# Stage M2: Continuous queries over trajectory robots data

Context: Autonomous robots move on plots to perform technical tasks such as mechanical weeding. They are programmed to perform these tasks by minimizing movement on plots, via trajectories planned, while avoiding potential fixed obstacles (such as a root or a pole) or mobile (human, animals, or vehicle) requiring a deviation to the trajectory predefined. Therefore, robots are the main sources of a trajectory data stream. This stream needs to be continuously queried in order to identify patterns and outliers. For example, every second the farmer could be interested to know if a possible collision among robots will happen. This kind of queries are implemented by Data Stream Management Systems (DSMSs). To the best of our knowledge, DSMSs have not been benchmarked for querying robot data, yet.

Contribution: In this project we will study the existing open source DSMSs that can be integrated with our supervision system LambdAgrIoT [1].

The study is conducted over two features, namely: querying capabilities and performance. Then, the best solution will be integrated into LambdAgrIoT.
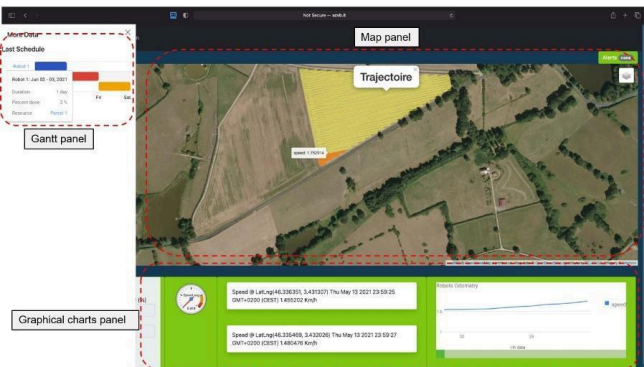
Work plan
- Analysis of existing work on trajectory stream querying (**Spark Streaming**, **Flink**, etc.)
- Benchmark existing DSMSs
- Integration in the LambdAgrIoT system of the chosen solution
- Implement the web interface to visualize the results

Skills: SQL, Java, Web programming, DSMSs (Spark Streaming, Flink, etc.) (optional), Kafka (optional), Grafana (optional)

Location: INRAE, Campus Cézeaux

Supervisors:
- Sandro Bimonte and Jean Laneurit, TSCF, INRAE, France
- Robert Wrembel, Poznan University of Technology, Interdisciplinary Centre for Artificial Intelligence and Cybersecurity, Poland

Date: 6 months from March-April 2024
To apply: Send an email with CV to sandro.bimonte@inrae.fr

[1] G. André, et al. : LambdAgrIoT: a new architecture for agricultural autonomous robots' scheduling: from design to experiments. Clust. Comput. 26(5): 2993-3015 (2023)