Case study on epistemic metadata in molecular modelling

Martin Thomas Horsch, Silvia Chiacchiera, and Björn Schembera

Based on a case study on epistemic metadata in molecular modelling [1-2], we discuss epistemic grounding with a focus on the disciplinary scientists' point of view. Both normative and reliabilist lines of reasoning are found to occur. As the call for abstracts states, even "in the philosophy of computational sciences and models, however, this difference has hardly been noted." Unsurprisingly, a conscious differentiation between the two is even less present among disciplinary scientists, and it can be complicated to classify individual instances according to such a schema. We have previously proposed to distinguish Type-1 grounding, based on characteristics of the outcome, and Type-2 grounding, based on characteristics of the underlying cognitive process [3]. Normative arguments generally belong to Type-1, reliabilist arguments to Type-2. This is part of a more comprehensive effort toward a standardized documentation of the knowledge status of data [3, 4].

Capturing epistemic grounding faces three methodological challenges that are particularly relevant to research data management in molecular modelling: First, the epistemic virtues that practitioners are most aware of have the tendency to oppose each other. Simplicity favours models with few elements and few adjustable parameters, whereas faithfulness in representing the underlying physics does the opposite. Second, it is complicated to draw a clear line between normative and reliabilist arguments. Norms are often in turn provided with a reliabilist justification; e.g., as we know from experience, models with fewer adjustable parameters are more reliable when extrapolating. Third, norms in operation within a disciplinary community are not usually explicitly written down; they therefore need to be analysed in terms of pragmatics [4]. However, we thereby risk to impute onto authors norms which have not been agreed upon and which these authors have not in fact invoked. Moreover, members of a community might disagree on norms and be unaware of their disagremment. Consequently, in reproducibility studies, secondary work may inadvertently hold the original work to a higher standard than the original authors intended. We therefore argue that reproducibility claims should be made explicit, and provide a formalism for this through the PIMS-II ontology for cognitive processes [4].

[1] M. Horsch, B. Schembera, "Epistemic metadata in molecular modelling: First-stage case-study report (10 cases)," doi:10.5281/zenodo.7516532, 2023.

[2] M. Horsch, S. Chiacchiera, G. Guevara Carrión, M. Kohns, E. Müller, D. Šarić, S. Stephan, I. Todorov, J. Vrabec, B. Schembera, "Epistemic metadata in molecular modelling: Second-stage case-study report (12 claims)," doi:10.5281/zenodo.7608074, 2023.

[3] M. Horsch, B. Schembera, "Documentation of epistemic metadata by a mid-level ontology of cognitive processes," in Proceedings of JOWO 2022: p. 2-CAOS, Aachen: CEUR-WS, 2023.

[4] M. Horsch, S. Chiacchiera, G. Guevara Carrión, M. Kohns, E. Müller, D. Šarić, S. Stephan, I. Todorov, J. Vrabec, B. Schembera, "Epistemic metadata for computational engineering information systems," in Proceedings of FOIS 2023, to appear, 2023.