

**Title:** Efficient detection of data quality flaws in social interactions of dairy cows

**Company/Institution:** SLU in collaboration with RISE Institute.

**Supervisor:** Lars Rönnegård, SLU

**Background:** In our research project, data from dairy cattle in free-stall barns have been collected since the beginning of 2020 using an ultra-wide band system where each individual cow position in the barn are updated every second. We collect data from two herds, one in Sweden and one in The Netherlands with around 250 dairy cows on each farm. All individuals' positions are saved in a separate file every day (around 1 Gbytes per farm and day) but most of our analyses require following individuals over several days and connecting the movement of every individual with their specific characteristics (such as age). Furthermore, the quality of the positioning data seems to vary between individual cows.

The overall aims of the research project are:

- To develop tools for summarizing animal movement and social interactions in dairy farms.
- To develop decision-support tools for minimizing disease transmission within dairy farms based on knowledge gained from animal movement and social interactions.
- To develop methodology for breeding on indirect genetic effects, ie inherited social effects.

The research project will run from 2020 to 2022 and is a collaboration between SLU in Uppsala, RISE, Växa Sverige and University of Copenhagen. Seven senior researchers, four postdocs and two PhD students are currently involved in the project. For more information see

<https://www.slu.se/en/faculties/vh/research/forskningsprojekt/not/precision-livestock-breeding/>

**Goals:** We would like the students of this project in scientific computing to: 1) assist in the construction of a data base such that the data can be easily accessible for analysis, and 2) to develop methodology and an interface for detecting cows producing poor quality data. Thus, the data should be efficiently stored and fast to access for analysis under different conditions, and it should be easy to trace changes in data quality.