Scientific Models as Works

Scientific models as works

By Anita S. Coleman

School of Information Resources and Library Science, University of Arizona at Tucson, AZ 85719, USA. Email: asc@u.arizona.edu

Abstract:

This paper examines important artifacts of scientific research, namely models. It proposes that the representations of scientific models be treated as works. It discusses how bibliographic families of models may better reflect disciplinary intellectual structures and relationships, thereby providing information retrieval that is reflective of human information seeking and use purposes such as teaching and learning. Two examples of scientific models are presented using the Dublin Core metadata elements.

Outline:

Background:

The current environment of scholarly information organization for retrieval in libraries is based on two important traditions:

1. Information handling tools like the library catalog are intrinsically different from bibliographic databases and indexes of journal articles. This is because library catalogs must accommodate information retrieval from both physical storage and conceptual content. Not so the databases or indexes, which are often concerned only with conceptual information retrieval only. Therefore, from the library perspective the two tools, periodical indexes (bibliographic databases) and library catalogs, provide bibliographic control of the universe of knowledge. From the user perspective, indexes and catalogs must both be consulted for information retrieval from the bibliographic universe of knowledge.

2. Information resources for inclusion in the library catalog are often chosen because they are bibliographically independent publications.[1] The Anglo-American Cataloging Rules, 2nd edition revised (AACR2R) specifies these as books, pamphlets and printed sheets, cartographic materials, manuscripts (including manuscript collections), music, sound recordings, motion pictures and videorecordings, graphic materials, computer files, three-dimensional artifacts, and realia (the exception to "bibliographic"), microforms, and serials.[2] In other words, these are the units of analysis, the item/object granularity level at which the library catalog functions. Typically, the whole item (book, serial, etc.) is described; individual book chapters are not cataloged though AACR2R provides for this.[3] Similarly, the periodical indexes have taken over the role of providing access to component parts that can also be considered independent units, such as journal articles.

Both of these traditions have been challenged as the practice of representing information on digital media continues to rise. Patrick Wilson notes that in the global, online, multimedia information world we can no longer take textual or conceptual stability for granted.[4] An important question then to investigate is this: How can the primary bibliographic tool, the library catalog, better reflect disciplinary knowledge structures? This paper investigates by using the notion of works for one class of intellectual (and disciplinary) creations, scientific models. Before proceeding to a discussion of scientific models as works, current practices in cataloging and indexing, assumptions, limitations and scope of this study are presented. There is a glossary, mostly drawn from Smiraglia, which defines terms used in this paper.[5]

Current cataloging and indexing practices:

Library catalogs and indexing databases focus on the subjects of disciplines (what topics and concepts are there within a particular subject or discipline) and not on disciplinary intellectual activities (at least not in the sciences, and for example, modeling) that result in creative products that can be indexed for information retrieval. Questions such as what are the intellectual products of disciplinary activity, how are such products (for example, models) represented in bibliographic entities, what are the component parts of such representation, how are the parts related, etc. appear to be beyond the scope of bibliographic organization. Therefore, current cataloging, indexing, and classification of models exists only for representations of the textual content of models, the written descriptions about models and the activity of modeling as recorded in published literature; these are usually found in text, items, and documents such as journal articles, scientific reports, theses, dissertations, books and chapters in books.

Bibliographic control of models and modeling that is reported in published literature as a scientific activity is enabled through three types of tools: library catalog, bibliographic utility, and periodical index (from henceforth the term index includes bibliographic databases and periodical indexes). Information retrieval in these tools is facilitated through description and subject analysis. Subject analysis includes classification. Resources about models are often classified as subjects and this entails the use of controlled vocabulary systems like thesauri, classification or subject heading lists. In the library catalog and in bibliographic utilities, the Library of Congress Subject Headings (LCSH) is the predominantly used controlled vocabulary list. Dewey Decimal Classification (DDC) provides the classification number for item location of the unit. The LCSH descriptor (preferred term) is Models and modelmaking, which may be subdivided geographically; there is Mathematical models which may be subdivided by object and narrower terms like Atmospheric models, Hydrologic models, Wind Tunnel models, etc..[6] Appendix 1 provides a list of the LCSH subjects under Models and modelmaking. Controlled vocabulary in the indexes is dependent upon discipline thesauri and each index usually selects and uses a different thesauri, subject heading or classification scheme. For example, GEOREF (an index for the geological sciences) uses the GEOREF Thesaurus while INSPEC (another index in the physical sciences and engineering) uses the INSPEC Thesaurus (see Appendix B for more details).

Assumptions and limitations:

This study is based on the following assumptions:

1) Information retrieval in libraries and through bibliographic tools must actively support, if not enable, end-user information seeking purposes such as exploratory learning and information uses such as teaching and learning. Therefore, our library tools must reflect disciplinary knowledge structures, products, and uses.
2) In the online world continuing to segregate tools such as indexes, catalogs, and bibliographies is inefficient. For information seeking purposes such as teaching and learning efforts should be made to merge the three tools for improved end user searching and information retrieval.
3) Boolean searching is end-user hostile. Tests as early as Cranfield have shown that Boolean searches do not improve retrieval performance significantly over other types of searches.[7] This study assumes that phrase searching, for example noun phrases like ‘tree rings’ are preferred user search strategies and hence bibliographic description should accommodate such information retrieval.
4) AACR2r defines models as "a three-dimensional representation of a real thing."[8] It also provides descriptive cataloging rules for cataloging models as physical objects. This is not the definition of scientific models as used in this study. A different definition, one that is grounded in how the word is used in the disciplines (sciences and social sciences) is proposed.

A major limitation of this study is that it does not completely include the representations of models in heterogeneous formats; it is limited to electronic formats only. Also, because modeling is a widespread activity, an attempt is made to identify important properties of scientific models only. Finally, this is part of a larger funded study that is developing a classification scheme for scientific models in one are, water quality, and building a prototype catalog of scientific models.

Definition of scientific models:

In every field of human endeavor, including the natural and engineering sciences, the word model can mean different things and conjure different images to different people.

The Oxford English Dictionary Online provides three major meanings for the word model: a representation of structure, a type of design and an object of imitation.[9] From the more than 15 meanings within the above three contexts, the following definition best fits an initial consideration of scientific models as works: a model is "a simplified or idealized description or conception of a particular system, situation, or process (often in mathematical terms: so ***mathematical model***) that is put forward as a basis for calculations, predictions, or further investigation" [emphasis original].

The Encyclopedia Britannica Online offered 4145 articles for a search on the word model.[10] This is in addition to the meanings for six words (model (n), model (v), model (adj.), animal model, role model, Watson-Crick model). The first article entry is titled "Model Construction from Operation Research" and has the brief introduction: "A **model** is a simplified representation of the real world and, as such, includes only those variables relevant to the problem at hand. A **model** of freely falling bodies, for example, does not refer to the colour, texture, or shape of the body involved. Furthermore, a **model** may not include all relevant variables because a small percentage of these**...**"[11] [emphasis original].

The Academic Press Dictionary of Science and Technology provides a similar definition for a model as used in the sciences.[12] It is "a pattern, plan, replica, or description designed to show the structure or workings of an object, system, or concept." More specific definitions are provided for the areas of Behavior, Computer Programming, Artificial Intelligence, and Photogrammetry.

It is clear that there are models in both social sciences and the sciences. It is equally clear that the activity of scientific modeling includes mathematical and computer-based is prevalent in many disciplines. For example, a search for the term model in Science Direct, a full-text index to Elsevier Journals subscribed to by the University of Arizona Libraries, yielded 1157 articles for the year 2002 only[1]. A search in the same database, all journals, 2001-onwards, for the phrase "computer models" within abstracts retrieved 2759 articles in journals such as *Chaos, Solitons, & Fractals*, *Muscle & Nerve*, *Computers & GeoSciences* belonging to science, medicine and social science disciplines.

Even in the sciences only we find that there are many different approaches that scientific disciplines take to models; they can be complex numerical models implemented on computers, mechanical analogs, performs of theories or restricted concepts within which basic dynamical aspects can be described and understood.[13] However, there appears to be widespread consensus that 'scientific models' reflect the intellectual activity and include computational and mathematical modeling. Hence the phrase scientific models is more accurate than mathematical or computer models.

[1] I was unable to replicate this search for other years as the search does not proceed because of the system limitation - retrieved too many hits – hence user is asked to modify the query.

Importance of scientific models for teaching and learning:

Integrating scientific, mathematical, and computational modeling, which includes model use and building, with teaching and learning new science concepts has been recognized since the 1980s as important for a number of reasons. Such integration, it is believed, provides the framework for assembling data and knowledge, for stimulating scientific reasoning, discovery, synthesis, analysis, and theoretical development skills in novice student learners.

NASA identified the kinds of prerequisites that are needed for integration of research and model use and building with pedagogical goals.[14]

1. The *acquisition of observations* (the model must exhibit a selective attitude to information) – student's ability to select relevant observables

2. *Analysis and interpretation* of the observational data (structured and pattern-seeking/replicating)

3. *Construction of and experimentation* with conceptual and numerical models (as analogies and prediction devices)

4. *Verification* of the models, together with their use to furnish statistical predictions of future trends (testing, experimentation and replication).

ThinkerTools, where researchers from the University of California at Berkeley and Educational Testing Services (ETS) collaborated with middle school teachers, is a scientific inquiry and modeling project.[15] They developed Morton Modeler, a computer agent, who walks users through the process of building good scientific models. Another noteworthy project is Modeling for Understanding in Science Education (MUSE) based in the University of Wisconsin at Madison.[16] MUSE includes middle and high school students and is focused on improving understanding about science as a modeling enterprise and scientific modeling skills.

Scientific models draw on fundamental laws and equations and are potentially rich sources for solutions to interdisciplinary problems. Hence, bibliographic tools that show disciplinary activity relationships associated with modeling, name, phenomenon, process, object relationships in the model (subject relationships), mathematical, computer, and use relationships (what kind of mathematical function does the model use, purpose of model, how many and what types of variables, what kind of computer, what kind of software) are important for disciplinary enlightenment during the information retrieval process that accompanies teaching and learning tasks.

Aspects of scientific models:

The specific area of water quality and the broad discipline of geography were studied to provide the preliminary aspects (facets) of scientific models. Appendix C provides excerpts of raw data used in this analysis. General properties of scientific models are:

1) Models reflect reality.
2) They are small representations of reality.
3) They are simpler than the process/phenomenon they study or model .
4) They are closed, not open, systems.
5) Any real situation can be analyzed if it can be described in terms of mathematical equations.
6) The most important features of reality are correctly incorporated; less important features are initially ignored.

Important aspects of scientific models are as follows:

1. Purpose or type of model. What is the purpose of modeling? Many classifications of models by purpose exist; classifications in geography, and

hydrology, environmental sciences (see Appendix C) and physics, biology. Scientific classification usually has different purposes from classification for information retrieval. ThinkerTools also provides a simple typology of models.

Example: quasi-realistic (simulation), cognitive (explanatory)

2. Object of study. What is the object or objects being modeled?

Example: wind tunnel

3. Process: These objects that the model studies, investigates or imitates, participate in a natural process. What process or processes does the model simulate or study?

Example: erosion

4. Phenomena. What phenomenon does the model simulate, study or seek to explain?

Example: clouds

5. Fundamental Law. Is there a fundamental law that the model is based upon? What?

Example: Bernoulli Law

6. Mathematical or statistical function. Models are represented through textual theory, law, and mathematical equations. Can the fundamental law be expressed mathematically? What is the equation? What are the mathematical functions that the model uses?

Example: differential functions

7. Variables. What are the conditions and variables studied, modeled, input, output, hypothesized or generated?

Example: soil properties

8. Spatial coverage. What is the spatial coverage of the model? Various schemes can be used here: geographic scale, named features, geo-references, etc.

Example: global models, regional models

9. Temporal coverage. What is the temporal coverage of the model? Again, various schemes can be used here: temporal scales, named geologic periods, historical periods, etc.

Example: micro, meso, 1-hr

10. Software. What software is needed to run the model? What documentation is available about software? Is the software available in executable code? Several components such as operating system and other computing requirements for the model can be included here.

Example: Fortran source programs

11. Hardware. What is the hardware environment of the model?

Example: PC

11. Person/group who proposed the original model/authored the paper, etc. Is there an original mathematical, computational, scientific model? Is there a theory? Whose? Who did the work? Is it possible to identify original models that continue to be revised, modified, updated? Or for which alternative solutions are proposed? What are the bibliographic relationships that exist between these models?

Example: Box-Jenkins, Streeter-Phelps

12. Discipline. What is the major discipline that can be determined for the model either by creator affiliations or other means? The disciplinary facet is an important one for promoting cross-disciplinary collaboration and information retrieval about models.

Example: hydrology

13. Replication. Has the model been replicated? This facet is related to the one about the person/group who did the original model. Model replications are important reports for continued model modification and use. Identifying varying types of replications may be useful too.

Example: IsReplicatedBy

14. Related Materials. What types of other related materials exist about this model?

Example: IsAnalysisOf

Much more work is needed before we regard this as the definitive list of scientific model properties and relationships to be used for information retrieval in a tool. But, this provides a good beginning for initial prototype design and subsequent user study.

Information retrieval of scientific models:

An informal and small survey of the information retrieval and use problems associated with constitutive models as currently represented in a library catalog, periodical indexes and bibliographic utility, and the WWW provided preliminary evidence for re-considering the cataloging of scientific models and investigating scientific models as works. The following two tables demonstrate the representation and problems associated with retrieving information resources about models in four sampling frames, index, utility, catalog, and the World Wide Web (WWW).

The survey focused on the following two types of terms, those advised by expert users and those selected from the LCSH. Controlled vocabulary terms from the LCSH include: Atmospheric models, Hydrologic models, Mathematical Models. User-suggested term is constitutive models. Constitutive models are an important class of models in civil and rock engineering. Constitutive models are based on constitutive equations, relations and laws. Engineering faculty at the University of Arizona suggested this as a class of scientific models that needed better information retrieval for novice learners, senior undergraduate and graduate level engineering students. Table 1 shows the search terms used, the sampling frame where the search was conducted, and the number of information items, hits, that were retrieved and the dates of the search on a particular class of scientific models. Appendix B provides the notes about the selected sampling frames.

Table 1: Representation and Number of Resources about Constitutive Models

Search Terms and Search Strategy	Sampling Frame	Number of Hits & Date of Search (in parentheses)
Constitutive models – Subject	OCLC	9 (Nov. 2001)
Constitutive models – Subject	Sabio	0 (Nov, 2001)
Constitutive models – Google	WWW	35,200 (Nov. 2001)
Constitutive Relations – Subject	OCLC	2 (Nov. 2001)
Constitutive Relations – Keyword	OCLC	33 (Nov. 2001)
Constitutive relations – Subject	Sabio	0 (Feb. 2002)
Mathematical models – Subject	OCLC	130,141(Nov. 2001)
Atmospheric models – Subject	OCLC	2018 (Feb. 2002)
Atmospheric models – Subject	Sabio	1 (Feb. 2002)
Hydrologic models – Subject	OCLC	1,075 (Feb. 2002)
Hydrologic models – Subject	Sabio	95 hits with 11 subject entries (Feb. 2002)
Models – Subject	OCLC	179,717 (Nov. 2001, Feb. 2002)
Models – Subject	Sabio	447 subject entries (Feb. 2002)
Models – Keyword	OCLC	241, 349 (Feb. 2002)

Table 2 identifies the controlled vocabularies used to index or catalog information resources about constitutive models. It shows how different and scattered the subject terms for constitutive models are. Often, the term model need never be used; at other times it must be used in conjunction with other subject subdivisions or topic ideas. There appears to be no easy provision for index displays of retrievals using model names, objects, processes, phenomenon; terms, subject headings, or descriptors for various subject concepts do not reveal precise subject relationships such as process/agent. Relationships that are important in modeling, therefore, are not revealed. Collocation, the bringing together of all 'works' on a particular subject, especially if we extend the notion to scientific models, is therefore complicated and made extremely difficult or incomprehensible for new students and people unfamiliar with the discipline or topic. For example, bibliographic records for models on runoff reveal nothing about objects such as water, hydrologic bodies, or processes, such as saturation and infiltration. When scientific models are cataloged with controlled values for objects, phenomenon, processes as parts different types of subjects, collocation is facilitated. Furthermore, current cataloging and indexing practices of models assigned one or two 'subjects' and not as creative artifacts, as works, reinforces the often-held views of non-scientific thinking; namely, that scientific models are physical objects, mechanical devices, or physical scale models only.

Table 2: Subject Information Retrieval (IR) & Constitutive Models

Name of Database	Name of Thesaurus	Subject Terms Used and Their Narrow Terms	Narrower Terms	Subdivisions	Notes
OCLC	Library of Congress Subject Headings	Engineering models Mathematical models	Acoustic models Beggs method Electromechanical analogies Hydraulic models	Bridges, Concrete – Models Mathematical models – Construction Industry	Neither constitutive models nor constitutive relations are preferred (used) subject headings.
OCLC	Local subject headings	Constitutive models			Used as local subject headings.
INSPEC	Inspec Thesaurus	Models	*Over 30 such as* *Band theory models* *Exchange models* *Brain models* *Elementary particle interaction models* *Quark models* *Sandpile models*	None	Neither constitutive models nor constitutive relations are preferred (used) as subject descriptors
EI COMPENDEX	Compendex Thesaurus	Models is not a descriptor; nor is constitutive models	*Note: A search from abstract/title/subject pulled up 6605 hits but majority of the hits picked the phrase from the title*		"constitutive models" is a common phrase in titles; possibly useful for index displays (using subject or other models relationships criteria) for IR that enlightens novice users
Materials Science (Cambridge Scientific) Abstracts (MSA)	Copper Thesaurus, Engineered Materials Thesaurus, NASA Thesaurus, Metallurgical Thesaurus	As above	*However 250 hits were retrieved for a search in title on constitutive models*		Many of the hits appeared to be chapters in books.

Table 3 shows descriptive bibliographic and brief subject information for the items about constitutive models that were found in OCLC.

Table 3: Information Resources in OCLC about constitutive models

OCLC #	Type	Language	Date	Format/Form*	# of Subject Headings
4552889	Book	English	1998	/Ph.D. Thesis	14
45523973	Book	English	1997	/Report	11
42890758	Book	English	1998	Microfom /Technical Memo	11
33345127	Book	English	1986	Microform/	7
33188617	Book	English	1986	Micorform/ Symposium proceedings	6
32866588	Book	English	1988	Microform/ Report	7
32617094	Book	English	1989	Microform/ Report	7
32115587	Book	English	1989	Microform/ Conference proceedings	4
25003695	Book	English	1990	Microform/ Technical Memo	11

*Format is what OCLC calls them sometimes; form is my analysis. Therefore, reading row one, we find that format is not given (hence just book or printed text) and form is a Thesis.

Using Tables 1-3, a partial list of information retrieval problems for scientific models can be deduced as follows:

1. For information resources and packages, such as books, theses, dissertations, and reports, in library catalogs, the preferred controlled vocabulary scheme is LCSH. LCSH uses a variety of subdivisions to enable the cataloger to assign subject headings; these include topical (other topics), form, time, and geography subdivisions. Smaller information packages such as journal articles, conference proceedings articles are covered by bibliographic indexing services and include indexing and abstracting sources.

For these, the preferred controlled vocabulary it can be seen varies from index to index. INSPEC uses the INSPEC thesaurus. EI COMPENDEX and MSA use their own home-brewed or multiple other schemes. Therefore, knowing the right vocabulary to search for constitutive models, or indeed any kind of scientific modeling, is a major problem.

2. Most information resources about models are theses, dissertations, and government or agency reports. These items are often the least cataloged in libraries and subject analysis is often cursory, maybe not even done. Yet, these are probably some of the richest resources about constitutive laws, equations and models. Table 3 is interesting in that it shows these types of resources richly cataloged in OCLC with local and controlled subject headings ranging from four to fourteen.

3. Subject cataloging principles such as specific, direct entry work only when there are specific subject entries and as we have seen there is no subject heading for "constitutive models." Even when direct entries are available such as mathematical models, simulation methods under which objects can be added as subdivisions the resulting sort criteria of date and publication type become meaningless in the context of actual use of scientific models. Table 3 provides one example of the lack of usefulness of current categorizations in the utility between type, form, and is probably illustrative of cataloging confusion about form and format (see columns Type and Format/Form).

4. With the increasing scientific activity of modeling, current descriptions of textual content about models are incomplete. It appears that the subject headings for texts about models do not provide much information on the underlying laws, processes, phenomena, type of model, computational requirements or mathematical functions. Yet, these are critical factors in disciplinary teaching, use, and research of models. They should be

included. Information retrieval may be improved if we consider scientific models as works.

Scientific models as works and their bibliographic families:

Are scientific models works? Smiraglia provides an operational definition for a work.[17] "A work is the intellectual content of a bibliographic entity; any work has two properties: a) the propositions expressed, which form ideational content and b) the expressions of those propositions (usually a particular set of linguistic (musical, etc.) strings) which form semantic content."

Using this definition, scientific models most certainly can be considered as works. Semantic content in scientific models includes mathematical expressions, formal propositions and hypotheses, and statements of laws. Ideational content includes ideas about objects, processes, and relationships, usually within or for specified spatial and temporal scales, and formally, semantically expressed as mathematical equations and algorithmic notation. The ideas include both observables (verified and expressed as measurements) and non-observables (hypothetical data, mathematical equations). MUSE researchers reinforce this view in their statement that "a scientific model is a set of ideas that describes a natural process" and that various "types of entities, namely representations, formulae, and physical replicas" are sometimes needed in the formation of scientific models.[18]

Works are bibliographic entities. Examining scientific models as bibliographic entities, we find that they have two properties, physical and conceptual. The physical components of a scientific model can be determined in terms of its form (what the instantiation is):

1. Textual works – includes articles, abstracts, bibliographies, reviews, analysis, software documentation.

2. Datasets – includes observations and measurements of the observed phenomenon, object, process reported as data, images, visualizations, and graphs.

3. Software – includes computer code, both source code and downloadable executables.

4. Services – includes interactive and other services (animation applets, databases, indexes, contact pages, submit forms, etc.)

Conceptual components can be determined in terms of the ideas the model expresses. Even more than just the ideas, the ideational (subject + other) relationships are important in modeling. Hence, conceptual components can also be called model concepts and relationships. They include:

1. Research foci - What is being modeled, or the object(s) being studied? Wind tunnels? Sediment? River? Are objects related? How?

2. Model type - What is the purpose of modeling – explain, predict, simulate, test?

3. Mathematic functions - This is probably the most complicated idea to abstract. The simplest mathematical submodel needs at least three types of variables and a set of operating system characteristics linking them. The three sets of variables are: input variables, status variables (the internal mathematical constant), and the

output variables (which depends on both input and status variables). Model strategy, the fitting and testing of the model by choosing the type of mathematical operations most suitable to the type of system one is trying to model. Finally, developing the algorithmic (computational notation) and testing the model to see if it works as planned before using it for prediction, etc.

4. Instrumentation - What relationships exist between observables, data collected, conditions, and instruments used to gather or generate data.

5. Fundamental theory, law, or hypotheses that drives the model.

6. Replication, revision, simulation and continued improvements, modification.

Examples of two prototypical scientific models investigated as works are the *Bernoulli Model* and the NASA GISS 1999 *Atmosphere Ocean Model*.

Bernoulli Model

The Bernoulli family in Switzerland was one of the most productive families in the field of mathematics in the seventeenth and eighteenth centuries. Table 4 shows the names of three generation of Bernoullis who have contributed to the literature of fields such as calculus and fluid mechanics. Searching through the library catalog, it is hard for the novice learner to clearly identify how the various concepts, which bear Bernoulli names, are related. There are Bernoulli numbers, Bernoulli equations, Bernoulli models, Bernoulli principles, Bernoulli law and the Bernoulli theorem. These have all traditionally been thought of as 'subjects' and cataloged in terms of the various physical formats these ideas were packaged in. There is only one LCSH descriptor Bernoulli

Numbers. This was used along with other terms to identify what, if any, items existed on the Bernoulli models. Can a bibliographic family be identified? Is there a relationship between the Bernoulli model and the Bernoulli equations that are an application of Bernoulli's Law and used in elementary physics learning for various things like curving baseballs and aerodynamic lift? What is the source of the Bernoulli model?

A bibliographic family is a "set that includes all texts of a work that are derived from a single progenitor."[19] It is therefore, "the tangible, and to some extent quantifiable, instantiation of the mutability of works." Bibliographic families usually are created by derivative relationships: one source is the progenitor. Derivative relationships are further classified simultaneous, successive, translations, amplifications, extractions, adaptations, performances.[20] Tillett's taxonomy of bibliographic relationships identifies other types of relationships besides derivative; the seven posited by Tillett and including derivative are equivalence, descriptive, whole-part, accompanying, sequential, and shared characteristic relationships.[21] Other bibliographic relationships that the library catalog tries to address include access point relationships and subject relationships, but Tables 4 through 6 try to show that these relationships are not clearly specified or made obvious to the user in current bibliographical tools.

Table 4 tries to show the searches to identify bibliographic families. It shows the term used, the type of search, and the number of hits the search retrieved in three sampling frames, OCLC (a bibliographic utility), Sabio (the University of Arizona Online Public Access Catalog), and the WWW (using the Google search engine). Sampling Frames & Dates of Search are: OCLC: 01/07/02; Sabio: 01/31/02; WWW: 01/31/02. Terms in Boldface type are Library of Congress subject headings.

Table 4: Bibliographic Families

Search Term	Type of Search	Number of Hits Retrieved In OCLC	SABIO	WWW
Bernoulli	keyword in Basic Search	604	80	125,000
Bernoulli	author in Advanced Search	648	43	N/a
Bernoulli, Daniel	author in Advanced Search	76	8	8,940
Bernoulli, Johann	author in Advanced Search	44	2 (see entries to Bernoulli Jean)	4,860
Bernoulli, Jakob	author in advanced search	107	3	2,530
Bernoulli, Jean	author in advanced search	133	5	6,550
Bernoulli law	keyword in (Basic search)	18	2	19,700
Bernoulli equations	subject words in Advanced search	4	0	21,800
Bernoulli numbers	Subject words in Advanced Search	21	2 subject search	22,800

Note: Bernoulli, Johann, 1667-1748 is same and entered in the library catalog as Bernoulli, Jean.

Table 5 shows subject headings and disciplines in selected records. OCLC# with a d1 or d2 indicates that they are bibliographic records for the same copy; Call number, if any indicates the discipline to which the items has been assigned (physical or online shelf browse number); SH stands for subject heading that is found in the record and the number indicates the position in the list of subject headings if an item had more than 1 subject heading. Subject headings in OCLC records for records with **Bernoulli Numbers** are indicated by an X. Therefore, reading the first row, OCLC record # 40937076 has another record for the same item (line below) and has no call number, three subject headings, with the subject heading Bernoulli number in the third position.

Table 5: Subject & Discipline Relationships

OCLC #	Call Number	SH #1	SH 2	SH3
40937076 (d1)	None	Education Research Methodology	Linear Models (statistics)	X
45340799 (d2)	150	Education Research Methodology	Linear models (statistics)	X
34844313	LB 1861 .C57	X	Numerical functions	-
37561894	515.5	Numerical functions	X	Euler numbers
34594846	QA 8.4	X	Dissertations, academic, Mathematical sciences	-
33001372	T171.G45x	X	-	-
36246557	Q 172.5	Chaotic behavior in systems	X	-
48189946	-	X	Graph Theory	-
25099001		Bernoulli Numbers - Bibliography	-	-
12220874		Euler's Numbers	Bernoulli shifts	-
15817046	QA 279.5	Bayesian statistical decision theory	Bernoullian Numbers	Numerical functions, Computer programs
11107209	QA 55	X	-	-
41382818	QA 246	X	Fermat's theorem	-
42909419 (d1)	QA 246 .F3	X	-	-
41780960 (d2)	QA 246 .F3	X	-	-
41777351	-	X	Series	Equations
3238858 (d1)	QA 246 .S2	X	-	-
48510451 (d2)	QA 246 . S2	X	-	-
05616123 (d1)	QA 246 .S73	X	-	-
05616122 (d2)	QA 246 .S73	X	-	-
43307081	QA 306. S3	Calculus, differential	Mathematics, Dissertation, Germany, Berlin	X

Table 6 tries to identify derivative and equivalence relationships and while it appears straightforward in some case, it is not always so. It is also not clear which of the various types of bibliographic relationships are found in scientific models.

Table 6: Bibliographic Relationships Matrix

OCLC #	Type (Format)	Language	Publisher or Place of Pub	Date of Pub	Form of Pub	Discipline & Sub-Discipline	Derivation Type	Equivalence. Relationship
40937076 (d1)	Book	English	Michigan State University	1998	Ph.D. Thesis	Educational Psychology	New	share an equivalence relationship with below (copy)
45340799 (d2)	Book	English	As above	1998	Ph.D. Thesis	150	As above	As above
34844313	Book	English	Eastern Illinois University	1996	M.A. Thesis	LB1861.C57	New	None
37561894	Book	German	Aachen	1995		515.5	Unknown	Unknown
34594846	Book	English	Tennessee State University	1995	M.S. Thesis	QA 8.4	New	None
33001372	Book	English	Georgia Institute of Technology	1995	M.S. Thesis	T171.G4	New	None
36246557	Book	English	University of Rhode Island	1995	M.S. Thesis	Q172.5	New	None
48189946	Microform	English	Georgia Institute of technology	1992	Ph.D. Thesis	Unknown	New	None
25099001	Book	English	Kingston Ontario (Queen's University)	1991	Bibliography	QA3 (510)	Revision	Updated bibliography is available on the WWW
12220874	Microform	English	NASA (Institute for Computer Applications in Science and Engineering)	1984	(NASA contractor) Report	Unknown	New	None
15817046	Book	French	Ecole Poytechnique de Montreal	1977	Report	519 Mathematics	New	None
11107209	Book	English	Lamar State College of Technology, texas	1968	M.S. Thesis	QA 55	New	None

41382818	Book	English	Philadelphia	1936	Society Report	Philosophy	New	None
42909419 (d1)	Book	English	Unknown	1925	Extracted from Messenger of Mathematics (July 1925)	QA 246	New?	Feinler Mathematics collection Shares equivalence below (copy)
41780960 (d2)	Book	English	Unknown	1925	Extract as above	QA 246	As above?	Equivalence relationship with above
41777351	Book	English	London	1914	Extracts from Quarterly Journal of Pure and Applied Mathematics (no. 181)	Unknown	New?	Unknown
3238858 (d1)	Book	German	Springer	1893	Photocopy	QA 246	Unknown?	Equivalence relationship with below (copy)
48510451 (d2)	Book	German	Springer	1893	Electronic reproduction	QA 246	Unknown?	Equivalence relationship with above
05616123 (d1)	Book	Latin	Unknown	1977 reprint of 1845 edition	Reprint of 1845 edition	QA 246	Edition?	Equivalence relationship with below + (Bound with another text) (copy)
05616122 (d2)	Book	Latin	Unknown	As above	Reprint as above	As above	As above?	Equivalence relationship (bound with another text)
43307081	Book	Latin	Berolini	1823	Dissertation	QA 306	New	None

Metadata for Scientific Models:

For purposes of clearly revealing only the important model concepts and relationships that need to be represented and further investigated, the metadata in the examples below are kept very simple. For example, Dublin Core is used as the content standard but a corresponding encoding scheme is not shown. Additionally, metadata for the English translations of the original works (in the case of the first example, Bernoulli model) are not included. DC facets are the one deviation from the Dublin Core standard elements, and again the purpose is to explore and demonstrate how analytical cataloging and faceted classification strategies can be combined for improving bibliographic control and information retrieval that reveals disciplinary structures. Index displays based on controlled values may be derived semi-automatically in the dc/type, dc/relation, and facets are the key to representing disciplinary and other structures of models. They will be described and demonstrated in a subsequent article and through the models prototype database, a classification-based catalog that is currently under development.

Elaboration about DC elements, as used in this study, are provided below; only deviations or elaborations are included and use of elements when consistent with DC 1.1 are omitted. These are DC-Description, DC-Contributor, DC-Date, DC-Format, DC-Language, and DC-Rights.[22]

- DC-Identifier: Universal Resource Identifier is the unambiguous reference identifier used for each of the items that make up the work
- DC-Title: Title (either cataloger assigned or creator's title)
- DC-Creator: author or other authority
- DC-Subject: Is not used in new records. Will be kept if found. Instead aspect or facet is proposed and many different ones identified; in keeping with the DC content standard model facets are optional and repeatable. Currently, standard classification schemes and thesauri like the ACM's Computing Classification System, American Mathematical Society's Mathematical Subject Classification, GEOREF Thesaurus, and ERIC Thesaurus can be used; but, in the larger study

they are being investigated, compared and will be drafted into a simpler models classification scheme based on the facets below.

- ❖ FacetConcept: is an idea, the traditional subject (for example, calculus of variations)
- ❖ FacetObject: the object studied in the model
- ❖ FacetDiscipline: the major discipline to which this model belongs (may be determined either through author affiliations or other means)
- ❖ FacetPhenomenon: the phenomenon being modeled
- ❖ FacetProcess: the process being modeled
- ❖ FacetMathRepresentation: the mathematical functions, equations used
- ❖ FacetSoftware: the software needed to run the model
- ❖ FacetFunLaw: the fundamental laws that the model is based upon
- ❖ FacetModelType: the type of model based on its purpose
- ❖ FacetVariable: number, types, conditions, and variables in this model
- ❖ FacetProblem: the problem the model is analyzing stated often as a question
- ❖ DC-Type: the type of resource is taken to be its form and until the models classification is fully developed, the model semantic unit and modified LCSH form subdivisions are shown as placeholders.
- ❖ DC-Relation: various types of bibliographic and model relationships to demonstrate work linkages are used and not limited to the ones in the DC Qualified list
- ❖ DC-Coverage: a distinction is made between spatial and temporal coverage; so there is DC-CoverageSpatial and DC-CoverageTemporal.

Figure1: Bernoulli Model – Metadata for sources + two members of the bibliographic family

dc/identifier: http://foo.bar.org/arsconjectandi
dc/title: Ars conjectandi
dc/creator: Bernoulli, Jakob
facetobject: Bernoulli numbers
dc/description
dc/publisher
dc/date: 1713
dc/type: textual works
dc/format: text/html
dc/source: Work of Johann Faulhaber
dc/relation: IsRelatedTo
dc/identifier: http://foo.bar.org/workofjohann

Bibliography

dc/identifier: http://foo.bar.org/bibliography
dc/title: Bibliography of Jakob Bernoulli
dc/creator:
facetconcept: Bernoulli numbers
facetmathfunction:
facetperson: Bernoulli, Jakob
dc/description
dc/date
dc/type: textual works (Bibliography)
dc/format:
dc/source:
dc/language:
dc/relation: IsUpdateOf

dc/identifier: madeupisbnnumber
dc/identifier: http://foo.bar.org/workofjohann
dc/title: Work of Johann Bernoulli
dc/creator: Bernoulli, Johann
facetobject: Bernoulli numbers
dc/description:
dc/publisher:
dc/date: 1742
dc/type: textual works
dc/format: text/html
dc/relation: IsRelatedTo
dc/identifier: http://foo.bar.org/arsconjectandi

Abstract

dc/identifier: http://foo.bar.org/abstract
dc/title: The significance of Bernoulli's Ars conjectandi
dc/creator: shafer, glenn
facetconcept: series, infinite
dc/date
dc/type: textual works (Abstract)
dc/format: text/html
dc/source:
dc/language:
dc/relation: IsAnalysisOf
dc/identifier: http://foo.bar.org/arsconjectandi

Figure 2: Bernoulli Model – Metadata for Other Forms, Members of the Bibliographic Family

This figure shows one added element under consideration, audience (appropriate audience level for the resource).

Exercise

dc/identifier: http://www./mtn-bern.pdf
dc/title: Numerical, graphical and symbolic analysis of Bernoulli equations
dc/creator: Bern, David
facetmathfunction: differential equations
dc/description:
dc/date: 2001
dc/type: textual works (Exercise)
dc/format: text/html
dc/relation: IsProblem
dc/identifier: http://foo.bar.org/bernulliequations
dc/relation: IsSupplementTo
dc/identifier: http://foo.bar.org/computercode
dc/audience: 9-12 (US)
dc/typical learning time: unknown
dc/coverage: not applicable
dc/rights: none

Sample Metadata for Atmosphere Ocean Model:

dc/identifier: http://aom.giss.nasa.gov/index.html
dc/title: Atmosphere Ocean Model
dc/creator: NASA GISS AOM Group

facetdiscipline: Meteorology
facetobject: sea ice
facetphenomenon: climate
facetprocess: river flow
facetprocess: advection
fascetprocess: atmosphere-ocean interactions
facetprocess: insolation
facetmathfunction: atmospheric mass equations
facetmathfunction: differential
facetsoftware: fortran source
facetsoftware: pc executables
facetfunlaw: mass
facetmodeltype: climate predictions
facetmodeltype: grid-point model
facetvariable: ocean entropy

dc/type: interactive service

dc/relation: References
dc/identifier: http://aom.giss.nasa.gov/publicaitons.html

dc/relation: ModelCode
dc/identifier: http://aom.giss.nasa.gov/code.html

dc/relation: Observations
dc/identifier: http://aom.giss.nasa.gov/observe.html

dc/relation: Personnel
dc/identifier: http://aom.giss.nasa.gov/people.html

dc/relation: Simulation
dc/identifier: http://foo.bar.org/12yearruns of the control simulation of 1995 version of CO23

dc/coveragespatial: global
dc/coveragetemporal: decade

As mentioned, the example metadata shown in Figures 1-3 continues to be under development for the models prototype database/classified catalog. The database is scheduled for completion by the end of the year 2002. There have been several other well-known efforts to catalog materials for learning, including the development of a content standard for computational models.[23] However, many of these efforts fail to consider scientific models as works, an instantiation with many entities and relationships and hence are not useful attempts to improve information retrieval beyond identification and location. This study is trying to do more; it attempts to map structures for scientific models using the bibliographic definition of works.

Conclusion:

Heckhausen, in the 1970s, discussed the growth of computer models and modeling as the analytical tools of a discipline.[24] However, models are products of scientific research, the creative artifacts reflecting the intellectual structures (mapping of relationships between objects, process, through bibliographic entities) of the discipline and the researcher. The activity of modeling conceals an incredible amount of intellectual relationships that traditional bibliographical tools (primarily the catalog and the index) neither capture nor describe from the texts, documents, and items about models. Should our tools do so? Yes. Our increasing awareness of conceptual and textual instability of electronic forms requires active investigation and experimentation with other types of knowledge organization and representation structures. Additionally, decades of research both in information retrieval and information seeking behavior complemented by the widespread success of Internet search engines has shown us that users tend to disregard Boolean searches, human indexing as opposed to machine indexing does not improve search performance significantly, and that users want a few relevant, good materials. Assessing relevance in terms of disciplinary structures has never been researched. Finally, access to published literature alone is insufficient and continued segregation of our tools, the catalog distinct from the index, is unproductive and outmoded as the goals of knowledge management merge symbiotically with information retrieval.

This paper has tried to show that scientific models may be described more meaningfully for information seeking purposes such as exploratory learning if they are considered as works. Dublin Core was used as the content standard for the examples of models cataloged as works in the paper and aspects of models and significant relationships continue to be investigated. Appendix D provides a brief graphical illustration of the status of scientific models as works and the research questions that should be empirically investigated.

Acknowledgements:

This paper benefited greatly from the comments of three anonymous reviewers who approved the funding for this models classification study. I also thank my former student, Olha Buchel, for many meaningful discussions on the cataloging of models for information retrieval that supports learning tasks and for her support in database and catalog design. Thanks to Professor Muniram Budhu, Dept. of Civil Engineering, Dr. William Rasmussen, Dept. of Agriculture and Hydrology, and Dr. John Kemeny, Dept. of Rock and Mining Engineering for their enthusiastic participation and support of the informal survey of models in the engineering disciplines. Thanks to Professor Smiraglia for writing 'Works' and all helpful feedback. They have helped to focus this article. This study is supported by the University of Arizona, Office of the Vice President for Research Faculty Small Grant, 2001-2002.

Glossary

Note: Many of the terms in this glossary are from Smiraglia (see Notes). Exceptions are noted. Definitions by the author are indicated by an asterisk.

Bibliographic Entity: A bibliographic entity is a unique instance of recorded knowledge (e.g., a dissertation, a novel, a symphony, etc.). Each bibliographic entity has two properties – physical and conceptual. A containing relationship exists between the two properties. All recorded and published representations of scientific models are bibliographic entities (including datasets, observational data, instrument and model-generated data).

Bibliographic Family: The set of all works that are derived from a common progenitor.

Bibliographic Relationships: A bibliographic relationship is an association between two or more bibliographic items or works. Four types of relationships are possible in the library catalog, bibliographic, access point, name, and subject relationships.

Source: Tillett (see Notes)
See also Tillett Taxonomy of Bibliographic Relationships, Smiraglia Taxonomy of Derivative Bibliographic Relationships.

Derivative relationships: Bibliographic families usually are created by derivative relationships: one source is the progenitor. See Smiraglia Taxonomy of Derivative Relationships.

Form: Form subdivisions indicate what an item **is** rather than what it is **about**. Items can be assigned form subdivisions because of:

* Their physical character (e.g., photographs, maps)
* The particular type of data that they contain (e.g., bibliography, statistics)
* The arrangement of information within them (e.g., diaries, indexes)
* Their style, technique, purpose, or intended audience (e.g., romances, popular works)
* A combination of the above (e.g., scores)

Source: "LCSH: Subject Cataloging Manual" [CD-ROM], *The Cataloger's Desktop* (Washington D.C.: CDS, 2001).

Item: The physical property, container which is the package for the intellectual part of the bibliographic entity.

Progenitor: A progenitor is the first instantiation of a work - the source - that has an original idea. This starts the propagation of other texts, items, and documents.

***Scientific model:** A class of models with mathematical and computational properties, which can be considered to be works. Like works scientific models have semantic content and ideational content. The following types of relationships exist in scientific models: 1) bibliographic (example, parent-child), 3) access point (not discussed in this paper) 3) name relationships (limited to model personal authors and informal/formal organizations and groups only) 4) subject relationships (to be determined for facets such as concept, object, process, phenomenon, discipline, mathematical representation, software, purpose, and coverage). The sum of these subject relationships is referred to as ideational relationships in works or as scientific model relationships and can be expressed through controlled value lists and classification schemes for improved information retreival.

Smiraglia Taxonomy of Derivative Relationships: Derivative relationships are simultaneous, successive, translations, amplifications, extractions, adaptations, and performances.

See Smiraglia (Notes) for complete definitions

Text: A text is the set of words that constitute writing. Includes textual works.

Tillett Taxonomy of Bibliographic Relationships: Bibliographic relationships include derivative, equivalence, descriptive, whole-part, accompanying, sequential, and shared characteristic relationships.

Services: Interactive and reference services, electronic.

Work: A work is an abstract entity; there is no single material object one can point to as the work. We recognize the work through individual realizations or expressions of the work, but the work itself exists only in the commonality of content between and among various expressions of the work. When we speak of Homer's Illiad as a work, our point of reference is not a particular recitation or text of the work, but the intellectual creation that lies behind all the various expressions of the work.

Source: International Federation of Library Associations, Study Group on the Functional Requirements for Bibliographic Records, 1998, *Functional Requirements for Bibliographic Records* (Muchen: K.G. Saur, 1998), 16-17.

Appendix A: Subject heading for Models in LCSH

Models and modelmaking (May Subd Geog)
[TT154-TT154.5 (Handicraft)]
UF Model-making
    Modelmaking
BT Handicraft
    Manual training
    Miniature objects
RT Modelmaking industry
    Simulation methods
SA subdivision Models under types of objects, e.g. Automobiles--Models; Machinery--Models; and phrase headings for types of models, e.g. Wind tunnel models
NT Architectural models
    Atmospheric models
    Engineering models
    Geological modeling
    Geometrical models
    Historical models
    Hydraulic models
    Hydrologic models
    Mannequins (Figures)
    Matchstick models
    Miniature craft
    Models (Patents)
    Patternmaking
    Relief models
    Ship models
    Surfaces, Models of
    Wind tunnel models
    Zoological models
--Motors
--Radio control systems (May Subd Geog)
[TT154.5]
BT Citizens band radio
        Radio control
        Models and modelmaking in literature
Models in art
    USE Artists and models in art

Models of surfaces
    USE Surfaces, Models of

Appendix B: Qualitative Analyses: Excerpts (classification of models in water quality and geography)

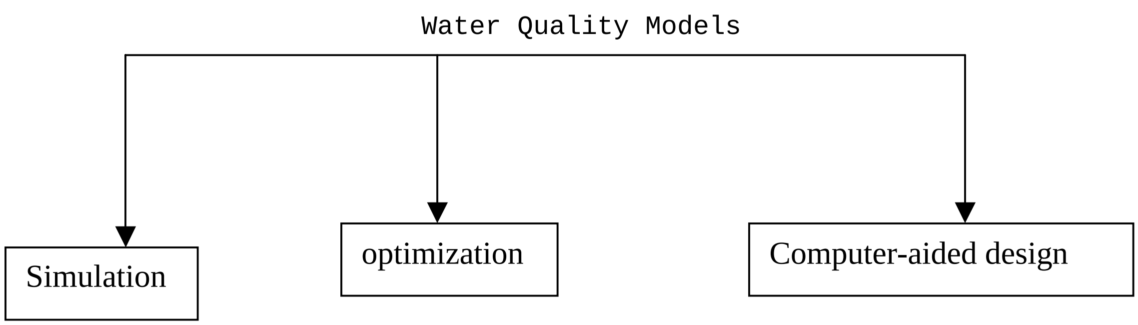
Five books in the area of water quality, three in the area of physical and human geography, and two thesauri (GEOREF and Water Abstracts) were analyzed to see how models are represented and what vocabulary/classification schemes exists for scientific models in a specific area like water quality and a broad discipline like geography. Excerpts are presented from the analysis of the books only.

The entire text of the book including tables of contents, preface, selected chapters, and the indexes of the five books were scanned for classification, names of models, objects, processes, phenomenon, and laws (only classification and named models are shown here). Here are the excerpts:

Text #1: An Introduction to Water Quality Modelling. Edited by A. James (Chichester, John Wiley, 1984.

Audience for book: Beginner. Based on courses in Civil Engineering Department introducing water quality modeling to scientists and engineers in water pollution control.

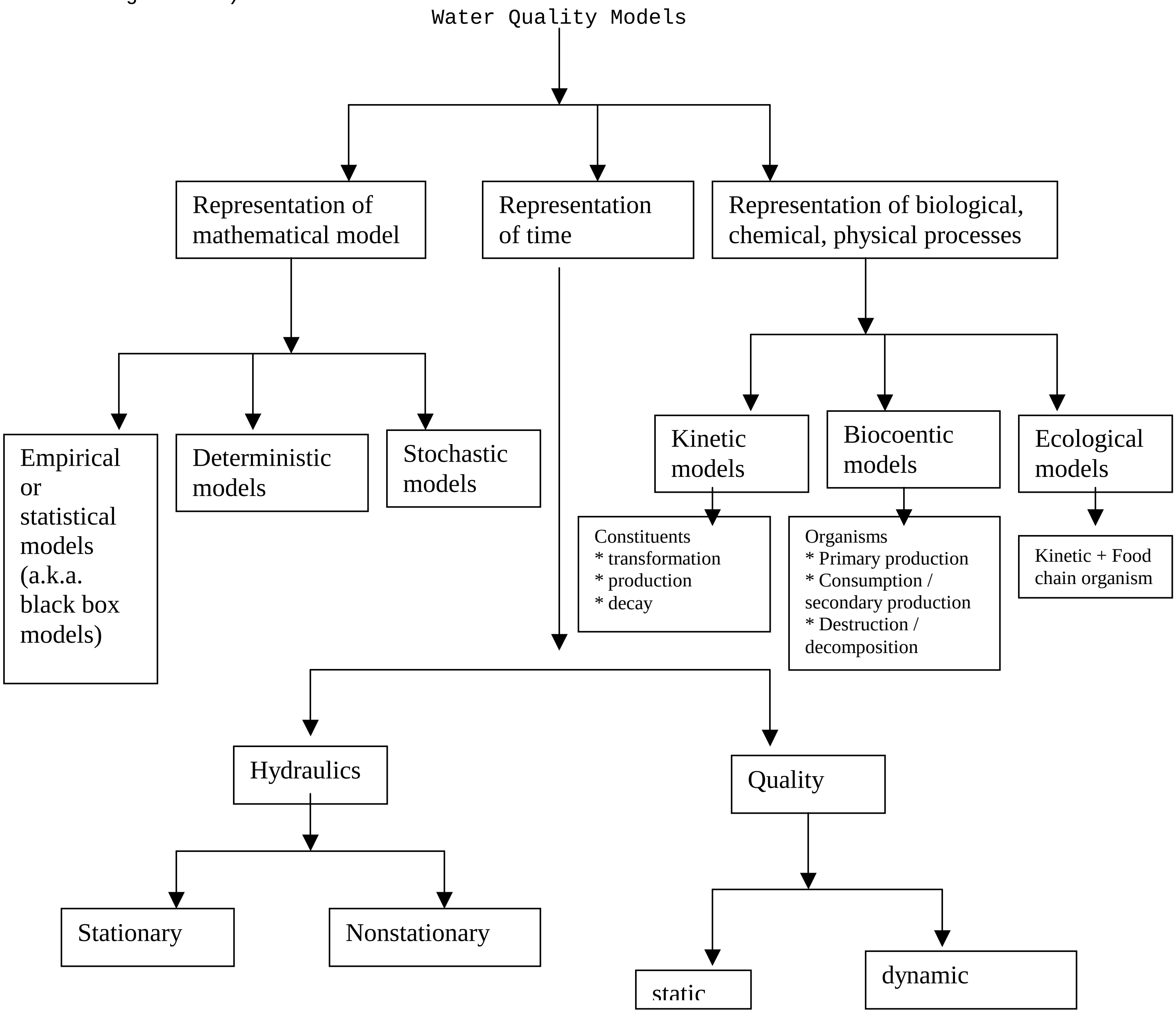
Classification Notes: Provides a classification of water models (see Fig. Below)



Models of Water Quality on Rivers:

1. BOD/DO models by Streeter and Phelps in the 1920s
2. Fick’s Laws of Diffusion
3. River models
4. Lagrangian models
5. Moving segment models
6. Models for discharge
7. Dispersion model
8. Block models

Text #2: Water Resources: Environmental Planning, Management and Development (New York: McGraw-Hill, 1996).
Analysis: Index, Table of Contents, Preface, Chapter 7
Classification Notes: Provides a classification of water models (see fig. below)



Named Models:
1. Advection-dispersion equation (Taylor, 1921 – classic cite)
2. Mass-balance modeling
3. Plug flow reactor models (Streeter & Phelps, 1925)
4. BOD/DO (Streeter & Phelps, 1925)
5. Oxygen-balance model (Streeter & Phelps, 1925)
6. Continuously stirred tank reactor (CSTR) modeling (Vollenweider, 1975)
7. Aggregate dead zone (ADZ) model
8. QUAL models (current version, Brown and Barnwell, 1987)

9. WQRRSQ models (Smith, 1986)
10. WASP (Ambrose et al, 1988)
11. MIKE 11 (Danish Hydraulics Institute)
12. Flow model
13. DRAINMOD
14. land-subsidence model

Text #3: Ajit K. Biswas, Editor, Models for Water Quality Management (McGraw-Hill, 1981).

Classification Notes: States that most of the Water Quality Models in use today (1981) are extensions of two simple equations by Streeter and Phelps (1925) – BOD and DO. Other modelers are: O'Connor, Thomann, DiToro, and Chen and Orlob.

Named models:
1. BOD/DO
2. QUAL 11

Geography – Qualitative Summary:

Minshull (1967) summarizes classification of models by geography scholars and offers a new classification:

Type I. Submodels of structure
Iconic or scale
Analogue
Symbolic
Type II. Submodels of function
Mathematical
Hardware
Natural
all three above can be used for a) simulation, and 2) experiment
Type III. Submodels of explanation or theoretical conceptual models (the causal factors of systems in the earth's surface)
Cause and effect models
Temporal models (including process, narrative, models of time or stage, models of historical processes), and
Functional models

Minshull further proposes new labels for describing model:

1. The nature of the model
hardware, symbolic, graphic, cartographic,
2. Functions of the model
descriptive, normative, idealistic, experimental, tool, procedure
3. Form of the model
static or dynamic
4. Operational purpose of the model
to store data
to classify data
to experiment on the data
5. Stage at which model is used:
a priori
concurrent
a posteriori (theory is proposed and verified before the model is made)

Content standard for computational models (CSCM) proposed by Hill et al. D-Lib Magazine, 2001. http://www.dlib.org/

CSCM is a descriptive standard. It identifies 165 elements in ten sections for computational (scientific models) in the environmental sciences. The ten sections are:

1. Identification Information
2. Intended Use
3. Description
4. Access or Availability
5. System requirements,
6. Input data requirements,
7. Data processing,
8. Model output,
9. Calibration efforts and validation,
10. Metadata source

APPENDIX C: Notes about sampling frames

Most investigations into the nature of works have established the library catalog and the bibliographic utility as the natural sampling frames. However, bibliographic databases (periodical indexes) are also a natural sampling frame, since the definition of a work imposes no limitation upon texts, manifestations, or other forms of derivative and other bibliographic entities that emanate from the progenitor. This appendix provides a brief description of the bibliographic databases (indexes) and their controlled vocabularies that were used in this study. The bibliographic utility OCLC is also briefly described.

Rationale for including indexes (bibliographic databases): Most information in the sciences gets outdated quickly. Articles, therefore, are timely in terms of providing information about models. Additionally co-citation has proved to be a valuable analysis in identifying disciplinary intellectual structures.

OCLC: OCLC was the bibliographic utility chosen. I tried searching in both CatExpress (the cataloging service) and WorldCat (the reference service) and found both equally frustrating to use. The distinction between subject words, subject phrase, and subject, searches were blurred in many searches.

INSPEC: While OCLC was chosen as the bibliographic utility for this investigation into the nature of scientific models as works, I did not limit myself to a single indexing database, though not all data has been included in this paper. I searched GEOREF, Water Abstracts, EI Compendex, and ISI indexes (note that these findings are partially reported here). INSPEC is a database that has over 5 million records and approximately 300,000 records are added annually. Coverage includes physics, electrical engineering, electronic, telecommunications, computers, control technology, and information technology.

The INSPEC thesaurus has broad subject terms under which constitutive models may be classed and indexed. They are:

1. A 4630 – Mechanics of solids
2. A4660 – Rheology of fluids and pastes
3. A5100 Kinetic and transport theory of fluids; physical properties of gases
4. A6200 Mechanical and acoustic properties of condensed matter
5. A 8140 Treatment of materials and its effects on microstructures and properties

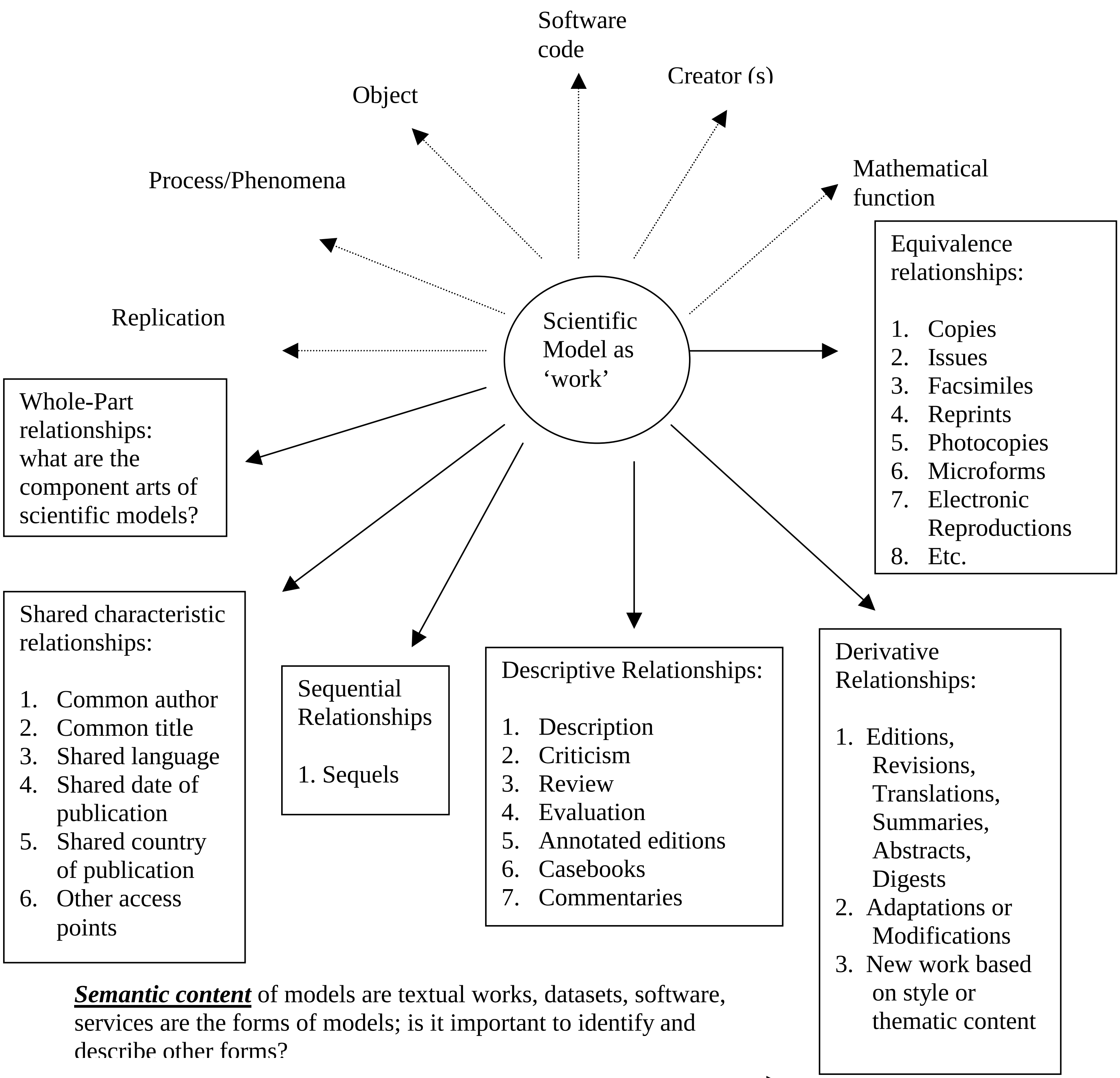
Once the user has browsed the thesaurus under these broad terms, she can search using more specific terms such as Viscoelasticity or an even more broader term such as mechanical properties. Examples of other narrower, specific terms are: anelasticity, creeps, cracks, deformation, stress-strain relations. Many of these terms are very good, and we consider that INSPEC had the best indexing, in terms of direct, specific entries, but it still does not identify constitutive models as a class of scientific models or as one category of works with many subject relationships, nor does it show specific bibliographic family relationships. These are critical aspects for successful information retrieval about scientific models. Additionally, these terms were established in 1995 and the field continues to change and evolve rapidly.

SABIO: This is the online public access catalog of the University of Arizona Libraries.

Appendix D: Scientific Models as Works and Future Research Questions

Model facets are indicated by arrows with dashed lines; bibliographic relationships that representations of models might possibly possess are indicated by solid arrows. Significant questions remain to be investigated. For example, if scientific models are identified as works, which bibliographic relationships do representations about model replications fit best? Can relationships in electronic resources be determined semi-automatically? What are important, generic name and subject relationships that are present in all scientific models? How can they be used to display and sort through results in library catalogs?

Ideational content of models includes identifying creator name and subject relationships. What are these?



Semantic content of models are textual works, datasets, software, services are the forms of models; is it important to identify and describe other forms?

Notes

---

[1] Patrick Wilson, "The Catalog as Access Mechanism: Background and Concepts," *Library Resources and Technical Services* 27 (1983): 7.

[2] *Anglo-American Cataloging Rules*: *2nd Edition, 1998 Revision with 1999 Amendments* [Electronic edition], (Washington D.C. Library of Congress, 2001), Part I.

[3] Op. cit. Chapter 13.

[4] Patrick Wilson, "On Accepting the ASIST Award of Merit," *Bulletin of the American Society for Information Science & Technology*28 (2002): 11.

[5] Richard P. Smiraglia, *The Nature of "A Work": Implications for the Organization of Knowledge* (Lanham, Maryland: Scarecrow, 2001).

[6] "Library of Congress Subject Headings," *The Cataloger's Desktop* [CD-ROM], (Washington D.C.: CDS, 2001).

[7] Cyril.W. Cleverdon, "The Significance of the Cranfield Tests on Index Languages," *Proceedings of the Fourteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrrieval, September 1991, Chicago, Illinois* (Baltimore, Maryland: ACM Press, 1991).

[8] *Anglo-American Cataloging Rules*: *2nd Edition, 1998 Revision with 1999 Amendments* [Electronic edition], (Washington D.C. Library of Congress, 2001) Glossary.

[9] "model" *Oxford English Dictionary*, Ed. J. A. Simpson and E. S. C. Weiner. 2nd ed. (Oxford: Clarendon Press, 1989), *OED Online*, (Oxford: Oxford University Press, 4 Apr. 2000), http://oed.com/cgi/entry/ 00149259 [Accessed 31 January, 2002].

[10] "model" *Encyclopedia Britannica Online*, http://search.eb.com/bol/search?type=topic&query=model&DBase=Articles&x=30&y=3 [Accessed 31 January, 2002].

[11] "operations research" *Encyclopedia Britannica Online*, http://search.eb.com/bol/topic?eu=109277&sctn=7, [Accessed 31 January 2002].

[12] "model" *Academic Press Dictionary of Science and Technology Online*, (San Diego, Calif.: Academic Press, 2001), http://www.academicpress.com/dictionary/ [Accessed 31 January, 2002].

[13] Hans von Storch. Models in Environmental Research. (Berlin: Springer, 2001), 17.

[14] *Earth System Science: A Closer View,* (Washington D.C.: NASA, 1988).

[15] *Thinkertools: The ThinkerTools Scientific Inquiry and Modeling Project* [online], University of California at Berkeley, http://thinkertools.soe.berkely.edu/ [Accessed 31 January 2002].

[16] *Modeling for Understanding in Science Education* [online], University of Wisconsin-Madison, http://www.weer.wisc.edu/ncisla/muse/index.html, [Accessed 31 January 2002].

[17] Op. cit. R.P. Smiraglia, 81.

[18] Jennifer Cartier, et al, "The Nature and Structure of Scientific Models" [online], National Center for Improving Student Learning and Achievement in Mathematics and Science (NCISLA), University of Wisconsin-Madison, http://www.wcer.wisc.edu/ncisla/ [Accessed 31 January, 2002].

[19] R.P. Smiraglia, op cit. 75.

[20] Ibid. 42

[21] Barbara A. Tillett, " A Taxonomy of Bibliographic Relationships," LRTS 35 (1991), 150-158.

[22] *Dublin Core Metadata Element Set, Version 1.1.: Reference Description* [online], http:dublincore.org/dces/ [Accessed 31 January 2002].

[23] Linda H. Hill et al. "A Content Standard for Computational Models," *D-Lib Magazine* 7 (2001), http://www.dlib.org/ [Accessed 31 January 2002].

[24] Heinz Heckhausen, "Disciplines and Interdisciplinarity," *Interdisciplinarity: Problems of Teaching and Research in Universities* (Paris: OECD, 1972), 85.