"Data Science for Sustainable Construction Materials" – Challenge for TUM Science Hack 2021

Alisa Machner[1] and Torben Gädt[2]

[1]Mineral Construction Materials, Department of Civil, Geo and Environmental Engineering
[2]Chemistry of Construction Materials, Department of Chemistry

1 Motivation

Modern cities are constantly being rebuild. Old structures are teared down, new structures are constructed. In our developing world, increasing urbanization will lead to massive growth of urban population and therefore building infrastructure. The main function of structural materials in general and construction materials in particular is to resist or support external loads. Due to the huge scale of modern construction, the annual production amount of cement and concrete makes them the largest man-made material by a large margin. The global CO_2 footprint of concrete amounts to approx. 8 % of global annual emissions. Additionally, the resource consumption of high quality aggregate is immense and contributes to beach loss and significant damage to river eco-systems. It was recently reported that human made infrastructure surpasses the entire biomass on earth.[1] It is therefore an urgent challenge to significantly improve the resource efficiency of concrete.

2 Outline of the Challenge

We (the Machner lab and the Gädt lab) are working on developing new concepts for sustainable binders which will enable ultra-low CO_2 building materials and facilitate additive manufacturing (i.e. 3D printing of concrete). In doing so, our labs at TUM are generating data every day. However, most of the data is stored in csv files or spreadsheets. Experience has shown, that the half-life time of this datatype is rather short. This wastes the opportunity of long term accessibility and utilization of precious data. TUM has developed an open source electronic lab journal called "TUM Workbench", which is based on the Python Webframework Django. The Workbench is a valuable tool but can currently not be used for customized data entry. The Workbench is therefore used for uploading csv files and spreadsheets and it ensures proper metadata management and long term accessibility.

Therefore, our proposed challenge is the development of a prototype web based database application (preferably Django and PostgreSQL), which would enable our staff and students to directly store their experimental data in a proper database. The prototype should be user friendly (i.e. simple graphical user interface) and should allow for easy upload of raw instrument data (i.e. csv files), filtering, sorting and also data retrieval. The most common data type are simple tables of instrument raw data.

Calorimetry Data A very important instrument in daily use in our labs is a heat flow calorimeter. This device measures the heat liberated by the the chemical reaction of inorganic binders, such as cement, as a function of time. Unfortunately, we are currently losing a lot of data, due to improper meta-data management and csv files, which are scattered over the individual computers of our staff and students. To succeed in developing ultra-low large-scale CO_2 binders, we need to leverage the wealth of all recorded data. Therefore, we need you to help us out by creating a database prototype.

[1]E. Elhacham et al., "Global human-made mass exceeds all living biomass", *Nature* **2020**, *588*, 442–444.

3 Deliverables

3.1 Database acceptance and user model

Torben Gädt has spent 10 years in an industrial research environment at BASF and has experienced that old habits are sometimes hard to break. People actually like their free-style spreadsheets. Hence, the new prototype web-app should be designed such that it is intuitive and welcoming but also enforcing data integrity and consistency. Additionally, a prototype should implement a user model, which enables clear data accessibility rules.

3.2 CO_2 efficiency calculator

The data would be stored together with the raw material composition of each mix tested. Based on this information and the recorded heat flow data, it is possible to predict strength of the binder material together with a CO_2 efficiency score. A very nice feature would be the calculation of such a score for each experiment in the database, together with a rough estimate of achievable mechanical strength.

3.3 Self-made semi-adiabatic calorimeter

As a bonus challenge - the team could set up their own (semi-adiabatic) calorimeter using a styrofoam opening, mixing fresh cement and creating a simple data-logger using an Arduino or a Rasperry Pi. This is only feasible if some micro-electronic tinkering is possible within the framework of the challenge.

3.4 Implementation

We would prefer a Python based web framework (ideally Django, but Flask would also be an option) together with a PostgreSQL database. The user-interface should be clean and simple. Data import should accept instrument csv files and store the raw data directly in a proper database structure. Data export should deliver flat csv files ready for data analysis.

4 Follow-Up Project Ideas

A successful prototype could be followed up in a master thesis, which would then complete the database architecture, obtain new instrument data on sustainable construction materials and potentially approach machine learning on the acquired data. Furthermore, funds would be available to finance additional development and maintenance as a student assistant within the framework. Ultimately, the data framework will facilitate machine learning approaches to identify new binder concepts (as alternative to pure Portland cement) with significantly improved eco-efficiency compared to commonly used cements.

- Contact Persons: Torben Gädt, Tobias Lange (Chemistry) and Alisa Machner (Mineral Construction Materials, BGU)
- Data: Instrument data and the respective metadata can be supplied in various formats or can be created during the Hack (see self-made calorimeter)