At Johnson-Matthey, most of data-usage is based on internal documents and literature. However, the company has made use of third party and proprietary databases in the past and they have paid for data. For instance, they had access to databases from NOMAD partners at the Technical University of Denmark, where there was data on a screening they were interested in and the access was granted as a result of being part of a consortium. The modeling group at Johnson-Matthey produces DFT simulation data. In general, they have access to open-source literature and data without any particular restriction. At the moment, the company are having discussions on how, and if, to start sharing or using the data they produce and thinking of possible solution for merging with open-source databases.

Johnson-Matthey are very interested in the use of machine-learning and big-data analytic tools and are open to ways of starting integrating such tools into their research practice. In general, they are mainly interested in applying the tools NOMAD offers on their own data both for curation and analytics. However, they are also thinking of possible ways of merging the data they generate with the data on the NOMAD Repository as well as experimental data and data from literature.

Data-usage from the NOMAD Archive in the current format would be successful for the company, as they have no problems with accessing open-source data. A consultancy-style platform would also be good as the company has have already interacted with consultants in the past although other issues related to licensing might arise for NOMAD interactions.

In terms of uploading their own data into NOMAD Archive, the company would prefer to keep the data on its own platform. One proposition would be to obtain the NOMAD Archive and tools as an external package that they can use internally on their own systems.

Johnson-Matthey would have no problem accessing data from NOMAD Archive in its current open-source format. In order for them to start integrating more NOMAD data-usage in their research practice, they would need data to be more specific to their research interest, such as surface data and 2D materials, classical molecular dynamics (MD). In general, traceability of the data they use (papers and authors) is important, as well as being able to have all needed information for reproducing the data (parameters and technical details). They do not have a specific practice for validation but it is quite crucial to have enough information for reproducibility.

As long as the data available and the tools developed are in line with the company’s research interests, they are very happy to allocate time to training, workshops and interactions with NOMAD. Indeed, a representative from Johnson-Matthey, Dr. Crispin Cooper, attended the third NOMAD Industry Meeting in Windsor in February 2018 and will attend NOMAD Summer in September 2018.

http://nomad-coe.eu/ @NoMaDCoE www.facebook.com/nomadCoE