## Mining Big Data Streams with Apache SAMOA

## Albert Bifet

LTCI, CNRS, Télécom ParisTech, Université Paris-Saclay, 75013, Paris, France

**Abstract.** In this talk, we present Apache SAMOA, an open-source platform for mining big data streams with Apache Flink, Storm and Samza. Real time analytics is becoming the fastest and most efficient way to obtain useful knowledge from what is happening now, allowing organizations to react quickly when problems appear or to detect new trends helping to improve their performance. Apache SAMOA includes algorithms for the most common machine learning tasks such as classification and clustering. It provides a pluggable architecture that allows it to run on Apache Flink, but also with other several distributed stream processing engines such as Storm and Samza.

Copyright © 2015 by the paper's authors. Copying permitted only for private and academic purposes. In: M. Atzmueller, F. Lemmerich (Eds.): Proceedings of 6th International Workshop on Mining Ubiquitous and Social Environments (MUSE), co-located with the ECML PKDD 2015. Published at http://ceur-ws.org