The illusion of objectivity
The invention of the modern catalog

Jens-Erik Mai
Faculty of Information
University of Toronto
Canada

International Symposium on the Future of Information Organization Research
National Taiwan University, October 4-5, 2010
Me - and where I come from...
Toronto: One of the most diverse cities in the world
Prelude
Basic, long-standing assumptions:

“value judgments have no place in the creation of bibliographic files” (Hagler, 1997, p. 43).

“no politics, no religion, and no morals” (Wilson, 1983, p. 190).
We want our catalogs to be:
We want our catalogs to be:

Neutral
We want our catalogs to be:

Neutral, objective
We want our catalogs to be:

Neutral, objective, unbiased
We want our catalogs to be:

Neutral, objective, unbiased, true.
This notion of the catalog is based in modernity’s philosophy.
Modernity?

“used in a famously ambiguous way” (Cahoone, 1996, p. 13)

“modes of social life or organisation which emerged in Europe from about the seventeenth century onwards and which subsequently became more or less worldwide in influence” (Giddens, 1990, p. 1)
Modernity:

“used in a famously ambiguous way” (Cahoone, 1996, p. 13).

“modes of social life or organisation which emerged in Europe from about the seventeenth century onwards and which subsequently became more or less worldwide in influence” (Giddens, 1990, p. 1).
For instance:

dualism -- separation of content and reader.

deb-traditionalization -- assuming documents can be represented independent of the activities that produces and uses them.

globalization -- assuming documents can be represented independently of time and space.
For instance:

dualism
For instance:

dualism -- separation of content and reader.
For instance:

dualism -- separation of content and reader.

de-traditionalization
For instance:

dualism -- separation of content and reader.

deh-traditionalization -- assuming documents can be represented independent of the activities that produces and uses them.
For instance:

dualism -- separation of content and reader.

de-traditionalization -- assuming documents can be represented independent of the activities that produces and uses them.

globalization
For instance:

**dualism** -- separation of content and reader.

**de-traditionalization** -- assuming documents can be represented independent of the activities that produces and uses them.

**globalization** -- assuming documents can be represented independently of time and space.
For instance:

dualism -- separation of content and reader.

de-traditionalization -- assuming documents can be represented independent of the activities that produces and uses them.

globalization -- assuming documents can be represented independently of time and space.
My aim for today:

Show that contemporary information organization research is based on modernistic assumptions.

Argue that we need to establish a late-modern alternative for information organization research and practice.
Agenda

- The problem of classification
- Three problems with modern classification:
  1. Classification vs. Categorization.
  2. Documents as Containers.
  3. Objectivity vs. Plurality.
- What to give up.
- What to do next.
The problem of classification

Three problems with modern classification:

1. Classification vs. Categorization.
2. Documents as Containers.
3. Objectivity vs. Plurality.

What to give up.

What to do next.
The problem of classification

Three problems with modern classification:
  1. Classification vs. Categorization.
  2. Documents as Containers.
  3. Objectivity vs. Plurality.

What to give up.

What to do next.
The problem of classification

Three problems with modern classification:

1. Classification vs. Categorization.
2. Documents as Containers.
3. Objectivity vs. Plurality.

What to give up.

What to do next.
The problem of classification

Three problems with modern classification:

1. Classification vs. Categorization.
2. Documents as Containers.
3. Objectivity vs. Plurality.

What to give up.

What to do next.

Agenda
The problem of classification

Three problems with modern classification:

1. Classification vs. Categorization.
2. Documents as Containers.
3. Objectivity vs. Plurality.

What to give up.

What to do next.
The problem of classification
• “the ‘putting together of like things’ is . . . the fullest and most exact” definition of classification” (Richardson 1935, 1)

• ”Classification…means putting together things or ideas that are alike” (Vickery 1975, 1)

• “classification is the act of bringing like things together” (Buchanan 1979, 9)

• “the purpose of classification is to bring related items together” (Wynar 1992, 317)

• “to group all works of a kind together” (Chan 1994, 259)

• “Classification brings like things together” (Svenonius 2000, 10)
What’s the problem?
A lawn mower and a plum, for instance, are alike in the sense that “they both weigh less than 10,000 kg, both did not exist 10 million years ago, both cannot hear, both can be dropped, both take up space and so on”
A lawnmover and a plum, for instance, are alike in the sense that “they both weigh less than 10,000 kg, both did not exist 10 million years ago, both cannot hear, both can be dropped, both take up space and so on”
A lawnmover and a plum, for instance, are alike in the sense that “they both weigh less than 10,000 kg, both did not exist 10 million years ago, both cannot hear, both can be dropped, both take up space and so on”

“similarity is too flexible to explain categorization” (Bryant, 2000, p. 57).
“likeness contributes little to an explanation of classification” (Broadfield 1946, p. 3).

“likeness is not a quality of things. It is a relation between them, not a ‘characteristic’ of things” (Broadfield, 1946, p. 2).
Two views of classification.
• One view of classification:

“...a system for organizing a universe of items be they objects, concepts, or records”

Dewey Decimal Classification
“...somewhere, somehow, we can, or should try to, produce the one best classification system that will serve all purposes”

Miksa, 1998, p. 81
But...
<table>
<thead>
<tr>
<th>English</th>
<th>German</th>
<th>Danish</th>
<th>French</th>
<th>Italian</th>
<th>Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tree</td>
<td>Baum</td>
<td>Træ</td>
<td>Arbre</td>
<td>Albero</td>
<td>Árbol</td>
</tr>
<tr>
<td>Wood</td>
<td>Holz</td>
<td>Skov</td>
<td>Bois</td>
<td>Legno</td>
<td>Leña</td>
</tr>
<tr>
<td>Woods</td>
<td>Wald</td>
<td></td>
<td></td>
<td>Bosco</td>
<td>Madera</td>
</tr>
<tr>
<td>Forest</td>
<td></td>
<td></td>
<td>Forêt</td>
<td>Foresta</td>
<td>Selva</td>
</tr>
</tbody>
</table>

Cultural relativity in word meanings
by Hjemslev, 1943, extended by Buckley, 2001
• Another view of classification:
  - We view the world from a specific point.
  - Our/my/your view of the world is unique.
Agenda

• The problem of classification

• Three problems with modern classification:
  1. Classification vs. Categorization.
  2. Documents as Containers.
  3. Objectivity vs. Plurality.

• What to give up.

• What to do next.
Problem #1

Classification vs. categorization
• *Classification* is the “orderly”, “lawful” and “systematic” classification of entities in “artificial and arbitrary” schemes (Jacob, 2004, p. 522).

• *Categorization* is “flexible and creative”, “draws nonbinding associations between entities” and “divides the world of experience” based on entities’ “immediate similarity within a given context” (Jacob, 2004, p. 527 + 528).
The distinction between “the cognitive processes of categorization and the formal process of classification … is an important distinction” (Jacob, 2010, p. 113-114).

because it will enable the establishment of a “science of classification” (Jacob, 2010, p. 11).
Here’s what the distinction does:
order, laws, systematic

cognition and language
Things as they are vs. thinking about things

Ontology <-> Epistemology
My sense is that we can’t separate the two this way.

Things and our thinking about the things are closely connected.
An example -
My neighbor and I are talking...

- Me: “I like the new *shrub* in your garden”

- My neighbor: “Oh, you mean the new *tree* I planted yesterday?”
Technically, in botanical terms:

- **shrub**: a woody plant, which have multiple stems and lower height, usually less than 5-6 m.

- **tree**: a woody plant, which has many secondary branches supported clear of the ground on a single main stem or trunk, typically higher than 3-6 m.
a shrub or a tree?
My neighbor and I drew in our conversation, in fact, on an established classification of woody plants.
• Technical terms for shrubs and trees are established lawfully, orderly and systematically in the botanical domain. [Classification.]

• However, we use the terms flexibly and creatively in nonbinding ways in a given context. [Categorization.]

• Are the terms ‘shrub’ and ‘tree’ a class or a category in our conversation?
The question doesn’t make sense.
Here is why:
Meaning is not based the cognitive and personal experience in the immediate context.

But is established and mediated through culture and the community’s “already there language.”
order, laws, systematics

cognition and language
Problem #1:

Classification and categorization is closely interrelated and interdependent.
Problem #2
Documents as containers
Documents are containers of information.

Information organization is about organizing the content of documents.
• information things (Buckland, 1991)
• “information bearing messages” (Svenonius, 2000, p. 8)
• “concepts are extracted from documents” (ISO, 1985, p. 2)

⇒ Information in documents exist “independently of whether it is perceived by any human being” (Furner, 2007, p. 147)
“Largely unacknowledged influence of the **realist view** on the activity of designers and users of knowledge organization systems.”

“It is difficult to find well-reasoned defenses of the realist view in the literature, yet most of us ... continue to act as if we accept the realist view as the correct one.”

Furner, 2010, p. 186-187
“... a well-established tradition of library- and information-science theory—not only in regard to bibliographical discussions of the work in cataloging, but also in regard to cognitive agency in information retrieval and information behavior—that understands ideas as being quasiempirical objects—generated in the minds of authors—that are contained in documents and that are sought by and transferred to the minds of information seekers or users upon reading, viewing, or listening” (Day, 2008, p. 1644).
“... the ancient definition of information as a process (e.g. the act of informing or being informed) was gradually translated by the rationality of the Enlightenment into the current popular definition of information as an entity (i.e. information as ‘noun’ rather than ‘verb’). The reification of information as a resource, which stands at the heart of the information society idea, is not therefore recent in origin, but parallels the development of modernity” (Black, 1998, p. 41).
“... the ancient definition of information as a process (e.g. the act of informing or being informed) was gradually translated by the rationality of the Enlightenment into the current popular definition of information as an entity (i.e. information as ‘noun’ rather than ‘verb’). The reification of information as a resource, which stands at the heart of the information society idea, is not therefore recent in origin, but parallels the development of modernity” (Black, 1998, p. 41).
“... information is the flow and exchange of a message, originating from one speaker, mind, or source and received by another” (Day, 2001, p. 38).
K [S] + ΔI = K [S + ΔS]

Brooks, 1980, p. 131
This idea that information is an external resource that is borne in messages and flows from one knowledge structure to another has dominated “current thinking of intellectual access to knowledge stores” and is based on “cognitive science theories” (Neill, 1992, p. 19).
This idea that information is an external resource that is borne in messages and flows from one knowledge structure to another has dominated “current thinking of intellectual access to knowledge stores” and is based on “cognitive science theories” (Neill, 1992, p. 19).
“... the work doesn’t ‘contain’ or reproduce ideas, but instead, these are generated by the work in relation to the world” (Day, 2008, p. 1650)
“... if meaning is embedded in the text, the reader’s responsibilities are limited to the job of getting it out; but if meaning develops, and if it develops in a dynamic relationship with the reader’s expectations, projections, conclusions, judgments, and assumptions, these activities (the things the reader does) are not merely instrumental, or mechanical, but essential, and the act of description must both begin and end with them” (Fish, 1980, p. 2-3)
Problem #2:

Catalogers’ and indexers’ challenge is not to extract documents’ content, but to note the moves the document makes in particular conversations, perspectives, and debates.
Problem #3
Objectivity vs. plurality
“... if we admit that the number of different perspectives from which the world can be viewed and described is endless, we shall expect that the library to contain competing, conflicting accounts of the world that cannot be incorporated into a single consistent story of the way things are” (Wilson, 1983, p. 165).
“... if we admit that the number of different perspectives from which the world can be viewed and described is endless, we shall expect that the library to contain competing, conflicting accounts of the world that cannot be incorporated into a single consistent story of the way things are” (Wilson, 1983, p. 165).
"... if we admit that the number of different perspectives from which the world can be viewed and described is endless, we shall expect that the library to contain competing, conflicting accounts of the world that cannot be incorporated into a single consistent story of the way things are" (Wilson, 1983, p. 165).
“... if we admit that the number of different perspectives from which the world can be viewed and described is endless, we shall expect that the library to contain competing, conflicting accounts of the world that cannot be incorporated into a single consistent story of the way things are” (Wilson, 1983, p. 165).
“... if we admit that the number of different perspectives from which the world can be viewed and described is endless, we shall expect that the library to contain competing, conflicting accounts of the world that cannot be incorporated into a single consistent story of the way things are” (Wilson, 1983, p. 165).
“... if we admit that the number of different perspectives from which the world can be viewed and described is endless, we shall expect that the library to contain competing, conflicting accounts of the world that cannot be incorporated into a single consistent story of the way things are” (Wilson, 1983, p. 165).
Conceptual obstacles
“... our analytic ability to decontextualize is admittedly a great intellectual achievement ... it clearly also entails some ‘context blindness’” (Zerubavel, 1991, p. 116).
‘Context blindness’ allows the separation of categorization and classification, of content and reader -- and allows for the assumptions of neutrality and objectivity.
We must give up, “our conventional rigid conception of classification by recognizing that any entity can be situated in more than one mental context” (Zerubavel, 1991, p. 121).
“...it is we ourselves who create categories and force reality into supposedly insular compartments” (Zerubavel, 1991, p. 76).
Classification:
not what’s there - but what’s created.
“Even at the scientific level, there exist different patterns and regularities which criss-cross the natural world—and different regularities can result in different divisions of the world” (Bryant, 2000, p. 112).
Pluralism
Pluralism?
A first take at pluralism:

“...yes, sure enough, the words or concepts of a given people or society are not absolute or God-given”

“...they are relatively stable across some identifiable (and expensive) unit of analysis; stable across a populace, say, or across a culture, or across a given historical period”

Smith, 1996, p. 109
“...pledge allegiance to a degree of cultural sensitivity, but at the same time to view that sensitivity, as it were, as extra-theoretical ... but then thereafter ignored.”

Smith, 1996, p. 112
A first take at pluralism:

“...yes, sure enough, the words or concepts of a given people or society are not absolute or God-given”

“...they are relatively stable across some identifiable (and expensive) unit of analysis; stable across a populace, say, or across a culture, or across a given historical period”

Smith, 1996, p. 109
What then?
“...a dynamic, day-by-day, in-the-rough, wrestling and struggling with the fit of one’s concepts and actions and thoughts into the world surrounding them.”

Smith 1996, p. 108
“...fit with the rest of one’s beliefs and hopes and desires, fit with ideas and desires of one’s fellows, fit with one’s community and history—and by no means least, fit with the subject matters they are about.”

Smith 1996, p. 108
Plurality is not something that can be set aside as simply something that has do with culture, society and language - but it is also something that has to do with the individual.
Pluralism in classification: any object, any document and any domain can be classified from *multiple equal correct* perspectives.
“...an attempt to achieve impartiality can become an insidious form of dogmatism. The classifier can seem to be fair to all, while under the cloak of consensus he is a reactionary seated in the centre of learning” (Broadfield, 1946, p. 78).
Problem #3:

A declaration of neutrality is a declaration that one assumes that one’s view is a view from nowhere, that one somehow holds a view that is superior to other’s views.
Agenda

• The problem of classification

• Three problems with modern classification:
  1. Classification vs. Categorization.
  2. Documents as Containers.
  3. Objectivity vs. Plurality.

• What to give up.

• What to do next.
Three foci in contemporary information organization research and practice that need to be given up.
focus on technicalities: although classification has political and social consequences -- and as such reaches beyond “pragmatic and technical questions” (Andersen & Skouvig, 2006, p. 316), the focus in much current classification research is on those technicalities.
#2
focus on objectivity: while many current classification theorists acknowledge the messy and interpretative nature of language and cognition, there is nonetheless a focus attempting to be objective by setting aside these issues in the “orderly”, “lawful” and “systematic” (Jacob, 2004, p. 522) classification of entities.
focus on globalization and standardization: classification has focused on finding or creating commonalities and generalities across domains, and general laws and principles (Svenonius, 1992); and to find or create the “the one best classification system that will serve all purposes” (Miksa, 1998, p. 81) - at the expense of the local and context dependent.
Agenda

- The problem of classification
- Three problems with modern classification work:
  1. Classification vs. Categorization.
  2. Documents as Containers.
  3. Objectivity vs. Plurality.
- What to give up.
- What to do next.
The future of information organization...
There are at least two alternative models proposed for the late-modern library:
#1: The virtual library.
The physical library will dissolve, there is no need for physical libraries in a world that is all digital. Libraries will only exist virtually -- all material are tagged by the community, in a Web 2.0 type fashion... people who are experts in the material’s subject matter will get to classify/index/tag it.

The need for pre-planned large-scale systems dissolves, as the organizing and representation emerge from the use and the users.
#2: The local library.
The physical library becomes rooted in its local community - and as such reflects the community's needs, philosophies, ideologies, make-up, and view of the world. These libraries are small, focuses on the localities, and are tightly connected to the particulars of the community and the people in that community. The organization and representation of material will be done with the community in mind - and with help from the community.

The need for pre-planned large-scale systems dissolves, as the organizing and representation emerge from the use and the users.
In both situations:

- no large scale, pre-planned classification systems.

- close collaboration with users.

- transparent systems.

- focus on the local.
A hope for future information organization research:

• register the diversity and complexity of the world.

• build theories of classification that match our interactions with the world.

• allow for plurality in classification systems.
Thanks

Jens-Erik Mai
individual.utoronto.ca/jemai
je.mai@utoronto.ca
For more, see:


Available at: individual.utoronto.ca/jemai
The illusion of objectivity
The invention of the modern catalog

Jens-Erik Mai
individual.utoronto.ca/jemai
je.mai@utoronto.ca
References
Fish, Stanley. 1980. *Is there a Text in This Class: The Authority of Interpretive Communities*. Cambridge, MA: Harvard University Press.


