Samuel Kaski
Helsinki University of Technology

Title: **Probabilistic retrieval and visualization of relevant experiments**

Abstract: Given a databank of data-intensive scientific experiments, a useful exercise is to find earlier experiments relevant to a new study. Our application is in modern cellular biology, where databanks of genome-wide measurement data from earlier experiments exist, and data content-based search is likely to produce novel insights. The goal is to retrieve experiments in which the same biological processes are activated. This can be due either to experiments targeting the same biological question, or to as-yet unknown relationships. We use a combination of existing and new probabilistic machine learning techniques to extract information about the biological processes differentially activated in each experiment, to retrieve earlier experiments where the same processes are activated, and to visualize and interpret the retrieval results.

This is joint work with José Caldas, Nils Gehlenborg, Ali Faisal, and Alvis Brazma