



## Learning from moments - Large-scale learning with the memory of a goldfish

Rémi Gribonval (INRIA Rennes, France)

**Abstract:** Inspired by compressive sensing, Compressive Statistical Learning allows drastic volume and dimension reduction when learning from large/distributed/streamed data collections. The principle is to exploit random projections to compute a low-dimensional (nonlinear) sketch (a vector of random empirical generalized moments), in essentially one pass on the training collection. Sketches of controlled size have been shown to capture the information relevant to certain learning tasks such as unsupervised clustering, Gaussian mixture modeling or PCA. As a proof of concept, more than a thousand hours of speech recordings can be distilled to a sketch of only a few kilo-bytes that captures enough information to estimate a Gaussian Mixture Model for speaker verification. The talk will highlight the main features of this framework, including statistical learning guarantees —obtained using tools from randomized low-dimensional projections and compressive sensing—, differential privacy guarantees, and open challenges.