Joseph T. Tennis. (2002). "Subject Ontogeny: Subject Access through Time and the Dimensionality of Classification." In *Challenges in Knowledge Representation and Organization for the 21st Century: Integration of Knowledge across Boundaries: Proceedings of the Seventh International ISKO Conference.* (Granada, Spain, July 10-13, 2002). Advances in Knowledge Organization, vol. 8. Würzburg: Ergon: 54-59.

PREPRINT

Joseph T. Tennis
University of Washington Information School

Subject Ontogeny: subject access through time and the dimensionality of classification

Abstract: Classification schemes undergo revision. However, in a networked environment revisions can be used to add dimensionality to classification. This dimensionality can be used to help explain conceptual warrant, explain the shift from disciplinary to multidisciplinary knowledge production, and as a component method of domain analysis. Further, subject ontogeny might be used in cooperative networked projects like digital preservation, online access tools, and interoperability frameworks.

1 Introduction

Knowledge changes through time. Classification schemes as tools for accessing knowledge undergo constant revision. It is impossible to claim that the ontology of subjects and their interrelationships, once established by a classificationist, remain constant within that scheme. As revisions to classification schemes emerge, so too do new subjects. These, new parts of the updated classification scheme are elements in a formal system – elements that represent the current interpretation of knowledge. The classes in a classification scheme are mutually exclusive and jointly exhaustive. Each class can be seen as a description of what the other is not, within this system. Thus, as the classificationist's interpretation of literary warrant changes, so too do the classes, their boundaries, and their interrelationships within the scheme. Normally, the result is a revised scheme absent of any record of what the knowledge landscape looked like before this current scheme. In a print medium, the record is set in type, allowing the curious to research the interrelationships and terminology of knowledge as it has grown. In a digital networked environment, a scheme can be constantly updated. And with each revision the palimpsest of the past paradigmatic cosmos disappears.

The following example illustrates how the DDC's interpretation of knowledge has changed over the last century. Eugenics, (a term with various definitions through the Twentieth Century) has been classed in the Dewey Decimal System alternately at 575.1, (with Genetics in before the 16th edition) and 363.92 (under Social Problems and Population). DDC 20 provides an index entry for Eugenics that looks like this:

Eugenics	363.92 [Population quality]
crime prevention	364.4
see also Crime prevention	
health	613.94

Joseph T. Tennis. (2002). "Subject Ontogeny: Subject Access through Time and the Dimensionality of Classification." In *Challenges in Knowledge Representation and Organization for the 21st Century: Integration of Knowledge across Boundaries: Proceedings of the Seventh International ISKO Conference.* (Granada, Spain, July 10-13, 2002). Advances in Knowledge Organization, vol. 8. Würzburg: Ergon: 54-59.

medical ethics	174.25
see also Medical ethics	
population control	363.98
social services	363.92
sterilization services	363.97

The entry in DDC 20 for Genetics is:

Genetics	575.1
animal husbandry	636.082 1
animals	591.15
humankind	576.139
social theology	291.178 365
Christianity	261.836 5
sociology	304.5
microorganisms	576.139
plants	581.15

The Relative index of the 16th edition reads as such:

Eugenics	613.94 [same as Eugenic practices—hygiene]
formerly	*301.323; † 575.1

The entry for Genetics in DDC 16:

Genetics

animal	591.15
general	575.1
human	613.9
plant	581.15

Above is an example of a subject ontogeny.

Current interest in the human Genome project will arouse curiosity in readers. They may ask how documents classed in 575.1, placed on the shelf, marked either as Eugenics or Genetics classed on this topic relate to what is being discussed in the news media, for example. It may also be true that Eugenics is still part of Genetics, as a discipline, in some relationship other than represented in the class numbers above.[1] What kind of access is granted by a classification system that shows how knowledge has changed, verses one that revises classes, denying access to the classificationist's interpretation of the change in knowledge? With each revision, a scheme for classification cuts itself off from its previous view of knowledge, building an artificial boundary of time. There are other rhetorical questions pertinent to time as it relates to subject access. For example, could one access the array of subjects in higher education that were taught during Plato's Greece? Through a classification scheme, can one collocate the works of proto-

anthropologists? These knowledges are not reflected in classification schemes, because each living scheme needs to be revised to be viable – thereby eliminating the fossil record of literary warrant.[2] To what degree do revised classification schemes blind us to how subjects change and are re-collocated through time? What can knowledge organization theory do to help the sophisticated user re-collocate knowledge through time? This can be answered by charting the development of a class in a classification system through time. In other words, this can be answered by charting the subject's ontogeny.

2 Classification Theory and the Dimensionality of Classification

Discussions surface constantly about the poverty of classification systems. Not necessarily the schemes themselves, but the conceptual accompanying material, the types of description they are not successful at, and on what philosophical ground they are based. Each of these complaints, outlined in further detail below, points to the social and documentary nature of classification schemes. That is, classification is a social and documented practice. Thus, as people interact (in a social way) with this information structure they are using, in a large way a document – the classification scheme. There is often a disconnect however, between the social dimension and the documentary dimension of classification. Cochrane, Beghtol, and Hjørland outline three examples of this disconnect.

2.1 Conceptual Warrant

In an effort to aid online searching Pauline Cochrane, asks for a tool, "some kind of management information system which would collect data about concepts indexed in our databases and provide some structured analysis" (Cochrane, 1995 p. 36) that will allow us to manage the semantics of the system. That is, she wanted to see a tool built for information professionals to manage the classification system as it grew and changed to reflect literary warrant. This tool would be used to fix the extension and intension (breadth and depth of coverage) of class numbers as warrant demanded. This same tool might prove helpful for us to examine how literary warrant is shaped through time, providing access to older conceptions of the universe of published materials.

2.2 Organizing Principles of Classification Systems

Clare Beghtol (1998) questions the organizing principles of knowledge organization schemes. They are: disciplinary structure, fiction/non-fiction distinction, and the document as a unit of analysis. The functionality of current classification schemes is questioned here because 1) they are not flexible enough to express meaningful relationships between, within and among disciplines, 2) they divide the universe into narrative and non-narrative dichotomies though literary warrant does not reflect this division, and 3) current schemes do not express the "internal elements of documents," (Beghtol, 1998 p. 2). However, each of these principles fit with the common interpretation of the universe of knowledge – at one time. They no longer fit the

interpretation of knowledge organization because of current literary warrant and information technology. Were mechanisms in place for subject ontogeny to adapt to these shifts, without rupturing other economic and social mores, for example in database displays, the expressivity and flexibility sought by Beghtol might be achieved. In this sense subject ontogeny is much like the need for multiple views outlined in Beghtol (1998) and Albrechtsen and Jacob (1997).

2.3 Domain Analysis

Related to the idea of subject ontogeny, Birger Hjørland (1998) establishes a rationale for classification to examine knowledge as it changes through time. This is the *historical method*, one of four basic kinds of methods (Hjørland, 2002). The *historical method* is one way of understanding the story of a domain. It is one way of understanding that the "classification of a subject field requires a concept or view of that particular field," that a "classification cannot be neutral regarding its approaches or theories about subject matter," (Hjørland, 1998 p. 164). This, like a shift from disciplinarity to multidisciplinarity as a design construct, can be documented in mechanisms that manage subject ontogeny. Thus, as a part of the tool of classification, subject ontogeny might be used as tool to refine classification via methods of domain analysis.

3 Mechanisms for Subject Ontogeny

The mechanisms for a subject ontogeny are at our disposal. Some are in place, though not used in a robust way. Authority files are one example of the way that revisions are managed now with subject headings. That technology is present. What is not present is the expertise and desire to shape this historical perspective in classification. What is needed is a simple tool that can be added to displays, that offers a representation of the subject ontogeny.

4 Applications

4.1 Metadata Preservation Models

The OAIS model (CCSDS, 1999) is a model, widely adopted in the digital preservation community. This model wraps digital data with metadata for future access. That future access, even if based on an established system, may hinge on the consciousness of subject ontogeny to the digital objects contained in this model. That is, in order to retrieve, through time, this kind of digital object, systems can be built that allow for updating and linking classes to the predecessors. By linking classes, as opposed to revising completely (that is deleting the old class), the information system can preserve relationships as they existed before.

4.2 Online Access Tools

Marcia Bates argues for a distinction between indexing and access for networked information resources, (Bates, 1998). The distinction she draws between indexing and access puts knowledge organization in the middle of the two. She says, "the user front-end can be designed

around the distinctive traits and evolutionary adaptations of human information processing, while the internal indexing describing the document may be different. Changes in our developing understanding of human cognition in information-seeking situations, and changes in vocabulary can relatively quickly be accommodated in a user-oriented front-end, without requiring the re-indexing of giant databases," (Bates, 1998 p. 1202-1203). That is, if the database is designed with the idea of subject ontogeny in place.

4.3 Interoperability

Current work in Digital Libraries is looking ardently at how collections, and specifically their metadata can be shared across the network – how they can interoperate. Many manual and automatic solutions are being suggested (Arms et al., 2002). What is clear from this work is coordination in the online environment requires policy that shapes knowledge organization structures' ability to evolve. Subject ontogeny, as a matter of design will aid in work load in such projects.

5 Summary of Classification's Potential Dimensionality

If we take a class to be an object in a system of objects that have relationships (hierarchical or otherwise), we can, by a number of warrants, chart its development through time. Subject ontogeny augments the dimensionality of classification schemes.

That classification is an interpretive process is not in question. Placing items in relationship to one another is an act of interpretation. However, the dimensions of those relationships need to manifest in information displays. Cochrane, Beghtol, and Hjørland call for an expression of dimensions. A user curious about the history of Genetics, clicking through a display using DDC numbers may seek an explanatory display about the history of this class, 575.1.

5.1 Encyclopedism

That classification needs dimensionality expressed in its construction, display and initial analysis, is evidenced in the literature cited above. By acknowledging this need, and by positing corrections and additions to this *document* that is a classification scheme, we are moving toward an encyclopedic structure. An encyclopedic structure is one that tells us many things about the nature of subjects. It tells us about their origins, how they change through time, and if they have ceased to exist. Again, this technology and interpretive stance is present in knowledge organization writ large. However, there has yet to be an expression of these needs in a display that allows us to access the social life of the class through a classification scheme or a class number.

Subject ontogeny, if added to classification, will address concerns classification theorists have expressed in the literature for a change in classificatory structure. Subject ontogeny could be used to express ideas of conceptual warrant and how those change through time (Cochrane,

1995). It can help us interpret multidisciplinarity, while recognizing the established institutions of knowledge production (Beghtol, 1998). Further, subject ontogeny will allow us to express the historicity of subjects in domains, (Hjørland, 1998). Most importantly, subject ontogeny is not a radical change, but an addition to existing systems. It is an interpretive layer that can serve the needs of information workers as well as users.

Notes

1 The social commentary offered by the DDC is well documented elsewhere. Any discussion of this context is beyond the scope of this paper.

2 However, because classification can be seen as a document (in the abstract – as a construct of an intellectual pursuit and in the concrete – the form of written schedules), it is therefore an artifact. Classification is a fossil in the sense that it is a record, no longer is it an access tool. Classification, in all its successive versions, could be considered both a record and an access tool, if we take into account the idea of subject ontogeny.

References

Albrechtsen, H. and Jacob, E. (1997). Classification Systems as Boundary Objects in Diverse Information Ecologies. In E. Efthimiadis (Ed.). Advances in Classification Research

Arms, W., Hillmann , D., Lagoze, C., Krafft, D., Marisa, R., Saylor, J., Terrizzi, C., Van de Sompel, (2002). A Spectrum of Interoperability: The Site for Science Prototype for the NSDL. In D-Lib Magazine (8)1 [Online] http://www.dlib.org/dlib/january02/arms/01arms.html

Bates, M. (1998). Indexing and Access for Digital Libraries and the Internet: Human, Database, and Domain Factors. In Journal of the American Society for Information Science, 49(13): 1185-1205.

Beghtol, C. (1998). Knowledge Domains: Multidisciplinarity and Bibliographic Classification Systems. In Knowledge Organization 25(1/2): 1-12

CCSDS [Consultative Committee for Space Data Systems]. (1999) Reference Model for the Open Archives Information System.

Cochrane, P. (1995). Warrant for Concepts in Classification Schemes. In Advances in Classification Research. Volume IV. (Medford, N.J.: Information Today). 35-46.

Hjørland, B. (1998). The Classification of Psychology: A Case Study in the Classification of a Knowledge Field. In Knowledge Organization 25(4): 162-201.

Hjørland, B. (2002). Epistemology and the Socio-Cognitive Perspective in Information Science. In Journal of the American Society of Information Science and Technology, 53(4): 257-270.