STUDY ON THE INTERCONNECTION AND INTEROPERABILITY OF INFORMATION SYSTEMS

By

Rachna Patnaik*

ABSTRACT

The development of computer and network technology has led to great changes in the field of information science. With the development of World Wide Web technology, the distribution, searching and browsing of information has led to revolutionary changes. Digital libraries calls for resource sharing, which is a convenient way for managing and distributing information among users. Heterogeneous systems have to be interconnected so that the user community can have seamless/precise flow of information as per need. Also, future systems and services must be interoperable so that there is no potential loss of precision or data in the data exchange. This paper is an attempt to study interconnection and interoperability of information systems. Keeping in view the future users requirements, new technologies in the next generation information systems are also discussed. It is essential that in future networks must be interconnected, the service must be interoperable and information must be shared.

---

* Scientist/Engineer SD, Library & Documentation Division, Space Applications Centre, ISRO, Ahmedabad

0. Introduction

In this information age, the new wave of technological innovations has caused a series of revolutions in almost all areas of human society. Recent advances in computer technologies make creation of a better information system, that can efficiently manage and distribute data to user communities. In order to build such a system, one has to manage and access distributed interdisplinary information sources; provide users with object-based content sensitive searches and access to the data.

1. Objectives and Basic description

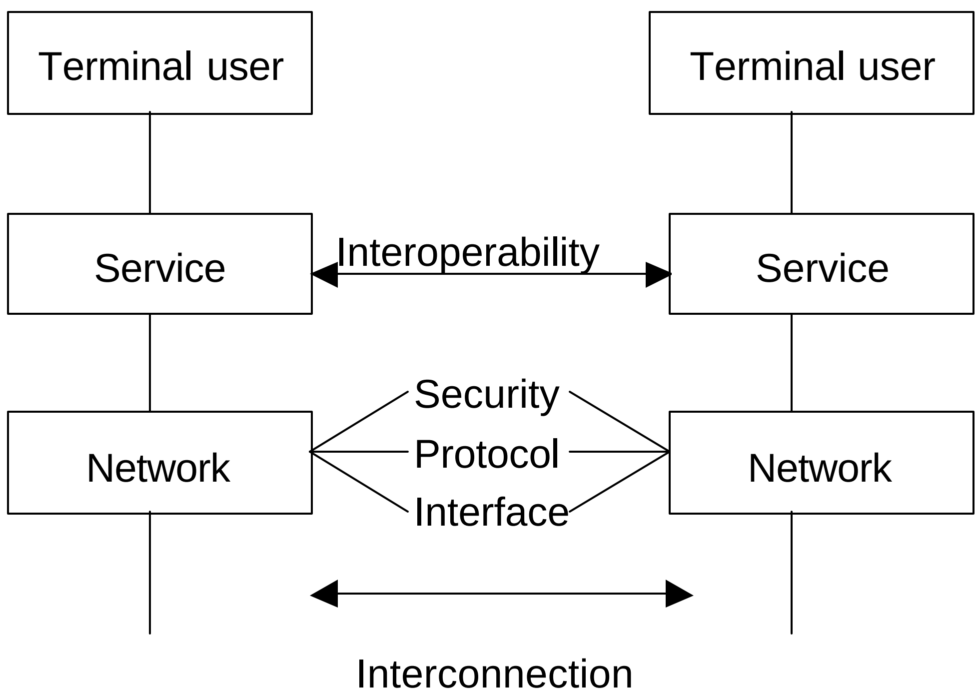
Interconnection is defined as communication between different networks. These kinds of communication deal with the lower protocol layers of Open System Interconnection (OSI) model i.e. physical layer, data link layer and network layer. Interoperability is a state, which interconnected systems can reach. Under that state, different terminal users can exchange their services directly and to their satisfaction. Specially, the degrees of interoperability should be defined and given to the users. Interoperability is the communication from one active process to another active process. It deals with layer above transport layer, when referring to OSI model. In other words, interoperability refers to the ability of two systems to exchange data:

- With no loss of precision or other attributes.

- In an unambiguous manner.
- In a format understood by and native to both the systems.
- In such a way that interpretation of the data is precisely the same.

There are two significant differences between interoperability and interfacing. The first is that, with interoperability, the exchange of data is performed without the need to translate to an intermediate format. This leads to the second difference, that, interoperable systems will produce exactly the same "answer" in the presence of identical data. Systems that are interfaced will not necessarily do so, because of the potential loss of precision or data in the data exchange.

As shown in fig. 1, from to point view of information systems users, the networks must be interconnected. The services must be interoperable, and information must be shared at a reasonable situation.



(Fig. 1)

Interoperability can be proposed at four levels as follows:

Level A : Universal --Virtual system: This level represents the ultimate goal of full interoperability. Universal interoperability is characterised by the ability to globally share integrated information in a distributed information space. Another way to view universal interoperability is a way to globally share systems.

Level B : Advanced –Integrated system: The advanced level of interoperability is characterised by shared data between applications, including shared data displays, and information exchange through a common data model. This level provides for sharing of information in a distributed but localised environment and for sharing of applications.

Level C: Intermediate—Distributed systems: This level is characterised by a client-server environment with standardised interfaces and distributed computing services that allow for exchange of heterogeneous data and advanced collaboration. This level of interoperability is achievable with implementation of "cut and paste" between applications through World Wide Web technology.

Level D: Basic—Discrete systems Interaction: A primitive level of interoperability characterised by peer-to-peer connected systems, that allows basic exchange of homogeneous data (e.g. email, formatted messages) and allows for basic collaboration. This level of interoperability is achievable by interfacing techniques described above and by use of standard office automation products that provide data import/export functions for handling data from another product.

So, in order to achieve higher level of interoperability from the lower, more advanced integration, modularity and stronger ability to share data are necessary.

2. Technical implementation

To achieve the goal of interconnection and interoperability of different information systems, a strategy of adopting common hardware, software, standards and specification architecture is the base. One type of these documents is Common Operating Environment (COE).

The COE concept encompasses:

- An architecture and approach for building interoperable systems.
- An environment for sharing data between applications and systems.
- An infrastructure for supporting mission-area applications.
- A rigorous definition of the runtime execution environment.
- A reference implementation on which systems can be built.
- A collection of reusable software components and data.
- An automated process for software integration.
- An approach and methodology for software and data reuse.
- A set of application programme interfaces (APIs) for accessing COE components.
- An electronic process for submitting/retrieving software, and data to/from the data link layer repository.

3. Vision, requirements and significance of the next generation systems

In order to build an ideal system, some fundamental issues have to be solved, including:

1. How to manage and access large, distributed, heterogeneous data and information resources over the Internet as an integrated, intelligent system in real-time.
2. How to extract domain-specific knowledge and information, intelligently and automatically, based on users requirements, from the data in such a system.
3. How to provide users with object-based, content sensitive search and access to the data, information and knowledge available in the system.

The next generation distributed intelligent information system will significantly enhance productivity and may lead to deep and far-reaching discoveries in library science.

4. New technologies available for such ideal systems

In recent years, several key technologies have advanced significantly and are mature enough for use to build next generation system. The technologies include:

4.1 Object-based distributed processing: Technological advancement has made the object-oriented distributed processing possible. The technologies include the object-oriented platform independent code system (e.g. Java) and distributed object infrastructure e.g. Common Object Request Broker Architecture (CORBA), Object Linking and Embedding (OLE). With these technologies, a truly distributed system with interoperable processing can be constructed.

4.2 Data interoperability standards and technologies: In order to make the processing interoperable in the distributed environment, the object-based distributed processing technologies are not enough. Domain specific objects and protocols have to be defined and standardised. In recent years, international bodies have developed metadata, catalogue and function standards. The major standard bodies developing interoperability standards include the American International Standard Institute (ANSI), and the ISO. Implementing these interoperability standards with object-based distributed processing technologies, will form the framework for the next generation data and information system.

4.3 Data mining technologies: Data mining deals with the discovery of hidden knowledge, unexpected patterns, and new rules from large databases. It is one of the key components in the process of information and knowledge discovery in databases. Examples of the technologies associated with data mining include the content-based search, statistical techniques, data and information visualisation, decision trees, neural networks and genetic algorithms.

Data mining technologies have been studied intensively in recent years for applications in the digital libraries and data warehouses. It is the right time for using these technologies in the next generation data and information systems for libraries.

4.4 Machine learning and artificial intelligence: With the exponential growth of data, it would be most useful to use intelligent methods for automatically extracting knowledge on a timely basis and distribute it to the people requiring it. Some current intelligent methods are artificial neural networks (memory-based and layered architecture), genetic algorithms (standard, fixed-length and messy) and fuzzy logic rule based systems.

It is expected that intelligent systems must be developed to facilitate the users in accessing, by themselves, the information or data they require. Such knowledge-rich systems must collect statistics for the current and projected demands for services. Such statistics can be used for automating the management of the data acquisition and delivery systems.

4.5 Advances in WWW, network and computer hardware: The next generation system has to be web-based, because of the popularity and efficiency of this technology. However, web technology is changing very rapidly. A number of new standards proposed by W3 consortium would have a significant impact on the development of data systems in the near future. The Extensible Markup Language (XML) will be the next generation language for web. XML helps to understand the data content standard. Extensible Style sheet (XSL) is another HTML standard that works with XML, which helps to render data in multiple ways. Resource description framework (RDF) will help in the organisation and description of metadata. RDF will help in improving data system design and search. There will be discipline specific standards such as ChemML for chemistry, MathML for mathematical representation of data.

These new web standards will improve the search results significantly. Manipulation and display of results in multiple ways on the web will become a reality. Intelligent search for information and complex queries can be implemented in a data system. Building links between various types of information will become possible. Data interoperability will significantly improve through Document Type Description (DTD's).

5. Conclusion

By integrating the above-mentioned technologies, the next generation system can be developed. Thus, the next generation systems should be intelligent enough to facilitate the users in accessing, by themselves, the information or data they require. Technologies such as data mining and data warehousing will help to discover hidden data from large database and will help in providing content-based searches to the users. It is the right time for using these technologies in the next generation data and information systems for libraries.

6. Acknowledgement

The author is thankful to Dr. A.K.S. Gopalan, Director, SAC, Dr. S.B. Sharma, Chairman, Library Committee, SAC and Shri R.S. Sharma, Head, Library & Documentation Division, SAC for their encouragement, suggestions, guidance and approval to present this paper at CALIBER-2001. The author is also grateful to Dr. P.S. Desai, Chief Scientist, SAC for screening and providing his valuable suggestions to complete this paper.

7. References

1. CEOS, 1996, Catalogue Interoperability protocol (CIP) specification-Release B. (ftp://harp.gsfc/nasa.gov/incoming/ptt/cip.b201.ps).
2. Liping Di Ken McDonald, 1999, Next generation data and information systems for Earth Sciences research in proc. Int. Symp. On Digital Earth held at China.
3. ANSI/NISO 1995 Information Retrieval: Application service definitions and protocol specification Z39.50. American National Standard Institute, New York.