Relevance-based Evaluation
the OAEI 2007 food task revisited

Willem Robert van Hage
Vrije Universiteit Amsterdam
wrvhage@few.vu.nl

Hap Kolb
TNO Science & Industry
hap.kolb@tno.nl

Guus Schreiber
Vrije Universiteit Amsterdam
schreiber@cs.vu.nl

Gather hot topics
- Expert Interviews
- Search Log Analysis

Gather documents
- FAO Newsroom

Manually Match Metadata
- FAO AGRIS/CARIS
- USDA AGRICOLA

Compute Recall on Relevant Sample
- Relevant
- Relevant
- Correct Found
- Correct Found
- Unbiased Recall
- Recall on Relevant Mappings

Results
- only considering equivalence mappings
- considering equivalence and broader/narrower mappings

Willem Robert van Hage
Vrije Universiteit Amsterdam
wrvhage@few.vu.nl

Guus Schreiber
Vrije Universiteit Amsterdam
schreiber@cs.vu.nl

Hap Kolb
TNO Science & Industry
hap.kolb@tno.nl

- The OAEI 2007 system ranking is consistent with the relevance-based second opinion.
- Especially “easy” mappings were needed to find the relevant mappings.
- The relevance-based results exaggerate the differences between the systems, because the best systems are good at finding all “easy” mappings first.
- It matters which correct mappings you find.
- A typical case is seldom an average case. Uniform sampling estimates average case behavior, not necessarily typical behavior.