MIREX 2005: What did we learn?

Dan Ellis
MIREX Panel
ISMIR 2005
2005-09-14
MIREX 2005 in review

- 10 contests
- 70+ submissions
- authoritative results
- algorithm submission!
- .. heroic effort
What did we learn?

• which labs did best

• overall spread of performances

• something about variety of approaches
  • from abstracts

• no standout techniques?
  • all top pairs differ < 10%
  • in 4 out of 10, differ by < 1%
What *didn’t* we learn?

- which *techniques* are successful
- impact of individual choices
  - e.g. features, classifiers...
- interactions of approaches and results
- the importance of diversity...
- the value of co-operation...
  - instead of competing
How to learn more?

• have more people look at detail of results
• have finer-grain breakdown of algorithms
• more ground truth / annotation
  • need more buy-in
  • no split responsibility: leaders = organizers
Access to detailed results

• participants run evaluations at home
  • even final eval, but not scoring
• separate development and test sets
• common basis for ‘progress’ reports
Algorithm breakdown

• provide common framework including default units
  • consensus on problem decomposition?
• participants can replace just one part, or whole set
• submitted components can be cross-combined
• sharing of code?...