Using a domain-specific language to model data pipelines

We would like to interview you about your experience using Jayvee, a domain-specific language (DSL), to model data pipelines. We investigate if it is viable to use a DSL based on blocks and pipes for creating data pipelines and how doing so influences the speed, quality and experienced challenges in comparison to a general-purpose programming language like Python with libraries like Pandas.

Contact: Philip Heltweg <heltweg@group.riehle.org>

Interview procedure

- The interview is voluntary and **has no effect on grades for AMSE / SAKI**. We are interested in your honest opinion and want to hear your challenges with Jayvee.
- The interview takes around 30 minutes.
- The interview is semi-structured. We aim to ask open questions, so you can include any experience you consider relevant. We attached the interview questions below if you want to prepare thoughts.
- We will audio-record the interview so we do not miss important points. We will ask at the start of the interview if we have the permission to do so.
- We handle the data confidentially and won't use them except for the research on the use of domain-specific languages to model data pipelines. We might publish the interview transcripts in the context of academic publications. The data will be published under a pseudonym by default.

Interview questions we will ask

I. Warm-up

1. Please very briefly let us know your background in data engineering / programming outside of AMSE/SAKI

II. Ease of use and quality

2. How easy was it to use Jayvee to create data pipelines?
   a. In general and compared to Python/Pandas
3. How would you describe the quality of data pipelines created in Jayvee?
   a. In general and compared to Python/Pandas

III. Challenges

4. Which challenges did you encounter when creating data pipelines with Jayvee?
   a. In general and compared to Python/Pandas
5. What worked well when creating data pipelines with Jayvee?
6. How did the blocks and pipes architecture of Jayvee change your approach compared to writing an imperative script in Python/Pandas?

IV Open topic

7. Is there any other topic regarding Jayvee or the comparison of DSLs like Jayvee to GPLs like Python/Pandas you consider important?