

Wrapup of MIR (Math Information Retrieval) Symposium

Michael Kohlhase & Petr Sojka

<http://kwarc.info/kohlhase>
Center for Advanced Systems Engineering
Jacobs University Bremen, Germany

9. July 2012, CICM/DML

Math Informatin Retrieval (MIR) Symposium 2012

- 25 participants, 10 talks (most importantly: lively discussions)
- Emerging Consensus: $MIR = \text{Math Search} + \text{Math Semantics Recovery}$
- MIR may be the “killer application for DML/MKM” but needs a lot of work. (↔ MIR happening)
- Will create MIR mailing list (ask me to be included)
- Invited Video Talk: Michael Trott (Wolfram Research) on Wolfram Alpha
 - Wolfram Alpha = hand-curated content + computation + NL query interface (complementary to MIR efforts)
 - Project Scope = 100 FTE \times 6 years
 - Math is only a 10% fraction of the content (economics, units, weather, geography)

MIR Happening

- **Definition 1** The **MIR Happening** is a friendly competition and public viewing of Math search technology in action.
Participants receive a search corpus ahead of time (for indexing)
Judges (Davenport/Ion/Meyer) provide queries, judge results in three categories
 - Formula Search (automated)
 - Document Search (automated)
 - Open Information Retrieval (semi-automated)
- Dry Run for the NTCIR-10 Math Track (Fall '12/Spring '13)
- **Lessons Learned:** (may also apply to other CICM disciplines)
 - Coming up with meaningful queries is hard (need your help)
 - What are sensible/comfortable/meaningful query languages?
 - How can users express their mathematical intuitions without programming?
 - still somewhat early for automated contents.