



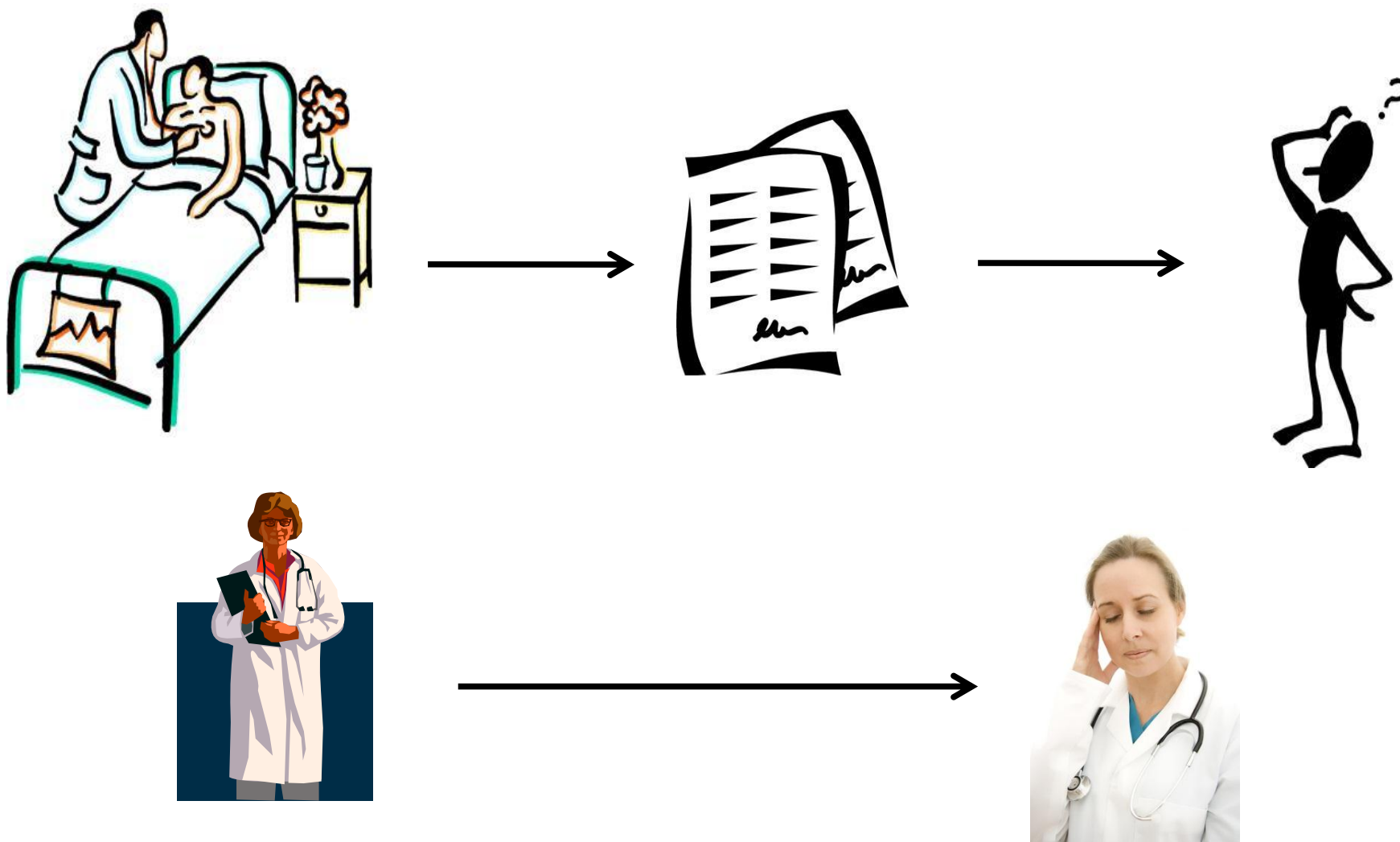
CLEF 2015 labs building on CLEF 2014 **and more!**

<http://clef2015.clef-initiative.eu/>

- ✓ **e-health:** *clinical texts IE and Patient-centred IR*
- ✓ **Image:** *annotation, medical application and prediction*
- ✓ **LifeCLEF:** *multimedia analysis & environmental challenges*
- ✓ **LL4IR: Living Labs for Information Retrieval Evaluation**
- ✓ **NEWSREEL:** *NEWS REcommendation Evaluation Lab*
- ✓ **PAN:** *Plagiarism, Authorship and Author profiling*
- ✓ **QA:** *Biomedical and Open Domain*
- ✓ **SBS: Social Book Search (Suggestion & Interactive)**

[Contact : eric.sanjuan@univ-avignon.fr](mailto:eric.sanjuan@univ-avignon.fr)

CLEF eHealth 2015 - Scenario





CLEF
eHealth
2015



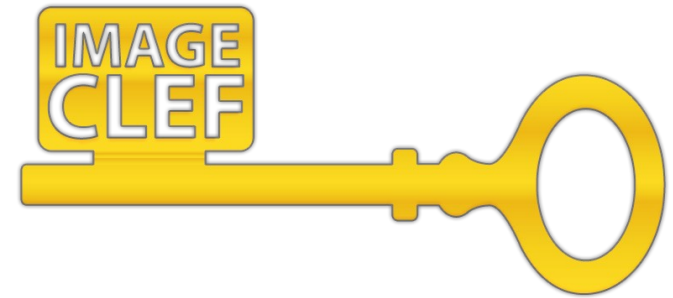
Task 1: Information extraction from clinical text

New: non-English languages, clinical spoken language

Task 2: Patient-centred information retrieval

New: queries, evaluation criteria, CLIR languages

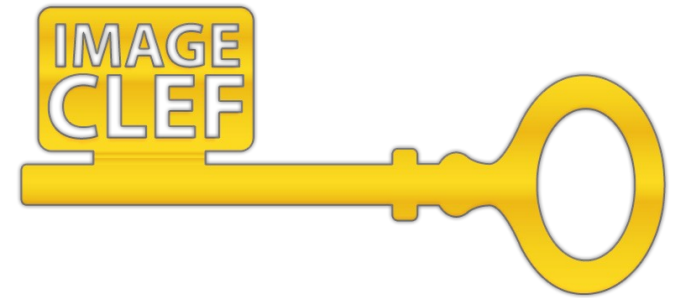
ImageCLEF 2015



Three tasks have been proposed:

- **ImageCLEFannotation**
 - Scalable detection and localisation of concepts
 - Plus generation of sentence descriptions
- **ImageCLEFmed**
 - Medical literature compound image identification, labelling and separation
 - Clustering of x-ray images of bone fractures
- **ImageCLEFliver**
 - Prediction of missing LiCO radiological annotations for liver CT image volumes

ImageCLEF 2015



Overall coordination:

Dr. Mauricio Villegas (Universitat Politècnica de Valencia)

Prof. Dr. Henning Müller (HES-SO)

Task organizers:

- **ImageCLEFannotation**

Dr. Andrew Gilbert (University of Surrey), *Dr. Luca Piras*
(University of Cagliari),

The ViSen consortium

- **ImageCLEFmed**

Alba García Seco de Herrera and *Dr. Stefano Bromuri* (HES-SO),

Dr. Ashraful Amin (Independent University Bangladesh),

Mahmood Kazi Mohammed (Sir Salimullah Medical College)

- **ImageCLEFiver**

Prof. Burak Acar and *Dr. Suzan Uskudarli* (Bogazici University),

Prof. Jose F. Aldana and *Prof. Maria del Mar Roldan Garcia*
(University of Malaga)



2015

The new lab of CLEF
dedicated to life media

Context & Objectives

- Context: biodiversity preservation
 - **the taxonomic gap** is a tricky problem to massively collect observations of plants and animals
- Objectives
 - Study, evaluate and boost state-of-the-art content-based **multimedia identification** methods (signals+metadata)
 - Assemble a **transdisciplinary** and **cross-media** community around the topic
 - Promote **environmental challenges** in the multimedia community





2015

The new lab of **CLEF**
dedicated to **life media**

Three tasks



BirdCLEF an audio record-based bird identification task
based on the **Xeno-Canto** social network
500 bird species from Brazil from hundreds recordists
around 15k **recordings**



PlantCLEF an image-based plant identification task
based on the **Tela Botanica** social network
500 plant species from France from hundreds
of photographers, around 50k **images**



FishCLEF a fish video surveillance task
based on the **Fish4Knowledge** network
30 fish species from the Taiwan's coral reef from underwater
cameras, 2000 **videos**, and 2 million **images**



Living Labs for Information Retrieval Evaluation (CLEF LL-IR 2015)

Academics $\Leftrightarrow$ LL-IR API $\Leftrightarrow$ Real Websites

3 user cases to start with :

Product search	Local domain search	Web search
regiojatek.hu	uva.nl	seznam.cz

Organizers :

Krisztian Balog (University of Stavanger, Norway, <http://krisztianbalog.com/>)

Liadh Kelly (Trinity College Dublin, Ireland, <http://www.computing.dcu.ie/~lkelly/>)

Anne Schuth (University of Amsterdam, The Netherlands, <http://www.anneschuth.nl>)

CIKM'14 short paper: Head First: Living Labs for Ad-hoc Search Evaluation

CLEF LL-IR 2015

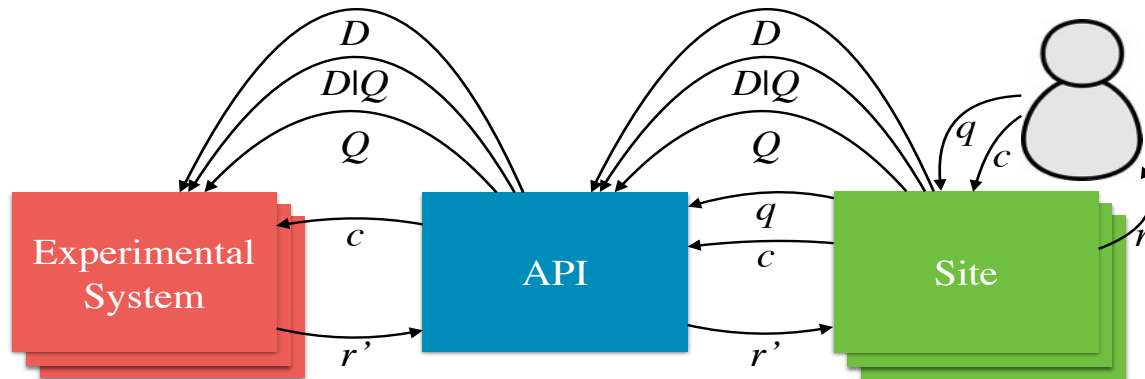
Living Labs for Information Retrieval Evaluation

Give us your ranking, we'll have it clicked!

- Overall goal: make IR evaluation more realistic
 - Evaluate retrieval methods in a *live setting* with *real users* in their *natural task environments*
- Three specific use-cases
 - **Product search** (on an e-commerce site)
 - **Local domain search** (on a university's website)
 - **Web search** (through a commercial search engine)
- Why is it interesting for you?
 - Access to privileged commercial data (search and click-through data)
 - Opportunity to test IR systems with real users in a live setting (not the same as crowdsourcing!)

How does it work?

- You're given: queries, candidate documents, historical search and click data
- Generate rankings for each query and upload them through an API
- (When any of these queries is fired) sites request rankings from the API and interleave them with that of their production systems
- You get detailed feedback on user interactions (clicks)
- Ultimate measure is the number of “wins” against the production system



For more information, go to living-labs.net!



CLEF NEWSREEL

Frank Hopfgartner, Technische Universität Berlin

CLEF NEWSREEL - Overview

- In CLEF NEWSREEL, participants can **develop news recommendation algorithms** and
- have them tested **in real-time** by millions of users over the period of a few months in a living lab (Task A)
- test them in **simulated real-time** using a novel bench\ -marking framework (Task B)



The screenshot displays a web browser window with a news article titled "BMW 6er Cabrio". The article text discusses the car's performance and features. To the right of the article is a sidebar titled "Das könnte Sie auch interessieren" (This might also interest you). This sidebar contains several recommendation cards, each with a small image and a text snippet. The cards include:

- "Daran erkennt man gute Autovermietungen" (From this you can recognize good car rentals)
- "VW Jetta" (VW Jetta)
- "Was bringt ein schadstoffarmer PKW" (What does a low-pollution car bring)
- "Beim Autokauf ein Schnäppchen machen" (Getting a bargain when buying a car)

At the bottom of the sidebar, there is a link that says "hier werben" (advertise here). The interface also shows a search bar at the top and a navigation menu on the left.

Source (Image): T. Brodt of plista.com

CLEF NEWSREEL - Tasks



Benchmarking news article recommendations

Task A: Living lab

- **Task:** Provide recommendations for visitors of **German news portals** within <100ms
- Various portals covering **different domains** (local news, sports, business, technology)
- Communication via Open Recommender Platform (ORP)

Task B: Simulation

- **Task:** Predict users' clicks on recommended news articles in simulated real-time
- Traffic and content updates of various German-language **news content provider** websites
- Use of **benchmarking framework** that allows for the simulation of data streams



@clefnewsreel

<http://www.clef-newsreel.org/>

PAN 2015

- **Plagiarism detection**
- **Authorship identification**
- **Author profiling...**

... **age + gender + personality...** in 

Big 5 personality traits

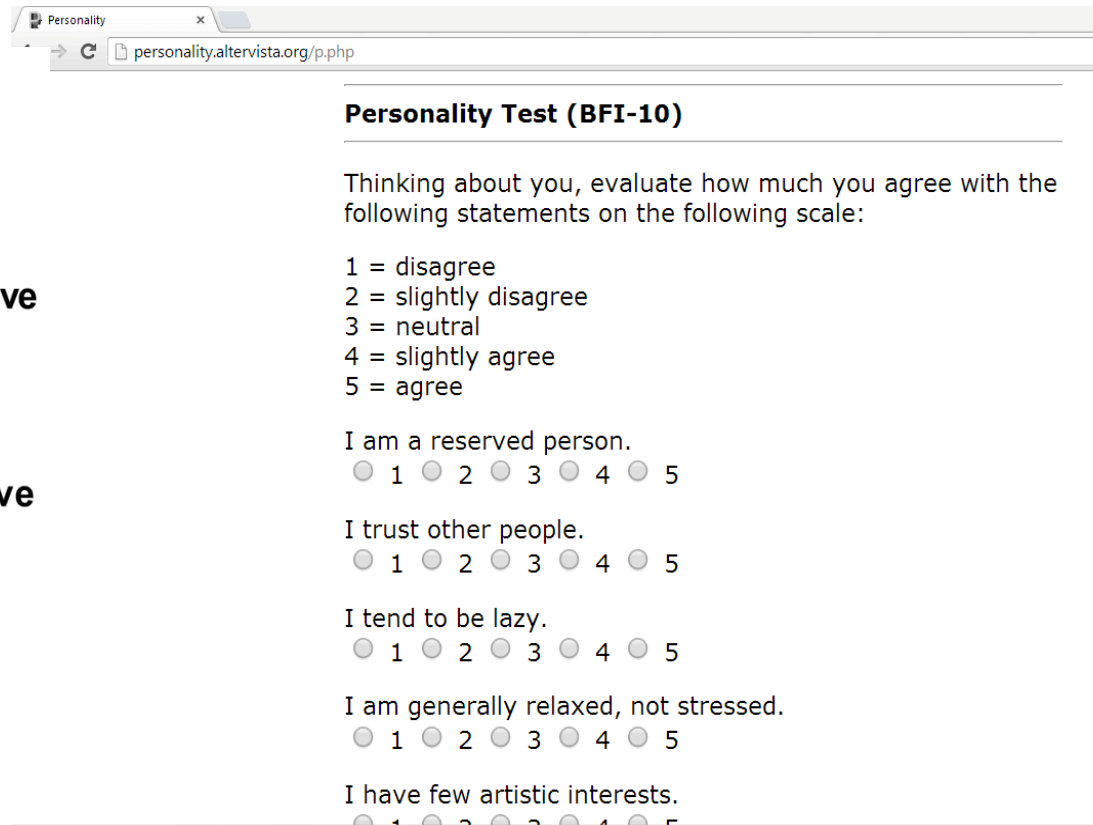
extrovert |-----| shy

stable |-----| neurotic

friendly |-----| uncooperative

organized |-----| careless

insightful |-----| unimaginative



The screenshot shows a web browser window with the address bar displaying "personality.altervista.org/p.php". The page title is "Personality". The main heading is "Personality Test (BFI-10)". Below the heading, there is a paragraph: "Thinking about you, evaluate how much you agree with the following statements on the following scale:". This is followed by a list of five statements, each with a corresponding scale from 1 to 5. The scales are represented by radio buttons.

Personality Test (BFI-10)

Thinking about you, evaluate how much you agree with the following statements on the following scale:

1 = disagree
2 = slightly disagree
3 = neutral
4 = slightly agree
5 = agree

I am a reserved person.
☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

I trust other people.
☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

I tend to be lazy.
☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

I am generally relaxed, not stressed.
☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

I have few artistic interests.
☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

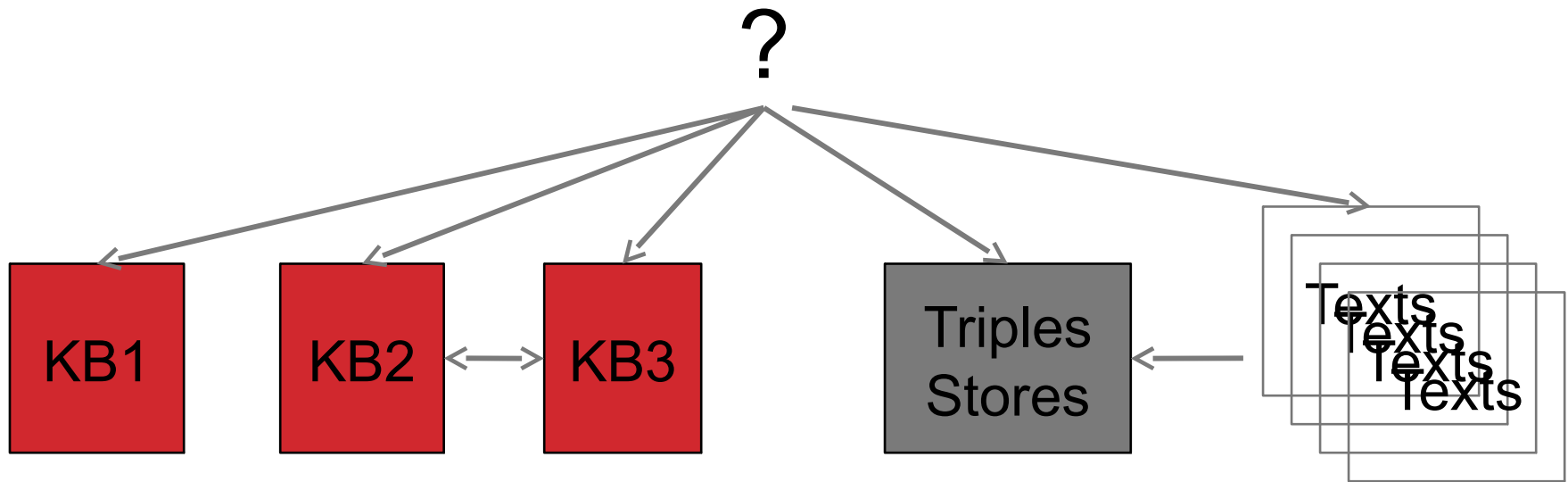
<http://personality.altervista.org/personalitwit.php>

CLEF QA

Track 2015

General Scenario

Natural Language Question



Some questions need querying KBs
Aggregations, logical inference, ...

Some questions need text processing
Textual inference, paraphrasing, ...

Some questions need both

Two instances

Biomedical

- Targeted to (bio)medical experts
- Medical Knowledge Bases / Ontologies
- Medical articles

Open domain

- Targeted to non-experts
- DBpedia, etc.
- Wikipedia articles, etc.

CLEF QA TRACK 2015

Tasks

1. Question Answering over Linked Data

- Phenomena addressed in different exercises in previous years have to be addressed in a single exercise/dataset
- The challenge: identify different phenomena and develop different appropriate strategies

2. Entrance Exams

- Systems improved results this year
- Same configuration next year, finish 3-year cycle

3. Large-Scale Biomedical Semantic Indexing

- Excellent participation, addresses a real problem with economic impact
- Systems still improving results

4. Biomedical Semantic Question Answering

- Systems still improving results
- Same configuration next year, finish 3-year cycle

... and beyond

Already thinking about the new tasks for 2016...

See you in Toulouse!

Social Book Search Lab at CLEF 2015

Marijn Koolen, Mark Hall, Hugo Huurdeman, Jaap
Kamps, Mette Skov, David Walsh

Professional metadata

Browse Books

Amazon Kindle

Kindle (Wi-Fi, 6")
Kindle 3G (Free 3G+Wi-Fi, 6")
Kindle DX (Free 3G, 9.7", Graphite)
Kindle Books

Books

Browse the [best books of 2011 so far](#), [best-selling books](#), [new and used textbooks](#), and [Kindle books](#)

Fall Reading
What's hot in Books this fall

Amazon Omnivoracious

Our Books editors' blog

[Read new posts](#)

Sell Back Your Copy

Receive an Amazon.com Gift Card for this item:



Modern Information Retrieval:
The Concepts and ...

Social Book Search

LibraryThing

Home | [Zeitgeist](#) | [Talk](#) | [Groups](#) | [Local](#)

English | Help

Search site



Become a Member?

Social content

Enter what you're reading or your whole library. It's an easy, library-quality catalog.

A community of 1,400,000 book lovers.

LibraryThing connects you to people who read what you do.

name

password

SIGN IN

or sign in with...



Tracks

- Social Book Search used to be part of INEX (Since 2007)
- **Suggestion Track**: system-centred, complex search tasks, mix of prof. metadata and user-generated content, user profiles
 - How should knowledge and opinion of user+friends/experts/others influence book ranking?
- **Interactive Track**: user-centred, focus on stages in search process, multi-stage search interface
 - How can interface support user at each stage of search process?