**DIAS: Data-Intensive Applications and Systems Laboratory**
School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne
Building BC, Station 14
CH-1015 Lausanne
URL: http://dias.epfl.ch/

**EPFL**

# *Real-time pattern recognition*

**Keywords:** stream processing, pattern recognition, real-time analytics

**Problem:** From social network analytics to gaming, fraud detection, and stock trading, streaming applications require the real-time processing of high-throughput, in-motion data. One of the challenges is to automatically and in real-time detect emerging patterns (e.g., hashtags that appear frequently in the same post, groups of products that customers tend to purchase together, etc.). Emerging pattern detection [1] allows the responsible to react as fast as possible and is valuable for a variety of applications, (e.g., recommendation systems) as well as system optimizations (e.g. prefetching).

**Project:** Pattern recognition is a very complex task and it has been widely addressed in the literature. However, different techniques make different assumptions and cover different application requirements. The first goal of this project is to compare existing approaches and categorize them based on their strengths and weaknesses. Additionally, the student will design an algorithm that, without assuming previous knowledge of the type of patterns, detects patterns that start becoming frequent as well as the moment when a pattern has stopped appearing often.

**Plan:**
1. Familiarize oneself with background work on stream processing and pattern recognition.
2. Compare existing approaches.
3. Design (& implement a prototype of) an algorithm that detects newly occurred patterns as well as patterns that stopped occurring.

**References:**

1. Fahed, Lina & Alfalou, Ayman. (2020). Emergent pattern detection algorithm for big data streams.

**Supervisor:** Prof. Anastasia Ailamaki, anastasia.ailamaki@epfl.ch
**Responsible collaborator(s):** Eleni Zapridou, eleni.zapridou@epfl.ch