LABORATORY FOR CHEMICAL TECHNOLOGY
Technologiepark 125, 9052 Gent, Belgium

Coach	**Supervisor**	**Funding**
Pedro Mendes	Joris Thybaut Pedro Mendes	UGent (BOF)

Recognition of key catalyst properties via data science

Aim

Develop a methodology to recognize the most relevant catalyst properties from experimental data in an automated manner.

Justification

The industrial implementation of sustainable, ultimately circular, processes requires the development of novel catalysts able to catalyse these new reactions efficiently. To design such catalysts, the first step is the **identification of the catalyst properties** that play a role in the reaction. Typically, reaction experiments with catalysts of diverse properties are performed at a laboratory-scale. Based on the his/her prior knowledge and experience, the researcher will focus on a few variables in the results (e.g. maximum selectivity towards desired product) and try to identify the properties that influence those variables the most. In short, the recognition of the relevant catalyst properties is often carried out intuitively based on a relatively small part of the whole information available.

This work aims, thus, at developing a methodology for **automated extraction of key catalyst properties** from a set of experimental (open) data. For this purpose, the tools of data science can be particularly useful.

Program

Develop a tool with the capability of extracting key catalyst properties from experimental data. In practice, this means the identification of cause-effect relationships between a set of catalyst properties and the catalyst performance on a given reaction. To do so, statistical tools, recently named under the umbrella term “machine learning”, can be employed. In particular, unsupervised classification (e.g. PCA, k-means clustering) and causal inference techniques are suited to evaluate such cause-effect relations without user intervention, i.e. enabling true automation [1]. Nevertheless, inspiration can also be drawn from the application of supervised learning to the extraction of kinetic information [2]. The developed algorithm will expand and complete a recently developed Python tool to extract automatically kinetic information from experimental (open) data [3].

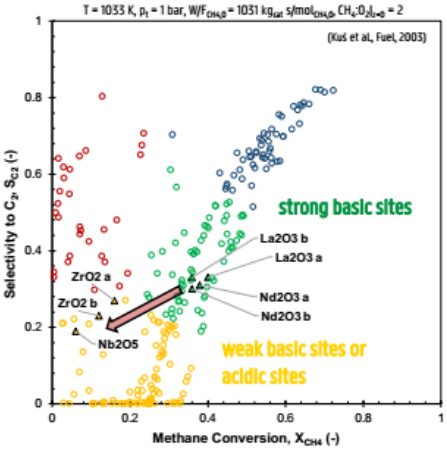

Figure: example of machine recognition of catalysts with similar performance (different colours). In this case, the identification of the key property (basicity) was still done by the researcher [4].

[1] L. Chiang, B. Lu, I. Castillo, *Annual Review of Chemical and Biomolecular Engineering*, 2017, **8**, 63.
[2] J. N. Cawse, M. Baerns, M. Holena, *Journal of Chemical Information and Computer Sciences*, 2004, **44**, 143.
[3] S. Siradze, Automated kinetic feature extraction from Open Access data, MSc thesis, 2019, Ghent University.
[4] L. Pirro, P. S. F. Mendes, S. Paret, B. D. Vandegehuchte, G. B. Marin and J. W. Thybaut, Catalysis Science & Technology, 2019, 9, 3109-3125.

GHENT UNIVERSITY
LCT DRIVING CHEMICAL TECHNOLOGY