
Glass Ontology and Machine Learning for Data-driven High-throughput Glass Development

Marek Sierka^{*1}, Felix Arendt¹, and Ya-Fan Chen¹

¹Friedrich Schiller University Jena – Germany

Abstract

Glasses are key components of current and future advanced technologies. Compared to other classes of materials, they offer constant adjustability of their chemical composition, which allows precise adjustment of physical and chemical properties within wide limits. Current challenges for glass as a material lie in the imminent conversion of melting technology to CO₂-neutral glass raw materials and a significant increase in glass strength. However, the resulting demands on the efficiency of glass development cannot be met at present. In addition to industrially manufactured products on a ton scale, glass is also produced in research largely empirically and at best partially automated. The manufacturing processes are time-consuming and not very reproducible, and deviations in the process parameters are unavoidable. Chemical formulation is mostly based on empirical values, empirical correlations or trial-and-error approaches. Data-based methods are not established.

The GlasDigital project (1), part of the German Platform MaterialDigital (PMD) initiative (2), addresses these problems and aims to develop an ontology-based digital infrastructure for data-driven high-throughput glass development. A glass ontology is a core component of this project. It was created by adapting the PMD core ontology to glass materials and includes material data as well as manufacturing and analysis parameters, methods and results. It also provides a common glass terminology that can be used for data extraction from literature or patents.

In addition to developing the glass ontology and creating a consistent knowledge base of glasses and glass-like materials on this basis, the project also aims to use the collected data to model glass properties using machine learning (ML) tools. To this end, we have developed descriptors based on ab initio simulations of glass components. These ab initio descriptors show very good performance in ML models trained on both large and smaller data sets that include a variety of oxides, chalcogenides, and metallic glasses. More advanced descriptors have also been developed using the glass structure information obtained from high-throughput molecular dynamics simulations, which are used to improve the predictive performance of the ML models. By treating the simulated glass models as 3D images, we applied image recognition methods such as b-variational autoencoders to extract compressed structural information. This combined approach of novel descriptors and consistent ontology for glass materials will be a valuable tool for accelerated glass design.

- (1) GlasDigital: Data-driven workflow for accelerated glass development, <https://materialdigital.de/project/4>
(2) Platform MaterialDigital: The Material Digitalization Platform, <https://www.materialdigital.de/>

^{*}Speaker

Keywords: glas ontology, machine learning, ab initio simulations, high, throughput glass development