# Summary Automatic Tagging : A State of the Art

- (Pour notre conclu) Tagging music can also be more than just a genre : a mood, an instrumentation, a style (humoristic…)

- Aggregate features or Bag of frames (cf article AdaBoost) : long time frames (~1s), which are subdivided in small windows, on which the "classical" features are computed (timbral ones), and then statistical values (mean, variance…) are computed over the whole long time frame

- Gathering good datasets : often either large and noisy (social tags on music streaming sites, Last.fm for instance) or small and clean (tagging games such as Magnatagatune, surveys…)
  - ➔ For genre classification into a dozen of genre, small and clean is preferable
  - ➔ Certain algs scale better or worse (typically, SVMs) with the dataset size

- Some tags/genre are much more common than others : do we try to have evenly distributed datasets, or do we keep them as they are ?

- Regression or classification ? Even music genre falls under this question : a lot of songs belong to multiple genres, it may be more useful to give them multiple tags (with a weight) rather than try to find the ONE corresponding tag (which is often not unique). Regression labels can be obtained by giving the percentage of people that attributed a certain tag

- Popular algs : SVMs (adaptable to regression), Boosting (typically decision trees), Gaussian mixtures (adapts really well to regression, but CPU-hungry, solved by HGMM ?), Neural Networks

- Evaluation is complicated : humans introduce a bias when they decide that a certain tag was or wasn't relevant ➔ evaluation is not 100% objective, still they are needed because all mistakes should not be penalized the same (labeling "rock" as "blues" is better than labeling it as "disco")

- Precision, recall, F-measure, accuracy, AROC ? Maximizing precision may be the objective ?

- Enormous amounts of data online : the best algorithm should work directly online

- Should handle sparse encodings (especially in time domain)