

# Spring25 Final Project – Lead generation data solution

## Problem statement

Currently the sales team is using their (extensive) networks and LinkedIn Sales Navigator as their basis of contacting potential customers. The annual meeting counts are high, several hundred potential clients met each year, but still higher quantities of potential leads would likely result in more potential customers, which would then convert to more successfully employed new data engineers. The importance of high quantity of leads is especially important in a bad economic situation and market environment like currently in winter 2025, where the conversion rates are lower than usual, and quantity is the key.

## Objective

To build a customer lead generation data solution to help our sales team tackle this problem.

## Deliverables

- Power BI app that can be shared to sales team
- Microsoft Fabric workspace with warehouse/lakehouse containing the data
- Documentation of the solution
- Presentation for the end of the final project

## Suggested steps

1. Setup Trello or Jira to track progress
2. Thinking and researching: Thinking what kind of data do we need for this? What is useful? What kind of sources exist?
  - a. We do have an access to <https://www.vainu.com/> which has an API that could be very useful for extracting plenty of financial information etc
    - i. <https://developers.vainu.com/docs/authentication>

- b. Job applications reveal something as well – we have previously done a solution that scrapes job applications from online sources – those can be used (Onni will share repository)
  - c. Be creative, there could be plenty of data that would help
- 3. Gathering the data to Fabric. Workspace called lead-generation has been assigned to you, and a Github repository will be assigned as well
  - a. Automated pipelines are preferred – so that data would keep updating
- 4. Playing around with the data – can we somehow rank the companies based on some assumptions? Could be something simple, for example, that bigger companies are likely to be relevant leads for data engineering needs. Or that certain industries are. Or that companies that have people with title CIO or CDO. Or something like that.
- 5. Sharing the analysis using Power BI
- 6. Iteration