Desirable Characteristics of Information Resource Dictionary Systems

Bijoy Bordoloi
Lingnan College, Hong Kong and University of Texas at Arlington, USA

Sumit Sircar
University of Texas at Arlington, USA

Bharat Lakhanpal
California State University at Fullerton, USA

The proliferation of Information Resource Dictionary System (IRDS) products spurred the issuance of national and international standards in 1988 and 1992. Even so, there are numerous features of IRDSs that are not covered by the standards. This paper seeks to provide a comprehensive set of criteria for evaluating the functional capabilities of the systems based on standards as well as the relevant published literature. The framework should benefit both potential adopters of IRDS systems who need to compare competing systems, as well as vendors who need to make design decisions on the functionality of their products.

It is now widely accepted that an organization’s information is one of its most valuable resources. The management of this resource, especially the organization’s computerized data bases, has become a critical function for its continued survival and viability. Data administration is the function that assists the organization in the management and control of data. This role has therefore achieved strategic importance in many organizations. Over the last few decades several concepts, tools and techniques have evolved to enable the data administrator to fulfill this role, culminating in the development of an automated facility called the Information Resource Dictionary System (IRDS), also commonly referred to as the Repository, or the Encyclopedia (Bruce et al., 1989). The IRDS incorporates a centralized “repository” of information about data relevant to the organization. The repository is an administrative database that allows storage and management of all the database and related information system definitions, referred to as metadata. It contains the attributes, domains, definitions, usage and relationships of the data in an organization. The central repository establishes a single source of metadata to be shared and reused by various users and tools throughout the lifecycle of an application or information system. It reduces or eliminates duplication of effort in creating and managing shared data. Furthermore, it promotes and enforces consistent definitions across interrelated application areas. These capabilities are required in order to provide all database users with the precise meaning of data, their availability and location, means of data storage, maintenance and usage, and information on who controls access to the data.

IRDS technology has evolved rapidly over the last several years, with a proliferation of commercial products. As a result, both the American National Standards Institute (ANSI) and the National Institute of Standards and Technology (NIST) began to work on data dictionary standards in the early 1980s, to promote compatibility among metadatabase management software implementations. They named this software “Information Resource Dictionary Systems,” and formulated a set of standards (X3.138) in 1988 (ANSI, 1988), which were extended in 1992 (ANSI, 1992). Apart from the
Rather voluminous standards committee reports, there was a paucity of succinct yet comprehensive, well-defined guidelines (in terms of journal publications) to enable IS managers to evaluate the functional capabilities of these products. In response to this, Bordoloi et al. (1994) proposed a fairly extensive set of desirable functional IRDS characteristics based on the 1988 ANSI IRDS committee report (ANSI, 1988) and related published literature.

Several authors (Bruce et al., 1989; Narayanan, 1988; Plotkin, 1992) had previously suggested criteria to select an IRDS, but they leaned towards the business decision-making aspect of choosing an IRDS for an organization. Their criteria, for example, include marketing-oriented features such as vendor support and vendor reputation. However, these authors did not focus in detail on the functional power of the IRDS and the overall functionality of the IRDS. The Bordoloi et al. (1994) study proposed a more detailed set of functionality-oriented criteria with which the capabilities of IRDSs can be compared.

Although the aforementioned framework was a useful contribution, it is now somewhat dated because it did not include the proposals included in the 1992 ANSI IRDS committee report (ANSI, 1992). Other developments have occurred since in the IRDS field. For example, criteria such as systems planning, systems performance analysis and export-import of IRD schemata and data were not included.

The objective of this paper is therefore to extend and update the Bordoloi et al. framework to make it more valuable to both academics and practitioners. We have retained the original format of separate subsets for “core” criteria and “additional” criteria. Indeed, as might be expected, the core criteria have not changed, and have been included here for the sake of completeness. However, a completely reengineered set of desirable functional IRDS characteristics is provided.

**IRDS Database Architecture, Standards and Core Criteria**

An IRDS is a database application that manages a centralized collection of data about all the relevant information resources within an organization. An IRDS is composed of a database component, the Information Resource Dictionary (IRD), and other components such as a query facility, and a report facility. The IRDS database (i.e., the IRD) is the heart of an IRDS. The IRD contains information about entities, data elements and their attributes such as size, type, where and how they are used, and their relationships with other entities or elements. This information is represented in the form of meta-entities such as table, record or element. This metadata should not be confused with user data, as metadata are used to identify, define, and describe the characteristics of user data.

The ANSI IRDS is a multidimensional model, which gives a view of the data ranging from extremely conceptual to actual physical storage. It is based mainly on the Entity-Relationship (E-R) model, which specifies information in terms of entities, attributes and relationships.

The ANSI IRDS Database can be viewed as a four-level architecture in which the information specified at one level describes (and potentially controls) the information stored at the next lower level. Thus, one level defines the types of “objects” which can be described at the next lower level, and that level contains the “instances” of those types. As illustrated in Figure 1, these four levels are:

1) IRD Schema Definition Level
2) IRD Schema Level
3) IRD Level
4) “Real World” Information Resources (or Production Data)

**The ANSI and ISO Standards**

Details of the ANSI standard model can be found in ANSI 1988 and 1992. For our purposes, it will be sufficient to state that the ISO model, although using different terminology, is quite similar to the ANSI model, as it is based on the evolving ANSI model. The ISO model is also a four-layered model just like the ANSI model, and the definitions of the layers are also essentially the same (ISO, 1990; Protocol Standards and Communications, 1989). The ISO model, however, contains some data modeling features which may be viewed as enhancements to the ANSI model. The major differences, which bear relevance to our proposed criteria, are discussed below.

One major difference is in the area of modeling constraints. For example, consider two entity-types table and row, and a relationship type, contains. Therefore, to represent “table-contains-row,” the model needs a way to specify values associated with objects, and on associations between objects, also known as cardinality. The ANSI model does not support...
Related Content

Evaluation of MDE Tools from a Metamodeling Perspective
www.igi-global.com/chapter/evaluation-mde-tools-metamodeling-perspective/39352?camid=4v1a

Protecting Data through Perturbation Techniques: The Impact on Knowledge Discovery in Databases
www.igi-global.com/article/protecting-data-through-perturbation-techniques/3292?camid=4v1a

General Strategy for Querying Web Sources in a Data Federation Environment
www.igi-global.com/article/general-strategy-querying-web-sources/3401?camid=4v1a

An Alternative Fit through Problem Representation in Cognitive Fit Theory
Hock Chuan Chan, Suparna Goswami and Hee-Woong Kim (2012). Journal of Database Management (pp. 22-43).
www.igi-global.com/article/alternative-fit-through-problem-representation/65540?camid=4v1a