Data Ideathon kick-off

Alain Borel &
EPFL Library Research Data team

February 20, 2023
EPFL Library’s 1st Data Ideathon
What is an ideathon?

- A brainstorming event where individuals from different backgrounds, skills and interests converge to diagnose predefined problems, identify the best opportunities and ideate the most viable solution.

In other words:
- Usually short & intensive
- Participants work in teams
- All teams must develop ideas to address designated issues
How would you use EPFL’s data (scientific, technical, administrative… you name it) to promote change? Which crazy ideas of yours could shape the campus of tomorrow?

Not an official call for projects. No official connection with the Solutions for Sustainability (S4S) Initiative. We’re here to have fun, but seriously good ideas are welcome as well!
EPFL Library’s data ideathon: how to participate

- Team up with 1 to 3 other enthusiastic people
- Imagine ideas based on data that you know, suppose or would like to exist at EPFL
- Prepare your pitch for the final event on Thursday, in the format of your choice: text, slides, images, illustrations, short videoclips, prototypes, theatrical performances, etc. *(actual implementation of the idea is NOT required)*
- Be creative! Mix the data sources! Think outside the box!
Q: Data? What data? A: It’s everywhere

- Scientific data, for example datasets uploaded to Zenodo or other repositories
Data? What data?

- IS-Academia: information about students, courses, etc.
Data? What data?

Statistiques Recherche

Voici les statistiques liées à la recherche. Veuillez consulter le glossaire recherche pour plus de précisions quant aux termes utilisés.

- Statistics, reports, surveys…
Data? What data?

**EPFL Data Center**
Eckhaus (VD)

Development of a data centre in a building volume with a minimum capacity of 200 racks and a power of 4 MW.

**Information**

- **Client**
  Ecole Polytechnique Fédérale de Lausanne, Lausanne

- **Services rendered by HRS**
  Sole contractor with full cost, quality and deadline guarantee

- **Architect**
  A&P|VIT|BER Architecten en ingenieursGbR, Brussels

**Contact**

HRS Real Estate SA
Rue du Centre 172
1025 St-Sulpice

+41 58 122 90 00
st-sulpice@hrs.ch

- Suggests operational data (power, water...), maybe also supercomputing
Idea example 1: Carbon footprint of EPFL publications

Data sources
- we know EPFL collects data to measure its carbon footprint: https://www.epfl.ch/about/sustainability/our-climate-your-campus-faq-en/#faq-item-55ff42df9d5d91ea780dac4d87f889
- we know EPFL produces scientific publications and records them by faculty, institute, laboratory: https://infoscience.epfl.ch/collections/?c=Infoscience%2FResearch
- we know EPFL laboratories can be located in buildings: https://search.epfl.ch/?filter=unit&q=EPFL

The idea
- Hypothesis 1: CO2 footprint data can be calculated at the level of a building (possible at least for power consumption)
- Hypothesis 2: producing scientific publications is EPFL’s main activity (wrong, but let's pretend just for fun)
- Hypothesis 3: each lab only produces CO2 within its own lab (probably wrong quite often)

- By combining the 3 data sources, one could produce a map displaying the average footprint of 1 publication by building.

Evaluation
- originality: average
- feasibility: probably OK
- presentation quality: weak in the current format 😞
- relevance: low (unlikely hypotheses + no clear outcome, sustainability-wise; should we just close the buildings with the highest per-publication footprint?)
Idea example 2: Daily commute optimizer

Data sources
- we know EPFL collects staff and student statistics
  https://www.epfl.ch/about/overview/fr/statistiques-institutionnelles/statistiques-personnel/ and
  https://www.epfl.ch/about/overview/fr/statistiques-institutionnelles/statistiques-education/; individual data include
  the permanent residence (and work place for staff; which campus
  is probably sufficient)
- we know EPFL performs yearly mobility surveys
  https://www.epfl.ch/campus/mobility/mobilite/mobility-survey/
- we know that several online solutions to calculate itineraries
  are available, some of them offer an API

The idea
- Hypothesis 1: we are allowed to obtain matching lists of
  residence and work/study places (personal names are not
  necessary, but still tricky due to privacy laws: precision should not
  be too high: work/study campus and postal number of residence,
  maybe?)
- Hypothesis 2: we are allowed to know how many days per year
  each person would travel to his/her destination campus

Using the API, one could obtain the yearly commute distance and
time for each person, ideally for all possible travel options which
would allow cost/carbon footprint/etc. estimations. A minimal
average over all EPFL members could be compared with the results
of the usual surveys.

Evaluation
- originality: not high, perhaps already done?
- feasibility: permissions will be a problem unless some
  work is delegated to someone who is already allowed
  to access the data
- presentation quality: too wordy but at least we have a
  picture!
- relevance: good?
How to join

- Teams of 2-4 participants
- Open to all EPFL (at least 1 team member; open to students, researchers, admin & technical staff). The Library staff and their families are excluded
- Register your group (names & e-mail addresses) by mail to researchdata@epfl.ch until Wed. Feb. 22 at noon (Lausanne time)
How to win

- Send us a title on behalf of your group until Feb. 23, 9:00 AM
- Present your team’s idea in 3 minutes during the ceremony on Thursday, Feb. 23, 12:00-14:00 (in French or English)
- Our jury will evaluate the ideas based on presentation, originality, feasibility, relevance to the themes… and pure fun
- 3 winning projects will receive rewards worth 500, 300 and 200 CHF (REKA vouchers [https://reka.ch/en](https://reka.ch/en))
Have fun!

Surprise us!

We look forward to your presentations on Thursday
Do not forget: sign up for Thursday, too

(if you want to participate or if you only want to join as a spectator)