# Scientific Presentations - Abstracts list
## Friday, September 2nd

| | |
|---|---|
| **Title** | An information theoretic approach to speaker diarization |
| **Speaker** | Fabio Valente (Idiap Research Institute) |
| **Schedule** | 08:30 – 09:00 |
| **Abstract** | The talk will discuss recent advances on non-parametric speaker diarization of meetings based on the Information Bottleneck (IB) principle. The IB system permits the robust integration of several sources of information with a very limited increase in computational complexity. Results obt.ained on data from the Rich Transcription campaigns show extremely competitive results while the system still runs faster then real-time. |

| | |
|---|---|
| **Title** | Natural Language Processing (almost) from Scratch |
| **Speaker** | Ronan Collobert (Idiap Research Institute) |
| **Schedule** | 09:00 – 09:30 |
| **Abstract** | We will present a single architecture, which excels in performance on various Natural Language Processing (NLP) tasks. Instead of handcrafting task-specific features, generic word representations are learnt from large unlabelled corpora. We will demonstrate how our architecture naturally applies to a wide range of tasks, from simple tasks like Part-Of-Speech, to complex tasks like Parsing. Analysis and comparison with existing NLP approaches will be given. The presentation will end with the introduction of simple standalone software, implementing all these tasks with blazing execution speed. |

| | |
|---|---|
| **Title** | Reality Mining for Real: Large-Scale Human Behavior and Smartphone Data |
| **Speaker** | Daniel Gatica-Perez (Idiap Research Institute) |
| **Schedule** | 09:30 – 10:00 |
| **Abstract** | Smartphones can constantly sense human location, motion, proximity, and communication, and represent one of the most accurate means of tracing human activities. All this data, as never before, is being generated at massive scales. In this talk, I will present an overview of work in my research group in this domain, which addresses mobile sensing, data analysis, and applications. I will first describe our experience with the collection of a rich corpus of everyday life data involving over 150 users and 15 months using smartphones. I will then discuss methods designed to discover a variety of patterns, including places of interest, social interaction types, trends of phone application usage, and personality traits. Our methods integrate observations coming from GPS, GSM, Bluetooth, accelerometer, and Wifi sensors, and logs of essentially every event on the phone, and have been tested on longitudinal, real-life data. |

**Title**        Learning and Sharing of semantic spatial concepts

**Speaker**   Barbara Caputo (Idiap Research Institute)

**Schedule**  10:30 – 11:00

**Abstract**  The ability of building robust semantic space representations of environments is crucial for the development of truly autonomous robots. In this talk I will show how to equip mobile robots with algorithms for semantic representations of indoor places. By learning visual features from the data, it is possible to recognize effectively place categories, achieving state of the art performance on three benchmark databases. I will then outline what are the conceptual challenges to exploit this knowledge for robots located in different locations, and show results in that direction.

**Title**        What's going on? Discovering temporal motifs from sensors logs

**Speaker**   Jean-Marc Odobez  (Idiap Research Institute)

**Schedule**  11:00 – 11:30

**Abstract**  In this presentation i will introduce a model for the unsupervised extraction of motifs from multivariate time series. We consider a particular kind of time series where observed values are caused by the superposition of multiple phenomena occurring concurrently and with no synchronization. The model will be demonstrated on single and multi-views surveillance videos from both relatively well-structured urban scenes and from more complex videos from the Torino underground, as well as signals from microphone array data.

**Title**        The MASH project: toward the collaborative design of hyper-complex learning software

**Speaker**   François Fleuret (Idiap Research Institute)

**Schedule**  11:30 – 12:00

**Abstract**  The MASH project is a European research initiative that aims at developing novel tools for the design of complex machine learning systems. The project's main objective is to create an open-source web platform allowing the collaborative development of complex families of image feature extractors for scene understanding and vision-based goal-planning algorithms. Performance will be measured on standard image classification and object detection data sets, and using a real robotic arm, and a 3d simulator for goal-planning tasks. I will present the motivations behind the project, the scientific and technical challenges to address, and the current status of the platform. I will then show how to participate and some early results.