Internship Position
Python functions through WebAssembly
internships@activeviam.com

"WebAssembly: a solution to execute Python in a Java application"

1 Introduction

ActiveViam has shown that its proprietary software Atoti is the leading technology to interactively analyse datasets from Gigabytes to Terabytes. Atoti can harness the largest machines of Cloud platforms and on-premise DataCenters to make all the CPUs work at a blazing speed on their Terabyte RAMs filled with client data.

To build projects, ActiveViam offers a Python API on top of its Java backend. This API allows easy adoption of the technology for all users. See https://docs.atoti.io/
Atoti Python API allows combining multiple functions to create new aggregated values from the data. However, it is not possible to pass Python code directly, using advanced syntactic features such as loops, list comprehensions or if statements. Hence, users are limited to the combinatory expressiveness of the existing functions.

The arrival of WebAssembly opens new doors. It defines a new instruction set, platform independent and targeted by many programming languages. Moreover, more and more languages offer runtime environments capable of running this instruction set. And this great portability is completed by runtime safety guarantees, preventing client code from accessing the rest of the program and performing unsafe operations.

In this context, ActiveViam looks for a way to offer to their clients new ways for defining calculation logic, directly through raw Python code. This code would then be sent to the backend and executed by the Java runtime via WebAssembly.

2 Expected work

The goal of this internship is to prototype a system to execute Python code from Java using the WebAssembly instruction set as an intermediary.

In an initial phase, the intern will explore ways to execute WebAssembly in a Java environment. From transpilation to WASM to direct execution of Python through a WASM kernel, the intern will test the different available techniques. This study can be completed with alternative approaches, like using cross-language integrations offered by the JVM GraalVM. Or compare WASM and LLVM for a target instruction set.
This step results will also provide the limits of each solution, listing the supported constructs, like loops or ifs.

In a second phase, the intern will study how to dynamically send the Python code to the backend. For a great user experience, ActiveViam wants to be able to let users write code and only code, without forcing side compilation steps. Particularly when working in a notebook environment, sending the function should be easy.

Finally, the intern will measure the performance of the solution. The main measurements will focus on the throughput for executing the functions, operating on billions of rows. These metrics may generate new designs for the function model, to maximise the performance of the prototype.
Additional metrics will report the duration of any translation phase, communication between the client and server, etc.

This prototype will be coded in Java, the core language of Atoti+, and Python for the definition of the calculations.

To turn the prototype into a production grade component and make it part of the Atoti+ solution, this internship may open up the opportunity to be continued as a standard position in the R&D team.

3   About ActiveViam

ActiveViam provides business users with instant insight into large volumes of fast-moving data for timely and context-aware decision-making.

Founded in 2005, ActiveViam employs over 150 people in its five offices of New-York, London, Paris and Singapore. We expect sustained growth in 2022 and will continue hiring the best talents from the top schools.