

Interactive Information Access (INS2)

Arjen P. de Vries

Information Science

“Search for the fundamental knowledge which will allow us to postulate and utilize the most efficient combination of [human and machine] resources”

- **M.E. Senko. Information systems: records, relations, sets, entities, and things. *Information systems*, 1(1):3–13, 1975.**

Key People

- **Lynda Hardman (HCI)**
- **Jacco van Ossenbruggen (SemWeb)**
- **Lloyd Rutledge (left to Open Uni)**
- **Frank Nack (left to UvA)**
- **Arjen P. de Vries (joined 2009)**

Key Projects

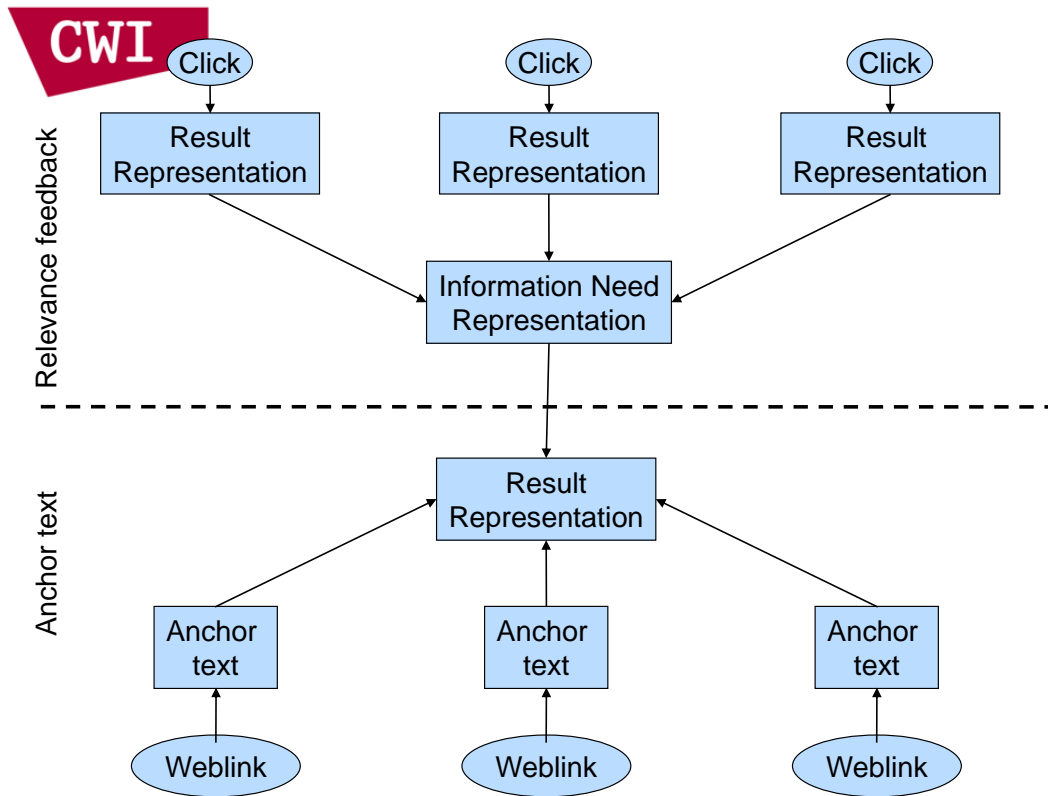
- **Recent**
 - K-Space NoE
 - Vitalas IP
 - MultimediaN: e-Culture, Semantic Search
- **New**
 - Fish4K STREP
 - LinkedTV IP
 - COMMIT: e-Culture, Entities
- **Via university relations:**
 - Delft: PuppyIR STREP, PetaMedia NoE
 - VU: Europeana

Core Questions

- **How to represent information?**
 - The information need and search requests
 - The objects to be shown in response to an information request
- **How to match information representations**
 - (Deductive) data retrieval, (inductive) information retrieval, or a mix?!

Interactive Information Access

- 1. Interaction improves information representation**



Vitalas IP – Belga logs

- **Training data for machine learning of multimedia ‘detectors’**
 - Theodora Tsirikika, Christos Diou, Arjen P. de Vries, Anastasios Delopoulos. *Image annotation using clickthrough data*. CIVR 2009
 - T. Tsirikika, C. Diou, Arjen P. de Vries, A. Delopoulos. *Reliability and effectiveness of clickthrough data for automatic image annotation*. International Journal on Multimedia Tools and Applications, 2010.

Vitalas IP – Belga logs

- **Exploit ‘semantic data’ to address sparsity in usage logs**
 - Vera Hollink, Theodora Tsikrika, Arjen P. de Vries, *Semantic search log analysis: A method and a study on professional image search*. JASIST 62(4): 691-713, 2011.

‘Co-creation’

- **Social Media:**
 - Consumer becomes a co-creator
 - ‘Data consumption’ traces
- **In essence: many new sources to play the role of anchor text**
 - Tags and/or ratings
 - Tweets
 - Comments, reviews

Tags & Ratings in Social Media

- **M. Clements, Arjen P. de Vries, M.J.T. Reinders. The task-dependent effect of tags and ratings on social media access. *ACM Transactions on Information Systems* 28, 21, 1–42, 2010.**
- **J. Wang, Arjen P. de Vries, M.J.T. Reinders. Unified relevance models for rating prediction in collaborative filtering. *ACM Transactions on Information Systems* 26, 1–42, 2008.**

Interactive Information Access

- 1. Interaction improves information representation**
- 2. Interaction can let user take over where machine would fail**

Present results grouped by facet of possible answers grouped by facet

Extra information on mouse-over

works showing matching person (100)

works showing matching person (1)

works showing event (8)

works titled (7)

Gogh, Vincent van (ULAN RKD Rijksmuseum people)
 Except for some brief periods of formal instruction, van Gogh was self-taught; he collected prints and reproductions to study and copy, especially those of Millet. His life and work are legendary in the history of art, making him the quintessential misunderstood, tormented, even insane artist, who sold only one work i ...

works with matching literature (16)

works created by a teacher of matching person (100)

works related to a teacher of matching person (3)

CWI

Science Results

- **First prize @ the Semantic Web Challenge at ISWC in Athens GA, USA**
- **M. Hildebrand, Jacco van Ossenbruggen, Lynda Hardman, G. Jacobs. Supporting subject matter annotation using heterogeneous thesauri, a user study in Web data reuse. *International Journal of Human-Computer Studies* 67, 888–903, 2009.**
- **A.K. Amin, M. Hildebrand, Jacco van Ossenbruggen, Lynda Hardman. Designing a thesaurus-based comparison search interface for linked cultural heritage sources. *IUI 2010* (14), 249–258, 2010.**

Interactive Information Access

1. Interaction improves information representation
2. Interaction can let user take over where machine would fail
3. Interaction can let user take over where system designer would fail

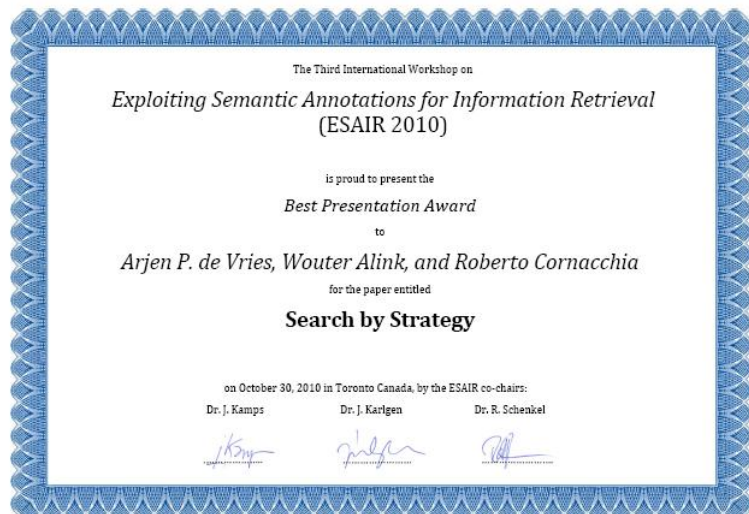


- **Search by strategy**

- Visually connect
- Each block that data

- **CWI spin-**

- Started on
- Follow-up



Search by Strategy

- **Based on unique architecture developed in close collaboration with INS1, e.g.,**
 - R. Cornacchia, S. Héman, M. Zukowski, Arjen P. de Vries, P.A. Boncz. Flexible and efficient IR using array databases. *VLDB Journal* 17, 151–168, 2008.
 - R. Cornacchia, Arjen P. de Vries. A parameterised search system. *Advances in Information Retrieval, 29th European Conference on IR Research, ECIR 2007*, 4–15, 2007. **Best Student Paper Award.**

Strengths & Weaknesses

- Unique mix of expertise:
Linked data + IR + UI/UX
- Excellent relations with humanities
- Limited success in acquisition of external funding

Opportunities & Threats

- Core research 'in vogue'
- Funding agencies preferring multi-disciplinary projects
- Prof. Smeulders @ CWI strengthening multimedia expertise
- Viewed as engineers, not scientists
- Growing competition in core research area

Strategy

- **Interactive: Value of User Actions**
 - Information Representation
 - How to exploit implicit (observed) and explicit (elicited) user effort?
 - Design methods for UI/UX
 - How to measure benefits of including designers and/or media artists in interaction/interface design loop?
 - Can explicit representations of information and/or communication roles drive UI/UX?

Strategy

- **Dataspaces**
 - “Big data” data sets in science and humanities, collected by many researchers for diverse purposes
 - How to assess the quality of large and heterogenous data sets for specific research goals?
 - How to transparently expose the provenance of each data item and propagate the various confidence levels?
 - The role of entities in exploratory search tasks?