The FRBR-Theoretic Library: The Role of Conceptual Data Modeling in Cultural Heritage Information System Design

Ronald J. Murray

Library of Congress
101 Independence Avenue SE
Washington DC 20012
rmur@loc.gov

Abstract

The use of digital technologies in support of Cultural Heritage missions has highlighted the need to create information modeling systems different from those that are used in conventional business and government. In addition, the practice of data modeling – and especially of the conceptual data modeling that engages cataloging theory and practice - must be urgently be brought up to date in order to develop the data models required to represent the desirable characteristics of both print and digital media.

Introduction

The use of digital technologies in support of Cultural Heritage (CH) missions has highlighted the need for information management systems different from those that are used in conventional business, government, and entertainment activities. In particular, a Cultural Heritage institution needs an information system that (a.) supports preservation of and access to both analog and digital content, and (b.) reflects that institution's customary understanding – its view – of the resources it possesses.

The Cultural Heritage community has evolved a number of well-established approaches to the description of resources created in a wide range of media. The hope within that community is that the long-standing theories, practices, and policies that inform the operations of libraries, archives, and museums and provide structure to analog as well as to some digital versions of actual Cultural Heritage content will extend even further to the vast quantities of resources available on the World Wide Web.

The Design of Information Systems Based on Cultural Heritage Concepts

Web-based resources require embodiment, organization, discovery, and access by an electronic information system. The design, implementation, and operation of globally-accessible Cultural Heritage inventory and discovery systems has benefited from collaborative efforts at standardization, with international information technology standards bodies playing a critical role in this effort. However, the crucial data and process modeling steps that lead to the creation of those systems have not received the same level of international attention.

The FRBR Conceptual Model as Exemplar

Since its introduction of the basic concepts underlying the Functional Requirements for Bibliographic Records (FRBR) conceptual model (IFLA 1998), the application of Entity-Relationship Modeling has achieved general acceptance by cataloging theorists. In addition to supporting cataloging theory formation, FRBR was intended to function as a guide for the description of bibliographic materials within and beyond the confines of a library.

However, the literature that details the intervening twenty-one year effort to come to terms theoretically with FRBR (and adapting the model to different media types or to archival records) suggests that deficiencies in or incompatibilities exist with the existing model. These disagreements with the FRBR conceptual model may simply result from problems with *data model quality*: (1.) the current conceptual data model lacks refinement; (2.) the data model reflects an individual modeling style that does not suit the task at hand, and (3.) the FRBR conceptual model's entity, attribute, and relationship definitions reflect mixed or inappropriate data modeling assumptions.

Other explanations for these disagreements are possible. For example, the third data model quality problem above may actually reflect what recent research would identify as an consequence of the complementary stances a data modeler can take relative to the bibliographic "Universe of Discourse" being modeled. More seriously, objections to FRBR may indicate that due to the complexity of the bibliographic universe (and to the numerous ways that interested parties seek to interact with it), there can be no single conceptual data model that will encompass all of the well-established perspectives evolved by archives, libraries, and museums.

The widely discussed and institutionally accepted FRBR conceptual data model can be taken as an indicator of the extent to which the Cultural Heritage community has adopted and the utilized data modeling methodologies that have evolved for purposes of information system design and implementation.

Intent of the Paper

This paper will explore the role that modern data modeling theory and practice has (or has not) played in the development of the FRBR conceptual data model. It also offers examples of modern data modeling approaches demonstrated with Cultural Heritage subject matter. The analysis is intended to provide guidance to parties attempting to further refine the FRBR conceptual model for theoretical purposes as well as for information system design.

Data Modeling Defined

Data modeling is the step in a database management system design process where things of interest to the enterprise are defined and their relationships delineated. Academic theory and professional educational materials describe data modeling as an interactive process that produces a textual and a diagrammatic representation of an enterprise's information at several levels of abstraction.

Data modeling begins with a review of information system requirements, continues with document reviews and user interviews and model building (in diagram and textual form) with feedback from users. The model may be subject to adjustment to improve performance and is then implemented in a specific implementation technology. Three key data model definitions apply (Hay 2006):

Conceptual Data Model – A description of a portion of an enterprise in terms of the fundamental things of interest to it. They are fundamental in that most things seen by business owners are examples of these.

Logical Data Model – The organization of data for use with a particular data management technology. For relational databases, these are tables and columns; for object-oriented databases, object classes and attributes.

Physical Data Model – The organization of data used to place it in specific storage media. This level refers to "tablespaces" and "cylinders."

Why a Data Model is Important

Because no database is ever built without a model, the question really becomes whether to model informally or formally, who will be involved, and how much effort will be spent in creating a good design. Data models possess three characteristics that make them essential to system design and implementation:

Leverage – As the data model provides a roadmap for the increasingly technical and implementation-specific representations, programming, etc. that follow, small changes in the data model can have major effects on the system being designed and implemented. A well-designed data model can minimize the need for model changes due to missed requirements and thereby reduce design implementation costs. If the things of interest to the organization are poorly modeled, the database implemented from the model will require more programming effort to effort to input and retrieve data.

Conciseness – Data models provide a compact specification of an information system's requirements and capabilities. Reviewing a data model takes less time than reading a lengthy functional specification document, and makes it easier to obtain an in-depth understanding of the kinds of information that are to be managed.

Data Quality – Problems with data quality (inaccurate data) can often be traced to inconsistency in defining and interpreting data and in implementing enforcement mechanisms for data definitions. Well-defined data model definitions (and enforcement mechanisms) of dates, addresses, and names preserve the common understanding of what is being recorded and minimizes the need for corrections or workarounds.

What Makes a Good Data Model?

Given that a *designed* data model (as opposed to a faithful description of what is "out there") is evaluated in terms of how well it meets requirements, a data model quality criterion of must apply. In the absence of a quantitative methodology, Simsion & Witt's criteria are helpful (Simsion & Witt 2005):

Completeness –Does the model support or can it generate the data as specified or implied by the requirements documentation?

Nonredundancy – Does the model preclude the possibility of storing the same fact in more than one place? At the conceptual modeling level, entities that contain the same data would indicate that the model is incomplete and that the model can benefit from the addition of a supertype, where the redundant data can find a home.

Enforcement of Business Rules – How well does the model embody and enforce the rules for handling the data? If data model elements do not allow for specification of all of the conditions that a business imposes on its data, business rules must be elicited and used to further document the model.

Data Reusability – If ways for usefully processing the data are discovered after the model is implemented, is the model flexible enough to permit this without modifying the database? Designing for data independence is very important because data that is organized around a particular application will be harder to adapt when the application changes or is replaced.

Stability and Flexibility – A data model is stable with respect to requirements if a change in requirements does not require changes in the data model. The model is flexible if it can be extended without difficulty to accommodate extensions to existing requirements. Depending on the application environment (e.g. where new media forms are being created, or where cataloging information is being acquired or updated continuously) taking the extra effort to design for stability and flexibility can pay off in reduced data model modification and reduced impact on other implementation levels.

Elegance – Elegance evokes the mathematical sense of the term, where consistency and relative simplicity in describing a model element can be discerned. Elegance in entity definition can be achieved by generalization, for example, when pragmatically compelling entities such as customer, employee, supervisor, security guard, supplier, etc., are generalized into a Party entity that represents these entities as subtypes within a logical, and often hierarchical – structure.

Communication – The ability of the data model to convey its content to technical and non-technical personnel is crucial to determining (a.) whether the model is an accurate representation, and (b.) whether the people who will use or manage the implemented system understand the full implications of the model. Unfamiliar terminology, new concepts, and high levels of complexity tend to render the model less comprehensible to its audience, so the modeler must organize and present the model with an eye towards maximizing its communicative potential.

Integration – How well does the complete system or system components fit in with what is already there? The ease of difficulty of fit may vary not only with the skill of the designer but also with the novelty of the requirements. For example, library catalog databases that never had to contend with online resources that change on a daily basis may face greater integration challenges than databases where incremental change in resource characteristics is common.

Towards Theory-Guided Design –

A problematic aspect of the data modeling process is that modeling efforts can be undertaken unaware of the *description*/*design* issue that underlies conceptual data modeling theory and practice. Is the data modeler describing things that are "out there" or is the modeler creating useful data structures that meet specifications? If the Universe of Discourse that is to be represented in the database being modeled contains its *own* unresolved description/design issues (as does the Bibliographic Universe), the result will be a data model where theoretically (or institutionally) compelling model elements become intermixed with elements designed to be useful to programmers and end users. The solution may satisfy the stated requirements rather well, but will please no one.

A good example of an institutionally compelled descriptive or design element (from cataloging theory as well as library tradition) is a hierarchical data structure, which some assert is "the most philosophically interesting of the semantic relationships." (Svenonius 2001) Notable counterexamples to Svenonius' interestingness assertion are the network structures that are regularly used to represent a wide range of current theoretical and pragmatic "things of interest" to Communications Theorists, Physicists, Information Scientists, and Political Scientists. (Monge and Contractor 2003; Watts 2003); Csermely 2006) Networks (i.e., graphs, the mathematical structures first described by Euler in 1735) are rarely mentioned in on cataloging theory, nor have they been invoked to describe or define data structures in FRBR.

Consider taking a database *design* perspective to another environment – in the person of a computer programmer at a Physics laboratory where decentralized teams build subatomic particle detectors and conduct research. Network-like structures would be a natural – even unavoidable – part of the intellectual landscape, beginning with a powerful diagrammatic shorthand for describing or hypothesizing particle interactions:[1]

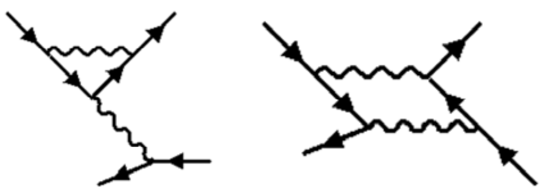

Figure 1: Feynman Diagrams of Electromagnetic Interactions

[1] Feynman Diagrams from http://www2.slac.stanford.edu/vvc/theory/feynman.html

The representation of relationships among information resources in that environment might take on an institutionally-compelled network flavor (Berners-Lee and Fischetti 1999), rather than the hierarchical one reinforced by cataloging theory and library institution administrative organization.

A *theory-guided design* solution would be one that evaluates theoretically compelling entities, attributes, and relationships from other fields of endeavor in addition to those originating from cataloging theory. The parties participating in the database management system design process would not be *compelled* to accept these elements, however, just because they have theoretical utility. The design process would benefit from a data model element review that engages a broadened theoretical base to include Social Sciences perspectives – in particular Anthropology, Psychology, Sociology, and Communication Theory.

Modelers initiating a *theory-guided design* strategy can also profit from recently published field research that indicates that professional data modelers depart in significant ways from espoused academic and professional educational teachings. modeling pathways. (Simsion 2007). Data modelers who possessed differing degrees of experience and training were surveyed and also participated in model design experiments.

The respondents were evenly split on the Description vs. Design issue, in spite of the topic never being discussed in the literature. In addition to differences in modeling stance, the data models produced by participants in the data model design experiments demonstrated an effect of experience and personal style on: the number and subtyping of model elements created (fewer used with less experience); the addition of elements and relationships not included in the requirements (more likely with experience and with the *design* stance); and the use of patterns from past modeling projects (more likely with age and the *design* stance).

A Critique of the Current FRBR Model –

FRBR from a modern data modeling perspective – The FRBR conceptual data model as advanced by IFLA raises a number of issues that may be grouped into four categories:

Modeling from legacy systems – Special attention must be paid to the consequences of developing a FRBR conceptual data model that borrows from or must otherwise be made to reflect the structure of legacy logical data models. The danger is that an implementation-specific feature (like a limit set on the number, attribute names or optionality of parties that play ***roles*** like Author, Editor, Publisher, etc.) will become a requirement to be met in the conceptual data model.

Accommodating legacy systems (in the sense of identifying the functions that were executed by the systems, and understanding the structure of the data in the system) can be made a requirement of an information system. But the design of the new system should not require that identical functions and data structures be created to accomplish this. In the literature, discussions of FRBR model characteristics using patron-oriented legacy displays and scenarios set up by researchers to test hypotheses – have been illustrated using catalog card-like or MARC record-type displays. (Taylor 2007) This indicates that legacy information structures and data presentation strategies continue to dominate designers' and researchers' thinking, irrespective of the changes in database design that should be replacing these legacy structures.

Efforts to represent, reason about, and display FRBR bibliographic data should focus more on the data *as it is understood within the new conceptual data model* rather than that of legacy systems. The fact that FRBR conceptual data models contain Many-To-Many network structure means that data modeling efforts should begin with generic network structures for the database design and display, and apply constraints to achieve legacy system hierarchical appearances where unavoidable.

Element use and skill level – The relatively small number of elements and lack of subtyping in the FRBR data model supports Simsion's finding pertaining to the products of beginning or infrequent modelers. In addition, the model reveals a Cultural Heritage data model documentation bias:

Reliance on the narrative/textual part of the model – By far, the substance of the IFLA FRBR conceptual model specification is textual description (with tables), and only a few diagrams. While these diagrams play a very small role in model documentation and presentation, they are what is used – naturally – to describe the model to the Cultural Heritage community and to the general public. It is difficult to appreciate the overall, emergent, characteristics of the FRBR conceptual data model – especially the more obvious interactions between model elements – from a reading of the text and then attempting to project that wealth of description into the few available diagrams. Especially interesting, but not modeled explicitly, is the means by which the very many neatly subtyped bibliographic relationships defined in the FRBR model text are represented – as attributes, relationships, and possibly even entities.

Simsions' research revealed that data modeling practitioners – like designers in other fields like engineering, architecture, graphic arts etc., – use diagrams: for their own benefit (contextual placement of model elements with the ability rapidly to modify the model in the face of user feedback, and to detect recurring or out of place patterns); as well to benefit clients (communication of overall model structure and its critical elements).

Missing elements that would improve model communication – Model elements that would make it easier to understand FRBR's benefits for bibliographic resource discovery are not provided. Also lacking is a distinction between a conceptual data model that is presentable to users (i.e, it is community-specific) vs. one more that is more expressive and accurate for the data modeler and the developers to follow (i.e., it employs data model design conventions and patterns).

Contextualization & Coexistence – The FRBR model at present does not situate its conceptual data model elements within what must be a larger environment of bibliographic and other information resources. The characteristics of information resources of various types, their descriptions, and the roles that institutions can/should play in creating and managing resource descriptions are therefore not addressed in the model. In a more contextualized model, FRBR and related IFLA data modeling products like name and subject authorities and identifiers – appropriately generalized – have play highly valuable roles to play in the bibliographic universe. These entities benefit from being modeled from a broader perspective.

In the broad context of *Resources*, where *Resource Descriptions* are created to describe the *Resources* that users want to discover and use, it must be stated whether FRBR-based resource descriptions can coexist with other descriptions produced by other institutions or individuals. This issue is not addressed in the present decontextualized model. Description coexistence has significant implications for the placement of resource identifiers, names, and some responsible party roles in the more broadly defined model.

FRBR is not a "Convergent" Conceptual data Model –
A *divergent* conceptual data model is one where entity names, quantities and relationships come from their specific user communities. Similarities in entities and relationships across different data models become difficult to see, and common conventions in data modeling practice are not present (like generalizing entities, standard entity and relationship names, and using patterns). Divergent conceptual data models are very useful however in that they capture a enterprise view that can readily be validated by users.

A *convergent* conceptual data model results when conventions in entity and relationship construction and naming are applied to the divergent model. This step may require the creation of additional elements and relationships based on the modeler's experience with the structures in the divergent model. (Hay 2005)

A very good indicator that FRBR is a divergent model is the use of community terminology for entity names. Even so, the use of the nondescript prefix "Group" to describe entities should be replaced by meaningful names given these groupings by users. Encouraging users to name data model entities and entity groupings/subtypes is a simple way to induce them to pay more attention to the conceptual data model.

Having commented on FRBR in terms of legacy system issues, element use and institutional preferences for data structures, and on model communication, we can now consider how data modeling can reconcile the desire to build systems that embody well-established intellectually, compelling cataloging concepts with the urge to create data structures that may lack theoretical resonance but get the job done. We propose that the modeling undertaken be guided by – but not be captive to – theory.

Improving on the FRBR Conceptual Data Model

To demonstrate the impact that modern data modeling techniques and conventions can have on increasing the understanding of a CH data modeling effort, the current FRBR conceptual data model – in the form of its diagrammatic representation – will be recast into a different form consistent with modern data modeling practice.

Figure 2 presents a conceptual data models for bibliographic information, names and identifiers, and subjects, respectively. Figure 3 presents a descriptive scenario for a continuing resource. Space limitations prohibit describing all of the data model elements in the diagram in detail: only a few top-level entities will be described in.

The Larger Context – The entities and relationships defined in FRBR function as descriptions of **Resources**[2] – analog or digital – that are of interest to one or more persons. For that reason the FRBR entities in this revised conceptual model be defined as **Resource** types. Refer to Figure 1 to clarify the relative placement and connectivity of the entities and relationships to be defined.

Note especially how the compact data model diagram elements expand into a very lengthy set of Business Assertions and comments. Note also that the data model presented in the diagram lies between being a divergent data model and a convergent data model. Some elements are modeled in a conventional fashion, while others remain specific to a Cultural Heritage perspective. This was done deliberately to keep the model somewhat familiar on one hand, but also to highlight design issues on the other.

Design – Beginning with the most general kind of information entity in our Bibliographic Universe, a **Resource** is defined as an information-bearing asset that is drawn upon to accomplish some function.

Commentary – Note the optional (and defined following standard modeling practice) ***One-to-Many*** relationship at the bottom right of the **Resource** entity. This indicates that a **Resource** may be composed of other **Resources**, of the same or differing **Types**. Defining the ability to "nest" **Resources** at this most basic level makes it possible for the data model elements that are **Resource** subtypes to "inherit" (depending on the conditions we define – we may permit nesting or we may not) the ability to contain sub-**Works**, **sub-Expressions**, **sub-Manifestations**, and **sub-Items**. As we will see, this design decision at the **Resource** level, in combination with judicious Business Rules, resolves a number of issues raised regarding Part/Whole relationships in FRBR entity definition.

Design – We now introduce the four primary **Resource** subtypes:

- **Institutionally Managed Named Resource**
- **Institutionally Managed Named Resource Description**
- **Institutionally Managed Find & Navigate Named Resource Description**
- **Institutionally Managed Find & Navigate Named Resource Assignment**

The characteristics of these different **Resource** types and their relationships with one another will be touched upon briefly. These above **Resource** subtypes make possible sophisticated grouping and referencing FRBR and FRBR-related data model elements.

About the Resources – An **Institutionally Managed Named Resource** is the actual **Resource** that users want to access and use. To efficiently find this **Resource**, an easily accessed *description* of the **Resource** can be consulted. An **Institutionally Managed Named Resource Description** is the means by which users can employ to find/navigate to, identify, select, access and use **Resources** of interest. An **Institutionally Managed Find & Navigate Named Resource Description** is an institutionally managed collection of identifiers, **Resource** names, people, place, concept etc., names, and types of possible relationships between **Resources**. It helps users (thanks to a library catalog some other analog or digital finding aid) to find to the **Resources** they want.

Keeping Track of the Connections – Finally, an **Institutionally Managed Find & Navigate Named Resource Assignment** is a **Resource** that consists of all of the "links" that have been defined between the **Institutionally Managed Find & Navigate Named Resource Descriptions**. Keeping track of the links makes it possible to take shortcuts to **Resources**, and to identify relationships that are not obvious without link information.

Commentary – The FRBR model is not currently defined in **Resource** terms. This makes the relationship between FRBR entity attributes and relationships and its referents in libraries difficult to discern. This decontextualization also makes FRBR entities seem more like *descriptions* derived from theoretical considerations rather than data structures designed to be linked to the actual analog or digital data desired by a user. While the idea of a **Resource** is still rather abstract, the relationship between a **Resource**, a **Resource** description, and the materials on library shelves or server hard drives, and the entries on catalog cards or screen displays is somewhat easier to understand.

An **Institutionally Managed Find & Navigate Named Resource Description**, by its name and by relationships defined in the data model diagram, signals a dependent, helping role with respect to the other two main **Resource** subtypes. The name indicates the role that institutions like libraries, archives, and museums can and do play in making resources (inc. non-bibliographic) easier to access. These institutions standardize names, defining subject headings, identify the persons and organizations etc., that may be sought, and also define the many relationships that exist between all of them.

A **Institutionally Managed Find & Navigate Named Resource Description** is in fact intended to act as a shortcut between what a user knows about – or has on hand in the form of a **Institutionally Managed Named Resource Description** attached to a **Resource** – to the **Resources** elsewhere with the same or similar description(s).

Design – Continuing, a **Resource** may be of one or more Types: a **Named Resource** and an **Other Resource**. A **Named Resource** is a **Resource** that is distinguished by the presence of a minimum of three **Institutionally Managed Find & Navigate Named Resource Descriptions**: an **Identifying Authority Resource Description**, a **Responsible Party Resource Description**, and an **Other Relationship Resource Description**.

Defining a Business Rule – It is useful to define the conditions under which a design element can be used. This definition is called a Business Rule, and is considered part of the data model. For a **Resource** to be managed effectively, we will define a Business Rule stating that the **Resource** must possess (a.) one or more unique identifiers, an optional name, and (b.) may optionally be related to another **Resource** in one or more defined ways via an **Other Relationship Resource Description**. This

[2] In this section, entity names are capitalized and in boldface.

Resource subtype contains the minimum of information required to distinguish it from other **Resources** (via the identifier), and to direct a user from that **Resource** to other **Resources**.

Named Resources are the Basic Units of Discovery and Access – The addition of identifying and relationship information to a basic **Resource** redefines it as a **Named Resource**. A party (identified by its **Responsible Party Resource Description**) can declare itself responsible for a **Named Resource** and then play one or more of several defined roles (e.g., Author, Creator, Publisher, Owner, etc. The institution will decide which *limited* set of possibilities can apply to this subtype) with respect to the **Resource**. A **Resource** that meets this additional responsible party requirement is called a **Managed Named Resource**. The remainder of the data model introduces new types of descriptions that correspond to customary institutional views line those held by librarians, archivists, etc.

What to Do About the FRBR Data Model and Data Modeling in General

Bibliographic information system efforts that rely upon conceptual data modeling can benefit significantly from an infusion of modern conceptual data modeling knowledge, abilities, and skills.

FRBR efforts need to be revisited, with an eye to ensuring that parties currently involved in model development (a.) appreciate the full implications of the original model, and of variations on the model such as the one presented here, and (b.) be prepared to change the model to reflect both improved model understanding and improved techniques for model construction and evaluation.

The talents of professional data modelers should be engaged to monitor data modeling activities taking place in Cultural Heritage institutions. Special effort should be made to have these parties to participate in community-initiated critiques of cataloging theory-based description/design data modeling methods.

A mutually acceptable institution should take leadership in advancing modern data modeling approaches like those introduced here by establishing a Web-accessible data modeling facility accessible to interested Cultural Heritage parties. This facility would (at a minimum) provide or promote training in conceptual data modeling, using well-accepted notations and documentation techniques. In addition the facility should endeavor to:

- Extend modeling activities to other needed areas in the Cultural Heritage realm.
- Design a professional education program and a collegiate curriculum.

This paper has presented a view of current theory and practice of conceptual data modeling, within the context of a unified model of database management system analysis and design. Particular attention was paid to how the FRBR conceptual data model has evolved, and how it differs from this and other alternative models. A conceptual data model that incorporates ongoing IFLA conceptual data modeling initiatives has also been presented and discussed, along with data model diagrams that address a wide range of content description scenarios.

Analysis of these data models supports a claim that the modeling approach used (*theory guided design*) can employ data modern modeling techniques, while at the same time incorporate the significant intellectual contributions of Cultural Heritage institutions in the realm of resource identification, description, discovery, selection, and access.

Acknowledgements

The author is grateful to Dr. Barbara Tillett for her editorial direction and her monitoring of the data modeling effort presented here.

References

IFLA 1998. International Federation of Library Associations and Institutions. Functional Requirements for Bibliographic Records. München : K . G. Saur München.

Simsion, Graeme. 2007. Data Modeling: Theory and Practice. Bradley Beach NJ: Technics Press.

Hay, David. 2006. Data Model Patterns: A Metadata Map. San Francisco: Morgan Kaufman.

Berners-Lee, Tim and Fischetti, Mark. 1999. Weaving the Web: the original design and ultimate destiny of the World Wide Web by its inventor. San Francisco : Harper San Francisco.

Simsion, Graeme C. and Witt, Graham C. 1994. Data Modeling Essentials. New York: Van Nostrand Reinhold.

Svenonius, Elaine. 2000. The intellectual foundations of information organization. Cambridge, MA: MIT Press.

Monge, Peter and Contractor, Noshir. 2003. Theories of Communications Networks. New York: Cambridge University Press.
Csermely, Peter. 2006. Weak Links: Stablizers of Complex Systems From Proteins to Social Networks. Heiderberg: Springer.

Taylor, Arlene. ed. 2007. Understanding FRBR: What It Is and How It Will Affect Our Retrieval Tools. Westport, CT: Libraries Unlimited.

Hay, David. 2005 Data Model Quality: Where Good Data Begins. The Data Administration Newsletter.
http://www.tdan.com/view-articles/5286