

# Impact of standardization work on the future of information technology

*Dr. T.W. Olle*

*T. William Olle Associates*

*2 Ashley Park Road, Walton on Thames, Surrey KT12 1JU, England*

*Tel. +44-1932-221224, Fax +44-1932-221227,*

*Email: 100010.3176@CompuServe.com*

## Abstract

This paper presents the way in which international standards for information technology are organized, and what are the driving forces behind such standards. The paper comments on the criteria for success of IT standards and suggests shortcomings in the current approach to standardization that need to be rectified to enable complete inter-operability in the future.

## Key words

IT standards, IT tools, Information Systems, Systems Integration, Inter-operability, Database technology, Electronic Data Interchange

## 1 OVERVIEW OF INTERNATIONAL STANDARDIZATION

There are many different kinds of standard. It is important to distinguish between de jure standards and de facto standards. A de jure standard is formally accepted as a standard because it has been authorized by a recognized standards making body such as the International Standards Organization (ISO) [1]. A de facto standard is informally recognized as a standard because it is widely recognized and accepted although it has not been through the standardization process of a formally recognized standards making body.

It is also useful to recognize "installation standards". Installations, namely organizations making use of information technology, often find it convenient to adopt standards for the way information technology is used in the installation. Some of these standards may be either the de jure or de facto standards referred to in the previous paragraph. However, there are many situations for which no such standards are available and in such cases an installation using information technology may find it convenient to develop their own installation standards.

Returning to formal “de jure” standards (which in the rest of this paper will simply be referred to as “standards”), the process by which such standards are developed is lengthy and complex. The process of standardization that preceded the advent of information technology has been largely adopted and applied to information technology.

Most countries in the world have a standards body, which is responsible for developing standards in that country. Some larger countries, such as the USA, may have more than one such body. However, only one such body from each country may be recognized as a member of the International Standards Organization which currently has more than 100 member bodies. Only a few of these bodies participate in development of information technology standards.

## **1.1 Role of ISO**

Most countries involved in information technology standards work tend to focus their work on the work of ISO. After a standard has been agreed by ISO, a country’s member body may choose to issue their own standard that is usually a carbon copy of the corresponding ISO standard. Since progression of standards in the international arena has been shown to be slower than within a single country, a country’s member body may choose to develop its own standard for a specific area and then submit it to ISO for progression through one of the available sets of procedures.

In 1986, ISO and IEC, the International Electronic Commission (IEC) merged and the ISO/IEC Joint Technical Committee 1 (JTC1) was created. This is a special committee with the name “Information Technology”. The scope of this committee is quite simply “standardization in the field of information technology”. One of the original ISO Technical Committees and one or two IEC Technical Committees became part of JTC1. Those which did not were considered to be concerned with standardization for one specific area outside the scope of information technology. The JTC1 work on the other hand was applicable across several such specific areas.

To give further insight into the scope of ISO/IEC JTC1, it is useful to include a list of the 19 subcommittees which are currently active in JTC1. These are as follows:

- JTC1 / SC 1 Vocabulary
- JTC1 / SC 2 Coded character sets
- JTC1 / SC 6 Telecommunications and information exchange between systems
- JTC1 / SC 7 Software engineering
- JTC1 / SC 11 Flexible magnetic media for digital data interchange
- JTC1 / SC 14 Data element principles
- JTC1 / SC 15 Volume and file structure
- JTC1 / SC 17 Identification cards and related devices
- JTC1 / SC 18 Document processing and related communication
- JTC1 / SC 21 Open systems interconnection, data management and open distributed processing
- JTC1 / SC 22 Programming languages, their environments and system software interfaces

- JTC1 / SC 23 Optical disk cartridges for information interchange
- JTC1 / SC 24 Computer graphics and image processing
- JTC1 / SC 25 Interconnection of information technology equipment
- JTC1 / SC 26 Microprocessor systems
- JTC1 / SC 27 IT Security techniques
- JTC1 / SC 28 Office equipment
- JTC1 / SC 29 Coding of audio, picture, multimedia and hypermedia
- JTC1 / SC 30 Open electronic data interchange.

The fact that several committee numbers are no longer in use must be taken as an indication that the subcommittees have either been disbanded or else are no longer part of JTC1.

## **1.2 Working of ISO subcommittees**

Each subcommittee has a secretariat which is typically one of the member bodies active in the work of that sub-committee. Each subcommittee also has two or more working groups who are responsible for carrying out the detailed technical work. Each member body involved in the work is expected to send a delegation to the plenary meeting of the subcommittee at which formal decisions are taken on the progression of the work.

Not all the countries represented in ISO participate in the work of JTC1. In fact only 26 countries participate actively. This means that the country participates in one or more of the 19 subcommittees. A further 34 countries have observer status in JTC1. The number of countries having a vote in the work any one subcommittee may be less than 26, while the number actually participating in the work may be even less. Countries such as the USA (ANSI), UK (BSI), Japan (JISC), Canada (SCC) play an active role in most JTC1 subcommittees.

## **1.3 Issues in international standardization**

The success of a de jure standard can be assessed by two criteria.

1. Are products being implemented based on these standards?
2. Are products conforming to these standards being used?

In all too many many cases the answer is distressingly negative.

One must take note of one of the major issues related to information technology standardization. This can be stated as follows:

Should an information technology standard be pre-emptive, in other words prepared prior to the availability of products indicating the need for the standard?

or

Should an information technology standard be post facto, in other words prepared after products indicating the need for the standard have become available?

It seems clear that there is no universally applicable way to resolve this issue. In some cases, one approach is better and in other cases it is the other approach. The issue is bedevilled by the rate in which new developments in information technology are emerging.

## 2 DRIVING FORCES BEHIND STANDARDIZATION

There are several driving forces behind standardization. Some are long standing while the technological aspects tend to change.

### 2.1 Portability

IT standardization has been in progress for about 30 years. During that time, the technological driving forces have evolved. In the early days of IT standards, portability was seen as one of the main objectives. Despite the fact that this has not been completely met, portability is much easier than it was in the early sixties when the first COBOL standard was being assimilated.

While COBOL has been frequently criticized on technical grounds by programming language experts, it stood the test of time commendably. To some extent, it can be categorized as a successful pre-emptive strike by the standardization bodies and their developmental associates (specifically CODASYL) who did not wait for a de facto standard before developing an approach which soon became de jure.

### 2.2 Interconnectivity

The need to connect computers was first recognized in the seventies and standardization work on open systems interconnection (OSI) started in earnest towards the end of that decade. At its peak, OSI occupied the work of the two largest of ISO's subcommittees (SC 6 and SC21). It was certainly a pre-emptive strike, but one can only question its success.

The emergence of the Internet and the unprecedented speed with which this technology was assimilated as a powerful force in linking together computers around the world, had a distinctly negative effect on the acceptability of OSI standards.

### 2.3 Interoperability

The main current technical objective for standards work is without doubt interoperability, as recently emphasized by an ISO JTC1 policy statement. Interoperability between any pair of computerized information systems can always be achieved - at a price.

The technological feasibility of interoperability should never be in doubt. The real problem for standards makers is the enabling of "ad hoc interoperability". This comes down to users (rather

than, for example, software houses) being able to make use of standards which enable interoperability between any pair of computer systems on an ad hoc basis - without breaching the security constraints of either.

Interoperability goes much further than interconnectivity. This can be illustrated by a simple analogy based on everyday voice telephony. Interconnectivity has made it possible for a person to direct dial any other person in any other part of the world. Interoperability would make it possible for any person to understand any other person, once the connection has been made.

To return to the world of IT standards, there are many aspects of interoperability, some of which appear to be recognized by ISO and some not. To give the complete picture, interoperability is not merely about computer systems being able to interoperate. One cannot start from the ability to inter-connect and add on what is needed to achieve the ability to interoperate. Software products, possibly but not necessarily based on standards, used in a computer system must also be able to interoperate.

The standards for the formats of messages transmitted from one computer system to another are being defined by various committees under the overall banner of electronic data interchange (in ISO, these committees are SC14 and SC30). The aspect of interoperability which seems to be overlooked by ISO is that which involves the way in which data is handled inside a computer system and the way in which the same data is handled when transmitted from one computer system to another.

There are many aspects of data which need to be taken into account. The concept of data has been broadened by the hardware technology which has become available. Data is no longer "files and records" as it has been for several decades, but now encompasses graphics (SC24), audio, and other forms of multimedia (SC29). Data has a representation form in computer storage (SC2, SC11 and SC14). Most importantly, conventional data (records in files, or more recently, rows in tables) has a semantics.

The developers of EDI standards have been concentrating for some time on the format of the messages which are needed to support electronic commerce. There is a growing realization that the semantics (meaning) of the data in the message must also be standardized. If this is not done, the meaning of the message to a receiver may well be different from the meaning of the same message to the transmitter.

Database technology is concerned to a greater extent than most people realize with the definition of data semantics. There is a common misconception that database technology is concerned only with the organization and representation of data in storage in order to enable efficient processing of that data.

As an intuitive reaction to the very justified "garbage in garbage out" assertions of several years ago, there has recently been an additional emphasis on defining the constraints (validation rules) which data has to satisfy (singly or collectively) in order to get into the database in the first place.

The main point here is that there has to be a problem of interoperability between the way the data (sometimes referred to as persistent data) is handled within a computer system and the way the data (sometimes referred to as transient or temporary data) is handled in the message transmitted between computer systems. The semantics of the data are just as important as the format. In fact mapping from one format to another is surely easier than mapping semantics.

### 3 COMMERCIAL INTERESTS

The commercial forces (not classified here as driving) associated with standards making have changed rather less over the years than the technical driving forces. Implementors and suppliers of IT products have always been concerned about protecting and expanding their market share. In the mid sixties, the IT tool of the era, namely COBOL, enabled portability of commercial applications programs.

Some of the concerns raised at the time in certain quarters about the technical merit of COBOL were raised by hardware vendors who perceived quite correctly that widespread acceptance of COBOL would make it easier to change supplier when the time came to upgrade their hardware. The vendors countered this threat by incorporating non-standard additions in their products. The forces in favour of the standard reacted by developing validation testing schemes to check for and flag such non-standard elements and by advising users not to use them if they wanted portability.

As software technology evolved, the battle lines moved to the users' databases. The way the database was handled became the way to "lock in" the customer to a vendor. (Many of the vendors had changed since portability was an issue, but the vendor intransigence is no different). The first SQL standard came along in 1986, quickly followed by the next in 1989 and by the 700 pages of what is popularly known as "SQL92" [2] in 1992.

To some extent, the worldwide acceptance of ISO's SQL standards as a means of structuring data has enabled portability of commercial databases. However, the battle lines have simply moved from database to the functionality associated with associated standards.

The importance of interoperability is recognized and to some extent accepted in the database standards community. A set of standards have been developed for Remote Database Access (RDA) [3] which make it possible for an interactive user at one location to query or update an SQL database at another location. These standards interface with standardized communications protocols (currently OSI).

Another set of standards is under development for the export and import of databases. The cornerstone of these is an Export/Import format for a file to which an exporter exports and from which an importer imports. Export/Import is seen as "bulk transport" compared with RDA's interactive querying.

While most commercial products support the kind of functionality defined in both Remote

Database Access and in Export/ Import, their acceptance of these standards by vendors is again cool. Importing a database which is built using a competitive product is desirable. Exporting to a competitor is much less desirable. (One can observe the same phenomenon with word processing packages.)

Another database related aspect of interoperability is to be found in the dictionary standards [4,5]. The importance of the semantics of the data has already been emphasized earlier in this paper. The most recent SQL standard [2] makes it effectively possible to define declaratively all of the constraints which the data must satisfy. In simple terms, one can regard this definition of semantics as being itself stored in a special kind of database which is here referred to as a dictionary. (It is also called a repository and occasionally an encyclopaedia.)

#### 4 IT TOOLS, CASE TOOLS AND OTHER TOOLS

The term "IT tools" is a broad term which may be interpreted in several ways. In the early sixties, a generalized payroll system might have been called an IT tool. However, the term "information technology" had not been developed and the use of the term "tools" had not yet come into vogue. Today, presumably word processing systems, database management systems, spread sheets, personal time managers, and systems to assist in the task of preparing income tax returns are all IT tools.

The term "IT tool" in this paper is interpreted as a computerized tool to assist a computerized process - the "IT" being related both to the tool and to the process being assisted. If one accepts that the future of IT is inextricably tied in with interoperability, then the future of IT tools is not so much a question of "whether?" - but "what?" and "how?".

The use of the term "tools" started with the advent of CASE tools, where CASE is an acronym with various interpretations. The decoded form preferred here is "Computer Aided Systems Engineering".

CASE tools have been intended to assist those persons who wish to create new information systems (or applications) by providing computerized assistance to the persons involved which is similar to that provided to writers and authors by a word processing package. Unfortunately, the life cycle of a document or book is much shorter and simpler than the life cycle associated with a computerized information system. Even in its simplest incarnation, the information systems life cycle consists of several stages using several different techniques.

Many of the CASE tools which have been marketed have provided computerized support to only one life cycle stage or for only one technique. The more expensive tools support more of the life cycle. The long honeymoon for CASE tools seems to be over and the bitter in-fighting of the marriage now prevails. Interchange of the specifications produced by CASE tools is not yet feasible - despite the praiseworthy efforts of an industry group known as CDIF (CASE Data Interchange Format).

CDIF has fallen headlong into the trap into which so many development consortia and standards bodies have fallen. They decided that the problem of defining the semantics of what they refer to as meta-models is somehow different from that of defining the semantics of application databases. They therefore developed their own way of defining meta-model semantics. The time taken to carry out the development of the definition technique can be approximately equal to the time taken to develop the meta-models.

It is not immediately obvious why the interchange of the specifications produced by CASE tools should be regarded as a separate and different technical problem from that of interchanging the data stored in application databases. While the interchange of application databases has its problems as discussed above in connection with interoperability, treating all levels of database in the same way would at least give some focus to the interoperability problem.

## 5 ELECTRONIC DATA INTERCHANGE

Over the years most IT standards have been business area independent, in the sense that the standards and products should be of equal benefit to business areas such as trading, banking, and transportation. More recently, the need has been perceived in many business areas for Electronic Data Interchange standards. These standards prescribe the format of the various messages needed when one enterprise is interoperating with another enterprise (for example, two retail banks) and exchanging messages represented in computerized format.

The main issue here is how to ensure that the data contained in a message is interpreted in the same way by the receiver of the message as by the sender - an issue of data semantics at both ends of the interchange.

There are three classes of approach to defining the semantics of data.

- A. The most obvious one is using natural language such as English which has the merit of being relatively easy to comprehend.
- B. The second is by using a declarative approach as is inherent in database technology and the closely associated quasi-discipline of data modelling.
- C. The third is by developing an executable procedure which may be used in a computerized information system for checking and enforcing the semantics of the data.

These three classes of approach are available to the definers of EDI standards. They may, for example, use approach A in the standards definition and assume approach C will be used in the source and target systems. If the source and target systems happen to make use of database technology for designing and constructing their information systems, then approach B may be used.



The main problem is that the semantics of the data must be the same in all three places, namely the source system, the transmitted message and the receiving system or systems.

## 6 CONCLUDING REMARKS

This paper has given a brief review of the operation of the International Standards Organization and identified various issues which ISO is facing with the ever increasing rate of development of new technology. Interoperability is noted as a major objective for standards work, but commercial interests clearly feel threatened by some of the standards being developed.

The paper notes the importance of standards for Electronic Data Interchange as a means of achieving a significant aspect of interoperability and discusses the importance of data semantics in the context of these standards. The problem exists of the semantics of the data in the message being defined in a different way from (and therefore possibly at variance with) the semantics of the data in the source system for the message and the receivers of the message. This problem could undermine the work on interoperability.

### References:

- [1] ISO pages on the World Wide Web (URL is <http://www.iso.ch>)
- [2] ISO/IEC 9075:1992. Information technology, Database Languages - SQL
- [3] ISO/IEC 9579:1993. Information technology, Remote Database Access
- [4] ISO/IEC 10027:1992 Information technology, Framework for Information Resource Dictionary Systems
- [5] ISO/IEC 10728:1993 Information technology, IRDS Services Interface

## 7. BIOGRAPHY

Dr. T. William Olle entered the computer field in 1953 as a research student at the University of Manchester in England. His graduate work involved extensive machine language programming. He subsequently worked for seven years in The Netherlands on scientific applications and for two computer manufacturers in the USA for seven years and finally in Norway for one year.

He returned to the UK in 1972 to establish his own international consultancy practice based on the then emerging field of database technology. He has been actively involved in the work of ISO since 1975 - mostly, but not exclusively, with ISO/IEC SC21/WG3 Database.