-down lists in the notebook. This enables standardized data capture, which consequently makes the data interoperable and reusable for analytics/data science. Using ontologies as a metadata input for data capture enables data to be ‘born FAIR’ (findable, accessible, interoperable, and reusable) which is significantly more efficient and less expensive than FAIRifying the data at later stages.</p> <p>This industrial and academic collaboration proved to be an effective means of achieving better-structured data with limited enterprise resources. The final PROCO ontology has been submitted to the OBO Foundry to broaden the development pool and usage of the ontology.</p>
Submitted	Jun 15, 14:42 GMT
Last update	Jun 15, 14:42 GMT
Decision	
Bidding and assignment information	
Assigned to	nobody
Conflict of interest	nobody
Bid	yes: none maybe: none

Authors

first name	last name	email	country	affiliation	Web page	corresponding?
Wes	Schafer	wes_schafer@merck.com	United States	Research and Development Sciences IT, Merck & Co		✓
Jan	Nespor	jan.nespor@merck.com	Czechia	IT Eng., Dev. & Integration, MSD		✓
Vincent	Antonucci	vincent_antonucci@merck.com	United States	Research and Development Sciences IT, Merck & Co		✓
Oliver	He	yongqunh@med.umich.edu	United States	University of Michigan Medical School		✓
Anna	Dunn	anna.dunn@merck.com	United States	Analytical Research & Development, Merck & Co., Inc		✓
Zachary E.X.	Dance	zachary_dance@merck.com	United States	Analytical Research & Development, Merck & Co., Inc		✓

