INEX-XER: Entity Ranking Overview Talk 2008

Gianluca Demartini
with Arjen P. de Vries, Tereza Iofciu, and Jianhan Zhu

Entity Ranking

- Many users search for specific entities instead of just any type of documents

Not relevant for XER…

- Articles *on topic* are not necessarily relevant entities
  - Actually, they are surprisingly often not!
  - INEX 2007 adhoc-derived XER topics show that only about 35% out of original relevant documents have been assessed as relevant

Example 2008 Topics

- Countries that have hosted FIFA Football World Cup tournaments: *countries; football world cup*
- Formula 1 drivers that won the Monaco Grand Prix: *racecar drivers; formula one drivers*
- Italian nobel prize winners: *nobel laureates*

…

Many examples on
http://www.ins.cwi.nl/projects/inex-xer/topics/

Entity Ranking

- Topical query Q
- Entity (result) type T_X
- A list of entity instances Xs

- Systems employ XML element text, structure, links

Topic 60

Q

Title
olympic classes dinghy sailing

Xs

Entities
470 (dinghy) (#816578)
49er (dinghy) (#1006535)
Europe (dinghy) (#855087)

T_X

Categories
dinghies (#30308)

Description
The user wants the dinghy classes that are or have been olympic classes, such as Europe and 470.

Narrative
The expected answers are the olympic dinghy classes, both historic and current. Examples include Europe and 470.

Topic 60
Title
olympic classes dinghy sailing
Entities
470 (dinghy) (#816578)
49er (dinghy) (#1006535)
Europe (dinghy) (#855087)
Categories
dinghies (#30308)
Description
The user wants the dinghy classes that are or have been olympic classes, such as Europe and 470.
Narrative
The expected answers are the olympic dinghy classes, both historic and current. Examples include Europe and 470.

Google Sets

Predicted Items
49er
470
europe
laser
optimist
finn
420
tornado
yngling
star
laser radial
29er
snipe
mistral
contender

2008 Tasks

- Entity Ranking (ER)
  - Given Q and T, provide Xs
- List Completion (LC)
  - Given Q and Xs[1..m]
  - Return Xs[m+1..N]

- Pilot: Entity Relationship Search (ERS, explained later on…)

XER Assumptions

Last night: 2 hours later...

- Entities (Xs) are still represented as Wikipedia pages
- Binary relevance, MAP (`xinfAP`)

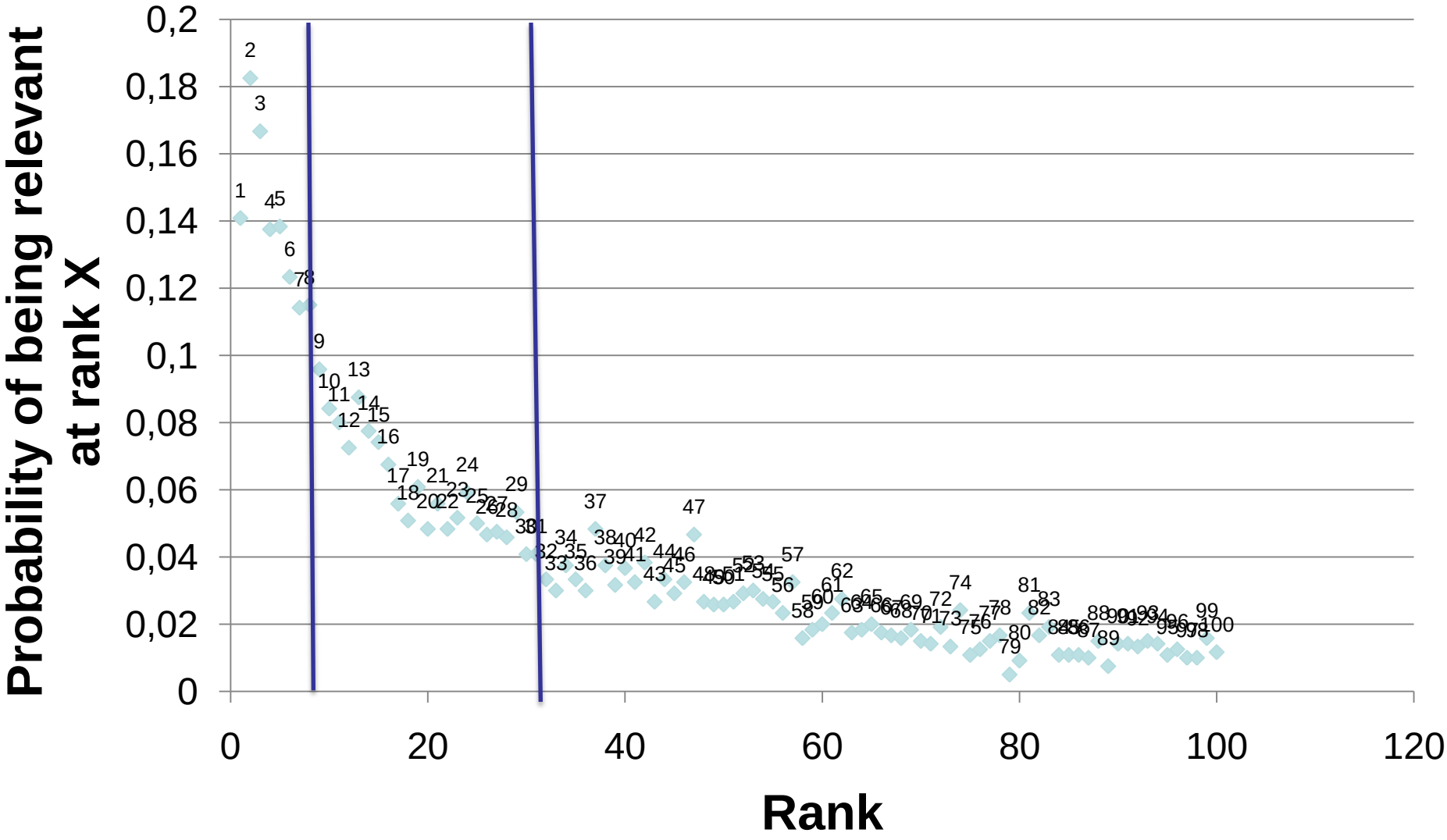
Runs

- Participation
  - >60 groups sign up
  - 11 groups submit topics
  - 6 groups submit 33 runs
  - 12 groups assess topics

Pooling by Sampling

- Approaches:
  - Random sampling
  - Relevance based sampling
  - Stratified sampling
- Collection
  - 24 XER2007 topics (pool size: 50)
- Comparison
  - IRSs ranking changes with less assessments

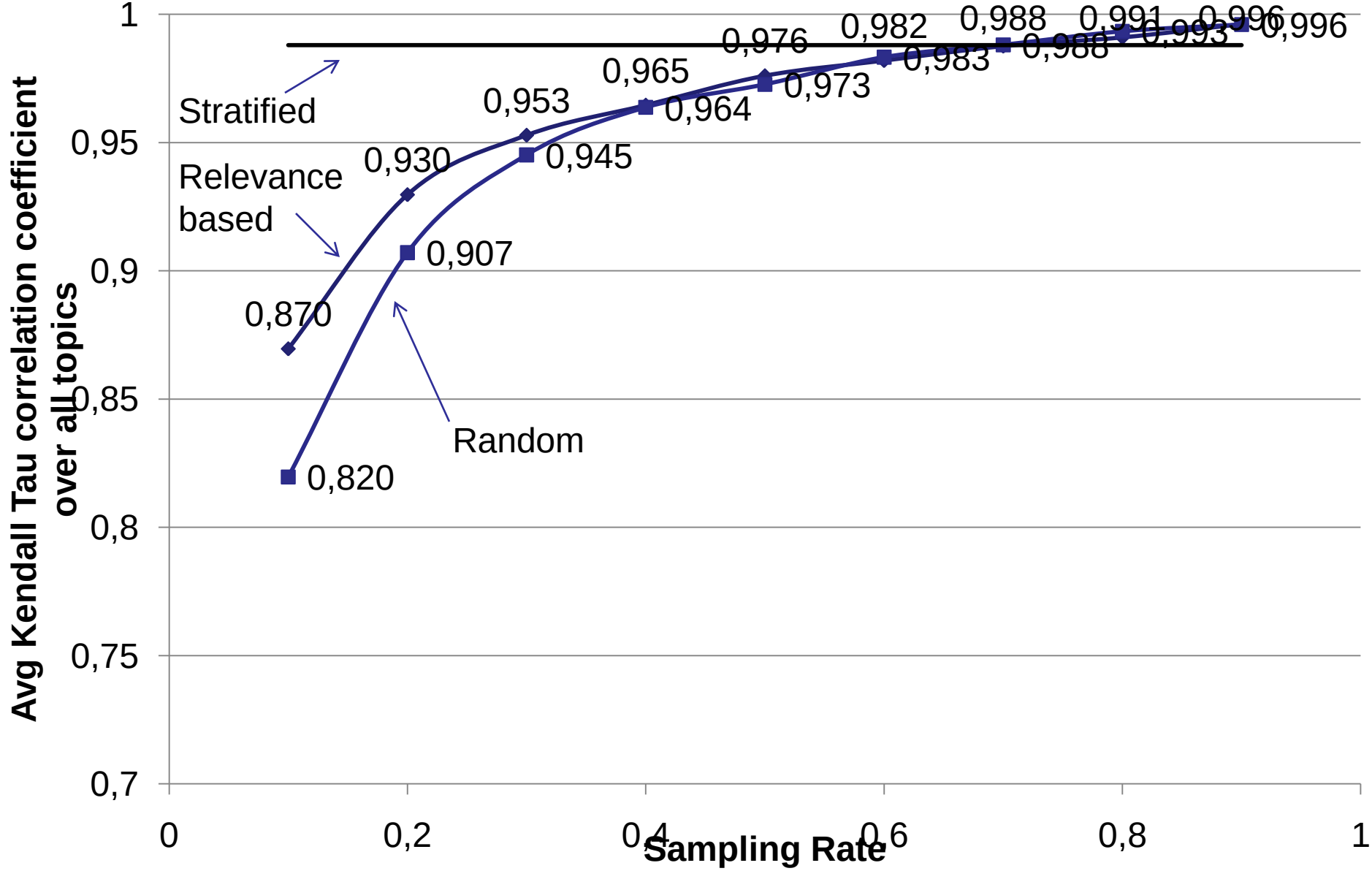
Relevance based Sampling



Stratified Sampling

- { 1,8 } 100%
- { 9,31 } 70%
- { 32,100 } 30%

IRSs Ranking comparison



Pool Contribution

- Random/Relevance based Sampling:
  - at 70%: 35 docs out of top 50
- Stratified Sampling:
  - 45 docs out of top 100 (30 docs out of top 50)

- XER 2007 pool: 50 docs

So… RESULTS!!!

Entity Ranking

- **ER/1_FMIT_ER_TC_nopred-cat-baseline-a1-b8: AP all 0.243326941228736**
- **ER/1_cirquid_ER_TEC_idg.trec: AP all 0.233202024909469**
- **ER/4_UAms_ER_TC_cats: AP all 0.226179832095729**
- ER/2_UAms_ER_TC_catlinksprop: AP all 0.224540266234725
- ER/1_UAms_ER_TC_catlinks: AP all 0.222165511518068
- ER/3_cirquid_ER_TEC.trec: AP all 0.198151482470439
- **ER/2_cirquid_ER_TC_idg.trec: AP all 0.195956981708847**
- **ER/2_500_L3S08_ER_TDC: AP all 0.189376490785161**
- **ER/1_CSIR_ER_TC_mandatoryRun: AP all 0.187349723227119**
- ER/1_L3S08_ER_TC_mandatoryRun: AP all 0.182930576395362
- **ER/3_UAms_ER_TC_overlap: AP all 0.180831755068589**
- ER/4_cirquid_ER_TC.trec: AP all 0.167776008664103
- ER/4_UAms_ER_TC_cat-exp: AP all 0.165395918333584
- ER/1_UAms_ER_TC_mixture: AP all 0.15835232612729
- ER/3_UAms_ER_TC_base: AP all 0.113264868377908
- ER/6_UAms_ER_T_baseline: AP all 0.078963024406816 5

List Completion

- **LC/1_FMIT_LC_TE_nopred-stat-cat-a1-b8: AP all 0.286897420671293**
- LC/1_FMIT_LC_TE_pred-2-class-stat-cat: AP all 0.272852692304109
- LC/1_FMIT_LC_TE_nopred-stat-cat-a2-b6: AP all 0.259142806119578
- LC/1_FMIT_LC_TE_pred-4-class-stat-cat: AP all 0.252080676791407
- **LC/1_CSIR_fixed: AP all 0.239955757701729**
- **LC/5_UAms_LC_TE_LC1: AP all 0.232243493200444**
- LC/6_UAms_LC_TEC_LC2: AP all 0.230389397906149
- **LC/2_UAms_LC_TCE_dice: AP all 0.228136165118533**
- **LC/5_cirquid_LC_TE_idg.trec.fixed: AP all 0.217632851524678**
- **LC/1_L3S08_LC_TE_mantadoryRun: AP all 0.205598133259403**
- LC/2_L3S08_LC_TE: AP all 0.204518922088608
- LC/5_cirquid_LC_TE_idg.trec: AP all 0.195453875103596
- LC/6_cirquid_LC_TE.trec.fixed: AP all 0.194585187775904
- LC/1_CSIR_LC_TE_mandatoryRun: AP all 0.183874665382473
- LC/6_cirquid_LC_TE.trec: AP all 0.177867887897393
- LC/5_UAms_LC_TE_baseline: AP all 0.094986866896919

Entity relation search pilot

- Tuple

  `<query, category, relation-query, target-category>`

- Two stages:
  - Entity ranking stage → main entities
  - Relation search stage → target entities
    - Retrieve further details about main entities
  - Relations:
    - 1 to 1, 1 to n (n>1), or n to 1 (n>1)
  - Results: pairs of main and target entities

Example of ERS

<title>Impressionist art in the Netherlands</title>

<description>I want a list of art galleries and museums in the Netherlands that have impressionist art.</description>

<narrative>Each answer should be the article about a specific art gallery or museum that contain impressionist or post-impressionist art works.</narrative>

<categories>

<category id="10855">**art museums and galleries**</category>

</categories>

<entity-relation>

<relation-title>**located in**</relation-title >

<relation-description>I want the cities where these art galleries and museums are located. </relation-description>

<relation-narrative>Each answer should be a city where a specific art gallery or museum that contain impressionist or post-impressionist art works is located. </relation-narrative>

<target-categories>

<category id="2917">**cities**</category>

</target-categories>

</entity-relation>

Evaluation of ERS results

- Pages for both main and target entities used for evaluators to judge
  - More difficult than entity ranking
- Two stage evaluation to simplify the process

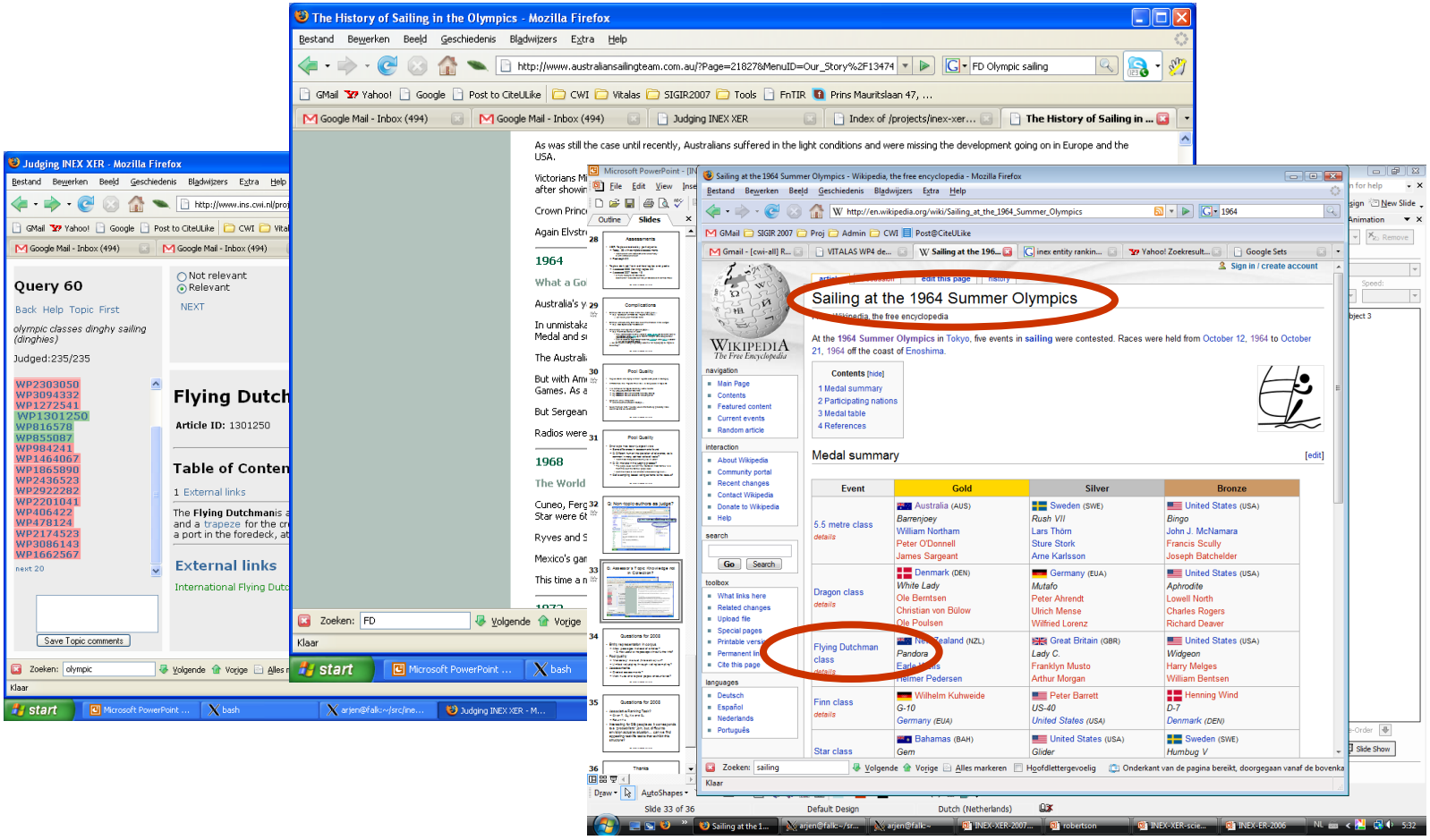
Evaluation of ERS results

1. Main entity judged relevant to original query
2. Main entity of correct category
3. Target entity of correct category
4. Relation match relation topic

Open Issues

- Other types of relation search
  - Relationships between main entities
    - Find pairs of impressionist artists who influenced each other
    - Find experts in an organization who worked together on a project
  - Issues:
    - How to define relations, e.g., “influence”
    - How to evaluate relations
    - How to define scope of relations

Q: Assessor's Topic Knowledge not in Collection?



INEX XER Overview 2008

What *could* be new in 2009?

- Entity representation in corpus
  - Allow passages instead of articles
    - Q: How useful is the passage without further info? Require support evidence?
- New tasks?
  - Leave out desired entity type
  - Identify the entities in the corpus that do not have their page yet