

# user studies

enquiry foundations and methodological considerations

Giannis Tsakonas

dbis/dlib/ionio

March 31, 2011 | 1st Workshop on Digital Information Management | Corfu, Greece



# structure

---

- introduction
- enquiry foundations
  - definitions
  - types of user studies
    - usage studies
    - behavior studies
      - fundamental theories and models
- methodological considerations

part a

enquiry foundations

# definitions

---

- User studies are research studies focusing on the use of information sources from “real” (end) users.
- User studies are extended on the channels of information, such as means (formats), networks (social, academic, workplace networks) or information providing systems (information retrieval systems, digital libraries, web search systems, intranets, etc.).
- In user studies the use of information sources is considered the dependent variable, while the various aspects of the users’ profile are considered the independent variable.

# variables

---

## **Dependent**

Usage

> what is used

> how is used

> why it is used

Users' opinion

Users' preference

## **Independent**

age

sex

scientific area

educational level

communities

location

...

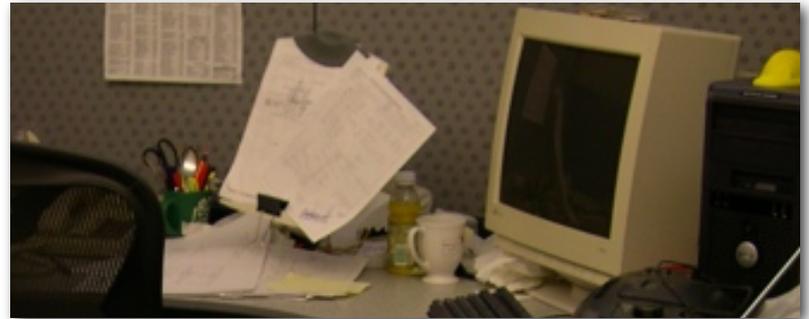
..

.

# definition of context

---

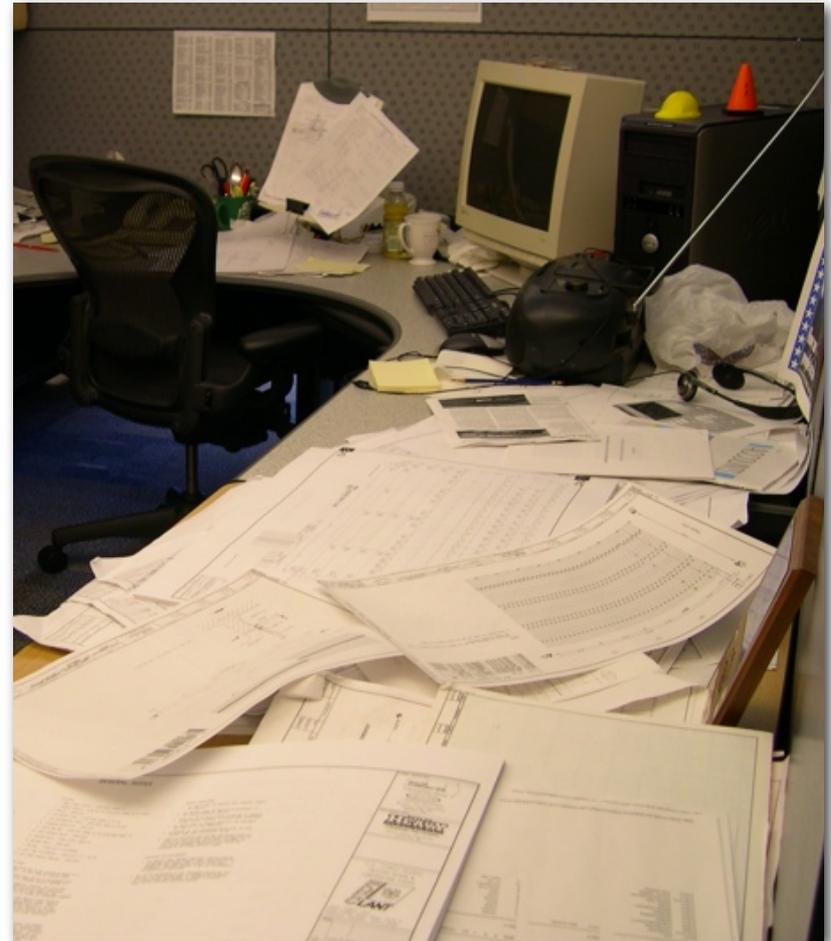
- A rather vague and unclear term
- In the present example of a context we identify:
  - a person (implied)
  - the mean (the computer and information system)
  - the content (which is supplied by the information system)
- An already rich image of a complex context.



# definition of context

---

- By expanding our view we can see:
  - the ability to interact with other information sources and means.
  - these means may pre-exist and -may again - have contributed to the formation of the user's context.



# definition of interaction

---

- Direct / personal
  - the user is interacting by his own means with an information source.
- Direct / collaborative
  - the user is cooperating with persons that understand the same or a similar work need.
- Indirect / intermediated
  - the user is supported by a third person for fulfilling his/her information need.

# definition of need

---

- Users needs are:
  - information needs
    - information needs are - usually - unclear and “messy”, while a part of them may remain unexpressed and unfulfilled.
    - information needs can be clarified or can be generated during information seeking.
  - work needs
    - work needs are more clear and their satisfaction is imperative and often takes place in pressing conditions.
- Obviously, information needs are included in work needs.

# definition of need

---

- Needs affect several expressions of information behavior.
- Needs are related to motives and constraints in information seeking.
- According to Taylor [1986] needs are:
  - Visceral
  - Conscious
  - Formalized
  - Compromised

# types of user studies

---

- User studies are distinguished in two broad categories:
  - usage studies
    - what, why and how much is used
  - behavior studies
    - how and why something is used

# usage studies

---

- Usage studies are important for the area of libraries and related information organizations.
- They have important practical significance and they can assist decision making.
- They focus on usage patterns.

# examples of findings

---

- *...one or two sessions per month, wide use of journals and articles, mostly social scientists and graduate students...*
- *...journal-focused (many sessions but concentrated on 4 or 5 specific journal titles and 50% full text, mostly scientists and graduate -students)...*
- *...topic-focused (searched less often and by subject rather than -specific journals, used many articles, mostly social scientists...*
- *Clinicians and biology researchers use e-journals differently: clinicians search online material (often abstracts) for educational and clinical purposes, while biology researchers use online material for research. Health professionals with an M.D. degree use abstracts rather than full-text articles to access treatment protocols and say they would go without an article rather than pay for online access. Those without the M.D. degree use e-journals even more often.*

Findings from SuperJournal and HighWire eJUSt (last bullet) studies as reported in Tenopir [2003]

# sources of data

---

- According to Franklin, Kyrillidou and Plum:
  - census counts
    - externally generated, e.g. COUNTER compliant statistics
    - internally generated, e.g. log files
  - sample counts
    - externally generated, e.g. LibQUAL+©
    - internally generated, e.g. MINES for Libraries©

# why?

---

- Cost-benefit decisions
  - subscription/operational costs
  - costs per use (download, view, etc.)
- Understanding users

# user behavior studies

---

- Behavior studies are known in information science as “information behavior” studies
- According to T.D. Wilson:
  - those activities a person may engage in when identifying their own needs for information, searching for such information in any way, and using or transferring that information.*
- Information behavior is inclusive of all stages of information lifecycle and is not constrained only in retrieval.

# user behavior studies

---

- Information behavior refers to the environment in one user is acting and his/her “primary” needs one attempts to cover.
- Various contributing areas:
  - psychology, with emphasis in cognitive psychology and behavioral psychology
  - anthropology and linguistics
  - communication-related sciences
  - computer science, with emphasis in human-computer interaction

# micro/macro behaviors

---

- Information searching: the sum of all interactions of the user with the system on two levels (a mental and a physical, i.e. the communication between a user and a machine's parts).
  - related concepts: search, browse, relevance, serendipity, keystroke level failures, multi-tasking, etc.
- Information seeking: the intentional and targeted seeking of information as driven by a primary level.
  - related concepts: filtering, sharing, collaborating

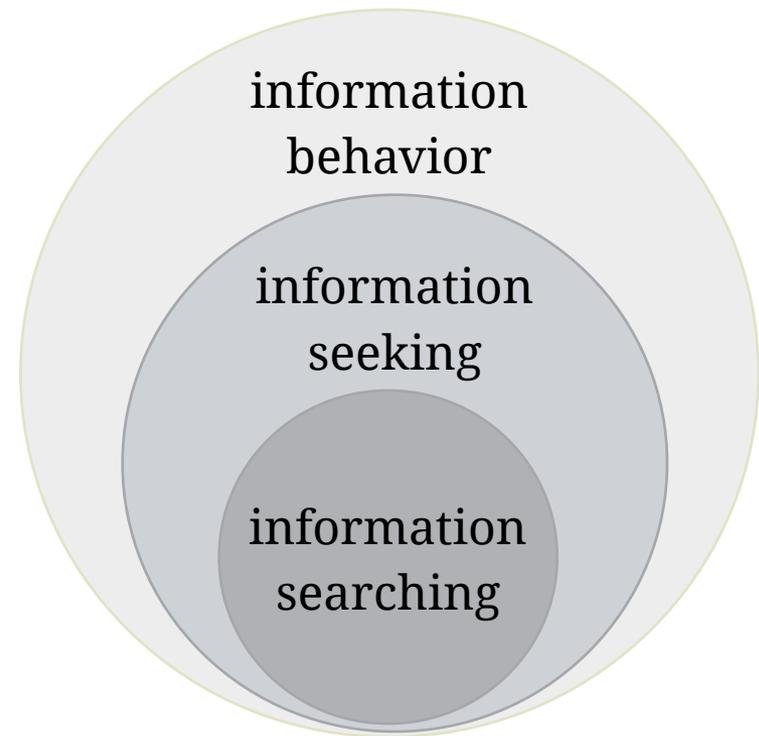


Figure by Wilson

# interaction ground

---

- The differences between the:
  - organized and certain state of a system.
  - unclear and uncertain state of user.
    - these particular differences express a certain paradigm, which is called the “bibliographic paradigm”
  - the advent of new information-organizing features, e.g. information architecture or knowledge organization features, has lessened these differences.

# fundamental theories & models

- Belkin's Anomalous State of Knowledge (ASK)
- Dervin's Sense Making
- Bates Berrypicking
- Ellis & Haugan's model for Information Seeking
- Kuhlhou's Information Search Process
- Wilson's extended model of IB

# anomalous state of knowledge

- Belkin's anomalous state of knowledge is based on the discrete mental models of the information system and the user.
- According to the model, there is an anomalous state of knowledge when the user can not resolve an information need based only on her existing knowledge. ASK reinforces uncertainty.

# ASK-coping IR systems

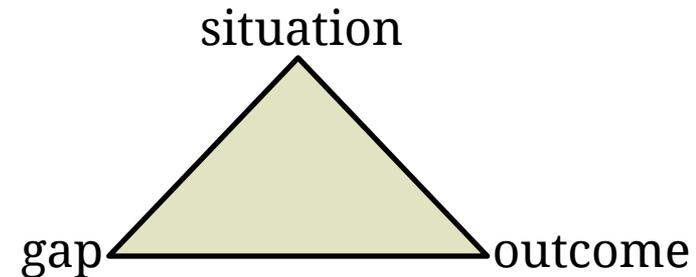
---

1. Statement of problem
2. Structural analysis of the statement
3. Choice of retrieval strategy according to type of ASK
4. Abstract presented to user simultaneously with explanations
5. Structured dialog between system and user to infer user's evaluation of
  - a. Method of choice
  - b. Suitability of document to problem
  - c. Whether need has changed
6. Modifications based on evaluation or conclusion
7. Return to 2 or 3 if needed

# Dervin's Sense-Making

---

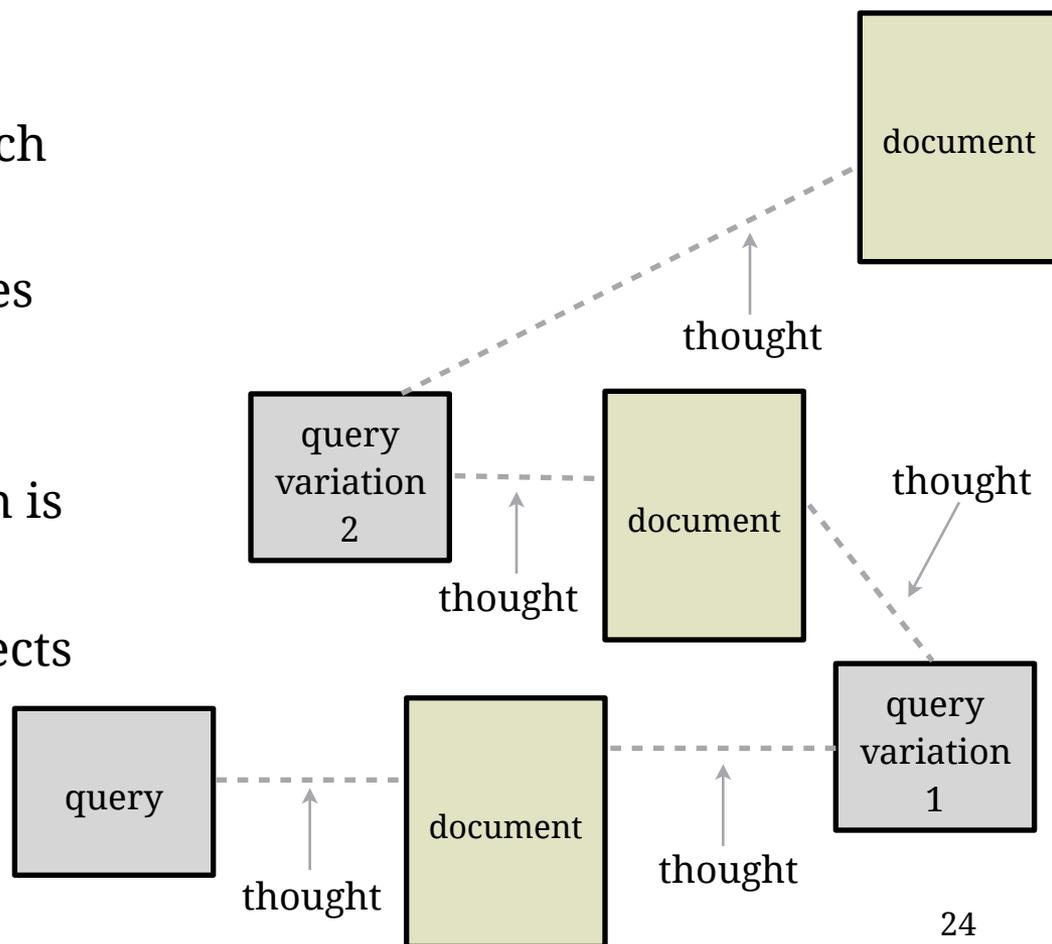
- Emphasized on the internal (cognitive) and the external (procedural) aspects of information seeking.
- A methodology rather than a model. Four methods:
  - Micro-Moment Time-Line Interview
  - Help/Hurt chaining
  - Close-ended Sense-Making Interview
  - Message Q/ing Interview
- Information seeking is a cognitive process evolved over time and space (a context) in support of resolving a knowledge gap for certain uses (desired outcomes).
- Each component has different variables.



# Bates' Berrypicking

---

- Differed from the traditional IR process:
  - Nature of the query.
  - Nature of the overall search process.
  - Range of search techniques used.
  - Information "domain" or territory where the search is conducted.
- Highlighted the dynamic aspects of queries.



# Ellis & Haugan

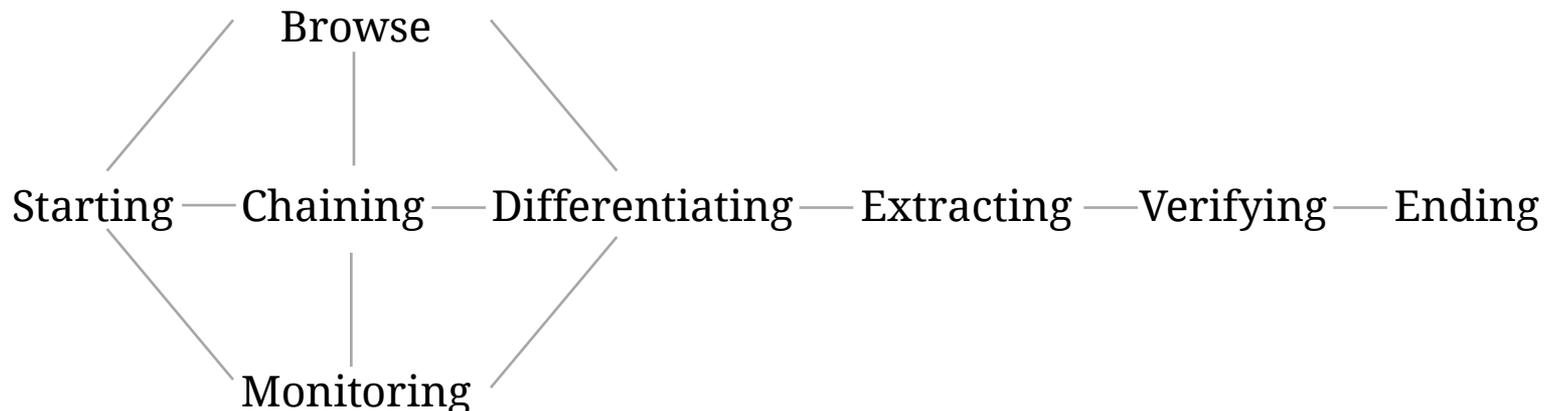
---

- Starting
- Chaining
- Browsing
- Differentiating
- Monitoring
- Extracting
- Verifying
- Ending

# Ellis & Haugan

---

- Ellis and Haugan do not refer to stages.
- They discuss on the characteristics of information seeking, which can be inter-related.
- These features refer to both micro and macro behavior level.
- Their model has a strong validation grounded on empirical studies.
- Wilson attempted to depict their model as a process.



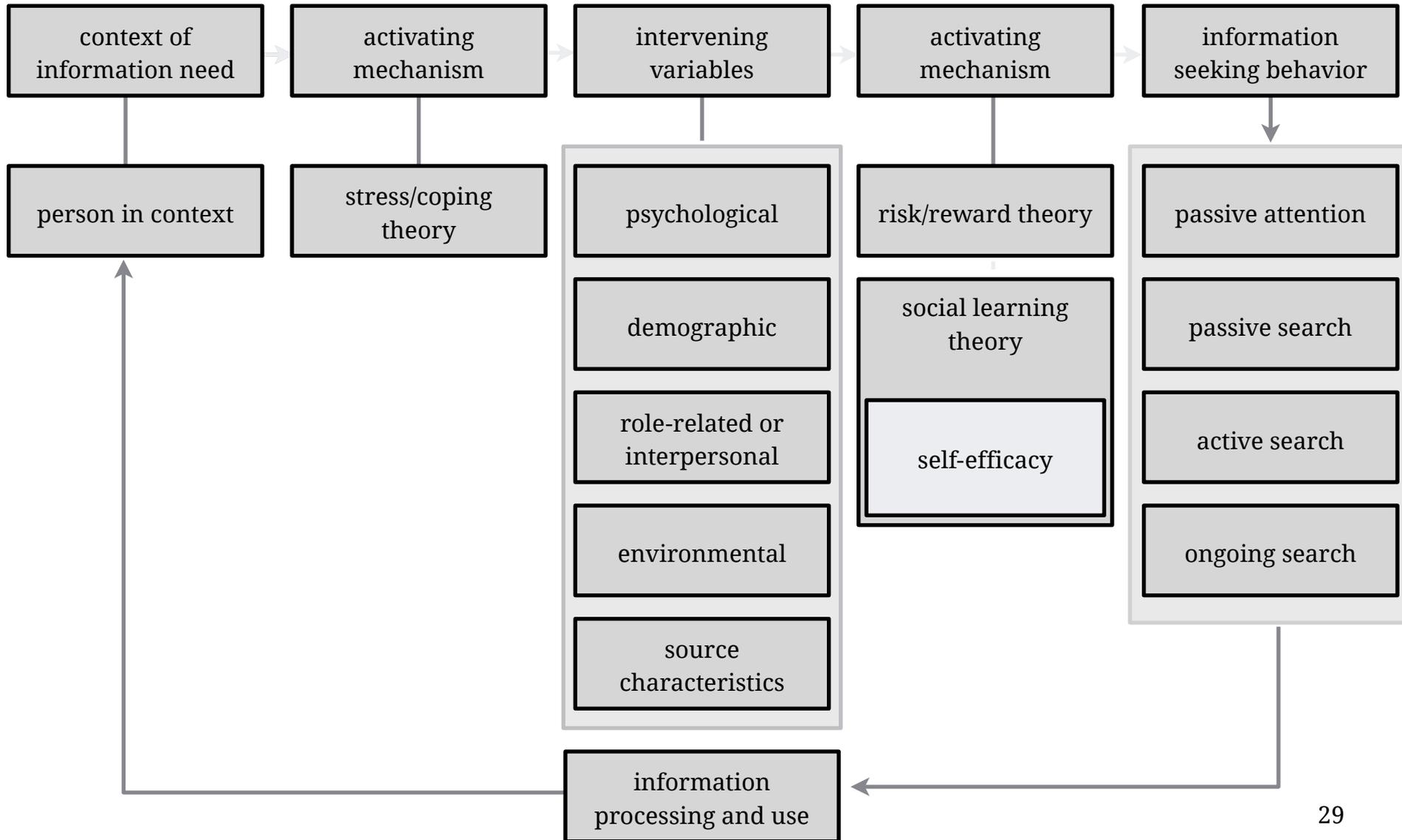
# Kuhlthau's Information Search Process

- The relation of cognitive and affective structures of user during the information search process (ISP).
- The rationale is:
  - at the early stages of information searching process there is a feeling of uncertainty (linked with the unclear state of user).
  - gradually the cognitive state of the user is improved.
  - during the later stages certainty gains ground.
- The cognitive changes are developed between the stages of resolution of an information need.
- The affective changes are developed according to the stages of searching.

# Kuhlthau's Information Search Process

Stages	Feelings	Thoughts	Actions
Initiation	uncertainty	vague	seeking relevant information
Selection	optimism		
Exploration	confusion, frustration, doubt		
Formulation	clarity	focused interest	
Collection	sense of direction, confidence	increased interest	seeking pertinent information
Presentation	satisfaction, dissatisfaction	clear	

# Wilson's extended IB model



# motives & constrains

---

- Personal features
  - cognitive dissonance, the disagreement of new information with the previous knowledge of user
  - selective exposure, the selection of information according to already known and familiar perceptions
  - physiological, cognitive and affective characteristics
  - educational level
  - the previously named as independent variables (demographic)
- Economic-related characteristics, time/cost
- Social characteristics
- Contextual characteristics, such as location, culture etc.
- Information source characteristics
  - access, physical and digital access to information
  - reliability, credibility, etc.
  - channel properties, such as readability, etc.

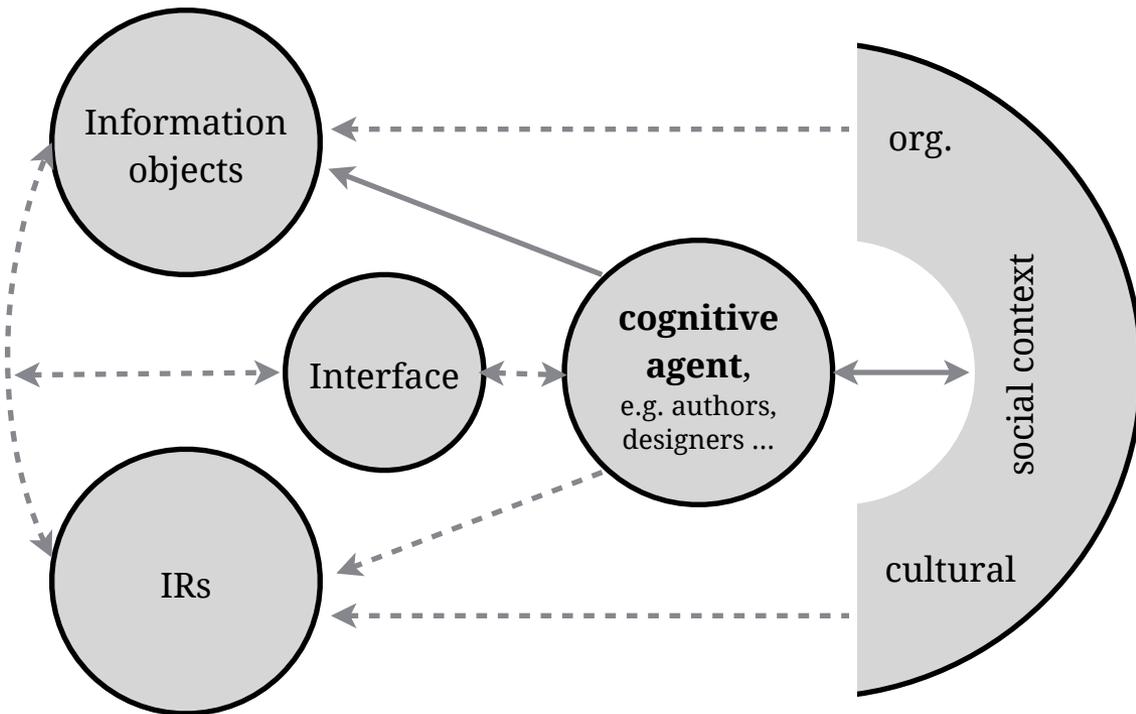
# a synopsis

---

- The previous models have established basic theories of information science and have fueled other models in IR.
- Some of them are linear, following a time-dependent sequence of actions.
- A few have empirical validation.
- The majority of these are generic models, which have been applied mainly in Information Retrieval system, due to various reasons (mainly proximity).
- The explanation of user behavior is useful:
  - support system design
  - support information literacy
  - support community practices

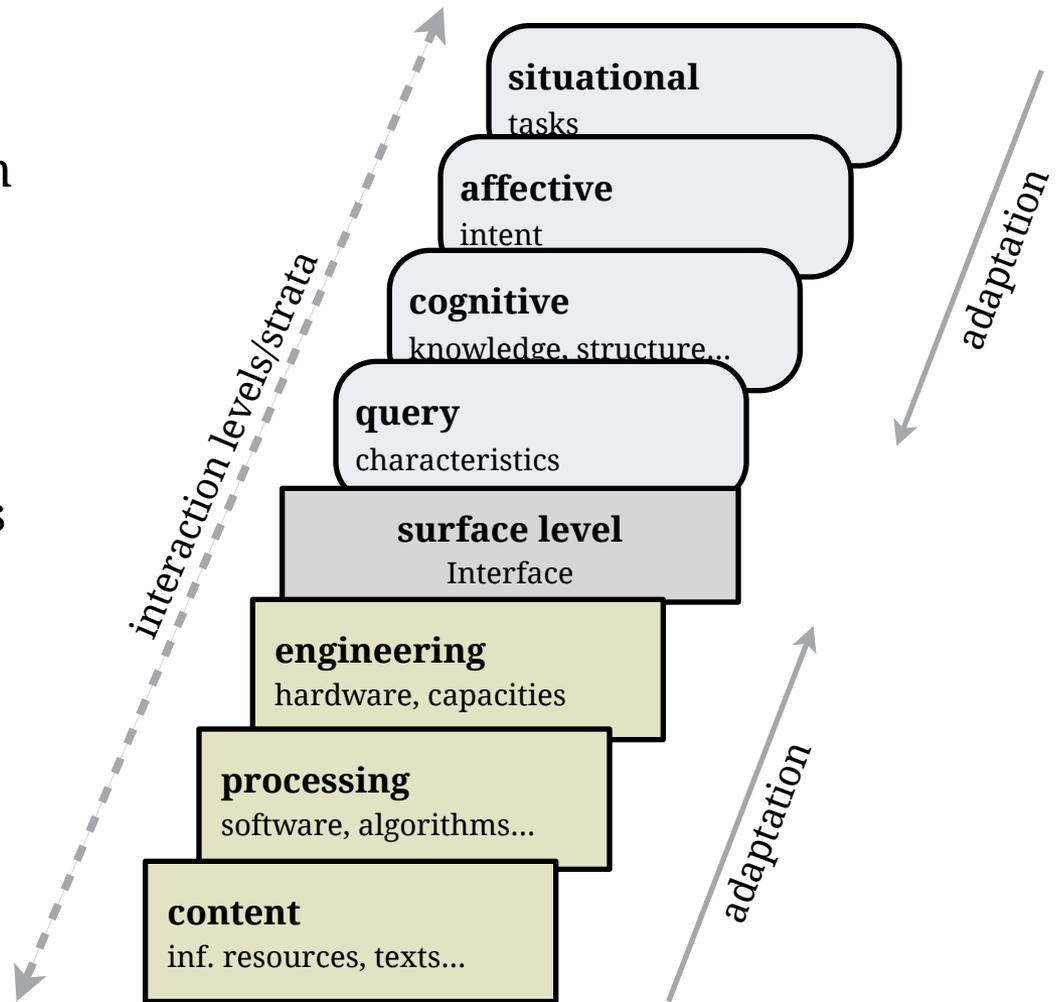
# Ingwersen's cognitive IR model

- Ingwersen analyzes cognitive information retrieval.
- He focuses on the interaction of mental models.
- He shows the short and the long-term change of these models, as well as their drivers.



# Saracevic's stratified model

- Saracevic in his model involves leveled features of the user and the system in order to portray challenges in IR interaction.
- In the model of Saracevic (and Ingwersen's) there is the expectation for the merging and adaptation between the two agents, i.e. the user and the system.



# ... and now metatheories

- All the models and the theories presented adhere to some conceptualization about human learning and cognition.
- The main approaches are summed at:
  - constructivism (cognitive) approach
  - collectivism (socio-cognitive/social constructivism) approach
  - constructionism approach

# nested models revisited

---

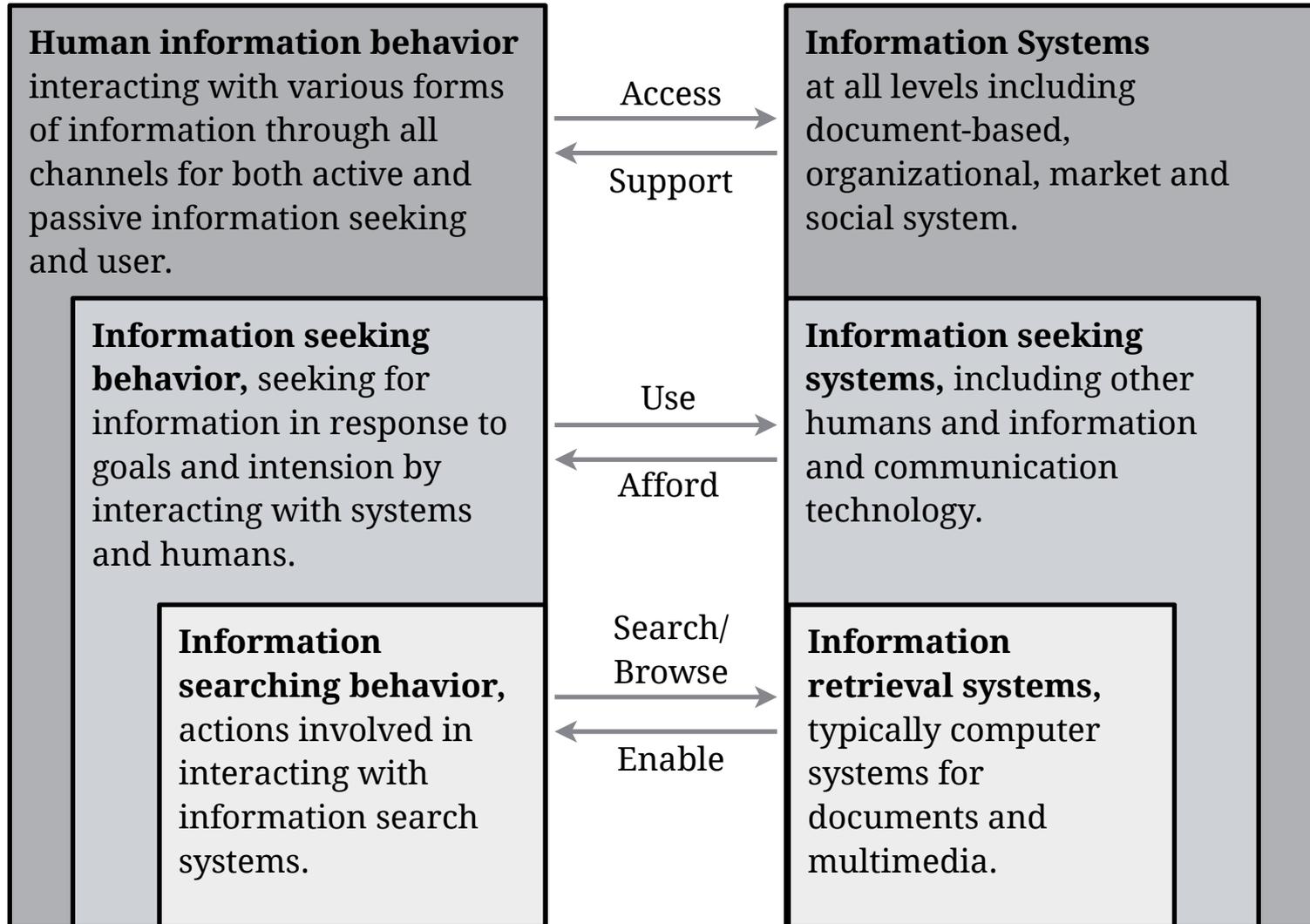
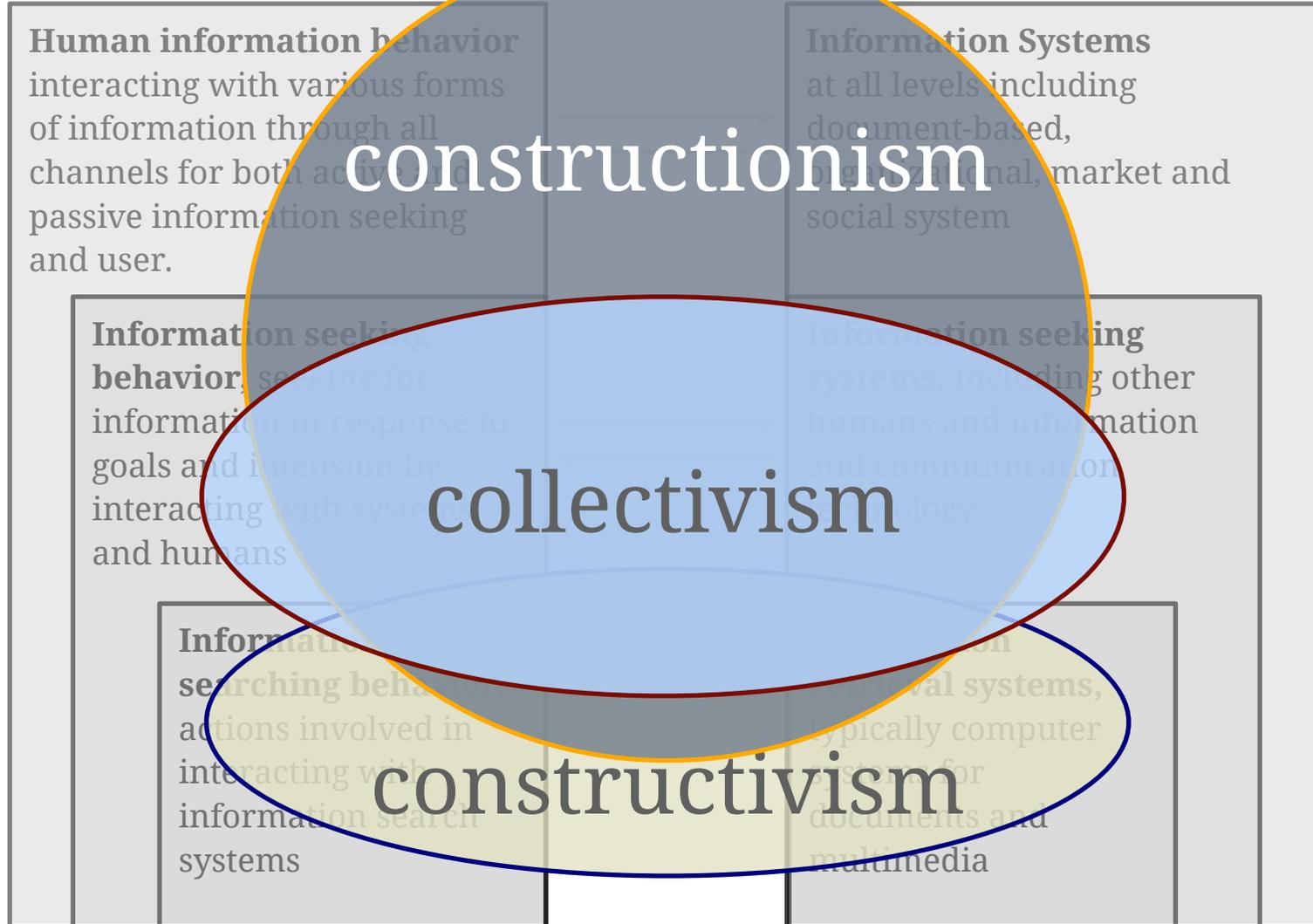


Figure by  
Jansen &  
Rieh.

# a synopsis (again)

---



part b

methodological considerations\*

\*some very few

# methods

---

- interviews
- focus groups
- surveys
- traffic/usage analysis
- logs/keystrokes analysis
- laboratory studies
- expert studies
- comparison studies
- observations
- ethnography/field studies

# approaches

---

- There is an issue between qualitative and quantitative research.
- According to Ford:

## **Quantitative, statistical**

Atomistic (focusing: component parts)

Sequential processing

Isolate variables

Precise analytic defn. + measuring

External observed behavior

Study discrete relationships

Concepts

Logic-mathematical

Convergent processing

Control

Predictability

Objectivity

Mechanistic

Trans-situational, generalizability

## **Qualitative, interpretative**

Holistic (focusing: whole)

Parallel processing

Preserve complexity life-as-situations

“Fuzzy” intuitive concepts

Internal cognition

Study complex interacting relationships

Sensitizing concepts

Intuitive-social

Divergent processing

Relaxation of precise control

Idiosyncrasy

Subjectivity

Constructivist

Context bound transferable “insights”

# tasks

---

- Steps of resolution of an information need.
- Tasks are fed by the work need.
- Tasks are influenced by the motivational or prohibiting powers of the user and her context.
- Tasks have to be based on usage scenarios; usage scenarios can be based on real features of persons and situations.

# further reading

---

<http://www.mendeley.com/groups/988571/user-studies-seminar-reading-list/>

?

here; now

later at [john@lis.upatras.gr](mailto:john@lis.upatras.gr)

Creative Commons License - Attribution 1.0 Generic

Build with Apple Keynote

Photographs from Morguefile.com

Fonts from FontSquirrel.com

