CHIST-ERA Projects Seminar
“From Data to New Knowledge”

C. Barras
Brussels, March 22-23, 2017
CAMOMILE project

Collaborative Annotation of multi-MOdal, Multi-Lingual and multi-mEdia documents

6 labs – LIMSI, IMMI, LIG (France), LIST (Luxembourg), UPC (Spain), ITU (Turkey)

Goals

Unifying use case
✓ Multimodal person identification
A *platform* for **collaborative annotation of multimedia documents**

we provide the back-end
you provide the front-end

authentication
permissions
notifications

generic data model
audio, text, image, video

https://camomile-project.github.io/
Major outcomes and achievements

- Person Discovery in TV Broadcast MediaEval 2015 and 2016 evaluation task
  - Unsupervised person identification

- Evaluation workflow relying on our CAMOMILE framework
  - A posteriori collaborative annotation of data set
  - Live leaderboard
Remaining challenges and needs

- Infrastructure is very important for research in artificial perception & machine learning area
  - Data collection and annotation (collaborative)
  - Evaluation of technology (competitive)
  - Reproducible research (both)
- ... long-term funding is difficult
  - Programs focus on innovation
- Possible work-arounds
  - Bottom-up from individual initiatives
  - Incremental effort
  - Shared across research communities
Discussion ?