



## HTAi Interest Sub-Group on Information Resources (IRG) - Business Meeting -

Radisson Blu Plaza Hotel, Oslo, June 15<sup>th</sup> - 13:30 to 14:30

# Classification scheme for Information Retrieval methods

### Project partners:

Sigrid Droste  
Institute for Quality and Efficiency in Health Care  
-IQWiG-, Depart. Quality in Health Care  
Cologne, Germany

Andrew Booth  
School of Health and Related Research –SchARR-  
University of Sheffield  
Sheffield, UK

## Information Retrieval is

- not new, but
- much easier than in former days ...



**“It's not the genius who is 100 years ahead of his time but average man who is 100 years behind it.”**

(Robert Musil, Austrian novelist (1880-1942))

# Modern Information Retrieval



- started after World war II
- Many recent IR methods have their roots in the 1950s and 1960s

## Examples for methods introduction

1957	Word frequency statistics
1967	Assessment on more than 50 IR methods
1969	IR evaluation framework
1971	Cluster searching
1978	Pearl growing
1989	Berry picking

**Mid 1990s: IR is an established discipline with sophisticated, internationally recognized methods, terms and definitions.**

## Background

### What's the problem?



- World-wide introduction of the **internet**
- **Easy (free) access** to many biomedical bibliographic databases
- **Easy to use** search functionalities and **search filter** introduction

### Consequences

- A multitude of **additional disciplines** (not / less educated in IR / informatics) entered the field of Information Retrieval
- New disciplines / institutions introduced their **own terminology**
- **New terms** were created although established terms and definitions already did exist
- **New or modified (wrong or incomplete) definitions** of existing terms were introduced.
- Attempts started to **define methodological standards** without knowing resp. ignoring the Mid 1990s state-of-the-art.

## Our aim?



### To recall some basics of IR

- Multitude of existing methods
- **What counts is the result, not the query / procedure.**
- High-quality IR results do require adequate IR methods and adequate IR competency
- Acceptance of IR competency is the result of applying and disseminating recognized expertise, incl. respecting the competency of precursors.

One important step to recover IR knowledge and to ensure high-quality IR is to collect and disseminate the existing methods, terms and definitions (share common language not creating a “Tower of Babel”). Thus we started preparing a

- Glossary of terms of interest and a
- Classification scheme for information retrieval methods

# What we did since the Washington Meeting



## First part of the project

### The Glossary

1. **Preparing a Glossary of terms** of interest currently incl. 350 Information Retrieval methods terms.

Details for each entry:

- Name of term,
- Definition,
- Source/s of definition,
- Synonym/s,
- Related term/s,
- Evaluations / applications (references).

2. **Reviewing and adapting the terms** to the ISO standards syntax requirement on preparing dictionaries.

Publishing the Glossary (ongoing)

The screenshot shows the HTA Glossary .net website. At the top left is the logo 'HTA Glossary .net'. In the top right corner, there is a 'Login' link. Below the logo is a search bar with a magnifying glass icon. To the right of the search bar is a 'Print' button. The main content area is titled 'external validity' and contains the following text: 'The ability of a research design to provide findings that can be generalised to other populations, contexts and periods.' Below this is the text 'Syn.: generalizability.' and 'Related concept: applicability'. At the bottom of the page, there is a navigation menu with the following items: 'List of all terms', 'Terms by field', 'About the English version', 'User's guide', 'Bibliography', and 'Links to other glossaries'.

3. Preparing a **list of search heuristics** (currently about 100)

# What we did since the Washington Meeting



## Second part of the project (ongoing)

### The Classification scheme

- comprehensive systematic literature search to identify publications on information retrieval methods.
- Building-up the hierarchical classification scheme (main categories and sub-categories) ⇒ PubMed MeSH

### Example

- Methods related to the information source
  - Databases
    - Fulltext searching
      - ...
    - Citations
      - ...
    - Factual data
    - ...

- Berry picking
- Chaining ((forward, backward)
- Citation searching (forward, backward, bidirectional)
- Citation tracking (forward, backward, bidirectional)
- Cited reference searching
- Co-citation retrieval, author
- Co-citation retrieval, work
- Known-item instantiation search
- Pearl growing
- Snowballing
- ...

## Outstanding activities

- **Defining keywords** e.g. applicable to the SuRe Info tool
- **Publishing** the Glossary and the Classification scheme
- Reporting and **publication writing** (e.g. IJTAHC or JMLA).



**Takk skal du ha**

### **Contact:**

Sigrid Droste  
Institute for Quality and Efficiency in Health Care – IQWiG –  
Department Quality in Health Care  
Mediapark 8  
50670 Cologne  
Germany  
sigrid.droste@iqwig.de