SHARED METADATA AND DATA FORMATS FOR BIG-DATA DRIVEN MATERIALS SCIENCE: A NOMAD-FAIRDI WORKSHOP

IRIS Adlershof, Berlin, Germany, July 8th to 12th, 2019

Specifics of metadata and design of experiment in heterogeneous catalysis research

Annette Trunschke

Department of Inorganic Chemistry, Fritz-Haber-Institut der Max-Planck-Gesellschaft, Berlin, Germany

Catalysis: systemic challenge

Theorist's simplified view (Paul Saxe):

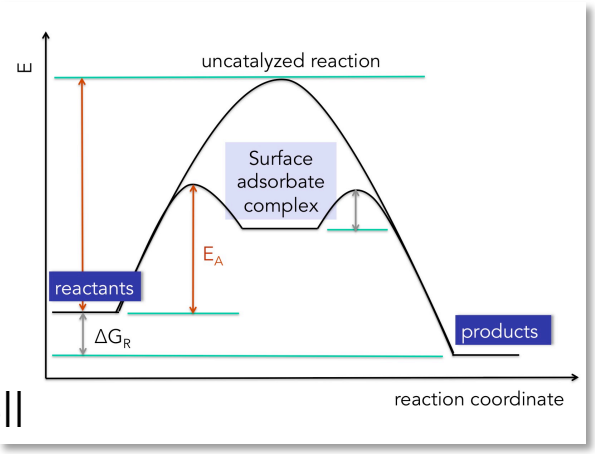
"Experiment has real problems:

- What is the sample, exactly?
- What was actually measured?"

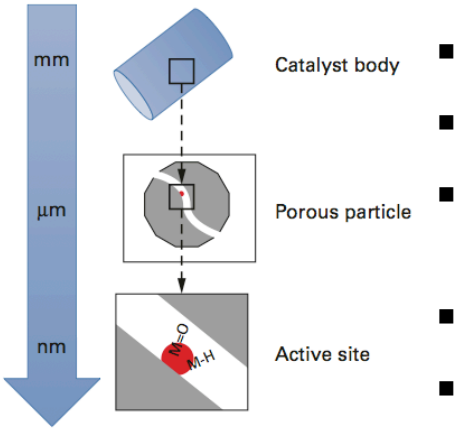
Situation in catalysis (or functional characterization) is even worse:

- The functional properties are determined by stochastic entities rather than stable features
- Materials properties need to be analyzed under operation
- Multidisciplinary approach necessary

Catalysis: systemic challenge



- The catalyst changes the reaction mechanism by interacting with the reacting molecule(s), but not the thermodynamic equilibrium: a phenomenon under kinetic control
- Mutual interaction causes changes of the catalyst as well
- Heterogeneous catalysis occurs at interfaces as functional unit: solid state versus molecular control?



- Multi-dimensional in space and time
- Selectivity requires control over complex reaction networks
- Chemical dynamics: interplay between kinetics and solid-state reactions
- Integrated workflow between theory - experiment – engineering
- Catalysis research requires collaboration

BasCat – FHI collaboration: Concepts in Heterogeneous Catalysis

Fritz-Haber-Institut
Inorganic Chemistry
Theory

BASF
We create chemistry

Technische
Universität
Berlin

unisyscat

Ambition:

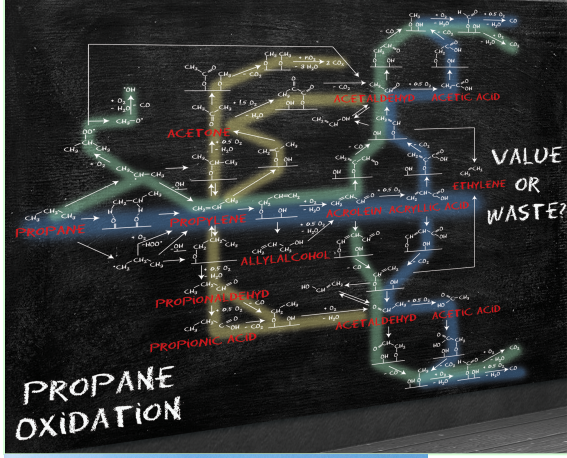
Predicting high performance based on physical insight

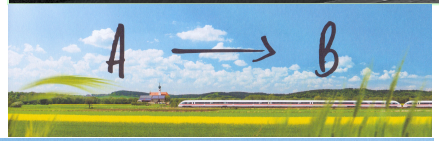
- Materials
- Reaction conditions
- Predictions by theory

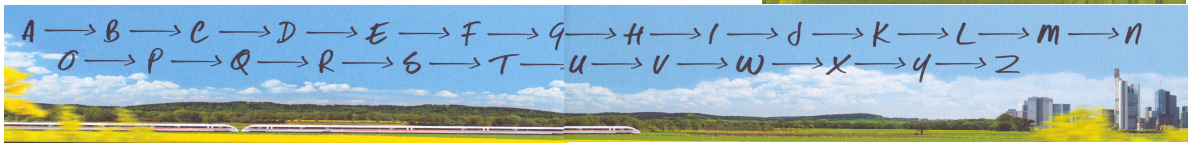
Descriptors in catalysis

- Identifying causal structure-function relations
- Target: high rate of formation of a desired product
- Property "rate" is a challenge as dependent on reaction networks and falsified by transport phenomena

$$X_i = \frac{n_{i,0} - n_i}{n_{i,0}} \qquad S_{k,i} = \frac{n_k}{n_{i,0} - n_i} \cdot \frac{|\nu_i|}{|\nu_k|} \Rightarrow r = \frac{1}{\nu_i} \frac{dn_i}{m_{cat} dt}$$





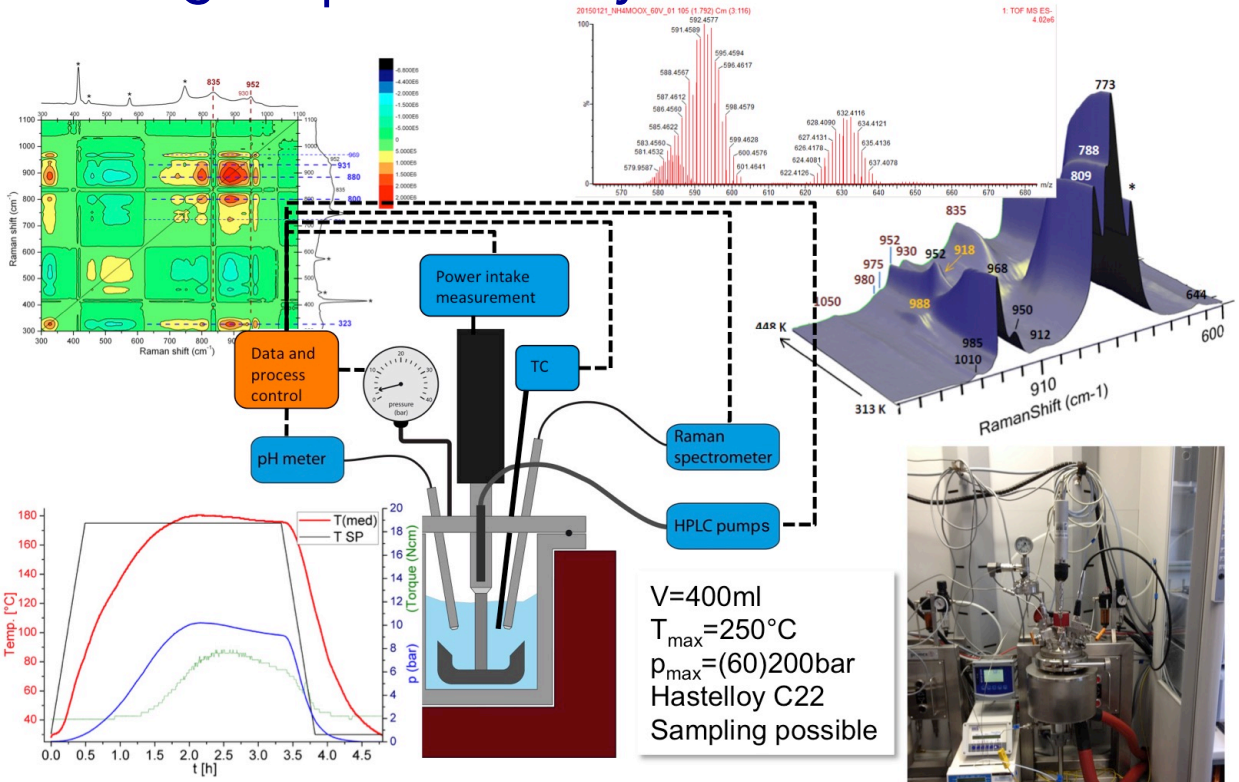


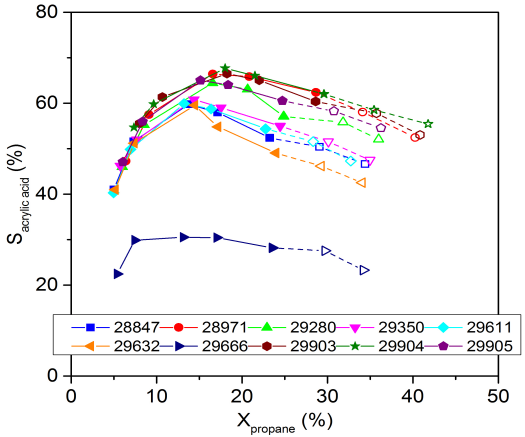
- Structure is also challenging as dynamical and with unknown number of "active sites"

$$TOF\left[s^{-1}\right] = \frac{N_i}{N_{sites} \cdot t}\left[s^{-1}\right]$$

- Difficult or impossible to measure in heterogeneous catalysis because N_{sites} is unknown or fluctuating

Challenge: reproducible synthesis



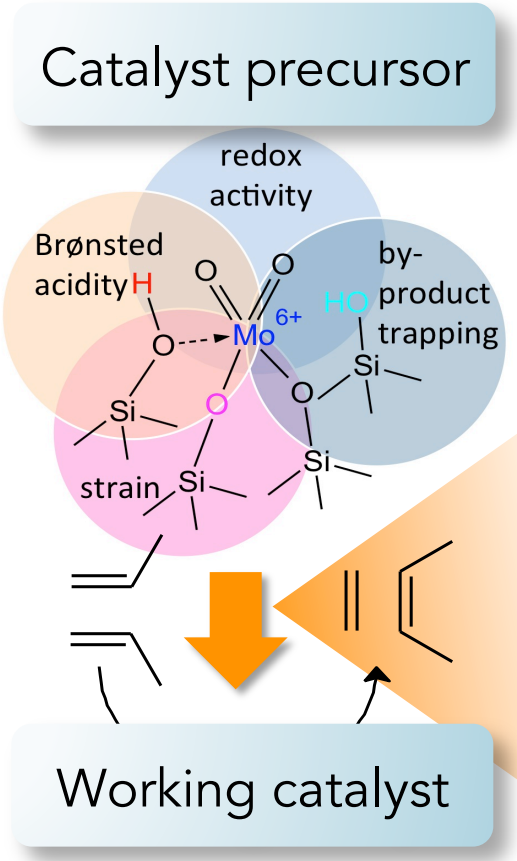


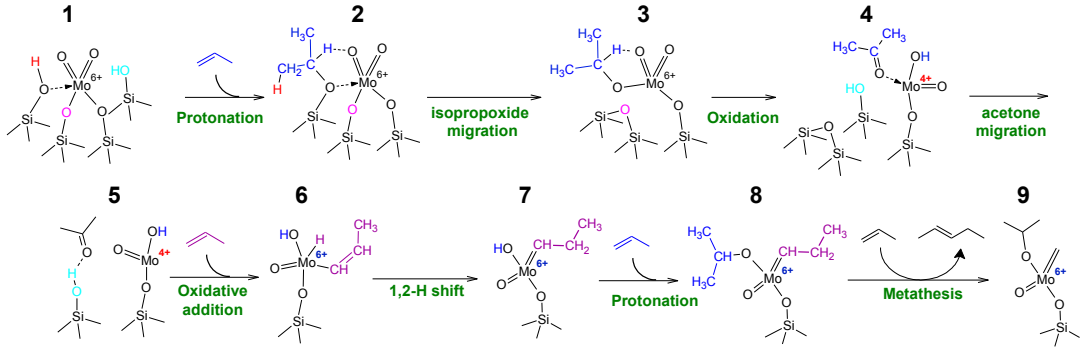
HPM-PT-040 (Premex Reactor GmbH)
Kaiser Optics Raman Spectrometer RXN1

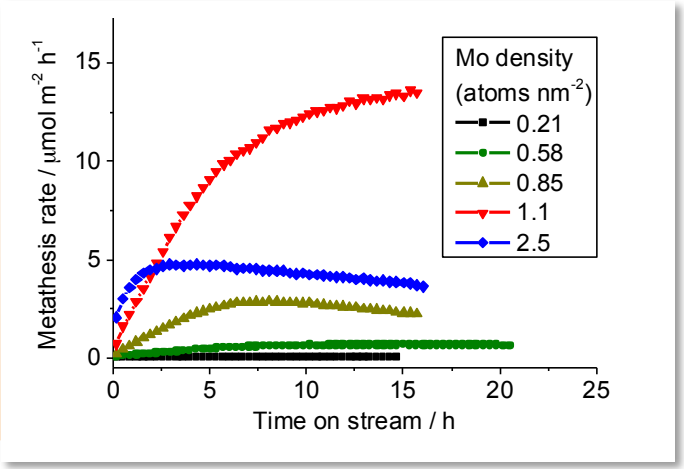
- Automated synthesis reactors deliver synthesis metadata
- Difficulties in automation:
  - Separation and washing steps
  - Activation phenomena

Zeitschrift für anorganische und allgemeine Chemie 2014, *640*, 2730-2736.

Challenge: activation – deactivation phenomena







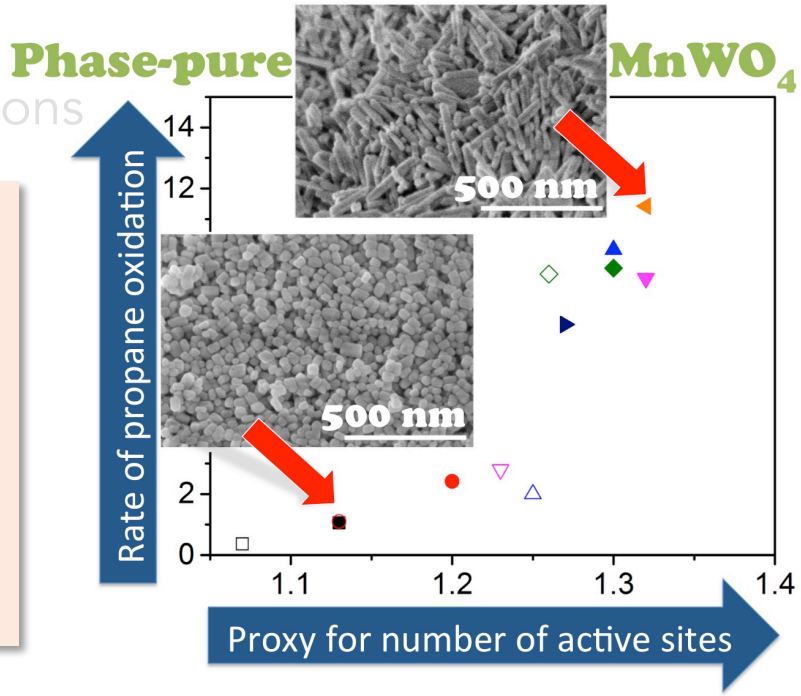
- Steady-state criterion
- Metadata will not help to detect violation if single rates are provided
- Agreements about procedure of kinetic studies necessary

JACS, **2012**, *134*, 11462-11473; *Angew. Chem. Int. Ed.* **2013**, *52*, 13553-13557. *ChemCatChem*, **2015**, *7*, 4059-4065.

Challenge: (Dynamic) interface

- Surface properties essential
- Dynamical surface nature
- Reaction network
- Activation energy is a function of conditions

- Same crystal structure of all catalysts in the series
- Identical chemical bulk composition
- Catalytic properties depend on surface composition
- The bulk crystal structure of an oxidation catalyst as the most popular descriptor in oxidation catalysis is not solely responsible for catalytic performance

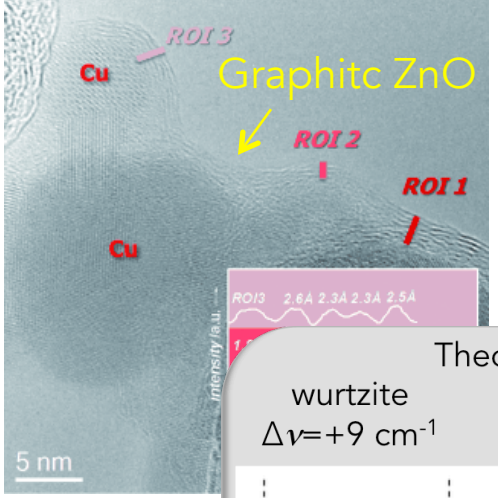


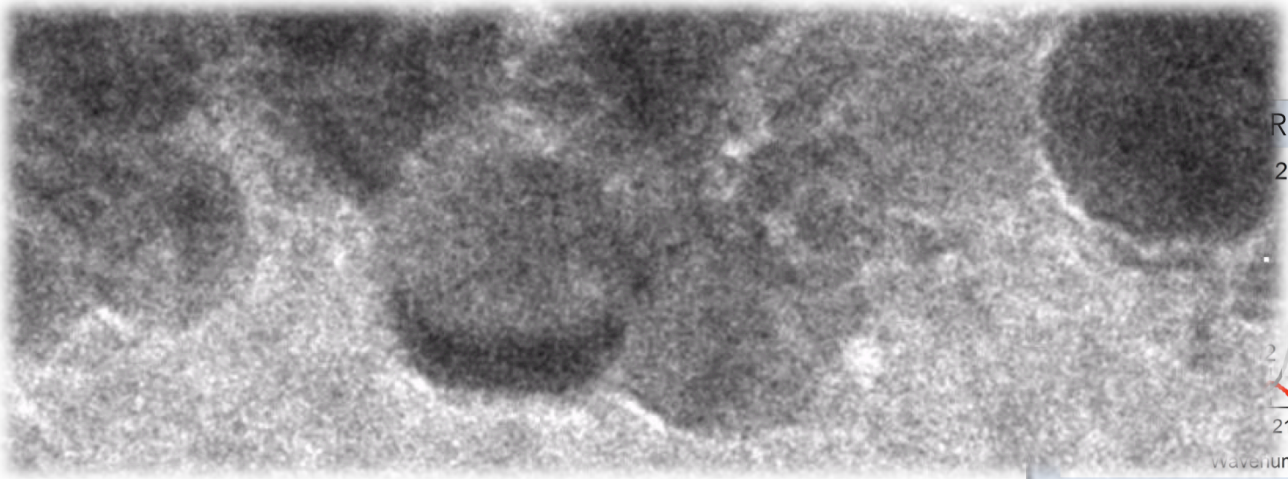
Chemical Science, **2019**, *10*, 2429–244 .

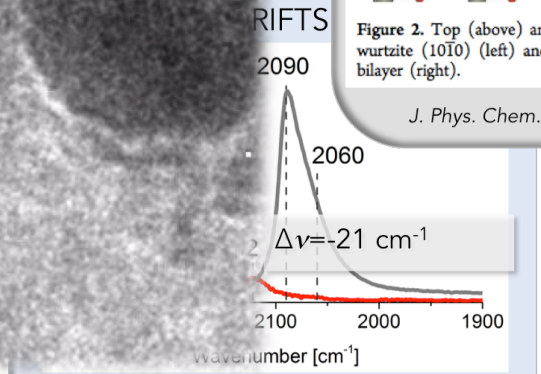
Challenge: (Dynamic) interface

- Surface properties essential
- Dynamical surface nature
- Reaction network
- Activation energy is a function of conditions

Reverse water gas shift
250°C,
1:1:8 CO_2:H_2: He
65 mbar







300 K, 0.5 ML

Top. Catal. **2017**, *60*, 1735–1743.

Theory

wurtzite	graphitic ZnO
$\Delta\nu$=+9 cm^{-1}	$\Delta\nu$=-23 cm^{-1}

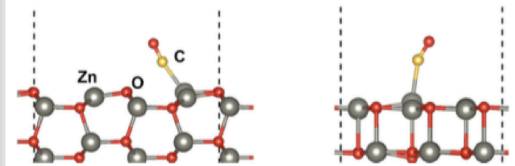


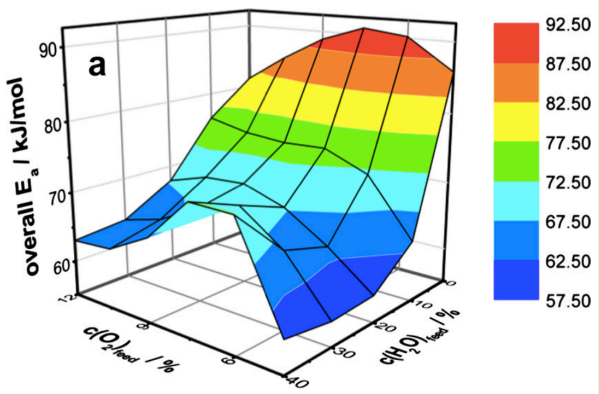
Figure 2. Top (above) and side views (below) of (a) CO adsorbed wurtzite ($10\bar{1}0$) (left) and (b) CO adsorbed on unsupported ZnO bilayer (right).

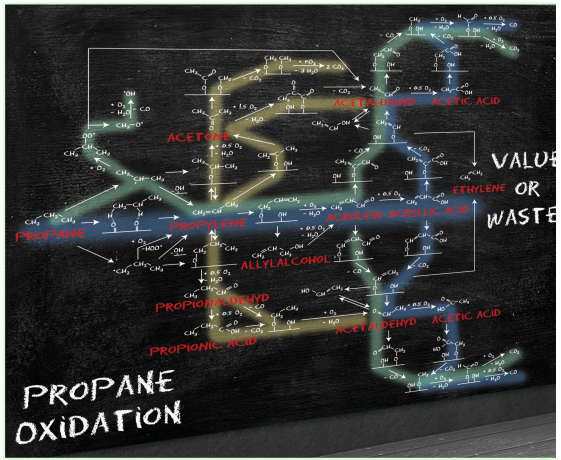
J. Phys. Chem. C **2017**, *121*, 27453–27461.

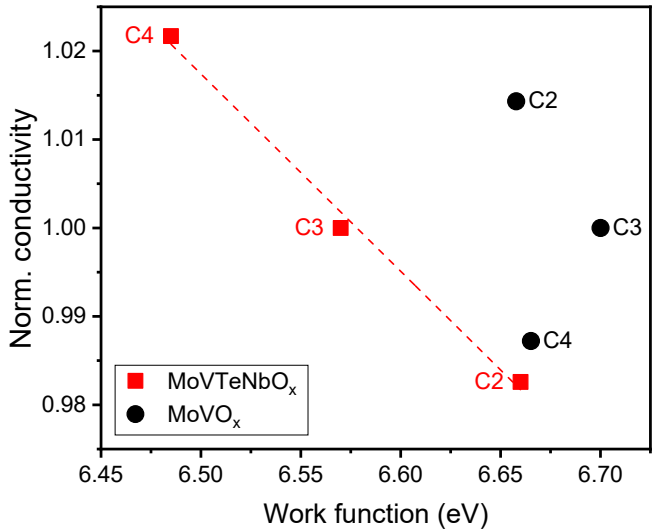
Angew. Chem. Int. Ed. **2015**, *54*, 4544–4548.

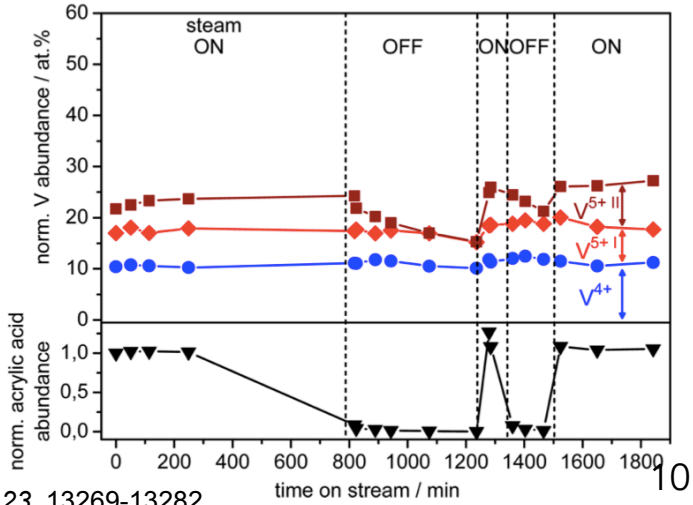
Challenge: (Dynamic) interface

- Surface properties essential
- Dynamical surface nature
- Reaction network
- Activation energy is a function of conditions









Challenges in catalysis

- The rate depends on the reaction network and may be falsified by transport phenomena, which is not always detectable, even if all metadata of the rate determination are stored
- The nanostructure and bulk real structures control the ability of a material to respond to the local chemical potential
- Active sites occur locally during the reactant-induced instability of the reactive surface phase
- We hope that properties of the precursor are reflected in the properties of the working catalyst – and we could check this with the help of machine learning
- But, we should better generate data and metadata of rate determination and physical characterization in one experiment simultaneously – in operando

The concept of clean data in catalysis

- Data will have to be
  - Well documented by meta data
  - Stored in a findable way
  - Documented and stored in a re-useable form
  - Clean enough to be interoperable with later and external analysis
  - Complete enough to allow reconstruction of the experiment
- This will eventually be a legal requirement coming with grants
- Users need to share much more information than in papers with a much better formalized organization
- The general availability is to the benefit of all as we can collaborate and test new hypotheses much easier than today

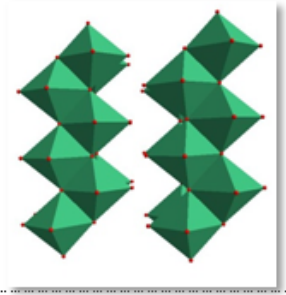
The concept of clean data in catalysis

- Standardisation of experiments
- A handbook of selective oxidation was developed
- A set of 10 structures was defined to be tested in three different (ethane, propane, butane) reaction networks
- This combination covers a wide space of structures and function comprises the necessary diverseness of data
- All kinetic operations are validated against material standards
- The kinetic test procedure and the total history of each sample is defined
- Data are stored in an in-house database
- Additional experiments are required to understand the relation results

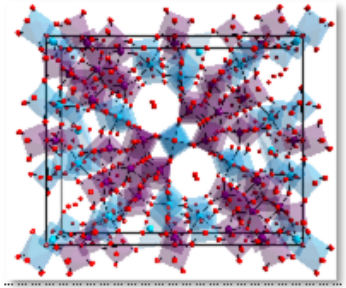
Data Science Project in Oxidation Catalysis

Generation of a Reference Data Set

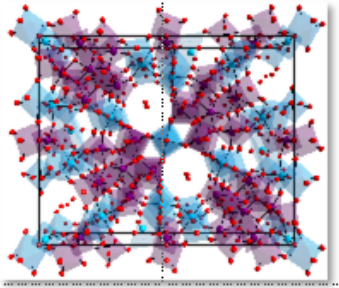
Catalysts



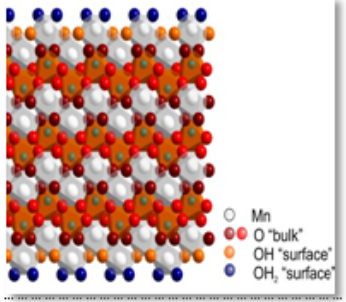
MoO_3



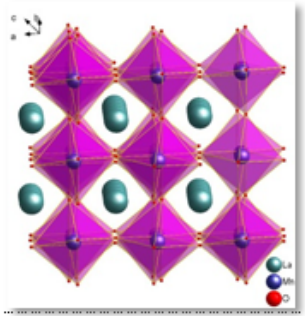
Mo-V oxide



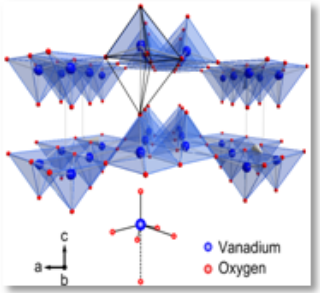
Mo-V-Te-Nb oxide



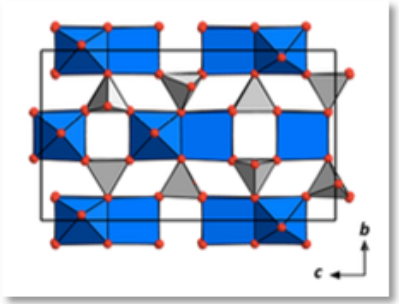
$MnWO_4$



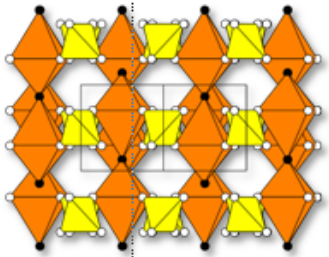
$SmMnO_3$



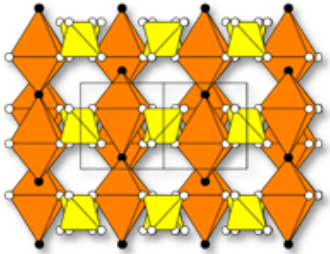
V_2O_5



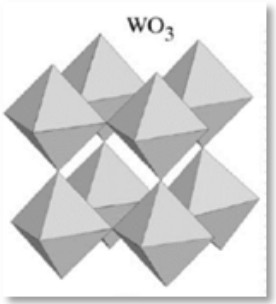
VPP/VPO



α_{II}-$VOPO_4$



α_{II}-$V_{0.8}W_{0.2}OPO_4$

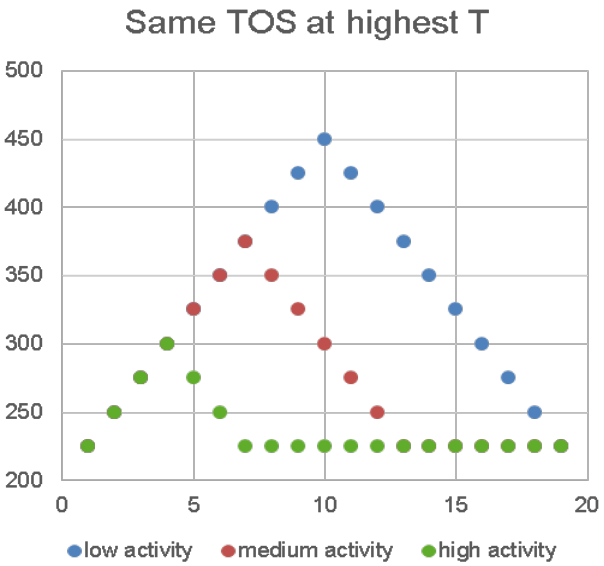


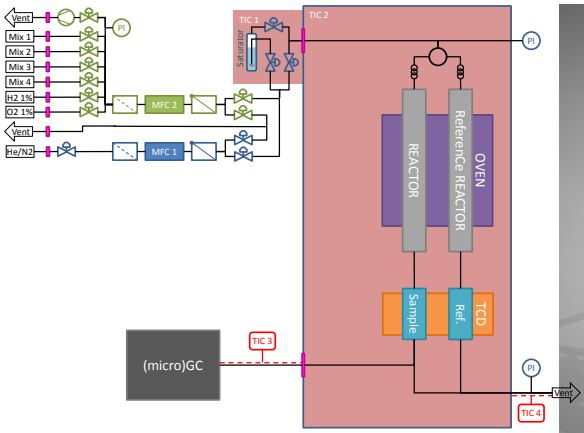
$V_{0.167}W_{0.5}P_{0.333}O_x$

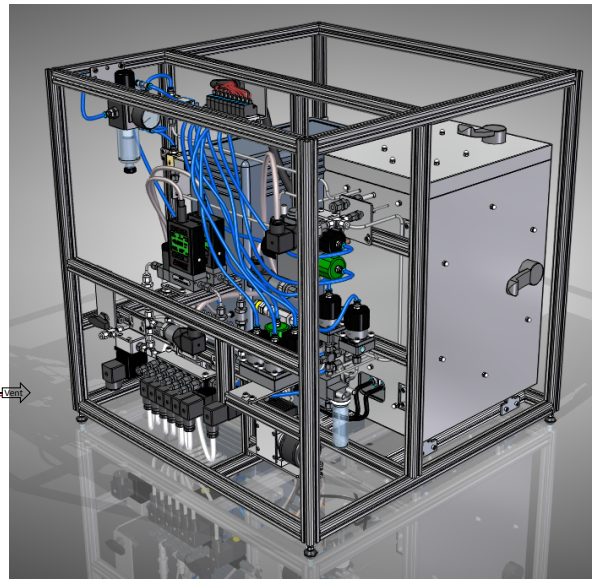
Catalytic testing

Oxidation of ethane, propane, *n*-butane

1. Temperature variation
2. Contact time variation
3. Determination of rates in different feeds







Labview based process control with real time CompactRIO controller

Catalytic testing

Catalyst ID	Feed	r (300°C) [mmol g^{-1} h^{-1}]	r (300°C) [mmol m^{-2} h^{-1}]	E_a [kJ mol^{-1}]	TR [°C] X=30%	r (TR) [mmol g^{-1} h^{-1}]	r (TR) [mmol m^{-2} h^{-1}]	r_{CO2} (TR) [mmol m^{-2} h^{-1}] X=10%	r_{CO2} (TR) [mmol m^{-2} h^{-1}] X=30%	r_{CO2} (TR) [mmol m^{-2} h^{-1}] X=60%
1	Standard									
	Less O_2	-	-	-	-					
	ODH	-	-	-	-					
	Fuel rich	-	-	-	-					
	5% water	-	-	-	-					
	10% water	-	-	-	-					
	20% water	-	-	-	-					
2	Standard									
	Less O_2	-	-	-	-					
	ODH	-	-	-	-					
	Fuel rich	-	-	-	-					
	5% water	-	-	-	-					
	10% water	-	-	-	-					
	20% water	-	-	-	-					
...										

Training values P_i - Descriptive features

Property	Physical value	Methods
Bulk crystal structure (ICSD#)	Lattice constants *a*, *b*, *c* [Å], α, β, γ [°] Unit cell volume *V* [nm^3] crystallographic density $\rho_{crystal}$ [g/cm^3]	XRD, TEM
Bulk chemical composition	c_M, c_O [at%]	XRF, ICP-OES, EDX
Surface composition	c_M, c_O [at%] surface active site density [nm^{-2}]	XPS
Binding energy, surface oxidation state	B.E. [eV], M^{n+}	XPS
Valence band onset, M nd onset, *E* secondary electron cutoff, work function, maximum band bending at the surface "s" (surface potential barrier), change in electron affinity	E_{VB}, E_{nd}, E_{cutoff}, Φ, eV_s, ΔX [eV]	XPS, NAP-XPS, ResPES
Absorption edge energy	O K-Edge, V L_3-Edge [eV]	NEXAFS

Training values P_i - Descriptive features

Property	Physical value	Methods
Specific surface area (BET) Total pore volume, mesopore volume	S_{BET} [m^2/g] V_{total}, V_{mp} [m^3/g]	Nitrogen adsorption
Heat of adsorption Ads. capacity of reactants (C_nH_{2n+2}, O_2)	ΔH_{ads} [kJ mol^{-1}] n [mol/g]	Microcalorimetry
Conductivity, real part of permittivity ε', imaginary part of permittivity ε'', Apparent activation energy of conduction	σ [S/m], ε', ε'', E_c [eV]	MCPT
Optical edge energy	E_{edge} [eV]	UV/Vis
Oxidation/reduction equivalents	Oxygen defect density [$V_o^{\cdot\cdot}/nm^2$]	TPR/TPD
Desorption temperature	T [K]	TPD
^{18}O exchange temperature as function of feed, M-O stretching frequencies	T [K] ν [eV]	Raman
Activity CO oxidation	T_{10}, T_{50} [K]	Fixed bed reactor

Conclusion

- Clean experimental data reveal the complexity of describing “the function” of a material
- Descriptors are needed that involve both, kinetics of reactants and catalyst material
- The handbook presented and the resulting clean data serve as input base
- In addition to metadata storage, some compulsory procedures need to be established and control mechanisms need to be introduced
- Potential benefit from AI applications provides motivation
- A reconstruction mechanism based upon the descriptor for interpolating missing data in literature reports can enhance and enlarge the database very significantly

Acknowledgements

FHI, Department of Inorganic Chemistry

Luca M. Ghiringhelli, Lucas Foppa, Matthias Scheffler (FHI, Theory)

Frank Girgsdies, Sven Richter, Sabine Wrabetz, Maike Hashagen, Jutta Kröhnert, Yuanqing Wang, Gregory Huff, Olaf Timpe, Detre Teschner, Spencer Carey, Michael Hävecker, Liudmyla Masliuk, Walid Hetaba, Thomas Lunkenbein, Andrey Tarasov, Pierre Kube, Ezgi Erdem, Giulia Bellini, Peter Kraus, Robert Schlögl (FHI, AC)

Raoul Naumann, Michael Geske, Stephen Lohr, Benjamin Frank, Frank Rosowski (BasCat)

BasCat
UniCat BASF SE JointLab

BASF
The Chemical Company