Working at (and pushing) the boundaries of IR: how other fields can influence your IR research.

Dr. David Elsweiler
Chair for Information Science
Faculty of Language, Literature and Cultural Sciences
david@elsweiler.co.uk
NO ENTRY
Ingwersen and Järvelin 2005
- different levels need different approaches, different evaluation methods
Information Retrieval

Information Science

Affective Computing

HCI

Data mining

Psychology

Sociology

Ethnography

Linguistics

Architecture

City Planning

Web science

Neuroscience

Information Seeking

Medicine

Statistics

Mathematics

Nutrition

Leisure Studies

Quantum Mechanics

Economics
When you move to the boundaries:

• Interesting problems

• Very challenging problems

• Standard IR methods are not enough (need to combine, be inspired by other fields)

• Examples from my own research
Personal Search

“Stuff I've Seen”

[Dumais et al., 2003]
Things they created
Things they received
Things they've read
Personal Search
Personal Search
Find out where my hotel is for FDIA

<table>
<thead>
<tr>
<th>Subject</th>
<th>From</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Re: Keynote talk at the 4th Symposium on Future Directions in Information Acc...</td>
<td>Silke Werger, C5 Ko</td>
<td>24/08/2011 10:54</td>
</tr>
<tr>
<td>Re: Keynote talk at the 4th Symposium on Future Directions in Information Acc...</td>
<td>Silke Werger, C5 Ko</td>
<td>19/08/2011 15:31</td>
</tr>
<tr>
<td>ESSIR 2011 - info request</td>
<td>Silke Werger, C5 Ko</td>
<td>19/08/2011 13:42</td>
</tr>
<tr>
<td>Re: Keynote talk at the 4th Symposium on Future Directions in Information Acc...</td>
<td>Silke Werger, C5 Ko</td>
<td>18/08/2011 17:28</td>
</tr>
<tr>
<td>Re: Keynote talk at the 4th Symposium on Future Directions in Information Acc...</td>
<td>Silke Werger, C5 Ko</td>
<td>11/08/2011 10:13</td>
</tr>
<tr>
<td>AW: AW: Keynote talk at the 4th Symposium on Future Directions in Information Acc...</td>
<td>Sergej Sizov</td>
<td>10/08/2011 10:07</td>
</tr>
<tr>
<td>AW: Keynote talk at the 4th Symposium on Future Directions in Information Acc...</td>
<td>Sergej Sizov</td>
<td>09/08/2011 21:59</td>
</tr>
<tr>
<td>Re: <a href="http://essir.uni-koblenz.de/">http://essir.uni-koblenz.de/</a></td>
<td>Ian Ruthven</td>
<td>06/01/2011 17:26</td>
</tr>
</tbody>
</table>
Personal Search

Guided by User Recollections
Memory is Important

Systems should take memory into account:
• Support what people are likely to remember / not remember
• Help people remember more

There isn't much IR literature on memory!
Cognitive Psychology

- 130+ years of literature
- Theories / models on (for starters):
  - Spatial recollection
  - Episodic recollection
  - Semantic recollection
  - Recollection for Texts
  - Cue-based recall
- Experimental Methods
Importance of Evaluation

- Lots of people have been building tools
  - Very few of these have actually been tested
- Major problem in the field (and related fields)
  - Boardman 2004; Capra & Perez-Quinones 2006; Cutrell et al. 2006; Elsweiler & Ruthven, 2007; Chernov et al., 2007; Elsweiler et al., 2011
- Few evaluations because it is difficult
#1 Evaluating in the wild ...

Dumais et al., 2003; Cutrell et al., 2006
#2 Evaluating in the lab ...
Lab-based approaches

- Elsweiler and Ruthven, 2007
  Task taxonomy for re-finding
  1. Split population into groups
  2. Perform investigatory studies (diary studies, tours, interviews)
  3. Derive task pools for each group
This has been used for:

- Recollection for personal information
- Learn about task perception / success
- Learn about user behaviour
- Evaluate system designs
What about systems IR experiments?
From: acm@sheridanprinting.com
Subj: your registration

Body: Dear Author,

Thank you for the submission of "Understanding Re-finding behavior in Naturalistic Email Interaction Logs" to ...

Azzopardi et al., 2007; Kim & Croft, 2009
Simulated Approaches

+ Ideal for testing algorithms
  Low cost
  Repeatable

- Do they really accurately represent user behaviour?
How can we make simulated approaches better reflect real-life queries?
Seed Simulations with User Study Behaviour

What do queries look like?

- Length
- Field
- Named Entities
- Spelling Error
- Advanced Operators

Do they change in different situations?

- Different kinds of user
- Different kinds of Task
- Different kinds of Collection

Do different retrieval models work better in different situations?
• Personal Search looks like standard IR problem
• Look deeper and we see it lies at the boundaries
• Creative solutions required
• Inspiration from other fields
  • Psychology, Ethnography,HCI
Casual-leisure Search
Loewe Project
What do people need?

How do they use existing systems?

What problems do they have?
Diary Study

- 1 week during Christmas holidays
- 38 participants (19 male, 19 female)
- Ages (10-72, avg. 39.5, sd=17.4)
- Mix of educational levels, occupations and living arrangements
- 381 recorded needs
Differences to our classical understanding of information needs

- Not in response to a gap in knowledge
- But in response to a mood or physical state
- To a need to be distracted
- To having some free time
Different emphasis on what is important for the user

- The information is not always what is important
- Experience is always crucial

- Success $\neq$ finding something (specific)
- It is the journey not the destination that is key!
Casual-leisure situations are important

- Many participants described escaping (monotonous tasks, stressful situations, boredom)
- Health (mental and physical)
Learning about search behaviour

I'm writing on a white board.
Learning about search behaviour

I think Sir Tim likes my idea
Learning about search behaviour
chrissysilvey Have spent the evening searching the net for a 'Dynasty' style dress for a murder mystery party !!!! Just need a wig now!
about 2 hours ago from Echofon

adicrazy Searching the net for birthday gift ideas – getting really dumb ideas.
about 14 hours ago from web

Cel22tmt I've been searching the net for time and I can't find ice cubes instrumental for 'Today was a bad day' some ones got it! Free dat up man lol
about 21 hours ago from UberTwitter

alfatweet Are you fed up with searching the net for an answer than watch these FREE videos! http://bit.ly/6cl53i
about 21 hours ago from API

claire98uh1 Waiting for the signing of Darren Bent.Searching the net for latest info
1 day ago from web
Missing Knowledge Gap

1)  ... I’m not even *doing* anything useful... just browsing eBay aimlessly...
2) to do list today: browse the Internet until fasting break time..
3)  ... just got done eating dinner and my family is watching the football. Rather browse on the laptop
4) I’m at the dolphin mall. Just browsing.
Experience over things found

1) exploring the neighbourhood with my baby!
2) What a beautiful day to be outside playing and exploring with the kids:)
3) Into the nineties and exploring dubstep [music] while handling lots of small to-dos
What does this mean for building systems?
How do you build an IR system that deals with the query “Entertain Me”? 
What does this mean for evaluating systems?
• We need to better understand what people want in various Casual-leisure situations
• How they behave to get this
• What we can do to provide assistance
Casual leisure information needs in mobile context

- Long nights of music, museums, science
- Evenings of entertainment, distributed over a city from 8pm – 3am
- Android app to support people find events of interest and the plan evenings / and routes to events
• Munich, May 2011
• ~160 bands / artists performed at 100 locations around the city (8pm – 3am)
• 20,000+ visitors with wide-ranging demographics
• We had over 500 downloads
• We logged user interactions
Learning about search behaviour

- Queries? Genres?
- Do they like to search for individual events of interest or
- Do they prefer to have routes prepared for them? If so do they edit these afterwards?
- We can learn a lot about how they think and what they want / need
How does search behaviour influence the evening

- How do they enjoy their evening?
- How many events do they visit?
- How long do they stay at the long night?
- How much time do they spend travelling between events?
- What kind geographical coverage do users have?

Long nights of Science and Museums in October
Health and Behaviour Change
In England nearly 1 in 4 adults, and over 1 in 10 children aged 2-10, are obese.

In England 2,338,813 are registered diabetic (5.4% of the population)

How can IR / IA help?
Self-efficacy is key to behavioural change
Individual can act on these (or not)

Collect data about an individual and his / her life

Give them appropriate tips or information

Present it to them in a way they can relate to

---

Thomas Goez's TED Talk, 2010

http://www.ted.com/talks/lang/eng/thomas_goetz_it_s_time_to_redesign_medical_data.html
Collecting Personal Data
How do you best present this information so that the individual can relate to it?

- Show with temporal context?
- Show with context of peers?
- Give warning feedback?

We have a PhD student working on this!
• Collection of articles from a German health magazine
• Medical Professionals
• Providing relevance judgements based on sensor values
• Providing extra documents they specifically think are relevant
Open questions and logistical issues

• How can you measure whether people have acted on information?

• Data collection issues! How much and how long do we need to collect data to detect behavioural change?
Food Recommender

http://m4bu.dyndns.org/bachelor/register/
Data Collection

- Collecting recipe ratings
- 5000 main meals and 500 breakfasts from 140k chefkoch.de
- 136 users have rated 3422 ratings after 4 weeks
- Reasons behind their rating
Our Idea

- Analyse the factors behind ratings
- Device strategies to move people towards rating healthier meals higher
Healthier food with tomatoes
Healthier meals that are quick to prepare

Just ideas!
Nutrition is incredibly complex!

What is healthy?

We have a nutritionist on the project

Data to show behavioural change toward healthy meals
Summary

• At IR boundaries you find interesting problems
• Challenging problems
• Creativity and inspiration from other fields can provide solutions
Some tips for young researchers

1. Read broadly
2. Talk to anyone and everyone
3. Build a network
4. Find a niche
5. Publish articles that count