The CHAMEO Ontology: Exploiting EMMO's Multiperspective Versatility for Capturing Materials Characterization Procedures

Pierluigi Del Nostro[1], Gerhard Goldbeck[1] and Daniele Toti[1,2]

[1]Goldbeck Consulting Limited, St John's Innovation Centre, Cowley Rd, Cambridge CB4 0WS, United Kingdom
[2]Catholic University of the Sacred Heart, Faculty of Mathematical, Physical and Natural Sciences, via della Garzetta 48, 25133, Brescia, Italy

Abstract
The characterization of materials includes a plethora of different methods and terminologies, and as such has historically been a challenging field for promoting effective mechanisms of interaction and interoperability among its communities. Within this context, the CHAMEO ontology has the purpose of providing a common modeling framework for materials characterization, in terms of a domain-level ontology under which method-specific, application ontologies in the field may be developed. CHAMEO, based on the CHADA documentation scheme, is part of a larger effort in standardizing materials modeling and scientific activities that has EMMO, the Elementary Multiperspective Material Ontology, as its core. In this regard, this work shows how CHAMEO, via its alignment with EMMO and the exploitation of the latter's constructs and perspectives, can effectively help scientists model the different aspects of the materials characterization processes.

Keywords
ontology, materials characterization, multiperspective

1. Introduction and Related Work

Material characterization is the process of measuring and determining physical, chemical, mechanical and microstructural properties of materials. The huge variety and complexity of materials led to the formation of multiple communities around the materials characterization field, establishing different terminologies which typically focus on specific application domains. This work represents a contribution to the creation of a common knowledge framework for the documentation of characterization methods, with the goal of facilitating reusability and transferability of knowledge across different communities and sectors.

The study, modeling and characterization of materials is a varied and complex field. This is why, over the years, several, different communities have arisen around it, each usually defining and setting up as many different term sets and nomenclature tailored to specific domains of

FOMI 2022: 12th International Workshop on Formal Ontologies meet Industry, September 12-15, 2022, Tarbes, France
pierluigi@goldbeck-consulting.com (P. D. Nostro); gerhard@goldbeck-consulting.com (G. Goldbeck); daniele.toti@unicatt.it (D. Toti)
https://materialsmodelling.com/ (P. D. Nostro); https://materialsmodelling.com/ (G. Goldbeck); https://www.unicatt.it (D. Toti)
0000-0002-5174-8508 (P. D. Nostro); 0000-0002-4181-2852 (G. Goldbeck); 0000-0002-9668-6961 (D. Toti)

application. The need of a commonly-acceptable, understandable and usable knowledge framework for the characterization of materials led to the development of a standard documentation of characterization methods, in order to improve and elicit the sharing and reuse of knowledge across the corresponding different communities.

In this regard, the CHAracterization MEthodology Ontology (CHAMEO) is meant to provide a harmonized and standardized representation of materials characterization methods and processes [1]. CHAMEO's inception is based on the CHADA document template [2], whose purpose was to provide a standard structure for documenting material characterization techniques.[1]

Specifically, CHADA used four concepts for classifying the steps of a characterization (workflow), namely "User case" (information about the sample, the testing environment, *etc.*), "Experiment" (characterisation process including probe, signal, detector, noise, *etc.*), "Raw data" (the output of the metrological process) and "Data processing" (analysis of the data for reaching the final shape). These concepts correspond to the sections of a CHADA document. This kind of document, while easily interpretable by humans, did not provide structured data for retrieving information on the characterization methodologies according to their different dimensions (*e.g.* material, probe, detector, properties). As such, the CHADA was used as a basis to build a more structured and shared knowledge, and thus the CHAMEO ontology was born. A more detailed description of the CHAMEO ontology can be found in [1].

As a consequence, CHAMEO is able to model the aspects shared by the varied characterization techniques by providing high-level, methodological definitions, paving the way to the development of further, more specialized ontologies at a finer-grained application level. This is also enabled by the intrinsic modularity of the ontological design, which places CHAMEO at the domain level of a larger ontological framework having the Elementary Multiperspective Material Ontology (EMMO) [6] as its upper layer. Due to this fact, CHAMEO is also able to exploit EMMO's versatility in terms of its multiple perspectives for describing processes and data according to different points of view. A more detailed description of EMMO is reported in Section 2, whereas Section 3 focuses on describing how EMMO's constructs and perspectives can be used to model the materials characterization processes.

2. Elementary Multiperspective Material Ontology (EMMO)

A number of initiatives have been carried out over the past ten years by the European Materials Modeling Council [7], with the objective of enabling and improving the interactions and cooperation among stakeholders from academia and industry in the area of Materials Modeling.

Within this context, one of EMMC's major accomplishments is the development of the Elementary Multiperspective Material Ontology (EMMO) [6], being the result of a multidisciplinary effort for defining a standard representational framework for applied sciences. EMMO is a top-level ontology based on physics, analytical philosophy and information and communication technologies, whose inception was brought about by practitioners in Materials Science in order to produce a framework consistent with scientific principles and methodologies.

[1]The CHADA document template was developed within the Horizon 2020 OYSTER project [3] and has been the subject of a recent CEN Workshop Agreement (CWA 17815) [4]. The development of the CHAMEO ontology has been carried out under the Horizon 2020 project NanoMECommons [5].

EMMO is in fact "Elementary", since it is a fundamental, top-level ontology, while at the same time assuming the existence of objects that cannot be divided further in space and time; it is also "Multiperspective", because it enables scientists to describe the world with it by resorting to different views or perspectives, thus providing a high expressive power and versatility.

EMMO's distinctive features and characteristics are the following:

1. EMMO has a minimal top level, which is based on the very few things one can say about the world based on fundamental science, in particular that everything is 4D, that there is a Universe object and a fundamental quantum 4D object and that everything in between cannot be specified a priori, other than to say it is either physical or a void. Everything else is a "Perspective".
2. There are no abstract objects in EMMO; everything must be described in terms of physical objects.
3. All relationships among objects can be of three types: topological (connectedness), mereological (parthood) and semiotic (signs that communicate meaning). In contrast to other ontologies that often have no or only a very loose way of organizing relationships, this provides a strong logic framework to support reasoning.
4. Semiotics is a core concept in EMMO. This is based on the semiotic theory by Charles S. Peirce [8], where the role of abstract objects are fulfilled by semiotic objects, *i.e.* real-world objects (*e.g.* a Symbol or Sign) standing for other real-world objects that are to be interpreted by an agent. These symbols appear in actions (semiotic processes) meant to communicate meaning by establishing relationships among symbols (signs).

EMMO's perspectives are the following:

1. Holistic: this perspective considers the importance and role of the whole as well as its parts without a specific, finer-grained decomposition, with classes like Whole (based on a given criterion) and Part (as it appears in relation to the Whole, also in terms of its role; thus, Role is an alternative label for Part, as is also used, for example, in theatrical plays.
2. Persistence: this perspective considers the persistence in time (process) or space (object).
3. Physicalistic: this perspective uses applied science concepts to provide meaning to objects (*e.g.* a material as a scientific object, interrogated by scientific means).
4. Reductionistic: this perspective focuses on a strict hierarchy of objects in terms of granularity levels (in space and/or time).
5. Perceptual: this perspective includes recognizable patterns in space and/or time such as sounds, languages, alphabets symbols, mathematics, graphics.
6. Data: this is a perspective in which entities (causal objects) are represented according to the variation (non-uniformity) of their properties, which can be recognized and interpreted according to the rules characterizing the data.

So far, EMMO has been applied in a number of European Projects, mentioned in the Acknowledgments section of its public repository[2], and a number of different domain and application ontologies have been created around it and aligned with it [9, 10].

[2]https://github.com/emmo-repo/EMMO#acknowledgement

3. CHAMEO and EMMO: Looking at Characterization Concepts through Different Perspectives

In order to allow the CHAMEO ontology to capture a wide range of aspects, a number of industrial domains(Organic Electronics, Avionics and Automotive, Additive Manufacturing, Energy and Chemicals) and different characterization techniques (*e.g.* Nanoindentation, Laser Flash Analysis, Differential Scanning Calorimetry, Geometric Phase Analysis, RAMAN) have been tackled. Competency questions have been collected from industrial partners to define the scope of the ontology, and a number of CHADA documents describing the different characterization techniques were collected (see [1] for additional details). By analyzing how the different CHADA documents were filled, it arose that the way the information is provided by different users is subject to interpretation, and this makes getting homogeneous information about the different characterization techniques rather hard. The CHAMEO ontology is meant to address this issue by providing a framework for defining a clear, machine-readable documentation, based on shared concepts and definitions. The main constructs of the CHAMEO ontology are here reported, highlighting how EMMO's multiperspective nature is exploited to model the characterization methodology concepts in a pluralistic way. This pluralistic approach supports different aspects and use cases including for example:

- The view of the overall characterization methodology as a system, *i.e.* an aggregate of things that interact, each with its own role, during the characterization process (persisting during that time -> *Persistence* perspective), and that is made of (sub-) objects as parts of a complex whole (*e.g.* Probe, CharacterizationMachine) (*Holistic* perspective).
- The procedural aspect which is important for operators of characterization, captured in terms of processes, stages and operations (*Persistence* perspective) and potentially with as a strict sequence (*Reductionistic* perspective).
- Representing the sample and the specimen as materials (*Physicalistic* perspective) as well as participants in the characterization method (*Holistic* perspective).
- Tracking the data used and generated, processing data using models and interpreting processed data (*Data* perspective) as properties of the material (*Semiotics* perspective).

The overall characterization process is modeled through the `chameo:CharacterizationWorkflow` class and is subclass of `emmo:Process` that is part of the *Persistence* perspective. A characterization workflow can be a composition of different methods, which are used to obtain the final characterization property. Methods are processes themselves, and are subprocesses of the overall workflow. They are represented by the `chameo:CharacterizationMethod` class which is subclass of `emmo:SubProcess`, under a combination of the *Persistence* and the *Holistic* perspectives, are processes themselves (*Persistence*), being a spatial part of the `chameo:CharacterizationWorkflow` (*Holistic* perspective). Each `chameo:CharacterizationMethod` can be divided into several `emmo:Stage`, following the nomenclature of the ISA-88 standard [11]. Since the CHAMEO ontology models the concepts that are in common across different characterization techniques, their most general stages are represented, namely:

- `chameo:SamplingProcess`: the act of extracting a portion (amount) of material from a larger quantity of material. This stage results in obtaining a sample that is representative of the batch with respect to the property or properties being investigated.

- `chameo:SpecimenPreparation`: the specimen preparation processes (*e.g.* machining, polishing, cutting to size, *etc.*). The preparation of a specimen can also implies the use of a holder (`chameo:Holder`). The way the holder is modeled is another example of usage of the EMMO perspectives: it is an `emmo:Object` (*Persistence* perspective), but also plays a role in the whole characterization process (emmo:Role in the *Holistic* perspective).
- `chameo:CalibrationMeasurement`: the act of measurement using a reference sample, with the goal to produce reference data.
- `chameo:MeasurementProcess`: this stage results in the determination of values using an equipment. It is subclass of the generic `emmo:Measurement`.
- `chameo:DataPreparation`: all the activities that are performed for preparing the raw data for the post-processing. Specific subclasses are added for data normalization `chameo:DataNormalization` and `chameo:DataFiltering`.
- `chameo:DataPostProcessing`: this stage refers to the application of a model (`chameo:PostProcessingModel`) through a software (`chameo:SoftwareProduct`) for the calculation of the final characterization property of interest.

These stages can be specialized in the ontologies related to specific techniques, and operations/actions (ISA-88) can be added. No specific temporal restrictions between stages are defined at the CHAMEO level, since they can be different in the various techniques. However, temporal sequences can be expressed through EMMO's *hasTemporalCause* object property. The characterization process runs trough a system (`chameo:CharacterizationSystem`), which is a subclass of `emmo:HolisticSystem` (*Holistic* perspective), defined as an object that is made of a set of sub-objects working together as parts of a mechanism or an interconnecting network (natural or artificial). It includes the probe and the characterization machine (which has the detector). The classes `chameo:Probe`, `chameo:CharacterizationMachine` and `chameo:Detector` are subclasses of `emmo:Object` (*Persistence* perspective). There is a distinction between the different types of data produced during the characterization process:

- `chameo:RawData`: The raw data is a set of (unprocessed) data that is the direct output of a detector, usually expressed as a function of time or position, or photon energy.
- `chameo:PrimaryData`: Data after pre-processing for corrections to normalize/harmonize (*e.g.* x/y axis correction, baseline subtraction, noise reduction) are called primary data.
- `chameo:SecondaryData`: The secondary data is the result of the application of a post-processing model (*e.g.* within the context of nanoindentation, the elastic modulus and hardness are obtained after the application of the Oliver-Pharr model).

All the three classes described above are subclasses of `emmo:Data` (*Data* perspective). The characterization properties of the specimen are modeled by the data, to which an observer assigns a meaning through a semiotic process (*Semiotics* perspective).

4. Availability

As detailed in [1], CHAMEO's axiomatization is stored in a GitHub repository available at: https://github.com/emmo-repo/domain-characterisation-methodology, and is published at the following URL: http://emmo.info/emmo/domain/chameo/chameo.

5. Conclusion

CHAMEO is part of a work to harmonize knowledge on materials characterization techniques, by representing their shared aspects. CHAMEO is placed within a modular structure, where the EMMO is on top as a reference for modeling knowledge in the area of Materials Science, and where CHAMEO is a common framework for the underlying ontologies to be developed for the specific characterization techniques. In particular, this work has highlighted how EMMO's perspectives can be exploited for modeling the different aspects of the characterization processes.

Acknowledgments

This work received funding from the European Commission via the Horizon 2020 projects "NanoMECommons" (Grant Agreement n. 952869) and "OYSTER" (Grant Agreement n. 760827).

References

[1] P. Del Nostro, G. Goldbeck, D. Toti, CHAMEO: An Ontology for the Harmonisation of Materials Characterisation Methodologies, Applied Ontology in press (2022). doi:10.3233/AO-220271.

[2] N. Romanos, M. Kalogerini, E. P. Koumoulos, A. Morozinis, M. Sebastiani, C. Charitidis, Innovative data management in advanced characterization: Implications for materials design, Materials Today Comm. 20 (2019). doi:10.1016/j.mtcomm.2019.100541.

[3] OYSTER, Open characterisation and modelling environment to drive innovation in advanced nano-architected and bio-inspired hard/soft interfaces, G. A. n. 760827, https://cordis.europa.eu/project/id/760827, 2017-2022.

[4] CEN, Materials characterisation - Terminology, metadata and classification, https://www.cencenelec.eu/media/CEN-CENELEC/CWAs/ICT/cwa17815.pdf, 2021.

[5] NanoMECommons, Harmonisation of EU-wide nanomechanics protocols and relevant data exchange procedures, across representative cases; standardisation, interoperability, data workflow, G. A. n. 952869, https://cordis.europa.eu/project/id/952869, 2021-2025.

[6] The EMMC Consortium, The European Materials Modeling Ontology (EMMO), https://emmc.info/emmo-info/, 2014-2021.

[7] EMMC, The European Materials Modeling Council, https://emmc.eu/, 2014-2022.

[8] A. Atkin, Peirce's Theory of Signs, in: E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy, Summer 2013 ed., Metaphysics Research Lab, Stanford University, 2013.

[9] M. T. Horsch, D. Toti, S. Chiacchiera, et al., OSMO: Ontology for Simulation, Modelling, and Optimization, in: CEUR Workshop Proc., volume 2969, 2021.

[10] G. Goldbeck, D. Toti, MAEO: An ontology for modeling Agents, Experts and Expertise within an open online Materials Modeling MarketPlace, in: CEUR Workshop Proc., volume 2969, 2021.

[11] A. N. S. Institute, I. S. of Automation, I. S. of America, ANSI-ISA-88.00.01-2010: Batch Control Part 1 : Models and Terminology, American national standard, ISA, 2010. URL: https://books.google.it/books?id=jbK0nQAACAAJ.