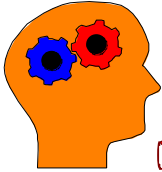


**From medieval philosophy to the Virtual Library: a descriptive framework for scientific knowledge and documentation as basis for document retrieval.**



Frances Morrissey,  
School of Information  
Management & Systems,  
Monash University

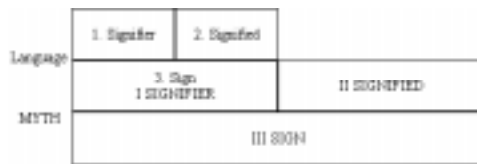
ACKMIDS 2000: 1

Epistemological and ontological foundations

- Q. What is knowledge?
- A. A level of meaning in a higher order semiotic system.

ACKMIDS 2000: 2

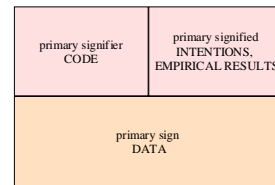
Semiotic systems (Barthes 1957, 1972)



(Barthes 1972: 115)

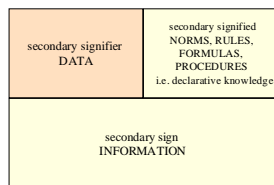
ACKMIDS 2000: 3

Data as a semiotic triad



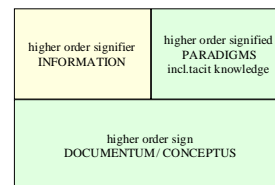
ACKMIDS 2000: 4

Information as a semiotic triad



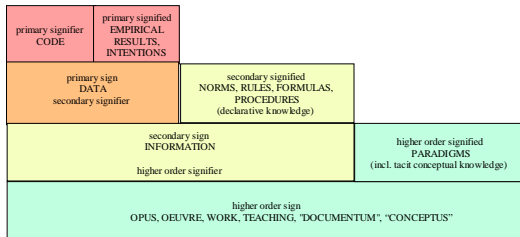
ACKMIDS 2000: 5

The semiotic triad for paradigm and conceptus



ACKMIDS 2000: 6

## Knowledge in higher order semiotic systems



ACKMIDS 2000: 7

## A dyadic construct for scientific documents

A scientific document, as part of Popper's Worlds 1 and 3, is composed of twin entities:

- the physical item (the medium associated with data or information, as first order and second order semiotic systems respectively)
- the metaphysical "conceptus" (the signified ideational content, at a higher order of semiotics, which in science is associated with paradigms or conceptual frameworks)

ACKMIDS 2000: 8

## Document retrieval functionalities

- AT DATA LEVEL
  - specification, characterisation and identification of the item or its data content in allocated or intrinsic metadata
  - lexical or alphanumeric content as item full text

ACKMIDS 2000: 9

### *Document retrieval functionalities (cont.)*

- AT INFORMATION LEVEL
  - item features as listed for data
  - subject representation through use of classification schemes or indexing (whether by literary warrant or user warrant)
  - structural features and context of item (genre)
  - citation indexing

ACKMIDS 2000: 10

### *Document retrieval functionalities (cont.)*

- AT CONCEPTUS OR OEUVRE LEVEL
  - features as listed for item, data and information
  - ideation of conceptus, e.g. as represented in author's or user's set of hypertext relations for the item (whether hierarchical, associative or chaining)
  - aspects of imprimatur or intellectual authority and history of conceptus, including citation trails, authorship, sponsorship, and publishing filters such as refereeing
  - details of intellectual property ownership, transfer, reformatting or reinvention not covered above

ACKMIDS 2000: 11

### *Document retrieval functionalities (cont.)*

- UNDISCOVERED PUBLIC KNOWLEDGE
  - literature-based discovery from syntaxis of item metadata or characteristics of information or of oeuvre, i.e. using combinations of any or all of the above.

ACKMIDS 2000: 12

## Kuhn on science

The three foci of science are:

- factual determinations or increasing the accuracy and scope of factual measurements
- investigations of norms and predictions, such as comparison of facts with predictions from the paradigm theory
- empirical work undertaken to articulate the paradigm theory
  - determination of universal constants
  - formulation of quantitative laws
  - elucidation of models

where the paradigm is the overarching conceptual framework.

*ACKMIDS 2000: 13*

## Ockham on concepts

- Mental propositions do exist, and are composed of understandings or concepts, rather than of things outside the soul.

## Price on concepts

- “Fundamentally a concept is a recognitional capacity, whatever else it may be besides.”

*ACKMIDS 2000: 14*

## Conceptualism

Concepts may be divided into 4 classes: mental representations (ideas serving classificatory purposes); brain states that serve the same function; general words and their usage; and the ability to classify correctly using one or more of these.

(Cambridge)

*ACKMIDS 2000: 15*

## *Price on concepts (cont.)*

Concepts are manifested through...

1. Sign cognition... understanding signs through mental and psychophysical response. This includes primary recognition and secondary recognition (where the latter can be mistaken). Also includes induction involving the operators **not**, **or** and **if**, and statistical probability (i.e. allowing for probabilities and degrees of instantiation).

*ACKMIDS 2000: 16*

*Concepts are also manifested through...*

2. production of quasi-instantiative particulars... such as production and recognition of images (i.e. “image thinking”, both quasi-instantiative and generic), production of replicas, and production of instantiative particulars (instantiation in action in real life).

*ACKMIDS 2000: 17*

*Concepts are also manifested through...*

3. Production of non-instantiative symbols such as words and codes, and symbolic operations. Also the ability to produce alternative verbal formulations (subject to degrees of actualisation).

*ACKMIDS 2000: 18*

## Knowledge organisation systems

Hodge (2000) lists 3 characteristics common to KOS:

- The KOS imposes a particular view of the world on a collection and the items in it
- The same entity can be characterised in different ways depending on the KOS used.
- There must be sufficient degree of between the concept expressed in a KOS and the real-world object or representation it applies to, for others to apprehend and apply the system with reasonable reliability.

ACKMIDS 2000: 19

## *Knowledge organisation systems (cont.)*

Hodge lists 3 main types of KOS:

- Term lists,
- Classifications and categories,
- Relationship lists,

which correspond directly with Ockham's division of concepts into the connotative, the absolute and the relative.

ACKMIDS 2000: 20

## Undiscovered Public Knowledge

The world of recorded objective knowledge contains implicit knowledge that has not to date been stated explicitly;

and contains increasing amounts of undiscovered public knowledge as the combinations and permutations of concepts exceed the number of concepts.

ACKMIDS 2000: 21

## Categories of UDPK

- Hidden refutations or qualifications of hypotheses;
- Inferences from transitive relations (partial syllogisms or inference chaining, found through cocitation analysis);
- Cumulative weak tests ( large sets of documents with weak evidence);
- Unrecognised or hidden analogies;
- Hidden correlations.

(Davies, 1989)

ACKMIDS 2000: 22

## Arrowsmith

Swanson and Smalheisers' method for disclosing UDPK operates in World 3, using syntaxis of concepts.

As a data mining approach, it utilises statistical analysis of title keywords from "complementary literatures", assuming

- the existence of an intermediate concept (which may be conjecture initially)
- that this concept can be deduced from putative co-citation of keywords, corresponding with Davies' UDPK category of "inference from transitive relations",
- and that hypotheses may be generated and/or substantiated through examination of documents retrieved using these keywords.

ACKMIDS 2000: 23

## Hjorland

Information seeking may concern

- the research object itself as phenomenon
- scientific theories and methods concerning the research object,
- scientific disciplines, traditions, and paradigms that deal with the research object,
- formal aspects of the communication process where knowledge about the research object is documented

Research on methodological aspects of information seeking can be broken down into two aspects, both covered today:

- problems concerning the conceptual structuring of the objects of information seeking and
- problems concerning the structure and the properties of the search tools that are available to the user.

ACKMIDS 2000: 24

## Parenthetical commentary re Kuhn and knowledge

Kuhn uses “paradigm” in two senses:

1. paradigm in the sociological sense, as shared beliefs, values, techniques, i.e. “components of knowledge...tacitly embedded in shared examples” in a discourse community. Here the discourse community is both producer and validator of scientific knowledge. “A paradigm governs, in the first instance, not a subject matter but rather a group of practitioners.”

ACKMIDS 2000: 25

*Parenthetical commentary re Kuhn and knowledge (cont.)*

2. paradigm in the epistemological sense, as theories or the “disciplinary matrix” possessed by the group of practitioners, such as  
“symbolic generalizations” or formalised syntax and expressions (language?)
  - “metaphysical paradigms” or beliefs in particular models, including preferences for particular analogies and metaphors
  - values, such as accurate quantitative predictions, problem solving, simplicity, self-consistency and plausibility (judgemental knowledge?)
  - sets of “exemplars” or concrete problem-solutions and formularies (procedural knowledge?)

Kuhn also discusses shared tacit knowledge, such as ways of perceiving and interpreting any experienced sense-data.

ACKMIDS 2000: 26

## Implications for documents as narratives of scientific knowledge

1. Documents may reflect the “bucket” theory of science, where historicity of data collections and intuitive leaps is disguised in publication as a phoney story (Popper).

The phoney story displays intertextuality with writings on current paradigm, and claims some novelty posing as logical or analogical progression (Berkenkotter & Huckin, Popper).

“Scientific facts are formulated in the denial and obliteration of their own historicity” (Latour).

ACKMIDS 2000: 27

## *Implications for documents as narratives of scientific knowledge (cont.)*

2. Documents may reflect the “searchlight” theory of science, where hypotheses direct and illuminate choices of experiment, and conversely, experiments support or falsify hypotheses (Popper).
3. Documentation and ongoing reference to collective memory are an integral part of scientific practice because of the threefold requirements for empiricism, performativity and articulation of paradigm. Latour termed ongoing documentation “literary inscription”, and regarded it as a central activity in science, at all levels from laboratory practice to publication of charts and figures.

ACKMIDS 2000: 28

## Acknowledgements

While undertaking the research covered in this paper, the author has been in receipt of a scholarship funded by the School of Information Management & Systems, Monash University; and by a SMURF grant to Dr Frada Burstein, SIMS (Knowledge Management Program for Information Communities).

ACKMIDS 2000: 29