

Collaborative Metadata Definition using FAIR Controlled Vocabularies, and Ontologies

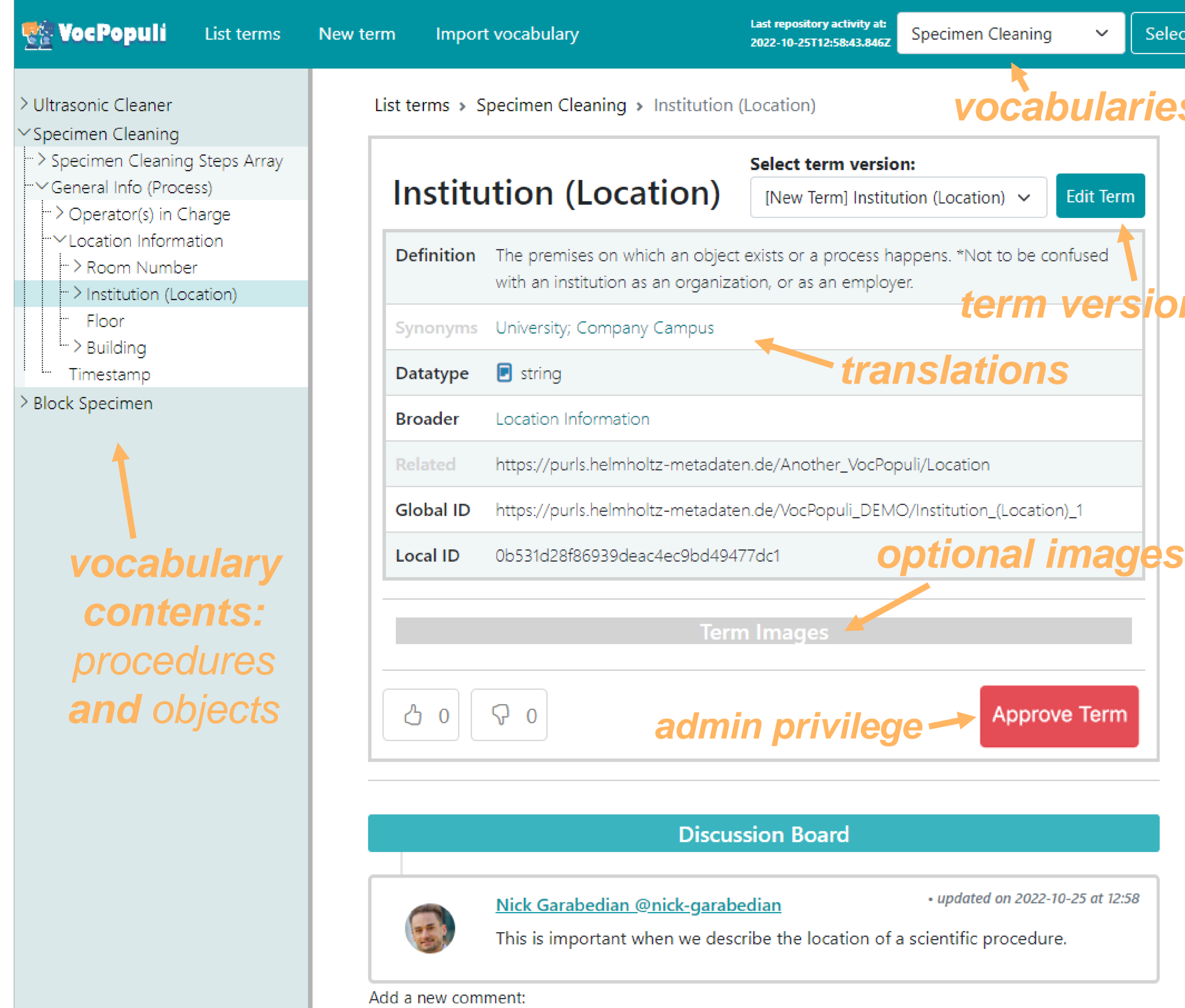
Nick Garabedian, Ilya Bagov, Christian Greiner
Karlsruhe Institute of Technology, Institute for Applied Materials, Germany

Motivation

- FAIR vocabularies are a prerequisite for FAIR data publication
- Easy tools for collaborative building of FAIR vocabularies are missing
- Vocabularies are an intuitive first step for any lab undergoing digitalization



VocPopuli



vocabularies

term version

translations

optional images

admin privilege

vocabulary contents: procedures and objects

- Ease of use by non-experts is first priority
- **Any resource can be digitalized:** Lab Procedure; Lab Specimen, ELN, ELN Export, Data Analysis
- GitLab login is used to associate user contributions with term versions and vocabularies
- Each term receives a git branch for editing
- GitLab Issues are used to collect comments
- Terms are voted on and approved or not approved
- PROV and SKOS are optional exports

Vocabulary IDs

Global ID: set at birth and never changes
Local ID: a hash of the terms' Local IDs
Version ID: a hash of all terms' Version IDs; changes when a term is modified
IRI (PURL): if published

Term IDs

Global ID: set at birth, never changes; it represents a self-contained concept
Local ID: denotes an approved version of the term
Version ID: changes for every unapproved version
IRI (PURL): if published

OntoFAIRCook

- FAIR vocabularies are a stepping stone towards FAIR ontologies
- In this ML-based app vocabularies are converted semi-automatically into FAIR data ontologies
- The conversion is actively improved as the ontologies grow
- Ontologies allow for scalable data-driven machine learning

Electronic Lab Notebooks (Results Management Systems)



- FAIR-Save Apps interact with vocabularies and ELNs
- Terms used within ELNs are linked to the IDs in the related vocabulary and ontology for an unambiguous definition
- An export of the defined schemes can be used to automatically create structured forms in the ELNs
- The output from the developed tools will be exemplarily first integrated into the ELNs Herbie and Kadi4Mat

Contact and Acknowledgements

This project is a close collaboration with our partners: Frederic Bock, Benjamin Klusemann, Catriona Eschke, Florian Wieland, and Martin Held at Helmholtz-Zentrum Hereon; and Karlheinz Weber, Nuoyao Ye, Floriane Bresser, Benedikt Stoll, Malte Flachmann, Jakob Biesinger at the KIT. It was funded by the Initiative and Networking Fund of the Helmholtz Association in the framework of the Helmholtz Metadata Collaboration project call.

