AN OPEN-SOURCE SPEAKER GENDER DETECTION FRAMEWORK FOR MONITORING GENDER EQUALITY

David Doukhan, Jean Carrive, Félicien Vallet, Anthony Larcher, Sylvain Meignier
ddoukhan@ina.fr, jcarrive@ina.fr, fvallet@cnil.fr, anthony.larcher@univ-lemans.fr, sylvain.meignier@univ-lemans.fr

Content-based gender equality estimation in media [1,2,3]

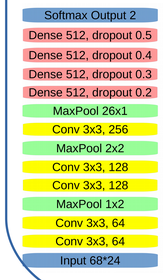
Presence rate: How many speakers?
Identification rate: Who are we talking about?
Expression rate: How long do speaker speak?

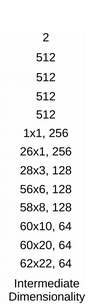
Roles (expert, interviewer, popular opinion)
Which topics are covered for each gender?

Realized with small amounts of data
Biases related to the context of collection!

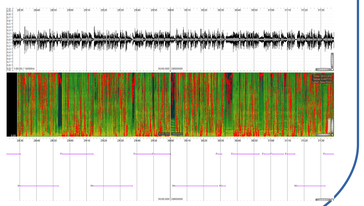
Speaker Gender Segmentation

http://github.com/ina-foss/inaSpeechSegmenter





MFC extraction with SIDEKIT [4]
Patch Normalization
CNN probability estimates
Viterbi post-processing



Training: Ina speaker dictionary [5]

Speaker Diarization

Optical Character Recognition

Manual Validation

Dictionary

Archives broadcast from 1957 to 2012

32000 segments

1780 mens (94 h)
494 womens (27h)

Evaluation Protocol

Material: REPERE challenge corpus (58 hours)
Street interviews, debates obtained from French news
1129 men (46 hours), 557 women (12 hours)

Baseline Models

Features: 19 MFCC + Δ + ΔΔ and feature warping (SIDEKIT [4])
GMM: 1024 gaussians and diagonal covariances
Ivector: Baseline HAC-BIC diarization with S4D [6]
I-vector extraction (Total variability matrix of rank 200)
Linear SVM classification

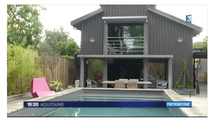
Frame-level results

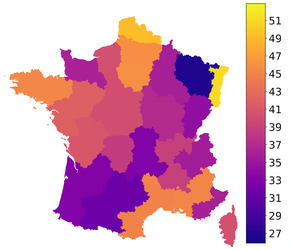
	Men Recall	Women Recall	F-score	Accuracy
GMM	97.25	94.28	95.74	96.63
I-vector	98.19	92.98	95.51	97.11
CNN	98.04	95.05	**96.52**	**97.42**

Men and Women Expression rates in France 3 regional news (2016) [7]

Prime-time (14 to 21% audience)

7 regions more or less equals
4 editions below 1 third
Moderate correlation between number of inhabitants and WSTP

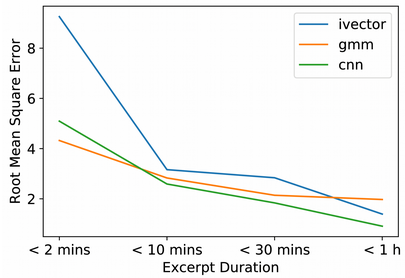




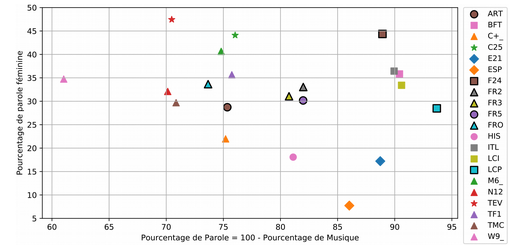
End-to-end evaluation

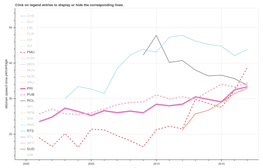
Women Speaking Time Percentage estimation (WSTP)

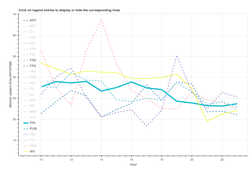
$$WSTP = \frac{100 \times Women\ speaking\ Time}{Men + Women\ Speaking\ time}$$



Men and Women Speech-Time percentage in TV and Radio [7]







References

[1] Macharia, S. et al (2015), Who Makes the News?: Global Media Monitoring Project 2015, World Association for Christian Communication
[2] Reiser, M. and Gresy, B. (2008), L'image des femmes dans les médias, Public Report
[3] Conseil supérieur de l'audiovisuel (CSA), La représentation des femmes à la télévision et à la radio - Rapport sur l'exercice 2016
[4] Larcher, A., Lee, K.A., Meignier, S. (2016), An extensible speaker identification sidekit in Python, ICASSP
[5] Salmon, F., and Vallet, F. (2014), An Effortless Way To Create Large-Scale Datasets For Famous Speakers, LREC
[6] Meignier, S. and Larcher, A. (2015), S4D: SIDEKIT for Speaker Diarization http://lium.univ-lemans.fr/sidekit/s4d/
[7] Doukhan, D., Poels, G., Carrive, J., (2018) Describing Gender Equality in French Audiovisual Streams with a Deep Learning Approach, Journal of European Television History and Culture