

# Two Central Research Frameworks in Information Retrieval: Drifting outside the Cave of the Laboratory Framework

Peter Ingwersen\* & Kalervo Järvelin\*\*

\* Royal School of LIS, Denmark ([pi@iva.dk](mailto:pi@iva.dk)) – and  
Oslo University College, Norway

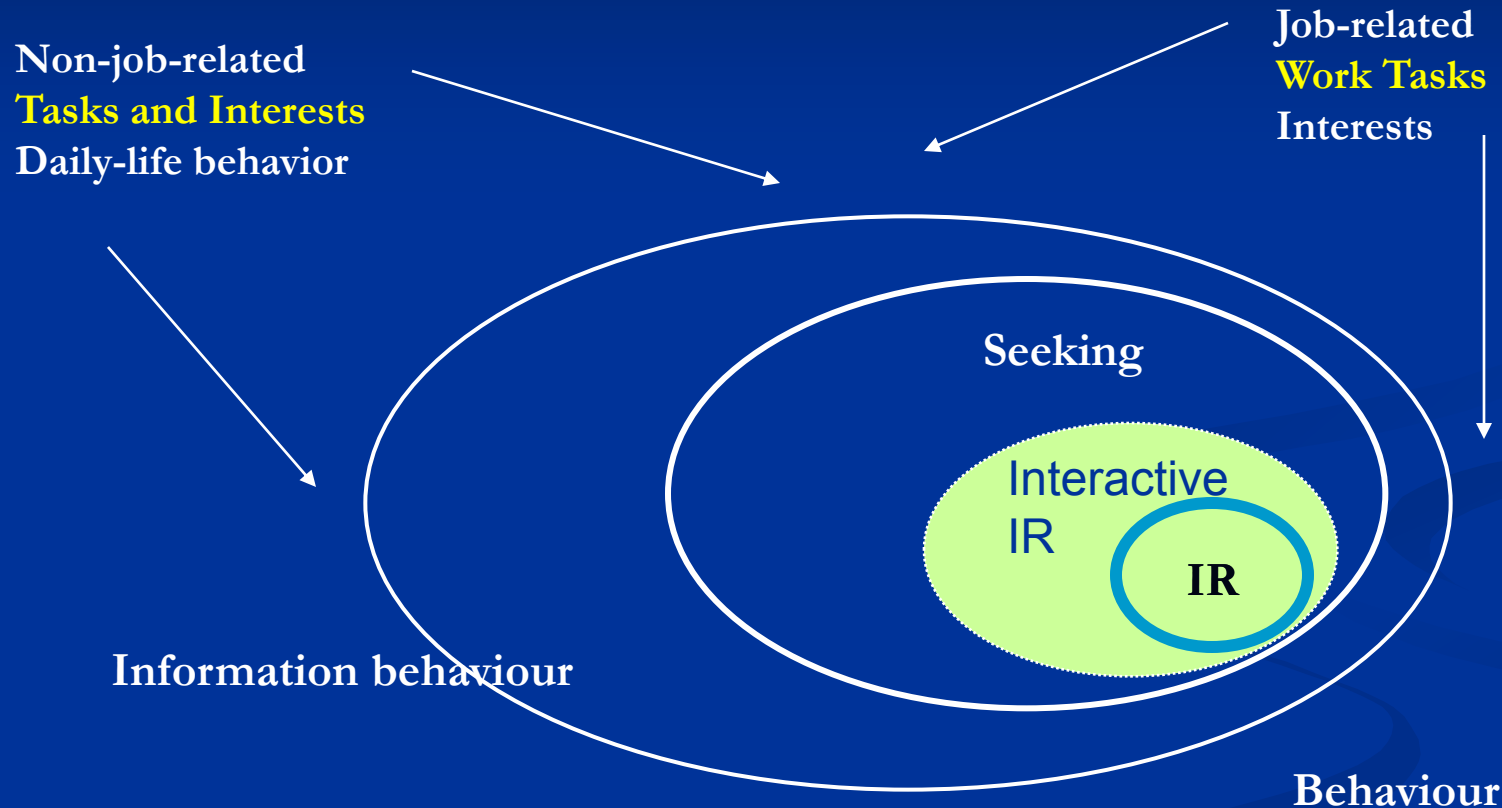
\*\* Tampere University, Finland ([kalervo.jarvelin@uta.fi](mailto:kalervo.jarvelin@uta.fi))

# Agenda

- 1. **Introduction** – Wilson Onion Model
- 2. **The Laboratory Framework for IR.**
  - Provides **system-driven evaluation**
  - The Framework – trapped in the Laboratory Cave
  - Drifting outside the Lab. Cave towards Context:
- 3. **Alternative ISR models leading to:**
  - **User-driven evaluation** – and
- 4. The **comprehensive** cognitive framework for research on Interactive IR
- **Contexts – Relevance – Interaction**

# Information behaviour and IR

T. Wilson's Onion Model, 1999 - extended:



# Information behaviour ... and other central concepts in Information Studies

## ■ *Information behaviour.*

- to create information – e.g., on the Net - blogs; human indexing, including social tags;
- to produce publications – e.g., as publisher
- to communicate – face-to-face; chat; e-mail
- to manage information sources – e.g. KM; selectivity

## ■ *Information seeking (behaviour)*

- Information behaviour with **desire** for Information
- **Information need** exist – even muddled
- Searching information sources – e.g. colleagues

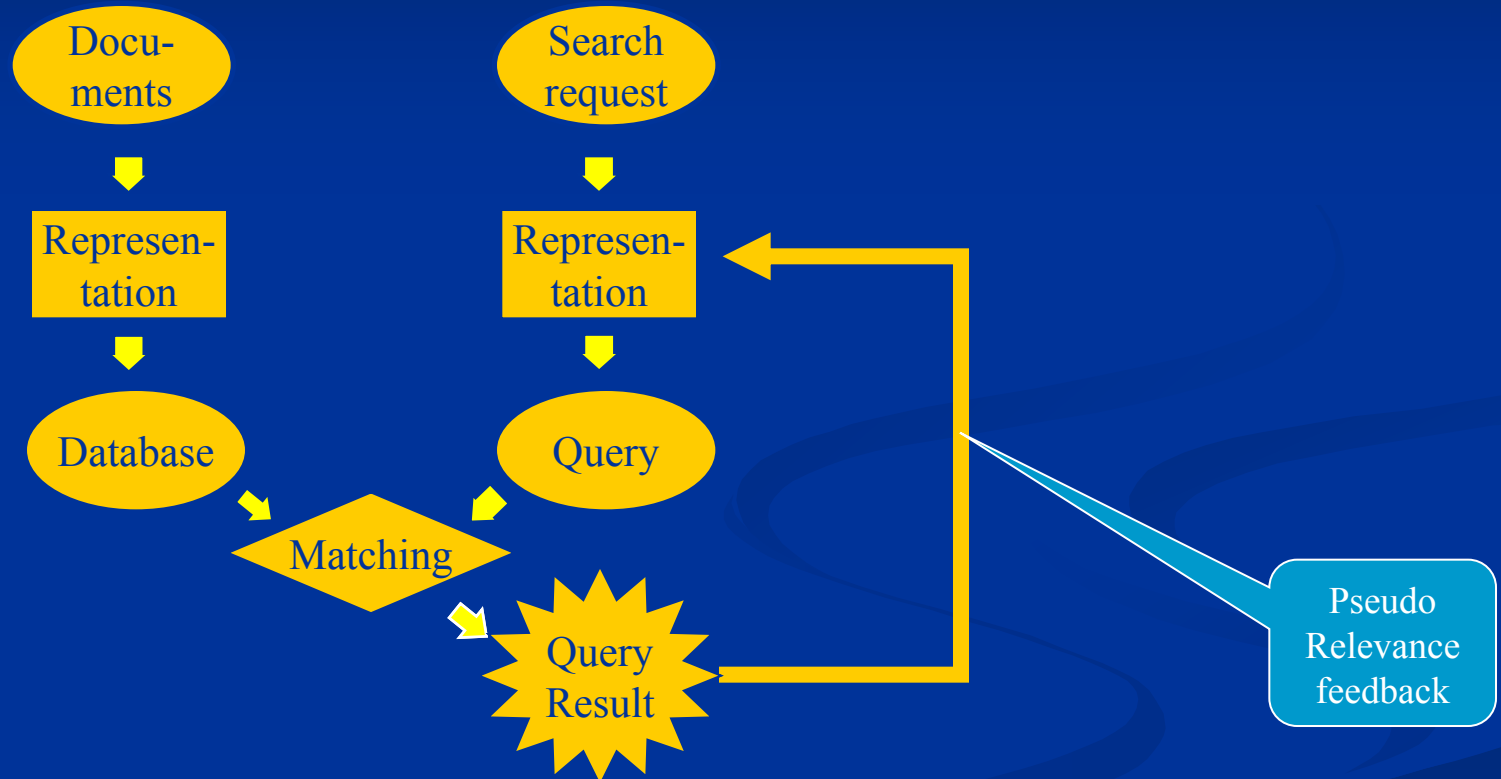
## ■ *Information Retrieval (I)IR*

- Searching information space via systems – **Digital Library & Assets** (interactive IR)
- Retrieval models; relevance feedback & ranking; query modification; auto indexing and weighting;

# The Laboratory Framework for IR

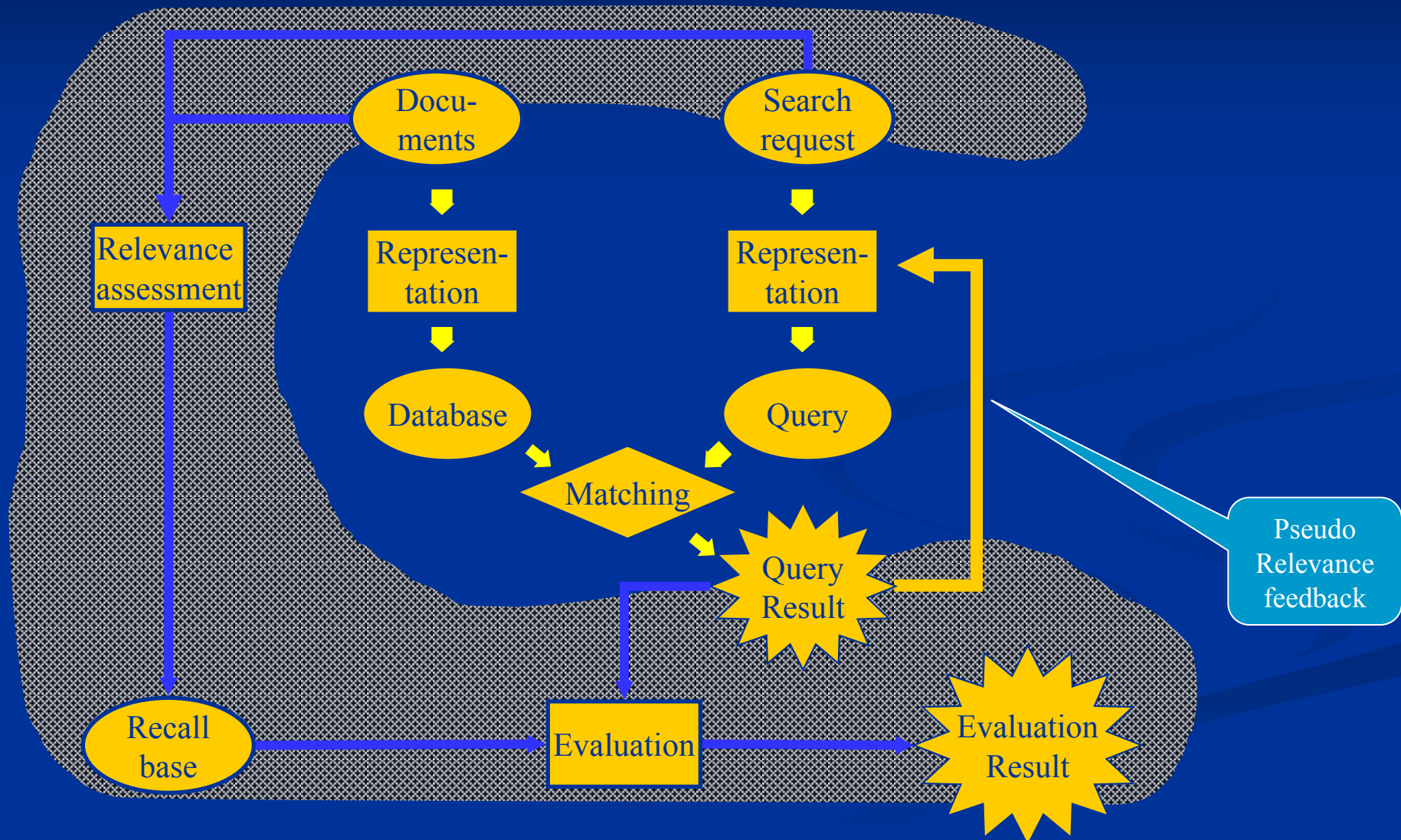
- One simplistic and robust framework
- Searchers not present
- Many competing retrieval models under one framework
- Few and well-defined variables
- Almost full control of experiments

# The Laboratory Approach to IR

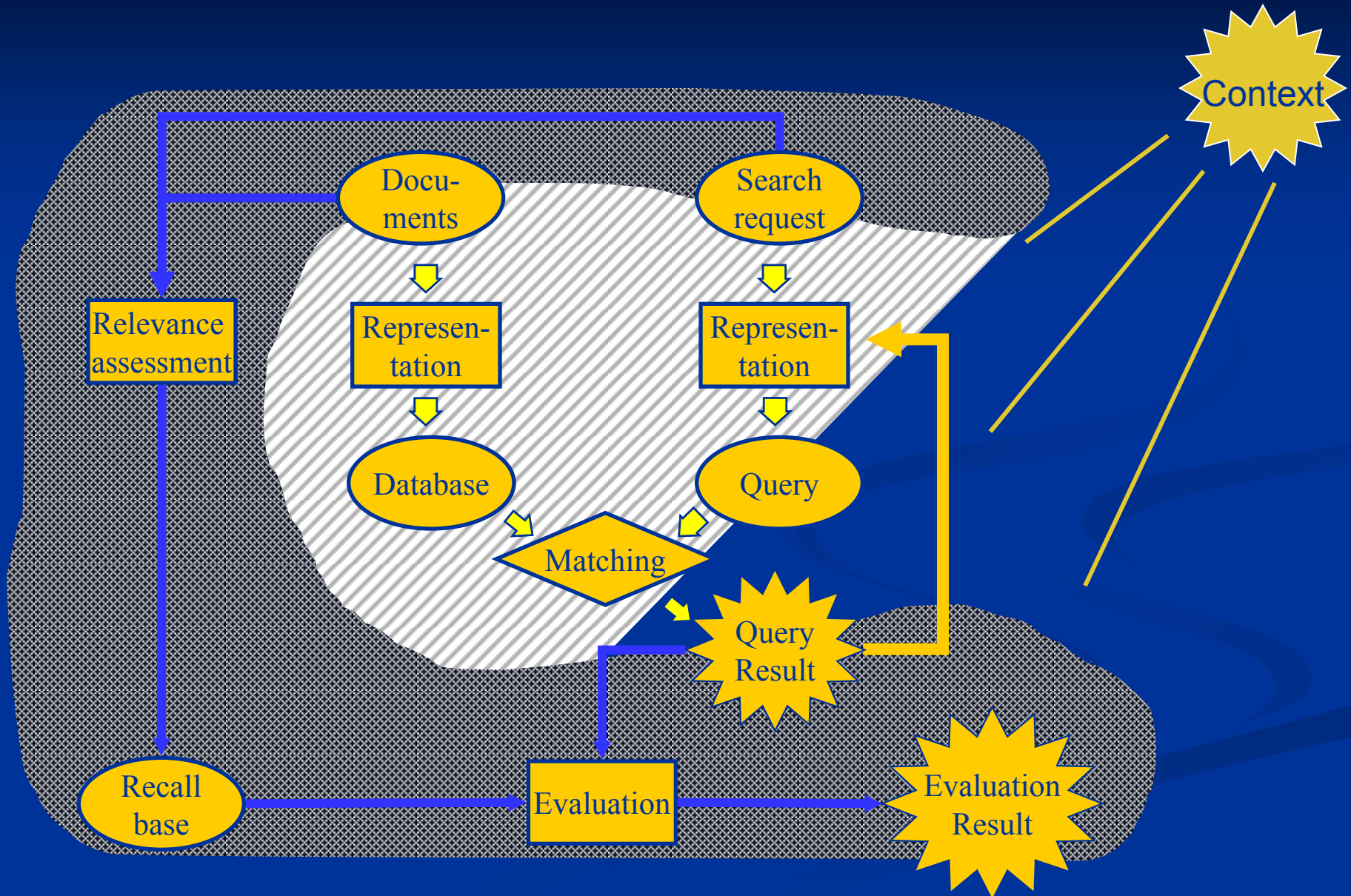


This is information retrieval, isn't it? But where is the lab?

# The Lab Included into a Framework

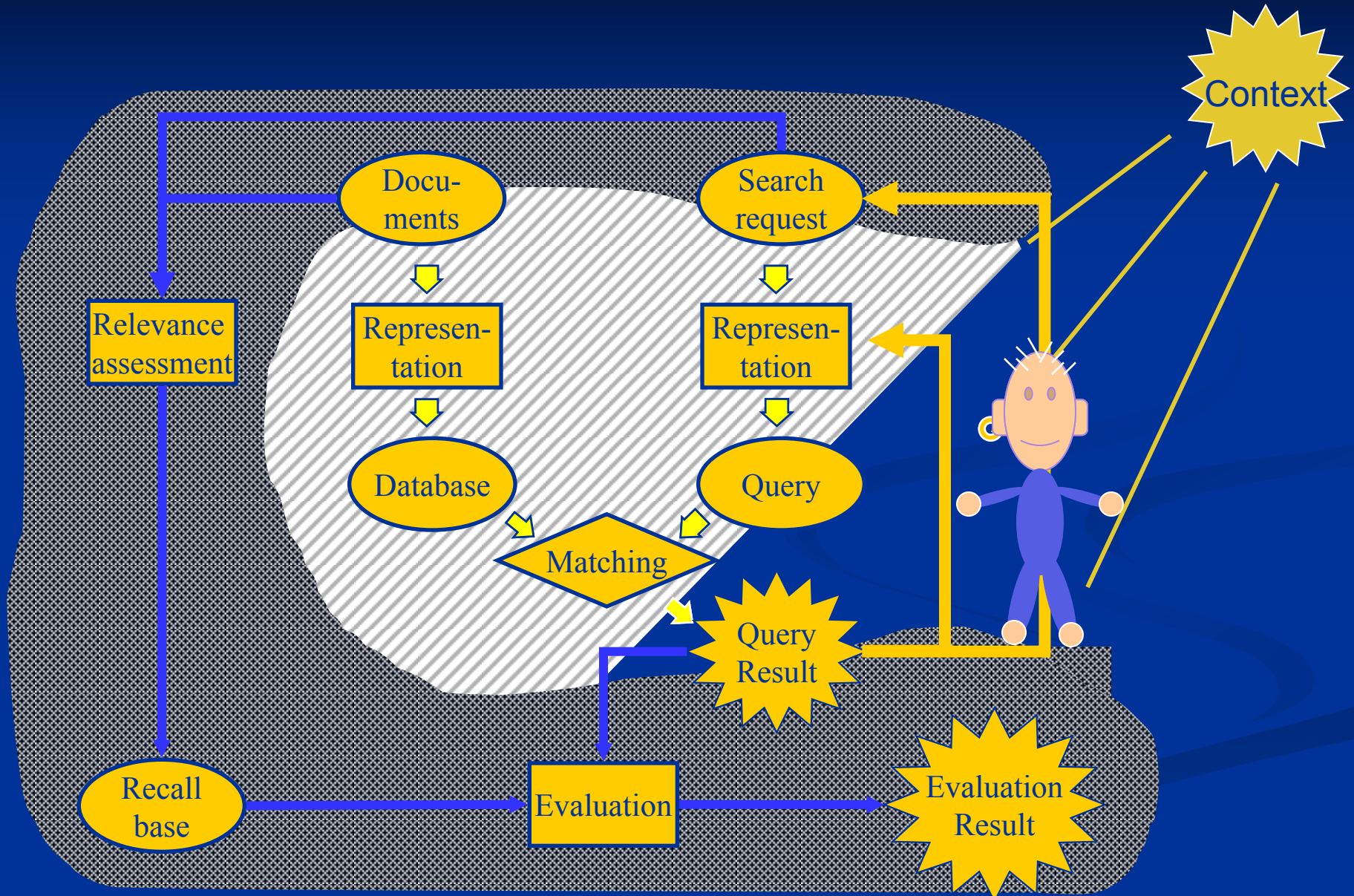


# The Lab IR Cave in Context





# The Lab IR Cave, with a Visitor



# LabIR: The Framework

- Searchers, “users”, lost
  - Have no interesting explicable attributes (**all-alike**)
  - **But nevertheless hiding** in the relevance assessments:
- Relevance assessments are rarely seen as problematic
  - Only related to the requests and documents
  - **The independence assumption**
  - Variations neutralized statistically
- Interaction:
  - Excluded: interface, searchers, search/seeking process
  - Regarded as a sequence of simple **independent** topical interactions; no saturation
  - **Only 1-2 runs allowed** (at least with Rel. Feedback in probabilistic model: **... user-driven?!**)
- Motivations:
  - Framework for the (algorithmic) IR phenomenon and IR system evaluation to support system design.

# LabIR: The Characteristics – 2

## ■ Documents & Rep

- Unstructured natural language news items - ‘just stuff’
- *independent indexing features*

## ■ Requests & Queries

- Unstructured natural language word bags; one, verbose & static *i*-need version
- ‘just stuff’

## ■ Matching and Results

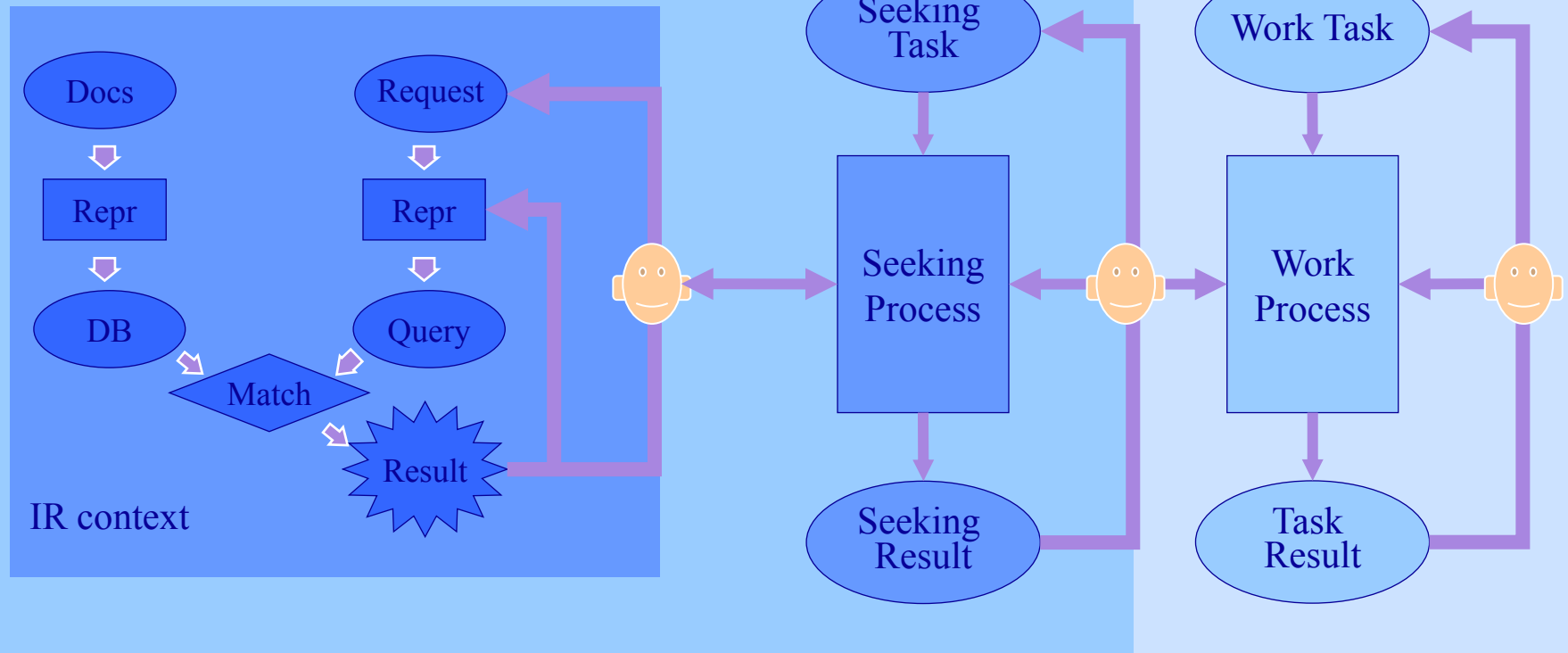
- Matching based on document and requests representations as guided by a retrieval model
- Results typically ranked lists of document reps; list items have rank, score and **binary relevance** (posteriori assessments)

# Nested Framework ... drifting into Contexts outside the Cave

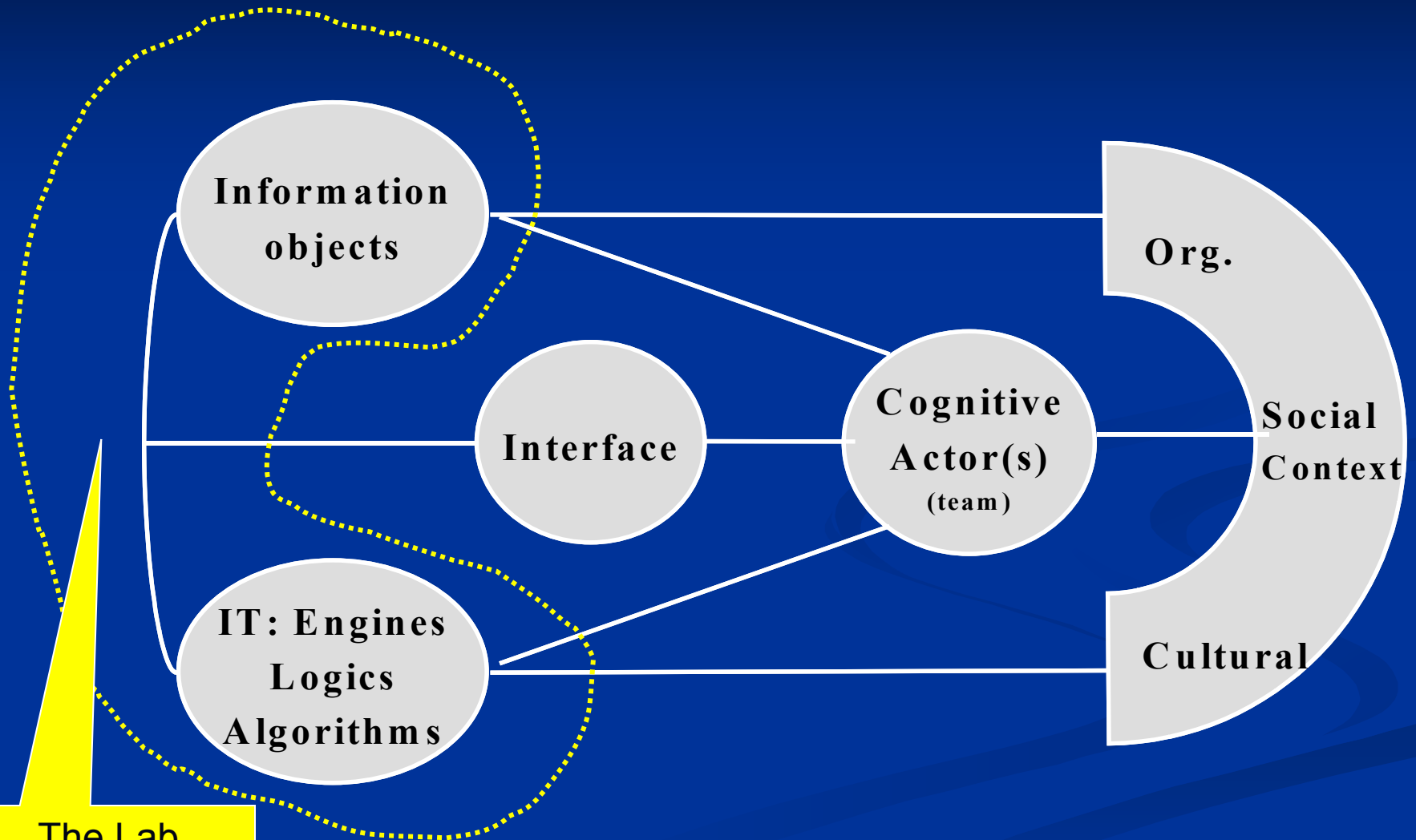
Socio-organizational & cultural context

Work task context

Seeking context



# The Integrated Cognitive Research Framework for IS&R– its basic model



The Lab.  
Framework

# Basic IR research approaches - 2

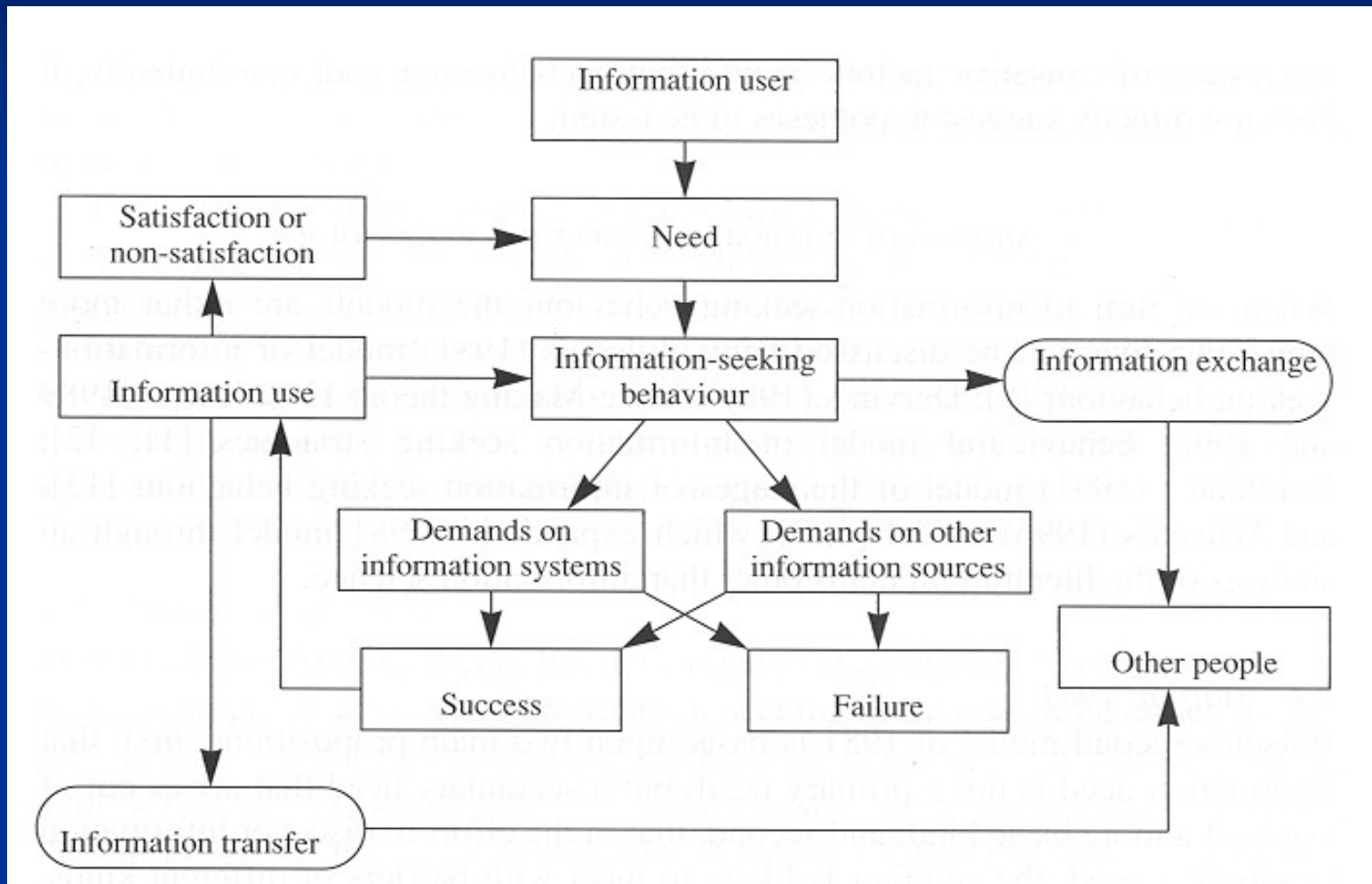
## ■ User oriented approach - 1970s...

- In operational settings (Boolean(like) systems)
- Scientific/technical information as object
- With users and often real information needs
- Information needs: variable over session time
- **Work and search tasks (reasons) not considered**
- Relevance assessments: by the users themselves
- Intermediary-end user interaction & behavior
- **Org., social or cultural context rarely involved**
- Measures: Recall & Precision; Satisfaction

# Information seeking studies in relation to user oriented and cognitive IR

- Commonly highly communication oriented
- Work tasks and system features rarely included
- **IS Theory foundation:** T. Wilson (1981); from 1986: Dervin & Nilan – but also Talja & Savolainen (2000)
- **There are exceptions who moved into cognitive IS&R** (Tom Wilson; Kal. Järvelin; Pertti Vakkari; Tefko Saracevic; Amanda Spink; Peiling Wang)

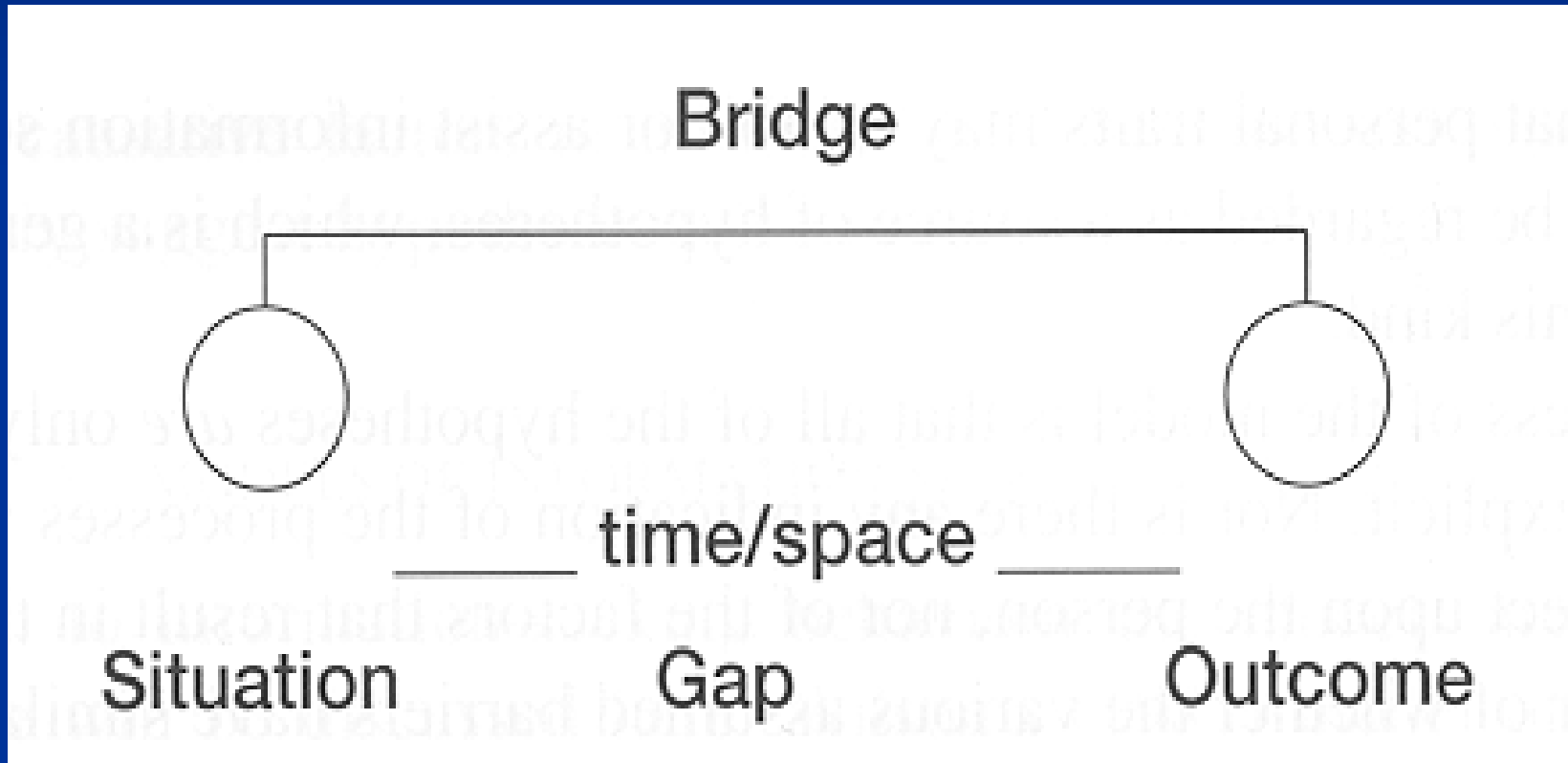
# Wilson's 1981 model of Information seeking





# Dervin & Nilan's sense-making (1986)

— (*The Turn*, p. 60)

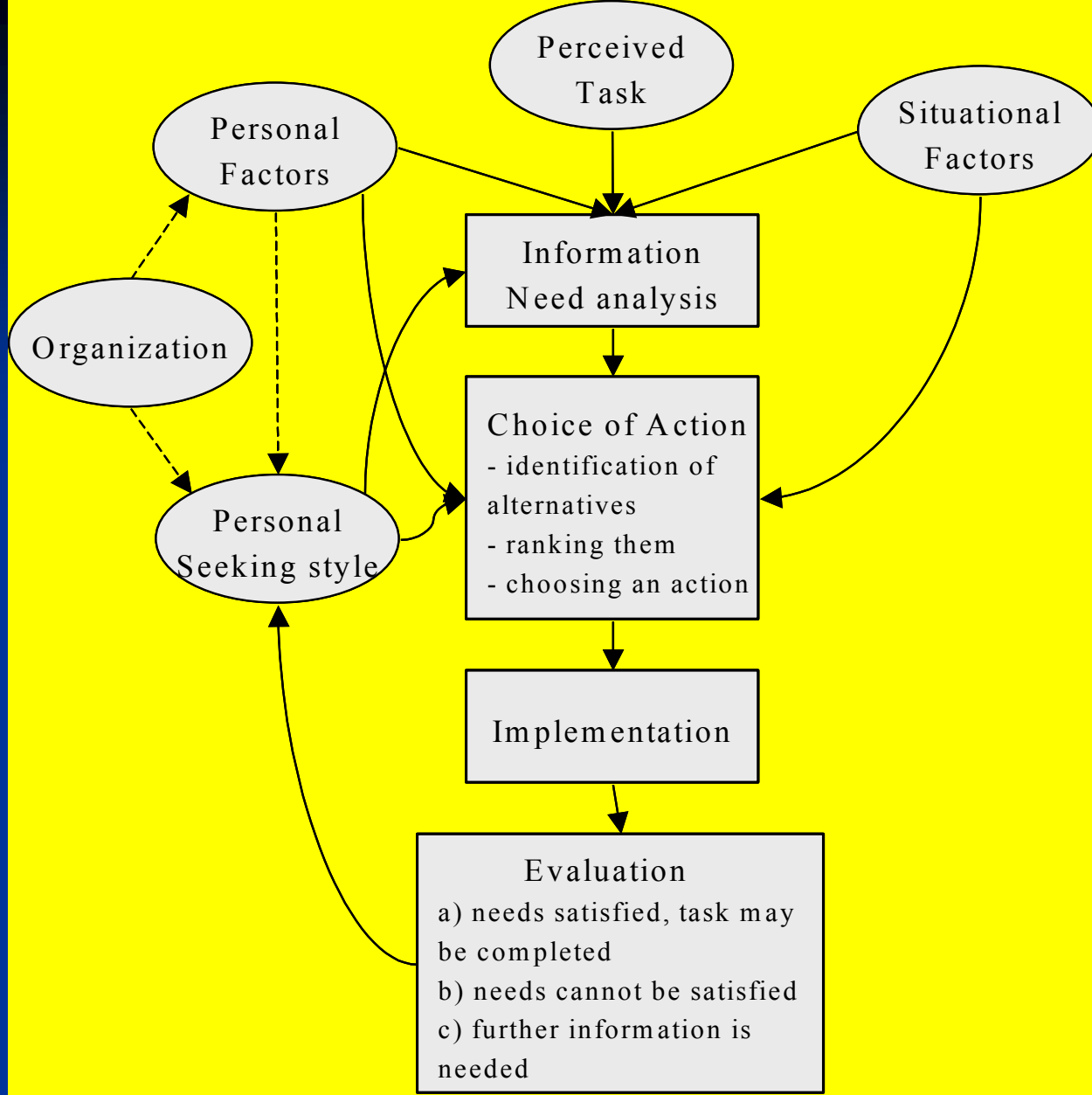


# Carol Kuhlthau's stage model – 1991/94 - (*The Turn*, p. 65)

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

# IS and (I)IR into IS&R

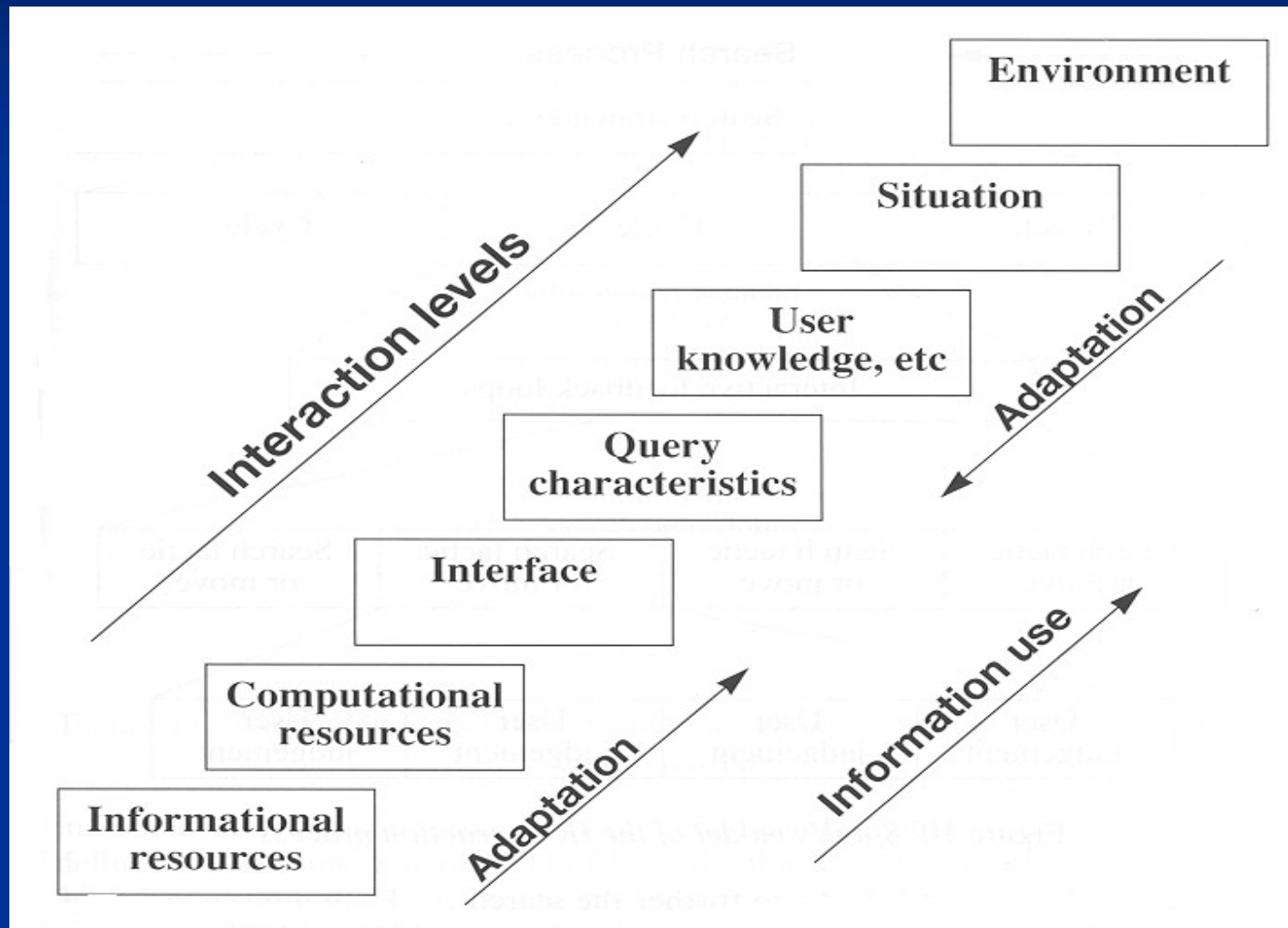
- Byström/Järvelin, 1995 – IS&R model
- Saracevic, 1996 – stratified model
- Ingwersen, 1996 – including contextuality
- Wang & Soergel, 1998 – assessing the retrieved/found document
- Vakkari, 2000 – IS into IS&R – model
- **Models become increasingly comprehensive and generalized to cover IR components too**



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

(From: *The Turn*, p. 69)

# Saracevic' stratified model for IIR (1996)



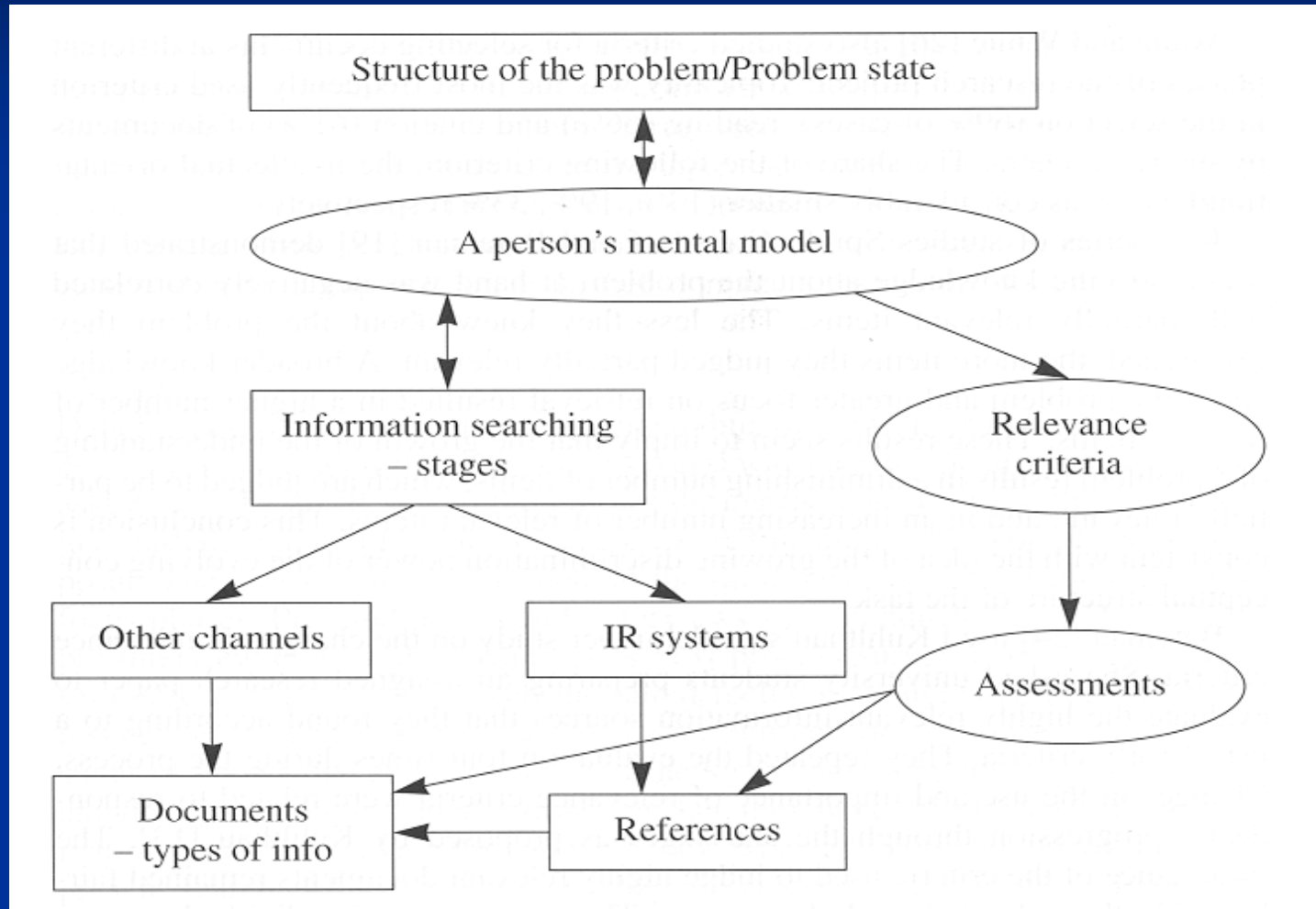
# Wang & Soergel 1998



**DIES: Document Information Elements**  
**Values: Document Values/Worth**

(From: *The Turn*, p. 201)

# IR and relevance in Seeking context – Seeking into IS&R: Vakkari 2000



# Task-based IS&R

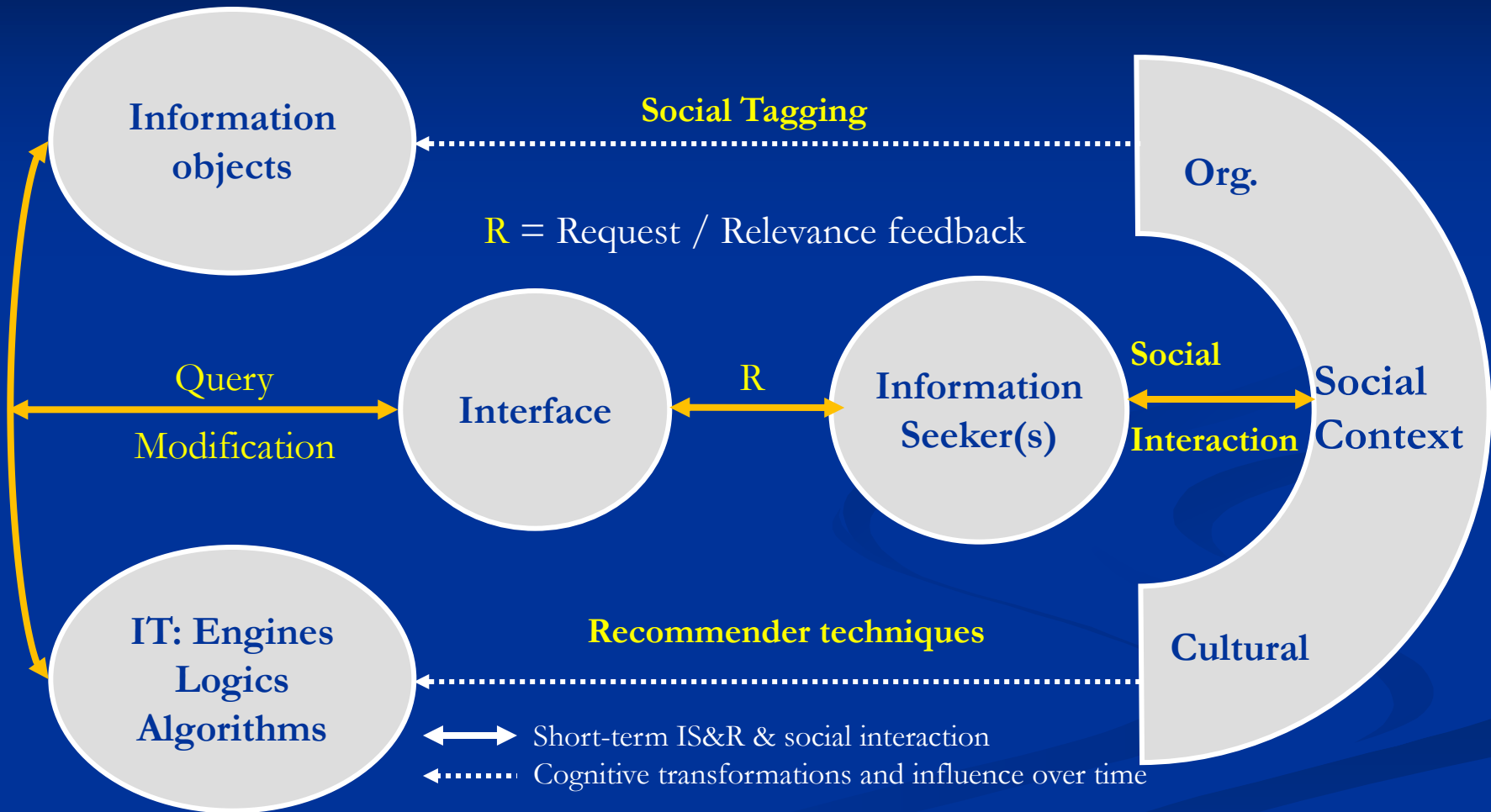
- Originates from Järvelin (1986) Ingwersen (1992) and developed empirically by Byström & Järvelin (1995) and Vakkari (2000; 2001), etc.
- **Task complexity** is one of several characteristics of work/search tasks to be investigated
- Leads to **Work task simulations** (cover stories) in IS&R investigations (Borlund Package, 2000 ...)



Situation in context > Work task > Perception >  
Uncertainty > Information Need

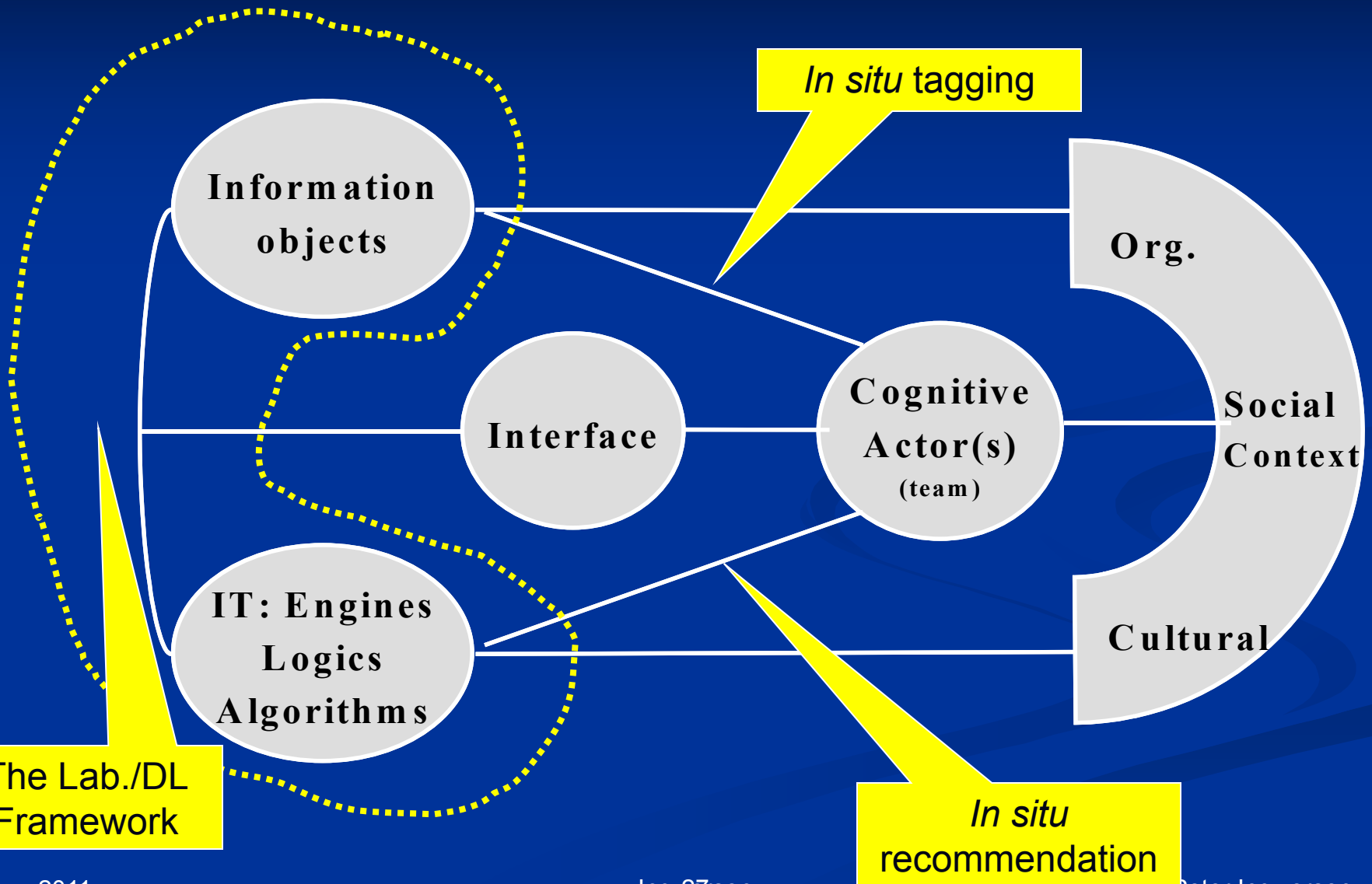
- **The more complex** the situation and work task - the greater the uncertainty and knowledge gap (Byström & Järvelin, 1995);
- The information need becomes increasingly ill-defined – **people** become knowledge sources
- *Recently in Lab. IR:* Situational (task) impact on search **behaviour** – relevance assessments: systems design should support cognition

# Simplistic model of ISR – short-term interaction – in context



# Central Components of Interactive IR

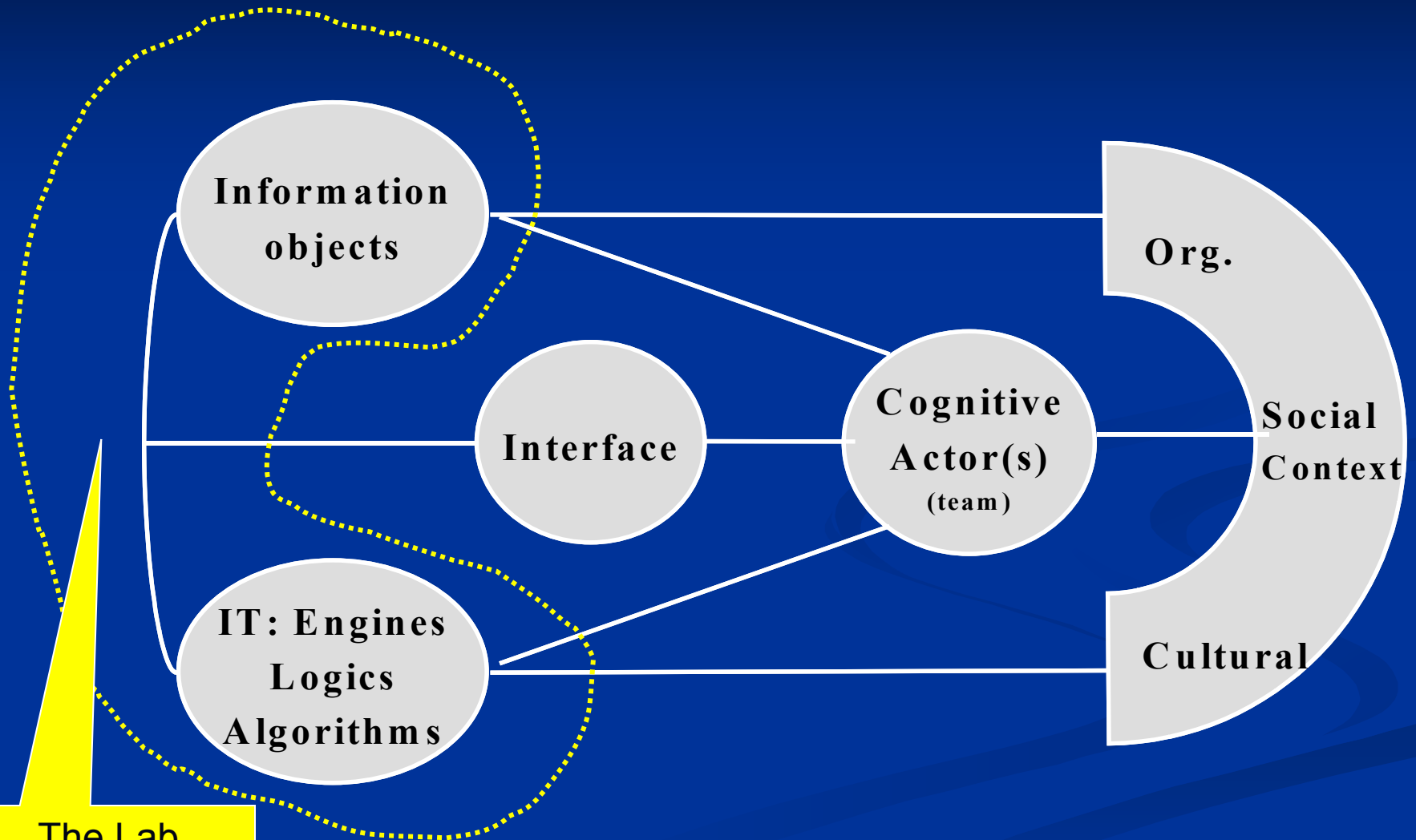
– the basic *Integrated Framework*



# Central differences between the Lab. and integrated cognitive frameworks

- Conception of Information – and hence:
- **Conception of Relevance**
- Task dependency (in Cognitive Framework)
- IR System Setting – **also seen as context to actors**
- **Role of Interaction – the central issue**
- Role of Intermediary – interface issues (not in Lab.)
- **Context characteristics**
- **Evaluation Approaches**
- Integrated perspective of **all** actors, processes and outcomes

# The Integrated Cognitive Research Framework for IS&R– its basic model



The Lab.  
Framework

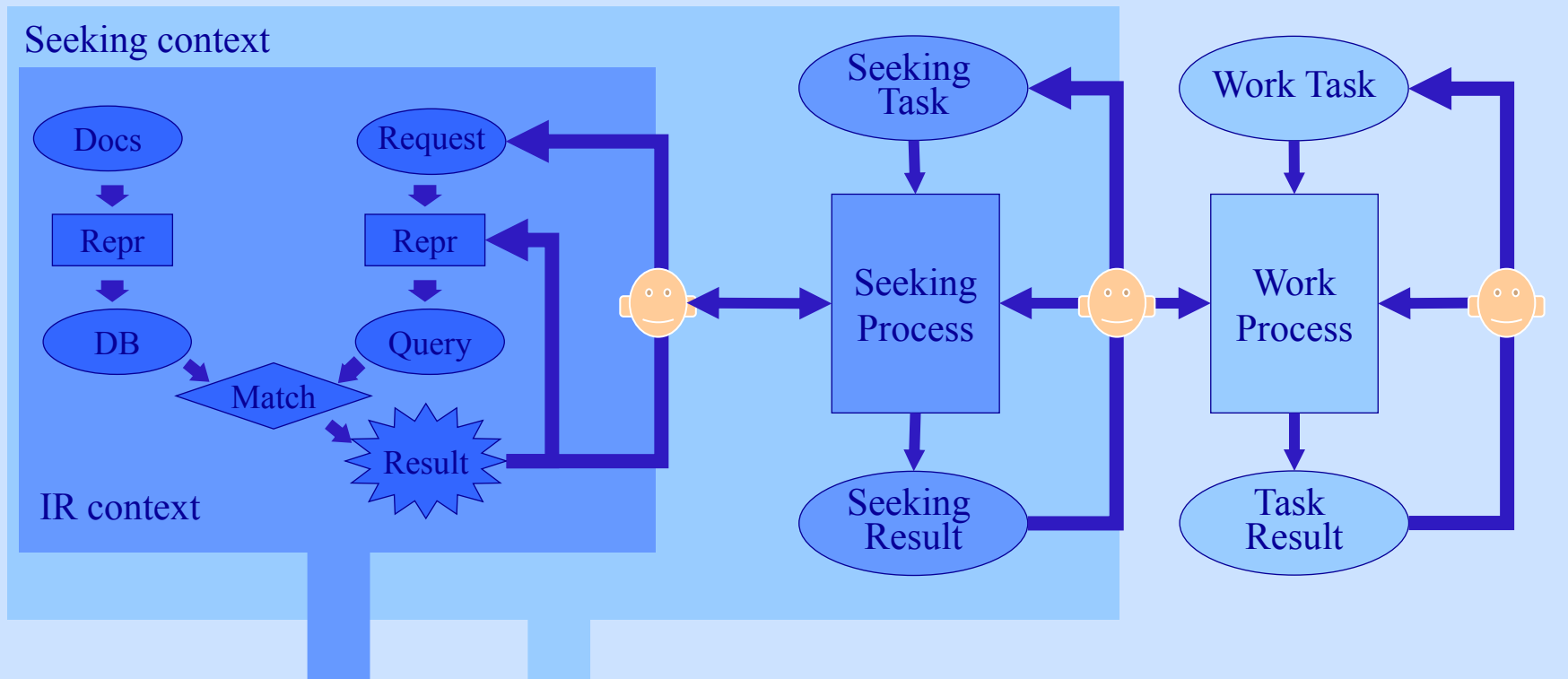
# The applications of the Model & the Cognitive Framework

- Illustrating the **roles of actors** in a variety of cases of information behavior, **like IR interaction**;
- **Pointing to core components** and information **processes** depending on (or influencing) such cases – i.e.,
- *Pointing to kinds of context – next slide*;
- *Pointing out central variables* involved in a variety of research designs – with a number of independent variables

# Cognitive Framework and Relevance Criteria

Socio-organizational & cultural context

Work task context



**Evaluation  
Criteria:**

**A:** Recall, precision, efficiency, quality of information/process

**B:** Usability, quality of information/process

**C:** Quality of info & work process/result

**D:** Socio-cognitive relevance; social utility; quality of work task result

# Relevance and Evaluation

- **Some information more relevant than other**
  - **Relevance changes over time**
  - **Major (horizontal in model) types of relevance:**
    - **Algorithmic** / System relevance (objective)

---

    - **Topical** (aboutness interpretation)
    - **Pertinence** (information need satisfaction – isness – authority of sources – novelty – currency) **Higher Order**
    - **Situational** (usefulness of objects to task/interest: refs.)
    - **Socio-cognitive/social utility** (group interpretation of objects – also over time: citations – recommender systems / collaborative filtering – web inlinks)(evidence exists)
- NB: Emotional (associated with all subjective **higher order relevance** types)**



# The Integrated Cognitive Research Framework suggests

- Applications of research designs
- Comparisons of retrieval (and seeking) in different types of collections
- Comparisons of experts and novices and other actor types by features
- Comparisons of simulated task types (degree of manipulation and semantic openness) – or real tasks – for experimental control
- **Consequences for IR performance**

# The Integrated Cognitive Research Framework informs about ...

- Central variables to combine as **independent ones**
- **Variables to be kept controlled** in a setting
- What kind of **variables that are hidden!**
- **Dependent variables** depend on the research goals

**Novel possible research designs, settings and measures ... there is a lot to do - really!**

THANK YOU!