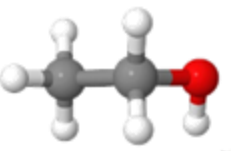
GENERATIVE AND PREDICTIVE MODELS FOR NEW MATERIALS

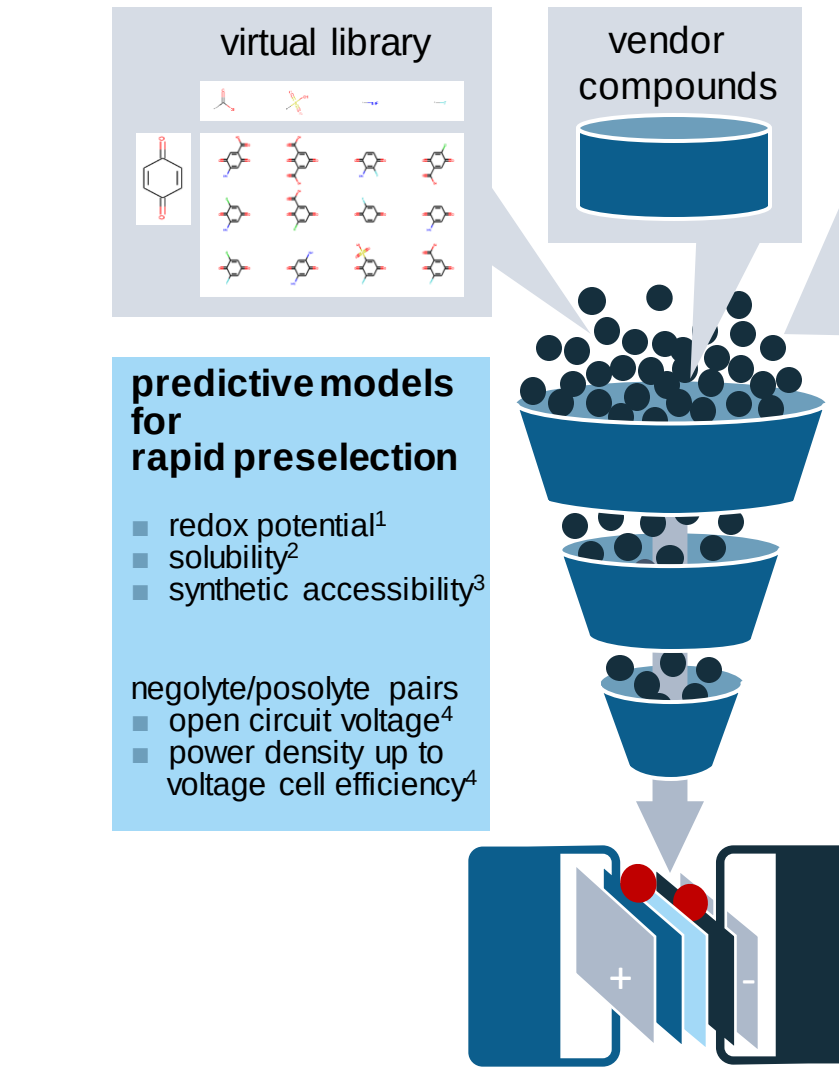
Niklas Dobberstein, Jan Hamaekers, Astrid Maass
Fraunhofer Institute SCAI, Schloss Birlinghoven,
53757 St. Augustin, Germany

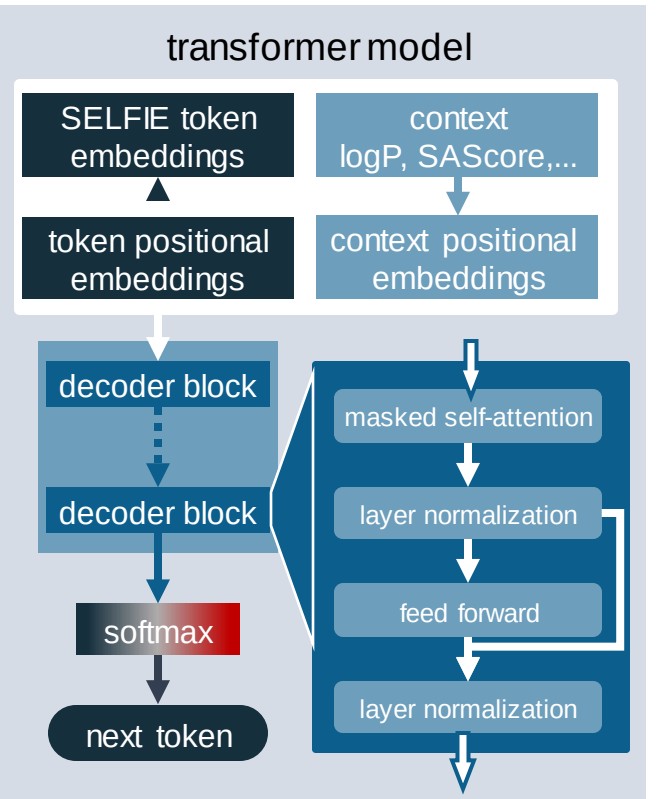


SELFIE: [C][C][O]

C	C	O
C	C	O
C	C	O

masked self attention matrix





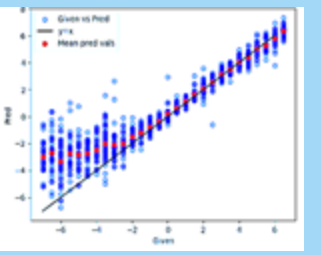
predictive models for rapid preselection

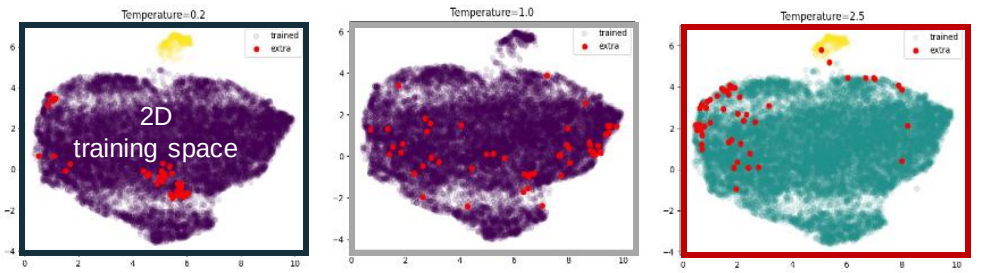
- redox potential[1]
- solubility[2]
- synthetic accessibility[3]

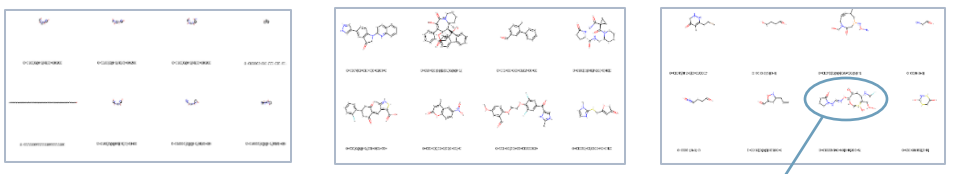
negolyte/posolyte pairs
- open circuit voltage[4]
- power density up to voltage cell efficiency[4]

generative models for new molecules[5,6]

- trained on 13M compounds[7-17]
- + temperature **control variability**
- + context **hit target property**



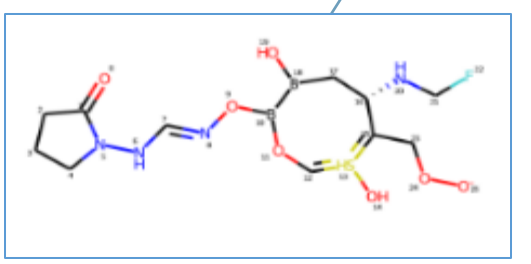




- flexible, python-based workflow, relying on **physical or data-driven models**
- selectable chemistry for redox and/or ageing reactions
- speed up by pre-trained models for PCET reactions
- adaptable sequence of filtering steps
- interface to other modules via JSON
- prototype: **redoxfox.scai.fraunhofer.de**

- outlook: **focus on relevant sub-space by multi-conditional prediction**

quality	%
novelty 10k	98.8
uniqueness 1k	98.7
uniqueness 10k	97.1
validity 10k	100



REFERENCES
1. Barker, Berg, Hamaekers & Maass, Batteries & Supercaps, 2021
2. Sorkun, Koelman & Er, IScience, 2021
3. Ertl & Schuffenhauer, Journal of Cheminformatics, 2009
4. Mourouga, Schaerer, Yang, Janoschka, Schmidt & Schumacher, Electrochimica Acta, 2022
5. Bagal, Aggarwal, vinod, Prijakumar, J.Chem.Inf. Mod., 2021
6. Vaswani, Shazeer, Parmar et al., arXiv: 1706.03762
7. Ramakrishnan, Dral, Rupp & von Lilienfeld, Scientific Data,2014,
8. Ruddigkeit, van Deursen, Blum & Reymond, J. Chem. Inf. Mod. 2012
9. Glavatskikh, Leguy, Hunault, Cauchy & Mota, J.Cheminf, 2019
10. Sterling & Irwin, J. Chem. Inf. Mod, 2015
11. Gómez-Bombarelli et al. ACS central science, 2018
12. Sorkun, Zhang, Khetan Sorkun, Er, Scientific Data, 2022
13. St John et al. J. Chem. Phys., 2019
14. Nakata, Shimazaki, Hashimoto & Maeda, J. Chem. Inf. Mod., 2020
15. Nakata & Shimazaki, J. Chem. Inf. Mod., 2017
16. Hachmann et al. J. Phys. Chem. Lett., 2011
17. Mendez et al., Nucl. Acids Res, 2019

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 875489.