Users in IR Context

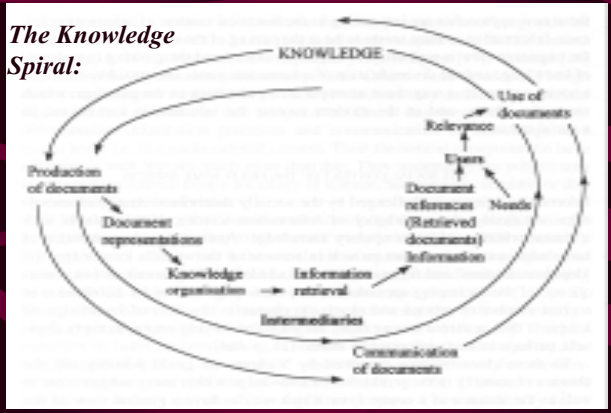
Information Seeking & Retrieval - Models and Perspectives

Peter Ingwersen
Department of Information Studies
Royal School of Library and Information Science
pi@db.dk - http://www.db.dk/pi

Table of Contents

The communication system for information science.
Ørom, A. in: J. Documentation, 2000, p. 12-26

The Knowledge Spiral:

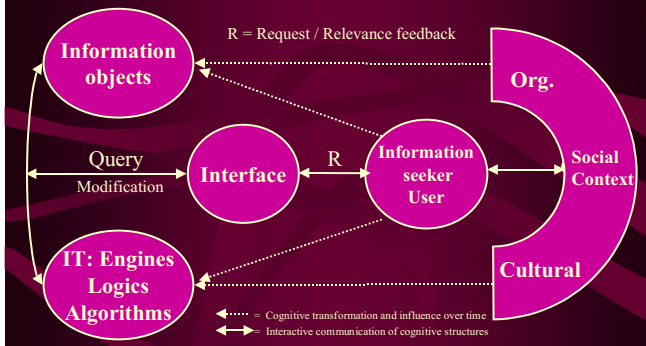


Users in Context - the disturbing variable in experimental IR - 1

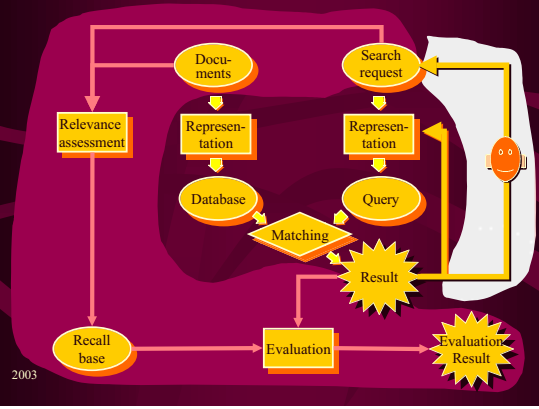
- Experimental IR:
- Non-interactive - system-driven - algorithmic
  - Goal: relative performance of engines
  - Means: one-run experiments; sets of queries; mono-dimensional assessor judgements of pooled objects

Simplistic model of longitudinal IS&R

The Turn. P. Ingwersen & K. Järvelin. Klüver, 2004 (forthcoming)



The Laboratory Model 2 (Kalervo Järvelin)



Users in Context - the disturbing variable in experimental IR

- Interactive IR - cognitive - user centred
  - Goal: understanding which engines, information structures & interface functionality that best suit/support information seeking behaviour in work (task) contexts
  - Means: iterative or longitudinal experiments; sets of simulated work task situations/real needs; multidimensional relevance assessments by users; info. needs are variable over session

II Retrieval models - overview 1

- IIR in context of Information Seeking Behaviour and Work Task Situations:
  - e.g. as part of scientific communication
  - or solving a (work / interest) problem
- Information seeking models:
  - T.D.Wilson´s models (1981…1996…1999)
  - Dervin & Nilan: sense-making (1986)
  - Kuhlthau´s phenomenological stage model (1991)
  - Byström & Järvelin, 1995 – Vakkari, 2000

II Retrieval models – overview 2

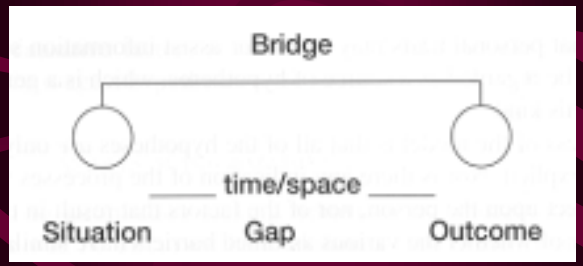
- Ingwersen´s model for IIR (1992/96)
- Saracevic´ stratified model for IIR (1996)
- Ingwersen/Järvelin cognitive model (2004)

– **The *relevance connection* - and the association to *information use* - in**

- **SITUATED CONTEXTS**

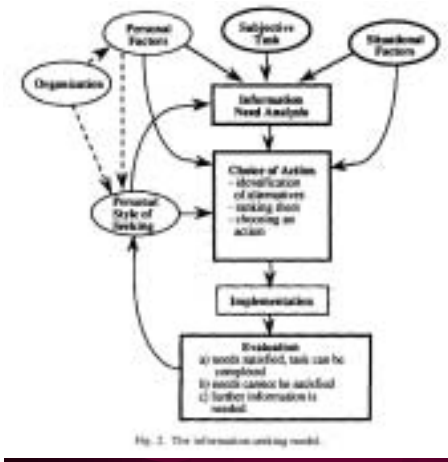
Wilson´s 1981 model of Information seeking

Dervin & Nilan´s sense-making (1986)



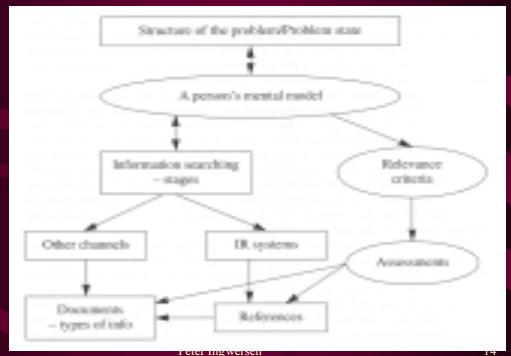
Carol Kuhlthau's stage model – 1991/94

Stages	Initiation	Selection	Exploration	Formulation	Collection	Presentation
Feelings	Uncertainty	Optimism	Confusion, frustration, doubt	Clarity	Sense of direction, confidence	Relief, satisfaction or disappointment
Thoughts	Vague			Clearer	Increased interest	Focused
Actions	Seeking background information		Seeking relevant information		Seeking pertinent information	
Appropriate tasks	Recognize	Identify, investigate	Identify, investigate	Formulate	Gather	Complete

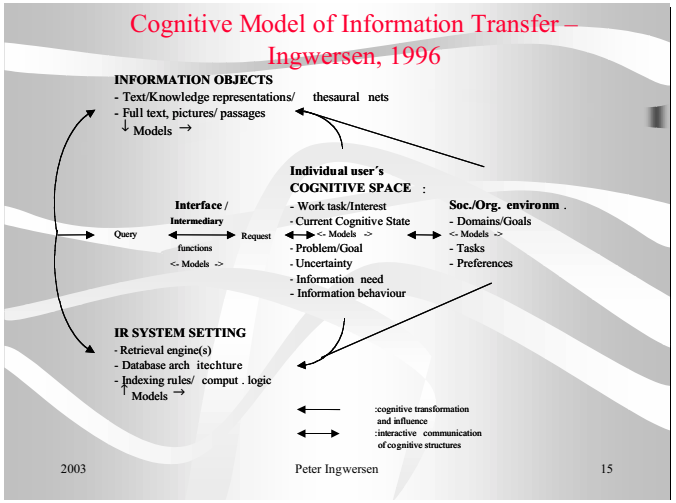


IS&R model, 1995: Byström & Järvelin, fig. 2

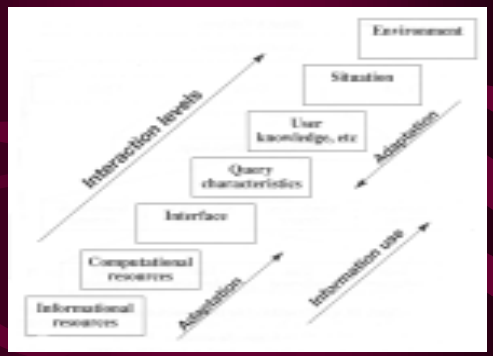
IR and relevance in Seeking context - Vakkari 2000



Cognitive Model of Information Transfer – Ingwersen, 1996



Saracevic´ stratified model for IIR (1996)



Resumé of models of ISR

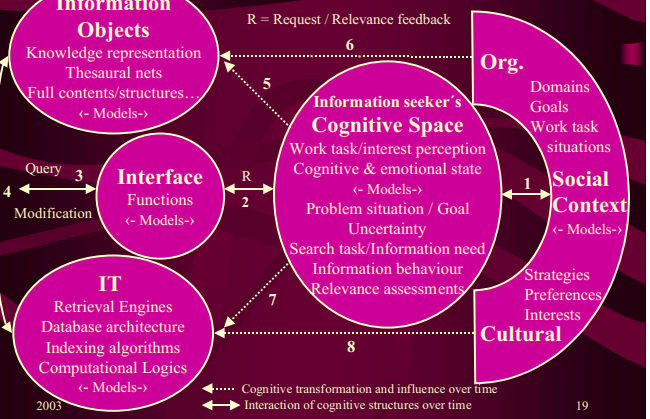
- ASK (1982) by Belkin et al.; Dervin & Nilan (1986); Ingwersen (1992, 1996); Byström & Järvelin (1995); Saracevic (1996);
- refer to the situation or context as the trigger of IS&R
- Kuhlthau (1991) concentrates on the mental process, like Wilson (1999) and Vakkari (2000).

The cognitive approach to IIR

- Ingwersen´s cognitive communication model (1996), based on Belkin (ASK) (1982) and earlier alike but simpler models
- 5 major components that act as context for one another during IIR
- Two kinds of tasks: **WORK & SEARCH**
- Two kinds of knowledge: **DOMAIN & IR**
- **Work task perception as trigger for search task performance, incl. information need perception**

Complex Cognitive model of IS&R – over time

The Turn. P. Ingwersen & K. Järvelin. Klüver, 2004 (forthcoming)



Work Task vs. Search Task

The Turn. P. Ingwersen & K. Järvelin. Klüver, 2004 (forthcoming)

- **WORK TASKs** or (socio-cultural) **Interests** may **exist objectively** (in environment) or **subjectively** (in the mind of the actor)
- Nonetheless: they **are perceived by actor** to be fulfilled or solved – **by MEANS of** action, i.e.
- **SEARCH TASKS** – instruments – the way **… *as means to an end***
- Search tasks are initiated by a perceived information need / gap & the searching process

Knowledge types in task performance

The Turn. P. Ingwersen & K. Järvelin. Klüver, 2003 (forthcoming)

Perception ⇨ / ***Task type*** ⇩	**.. of *passive* features of task**	**.. of *activity* features of task**
Work task** or **interest	Cognitive & emotional task knowledge (1)	Problem & task solving knowledge (2)
***Search task** – interactive IS&R*	Information source & system knowledge (3)	IS&R process & activity knowledge (4)
***Search task** – inter personal communication*	Person & group knowledge (5)	Social interaction skills (6)

Work & Search Task Categories

- **According to COMPLEXITY in 3-5 categories:**

→Decision tasks (genuine or known)
→"Normal" tasks (decision or info. processing)
→Routine / automatic tasks (information processing)

→**Complexity depends on amount/type of information required on:**

Task contents - domain knowledge - task solving

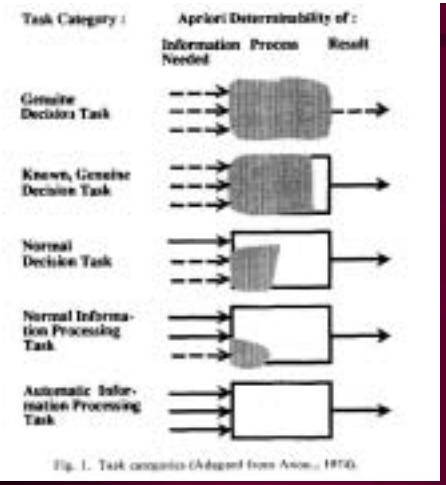


Fig. 1. Task categories (Adapted from Anon., 1976).

Work task complexity

Byström & Järvelin, 1995, IP&M

⟶ Task solving info
⟶ Domain info
⟶ Task contents
⇢ Info. is missing

Situational context > Work task / Interest > Perception > Uncertainty > Information Need

- **The more complex** the situation and work task - the greater the uncertainty and knowledge gap (Byström & Järvelin, 1995);
- The information need becomes increasingly ill-defined – people required as sources
- Impact on search task complexity & behaviour - relevance assessments: systems design should support cognition

8 types of intrinsic information needs

The Turn. P. Ingwersen & K. Järvelin. Klüver, 2003 (forthcoming)

IS&R goal ⇨ / Known data type ⇩	**Searching for information source contents as**	**Searching for informative data entity(ies) as**	**Quality of prior knowledge**
Passive IS&R knowledge (e.g. structured bibliographic or relational data)	***Known item*** (1) ***Muddled item*** (5)	***Known data element*** (2) ***Muddled data element*** (6)	**Well-defined** vs. **Ill-defined** (exploratory)
Active/passive work task knowledge (e.g. aboutness or contents data)	***Known topic or contents*** (3) ***Muddled topic or contents*** (7)	***Factual*** (4) demand ***Muddled factual*** (8)	**Well-defined** vs. **Ill-defined** (exploratory)

INFORMATION NEED TYPES – 1

- **Given a STABLE perceived TASK**
  - TO VERIFY/FIND entire INFORMATION OBJECTS WITH KNOWN (structured) DATA = "known item" retrieval
    - Full object: by known meta data
  - TO VERIFY/FIND data elements with known (structured) data
    - Bibliographic records – client address: by known meta data
  - Information need intrinsically STABLE

INFORMATION NEED TYPES – 2

CONSCIOUS TOPICAL or CONTENTs NEEDS

3) TO CLARIFY, REVIEW OR PURSUE INFORMATION in known subject matter, domain or content (by unstructured data)
4) TO FIND TOPICAL FACTS in known subject matter, domain or content (by unstructured data)

- **WITH or WITHOUT "LABEL FFECT"**
  - Information need intrinsically STABLE or VARIABLE

INFORMATION NEED TYPES – 3

ILL-DEFINED or MUDDLED (TOPICAL or VERIFICATIVE) NEEDS

5-8) TO EXPLORE NEW CONCEPTS AND RELATIONS OUTSIDE KNOWN DATA STRUCTURES or SUBJECT MATTER or DOMAIN – exploratory information needs

- **ALWAYS "LABEL EFFECT"**
  - Information need intrinsically VARIABLE

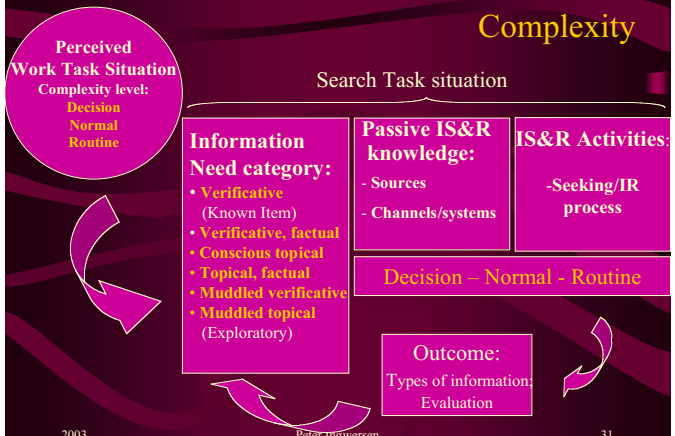
The Label Effect

- Users DO NOT ACT RATIONALLY:
- People act more randomly, becomes uncertain, even when knowing about subject matter or contents – due to:
  - Previous search task expectations & assumptions
  - influence of the domain and situation in context
- short compromised statements - labels
- known empirically since 1982 - Web IR

Label Effect implications

- Labels do not provide context
- Labels are unsuitable for ranked IR
  - relevance feedback is hence non-informative or highly uncertain at initial stages of IIR
  - query modification may help machines, if user has a rich cognitive state
- Labels make distinction between well-defined and ill-defined needs difficult

Search Task Complexity



Levels of Relevance Types - Saracevic, 1996

- **Lower order of relevance:**
- **System or Algorithmic relevance**:
  - query-object (objectivity)
- **Topical relevance**: aboutness relation of
  - query-object (interpretation/subjectivity?)

Levels of Relevance Types - Saracevic, 1996

- **Higher order of relevance:**
- **Pertinence:** perceived correspondence of
  - information need-objects
- **Situational relevance**: relation as perceived
  - between task, situation or problem and objects
- **Socio-cognitive relevance:**
  - group/peer perception of object (Cosijn & Ingwersen – 2000)
- Affective relevance (Saracevic '96): In all subjective assessments

System or Algorithmic relevance

- The ranked output of information objects - ranked by engine's relevance scores
- Commonly judged against expert assessor´s binary relevance assessments of the pooled documents
- Assessor´s judgement seen as topicality & objective - is of course intellectual (subj.)

Topical relevance

- Contains **interpretations**
- **Problem of nature of aboutness**
- Inconsistency among several assessors (yet: see Sigir 98 paper by Vorhees)
- Used for relative performance indications
- Why not simply apply the mean of the algorithmic output from sites (TREC) ?

Pertinence

- **Requires knowledge of intrinsic information need for an observer - difficult to obtain**
- **Is the domain of the seeker!**
- **May not be achieved in case of ill-defined needs**
- **Involves other facets of objects than simply topical ones (novelty – authorship – cognitive authority of journal/inst.)**

Situational relevance

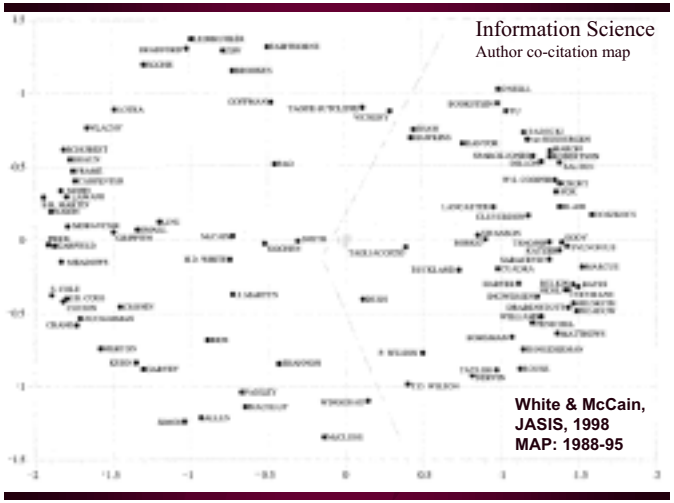
- **Relates to the WORK TASK (interest) SITUATION (e.g. a peer review task – or giving credit (references / links on a list of refs.) – expressing satisfaction - direction)**
- **Work tasks are NOT search tasks**
- **Individual relevance assessment in context**
- **Can be based on simulated work tasks (Borlund, 2002) - and observed**

Affective/motivational Relevance

- **May be found in all subjective & higher order relevance types**
- **Motivational Relevance is rather an attribute to the other relevance types**

Socio-Cognitive Relevance

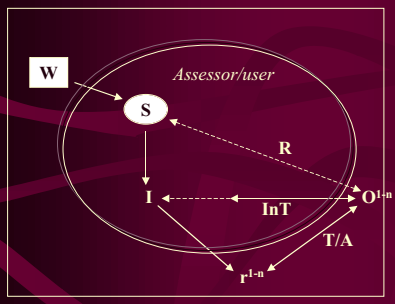
- Proposed by Ørom (JoD, Jan. 2000) as associated to the social context.
- **Discussed by Cosijn/Ingwersen (IPM, May 2000) as possibly related to organisational or social strategies, conventions & perceptions (group decisions = peer reviews (journal) & decisions at conference PC – or in domain over *time*)**
- Can be observed and measured, e.g. by (author) co-citation analysis or inlinks over time.
- **Can be used as a presentation tool of expertise:**



Such presentation tools from informetrics and visualisation

- Useful as entry for IS&R in novel area to users
- Providing relationships – knowledge sharing visualisation between:
  - People – experts – workers in knowledge-rich org.
  - Journals – and other carriers of information
  - Institutions within a region or country
  - "Everything" that can be represented by features
- Can be made dynamic over real-time

Illustration of types of relevance:



Legend:
W: Work task
S : Cognitive perception of W
I: Information need version(s)
O: Information object(s)
r : Request version(s)
T/A : Topical/Algorithmic relevance
InT : Intellectual topicality
R: Situational/socio-cognitive rel. : Relevance assessment(s) or interpretation(s)
: Transform.
: User's cognitive space

Illustration of algorithmic relevance, intellectual topicality, situational and socio-cognitive relevance. (Modified version of : Borlund & Ingwersen. Journal of Documentation, 1997).