

Data-centric materials science: identifying "materials genes" with AI

Lucas Foppa^a

^a*The NOMAD Laboratory at the Fritz Haber Institute of the Max-Planck-Gesellschaft and IRIS-Adlershof of the Humboldt-Universität zu Berlin, Faradayweg 4-6, 14195 Berlin, Germany.*

The performance in heterogeneous catalysis is an example of a materials function governed by an intricate interplay of multiple underlying processes, such as the surface reaction networks, the material restructuring under the reaction, and the transport of reactants and products. In this talk, the combination of experimental and theoretical data with artificial intelligence (AI) is presented as an approach to model catalysis and determine the key physicochemical descriptive parameters ("materials genes") reflecting the processes that trigger, facilitate, or hinder the materials performance.[1,2,3] The symbolic regression and subgroup discovery AI approaches leverage the small number of materials that can be accessed experimentally and identify nonlinear correlations that can be exploited for the design of materials with statistically exceptional performance.

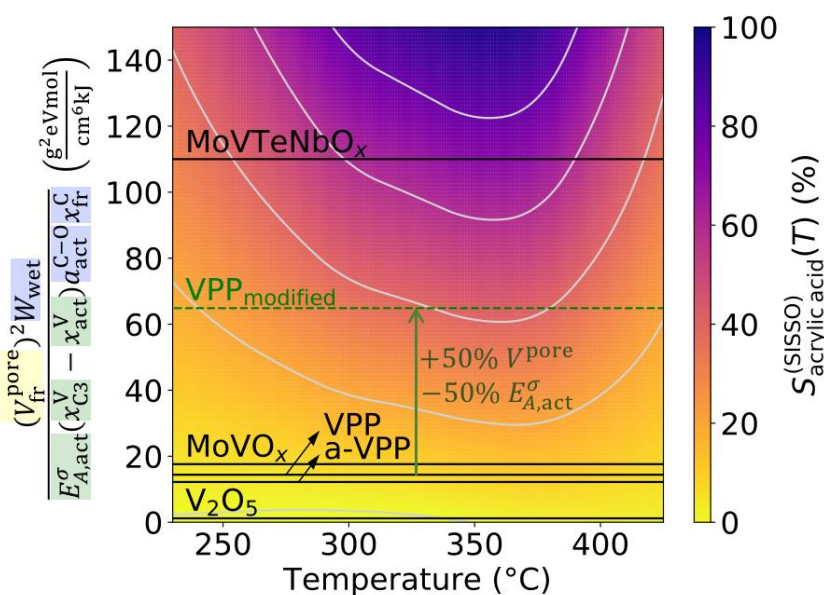


Fig. 1: Map of potential catalysts indicating the selectivity towards acrylic acid in propane oxidation catalyzed by vanadium-based materials. The black horizontal lines show the nine catalysts used for deriving the map via the symbolic-regression SISO approach. The expression on the ordinate contains the key descriptive parameters reflecting the processes governing oxygenate formation. Adapted from [1].

1. L. Foppa, L. M. Ghiringhelli, F. Girgsdies, M. Hashagen, P. Kube, M. Hävecker, S. Carey, A. Tarasov, P. Kraus, F. Rosowski, R. Schlögl, A. Trunschke, and M. Scheffler, *MRS Bull.* 46, 1 (2021).
2. L. Foppa, F. Rütther, M. Geske, G. Koch, F. Girgsdies, P. Kube, S. Carey, M. Hävecker, O. Timpe, A. Tarasov, M. Scheffler, F. Rosowski, R. Schlögl, and A. Trunschke, *J. Am. Chem. Soc.* 145, 3427 (2023).
3. L. Foppa, C. Sutton, L. M. Ghiringhelli, S. De, P. Löser, S.A. Schunk, A. Schäfer, and M. Scheffler, *ACS Catal.* 12, 2223 (2022).