

CHAPTER 2

DEFINITIONS OF DESIGN AND DATA MODELING

And why beholdest thou the mote that is in thy brother's eye, but considerest not the beam that is in thine own eye?

– Matthew 7:3-5

This short chapter provides an initial review of the design / description issue by comparing published definitions of data modeling with generic definitions of design. As the opening quote implies, data modelers, who regularly extol the value of common definitions, have failed to establish agreed definitions for some of the key terms used in their own work, including *data model* and *data modeling*, and the differences appear to reflect different perceptions of the nature of modeling.

The final section establishes *description* as an appropriate antonym for design in the context of data modeling.

2.1. DESIGN IN INFORMATION SYSTEMS

Before turning to the generic literature on design, we draw attention to some ambiguity in the use of the word *design* in the context of information systems development. *Design* is frequently used to name or describe a broad stage in systems development whether or not it conforms to a formal – or indeed plain English – definition of design. Typically, but by no means universally (Larman 1998 p14), *design* is used to name the later, more technical, stages in systems development (Hirschheim, Klein et al. 1995; Hay 2003). In contrast, *analysis* is more often used to name the earlier, business-focused stages.

This use is perhaps a legacy of early information systems practice in which the goal was automation of manual processes, and effective use of the new technology was the major creative challenge: the “analysis” stage described existing processes (the “requirements” or “problem”); the “design” stage that followed created computerized systems to perform them (DeMarco 1978; Davis 1993).

This usage carries into the information systems sub-discipline of data modeling. Atkins (1997) provides some examples of the resulting inconsistency and

Chapter 2 – Definitions of design and data modeling

ambiguity, and cites the following textbook extract (Hawryszkiewicz 1997 p182) as a compact illustration of different uses of *design* (and indeed of *analysis*) in the context of data modeling:

“Data modeling is part of the development process. In the linear development cycle, it is used during the system requirements phase to construct the data component of the analysis model. This model represents the major data objects and the relationships between them. It should not be confused with data analysis, which takes place in the systems design phase. System design organizes data into good shape. Usually this means removing redundancies, a process often called normalization... Most designers develop only the high-level conceptual model in the system specification phase. The more detailed analysis using normalization is carried out during design.”

Even more compactly (and again in the context of data modeling): “The analysis phase includes the creation of a logical systems design” (Rob and Coronel 2002 p324).

Some authors do try to use the terms *analysis* and *design* to capture much the same distinction as we seek in using the terms *description* and *design*. Olle et al. (1991) characterize the difference as one of *description* (analysis) vs. *prescription* (design) – and classify data modeling as analysis. Veryard (1984 p1) states that “data analysis is a branch of systems analysis and therefore shares its principles. Of particular relevance are [sic] the separation of analysis from design...” His position on data modeling¹⁶ would appear to be implicit in the term *data analysis*. Witt (1997) divides data modelers into *analysis modelers* who see data modeling as determining, understanding, and communicating business information requirements, and *design modelers* who focus on producing a database blueprint. Here, the difference is portrayed as one of personal preference or style rather than as relating to different stages in the data modeling process or different contexts / purposes.

The different uses of *design* and resulting imprecision in meaning created two problems in reviewing the literature and designing the research described in this book:

¹⁶ Veryard’s later book *Information Modelling* (1992), as the title indicates, opts for the more neutral term *modelling*.

Chapter 2 – Definitions of design and data modeling

1. Publications on data modeling do not articulate a position on the description / design question as clearly as they would if the term *design* was employed with more precision. The word *design* in a task name, job role, or artifact cannot reliably be interpreted as meaning that it entails design in a strict sense. Similarly, use of the word *analysis* does not preclude it describing a design task, role, or artifact.
2. The description / design question itself may be misunderstood; *design* might be interpreted as locating data modeling in a particular stage of systems development rather than as describing the nature of the task. Preliminary discussions with data modeling researchers and feedback from practitioner-oriented articles (Simsion 1996; 2005c) that presented the question as a choice of *analysis* or *design*, were dogged by confusion of this kind. Substitution of the word *description* for *analysis* to clarify the dichotomy reduced the problem, at least to some extent.

While *design* and *analysis* may be used loosely in information systems work, we could expect them to carry some flavor of their original meanings, particularly for non-specialists involved in the development of information systems. *Data analysis* has historically been used as a synonym for data modeling (Veryard 1984; Howe 2001; Avison and Fitzgerald 2003 pp75-77), and may well carry the implication to business participants that process is one of investigation and description. The term *modeling* is more neutral: one can construct a model of an existing object (description) or of a proposed object (design) (Barker 1996).

2.2. DEFINITIONS OF DESIGN

Despite (or perhaps because of) the substantial body of literature on the subject, there is no single accepted definition of design. The following two, however, are representative of definitions that seek to embrace design across a wide range of disciplines:

“To initiate change in man-made things”

(Jones 1970)

“The conception and planning of the artificial”

(Buchanan 1990)

Chapter 2 – Definitions of design and data modeling

Gero and Maher's (1993) preface to a book on creative design is consistent with these:

“Design... creates a new, artificial world for us to inhabit...”

The dictionary¹⁷ definition conveys the same flavor of design as a vehicle for creating something new:

“1. A plan or drawing produced to show the look and function or workings of something before it is built or made. 2. The art or action of producing such a plan or drawing.”

2.2.1. Lawson's properties of design

Lawson's list of design characteristics, synthesized from the design literature, identifies some of the important characteristics of design problems, processes, and products. The list (Table 2–1) is organized into three groups (corresponding to *Problem, Product and Process*) of numbered items with descriptions. Lawson uses the word *characteristics* for all three levels; for clarity, we will use the words *Dimensions* for the groups, *Properties* for the numbered items, and *Characteristics* for individual statements about the Properties.

¹⁷ Compact Oxford Dictionary (Soanes and Hawker 2005)

Chapter 2 – Definitions of design and data modeling

Design Problems
1. Design problems cannot be comprehensively stated
2. Design problems require subjective interpretation
3. Design problems tend to be organized hierarchically
Design Solutions ¹⁸
1. There are [sic] ¹⁹ an inexhaustible number of different solutions
2. There are no optimal solutions to design problems
3. Design solutions are often holistic responses
4. Design solutions are a contribution to knowledge
5. Design solutions are parts of other design problems
The Design Process
1. The process is endless
2. There is no infallibly correct process
3. The process involves finding as well as solving problems
4. Design inevitably involves subjective value judgments
5. Design is a prescriptive activity
6. Designers work in the context of a need for action

Table 2-1: Lawson's properties of design

A notable omission from Lawson's synthesis is that creativity is not treated as a Property in its own right²⁰. Creativity is widely seen as playing a central role in design (Archer 1965; Jones 1970; Akin 1990; Owen 1992; Kepner 1996 p3). Lawson's framework does recognize creativity under Process Property No. 3: *The process involves finding as well as solving problems* where he states that "we must expect the design process to involve the highest levels of creative thinking." We also note one of Willem's (1991) properties²¹ not covered explicitly by Lawson: *Design solutions occur in terms of media (rather than as pure knowledge)*.

¹⁸ We refer to Lawson's Solutions as Products in the 5Ps framework.

¹⁹ This minor point of grammar is corrected in further references to this Property in this book.

²⁰ Lawson clearly recognizes the importance of creativity by devoting a chapter of his book to it.

²¹ The other three properties in Willem's framework correspond reasonably well to properties covered by Lawson.

Chapter 2 – Definitions of design and data modeling

We review data modeling against each of Lawson’s Properties in Chapter 4 (theory) and Chapter 10 (practitioner perceptions).

2.2.2. The antithesis of design

When the generic design literature seeks to contrast design with other activities, the usual choice of antithesis is *science* or (more precisely) *scientific investigation*. The defining principle of design is a concern with the development of *things*, whereas science is concerned with the development of *knowledge of the world* (Willem 1991). Designers suggest how the world might be; scientists *describe* how it is (Lawson 1997 p113). Gat and Gonen (1981), writing on design in general, suggest an alternative dichotomy of *design* and *planning*, but the same flavor is evident in their use of the terms *pure will* and *pure inquiry* to characterize the two extreme positions.

Unfortunately, the word *science* can simply imply rigor in a process rather than investigation. Thus far we have used the term *description* to denote the antithesis of design, and this would seem to be appropriate in the context of data modeling²², while retaining the spirit of the design literature dichotomy.

2.3. DEFINITIONS OF DATA MODELING

Representative definitions of data modeling²³ and data models²⁴ are well-removed from even the very inclusive definitions of design in the previous section; on the contrary they clearly reflect the opposite. A number of examples from the research, teaching, and practitioner literature are listed (in chronological sequence) to illustrate the pervasiveness of the descriptive characterization. Chapter 1 (Section 1.1) provides some further examples.

²² The word *science* has occasionally been used to describe data modeling (Wood-Harper and Fitzgerald 1982) and that characterization is implicit in such statements about data modeling as “a key criterion for any real discipline is replicability” (Warner 1996).

²³ Some of these definitions apply specifically to conceptual data modeling, but, as discussed in Section 3.6, there is usually an assumption that the conceptual model can be mechanically transformed into a conceptual schema.

²⁴ *Design* is, of course, both a verb and a noun. The definitions need to be matched with *data modeling* or *data model* as appropriate.

Chapter 2 – Definitions of design and data modeling

“A model²⁵ is a basic system of constructs used in describing reality. It reflects a person’s deepest assumptions regarding the elementary essence of things”

(Kent 1978 p93)

“formalizing and representing the data structures of reality”

(Shoval and Frumermann 1994 p28)

“Conceptual modeling is a process of extracting and representing knowledge.”

(Wand, Monarchi et al. 1995 p290)

“a representation of the things of significance to an enterprise and the relationships among those things.”

(Hay 1996a)

“an attempt to capture the essence of things both concrete and abstract...”

(Keuffel 1996)

“an abstract representation of the data about entities, events, activities, and their associations within an organization”

(McFadden, Hoffer et al. 1999)

“the activity of discovering and documenting information requirements.”

(DeAngelis 2000 p7)

(Conceptual modeling is) “motivated by a single goal – namely to provide an accurate, complete representation of someone’s or some group’s understanding of a domain.”

(Bodart, Patel et al. 2001)

“The core idea underlying all the definitions is the same: a data model is used for describing entities and their relationships within a core domain.”

(Topi and Ramesh 2002)

If these definitions of data modeling were universally accepted, then it would seem that data modeling was description, *by definition*. But “almost every company has its own understanding of ‘data modeling’” (Maier 1996) and there are alternative descriptions in the research, teaching, and practitioner literature that convey at least the flavor of design:

²⁵ Kent is using *model* as Model (formalism)

Chapter 2 – Definitions of design and data modeling

“Data modeling is generally viewed as a design²⁶ activity”

(Srinivasan and Te'eni 1990)

“an activity that involves the creation of abstractions...”

(Davydov 1994)

“the development of a conceptual data model is a design activity...”

(Shanks 1997b)

“the art and science of arranging the structure and relationship of data...”

(McComb 2004 p293)

“data modeling is a design discipline”

(Simsion and Witt 2005 p7)

Stephens and Plew (2001 pp58-59) cite dictionary definitions of design and use them to characterize data modeling. Also at odds with descriptive characterizations is the use of design disciplines as references for data modeling. Engineering and architecture have long been employed as paradigms for information systems development in general (Lee 1991a; McDermid 1991 pp83-106; Avison and Fitzgerald 2003 pp83-106, 395-407), and data modeling in particular (Finkelstein and Martin 1981; Zachman 1987; Marche 1991; Martin 1991; Hoberman 2002 p19). *Architect* is a common job title for data modelers²⁷ (Hay 2003 p58).

To summarize: definitions of data modeling commonly characterize it as description, but there are dissenting views and common metaphors which align better with the design characterization.

²⁶ The authors make it clear that they are using design in a strict sense by adding, “however, research on the data modeling process has not drawn from research about design in general”.

²⁷ See also participant occupations in Section 7.6.