CWI

Interactive Information Access (INS2)

Arjen P. de Vries

CWI

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.**

CWI

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)**

CWI

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

CWI

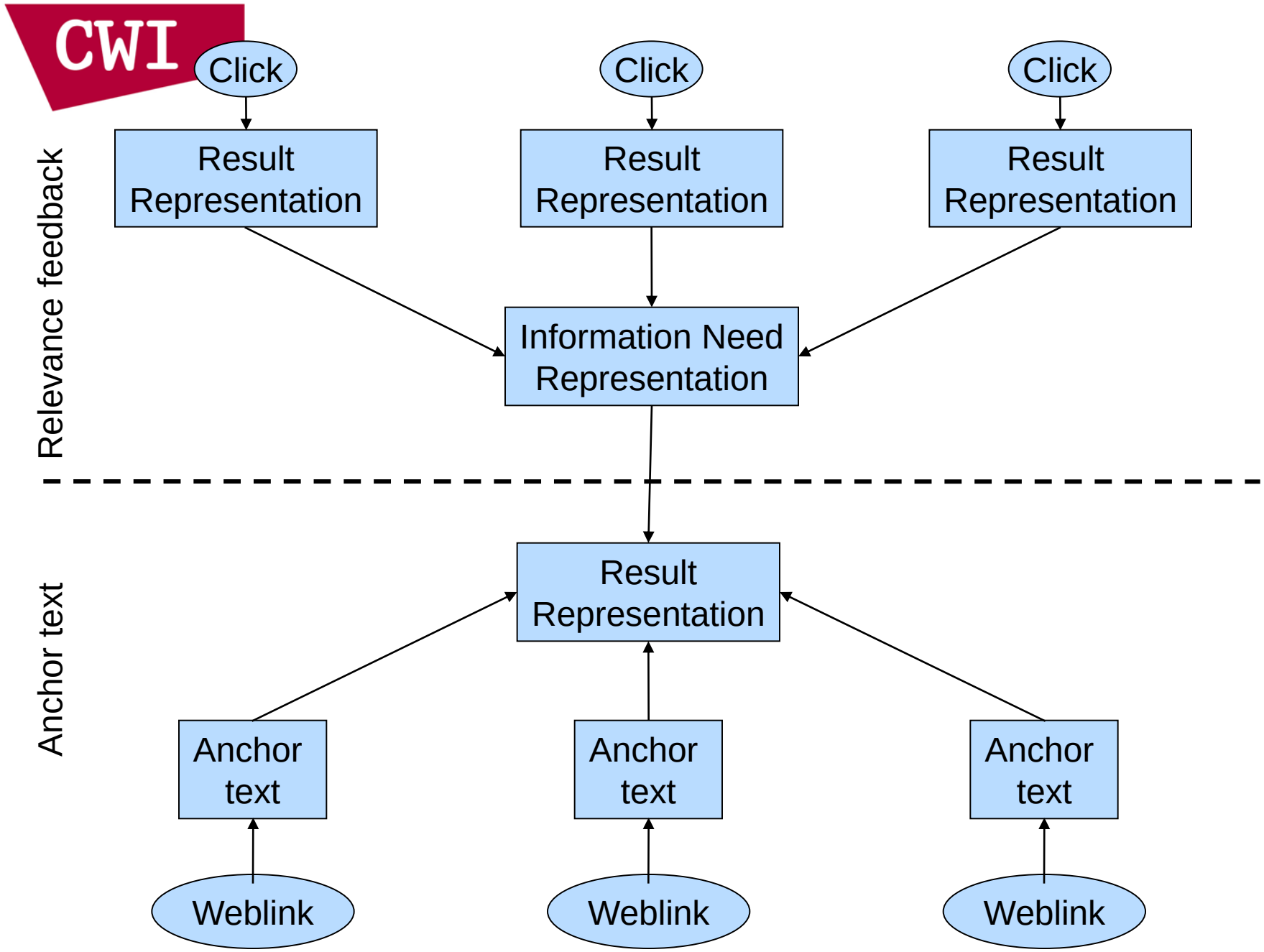
Core Questions

- **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?!

CWI

Interactive Information Access

1. **Interaction improves information representation**


CWI
Click
Click
Click
Result Representation
Result Representation
Result Representation
Relevance feedback
Information Need Representation
Result Representation
Anchor text
Anchor text
Anchor text
Anchor text
Weblink
Weblink
Weblink

CWI

Vitalas IP – Belga logs

- **Training data for machine learning of multimedia 'detectors'**
  - Theodora Tsikrika, Christos Diou, Arjen P. de Vries, Anastasios Delopoulos. *Image annotation using clickthrough data.* CIVR 2009
  - T. Tsikrika, 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.

CWI

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.

CWI

‘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

CWI

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.**

CWI

Interactive Information Access

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

europeana
think culture
login | help | English
Van Gogh
artefact related to keyword: "Van Gogh"
show facets
artefact view all 19 results
Interieur van het huis van Andri...
Moord op Theo van Gogh
omslag catalogus Vincent van Gogh
concept view all 9 results
van Gogh
painter
De broer van Vincent van Gogh H....
In 1861 trad de schilderijenexpe...
event
Moord op Theo van Gogh
location
Van Gogh Museum (Amsterdam)
person view all 32 results
Gogh, Marie van
Gogh, Vincent van
Gogh, Ank van
Present results grouped by facet
Present types of possible answers grouped by facet
works created by matching person (100)
Van Gogh's slaapkamer in Arles
Gogh, Vincent van
Van Gogh's woning aan de
Gezicht vanuit de kamer van
Zelfportret
works showing matching person (1)
De opening van de Van
Crone, Werry
Extra information on mouse-over
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 showing event (8)
works titled (7)
view all
Naakt voor van Gogh's
Israëls, Isaac
Ex-Libris Vincent van Gogh
Bauer, Marius
Ex-Libris Vincent van Gogh
Bauer, Marius
Vrouw lezend naast 'De
Israëls, Isaac
omslag catalogus Vincent van
Roland Holst, Richard
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.**

CWI

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**

CWI

spinque

- **Search by strategy**
  - Visually cc
    connectinç
  - Each block
    that data
- **CWI spin-**
  - Started op
  - Follow-up

The Third International Workshop on

Exploiting Semantic Annotations for Information Retrieval
(ESAIR 2010)

is proud to present the

Best Presentation Award

to

Arjen P. de Vries, Wouter Alink, and Roberto Cornacchia

for the paper entitled

Search by Strategy

on October 30, 2010 in Toronto Canada, by the ESAIR co-chairs:

Dr. J. Kamps Dr. J. Karlgen Dr. R. Schenkel

CWI

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.**

CWI

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

CWI

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?

CWI

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?