論文再録

- Development of charged particle nuclear reaction data Retrieval System on IntelligentPad: CONTIP -

Journal of Information Science誌 Volume 26, Number 1, 2000, pp.29-37.

北大知識メディアラボラトリー
大林由英

昨年度の年次報告書掲載の、「CONTIP(experimental system) Users Manual」にも紹介しているが、我々はデータの検索や再利用の基盤としてIntelligentPadというソフトウェアを用いた、核反応データベースの新たな利用システムを開発している。本稿は、我々の今後の開発の基本的な概念をまとめ、Journal of Information Science誌に掲載された論文の再掲である。ご意見、ご感想をいただければ幸いである。

CONTIP(Creative, Cooperative and Cultural Objects for Nuclear data and Tools on IntelligentPad)とは、メディアシステムIntelligentPadをベースにした、荷電粒子核反応データベースNRDFの新しい利用システムの名称である。このシステムの開発は現在も引き続いて実用化への改良や新機能の導入などを進めているが、その基本姿勢は簡単にまとめると、

I)検索データの目的に応じた、複製、再編集、分析が可能なツールの実装。

II)ネットワークを通じたデータやツールの流通、それを通じた機能発達の促進基盤形成。

III)大量のデータからの知識発見の支援

である。本論文は、この基本姿勢を如何に将来我々が手にするであろう計算機技術の恩恵を元に実現できるかの考察を試みたものである。今後の開発を通じ求められるのは、我々が本論文に示したビジョンをより、明確にし、具体的にした形で、実際のシステムに反映させることであると思う。今後、本論文の考察をより掘り下げた形で、論文やシステムの発信を行っていきたい。

Development of Charged Particle Nuclear Reaction Data Retrieval System on IntelligentPad: CONTIP

Yosihide OHBAYASI [a),*], Shigeyoshi AOYAMA [b)], Hiroshi MASUI [a)],
Kiyoshi KATÔ [c)] and Masaki CHIBA [d)]

a) Meme Media Laboratory, Hokkaido University, Sapporo, Hokkaido 060-8628

b) Information Processing Center, Kitami Institute of Technology, Kitami, Hokkaido 090-8507

c) Division of Physics, Graduate School of Science, Hokkaido University, Sapporo, Hokkaido 060-0810

d) Division of Social Information, Sapporo Gakuin University, Ebetsu, Hokkaido 069-8555

Abstract

A newly designed retrieval system of charged particle nuclear reaction data (CPND) is developed on the IntelligentPad architecture. NRDF (Nuclear Reaction Data File), which is a kind of CPND compilation, is applied as an application example. We designed the network-based (client-server) retrieval system. The server system is constructed on a UNIX workstation with a relational database, and the client system is constructed on Microsoft Windows PC using an IntelligentPad software package. Our system is called CONTIP, which is an abbreviation of "Creative, Cooperative and Cultural Objects for Nuclear data and Tools". We will develop CONTIP to realize effective utilization of nuclear reaction data: I. "Re-production, Re-edit, Re-use", II. "Circulation, Coordination and Evolution", III. "Knowledge discovery".

* Corresponding author, Email: oba@nrdf.meme.hokudai.ac.jp

1. Introduction

"Nuclear data compilation" generally means an activity to accumulate physical quantities which are experimentally measured using reactors or accelerators. It is one of the most important bases for the application of atomic nuclei to science and technology. Various nuclear data compilations exist in the world [1]. As a consequence of continuous activities of such compilations, the utilization of nuclear data is expanding from nuclear engineering and nuclear physics to the other related fields. The nuclear data is roughly classified into reaction data and structure data. The nuclear reaction data is provided with the information of the nuclear reaction itself, and nuclear structure data is derived from the nuclear reaction, decay observation and so on. Especially, focusing on the nuclear reaction data, the interests and the needs toward the "charged particle" nuclear reaction data (CPND) have been increased in recent years. Their utilization is extended to many fields including, e.g., medical uses such as a radiation hazard [2], and the elucidation of the universe (nucleo-synthesis) [3].

For every data compilation, development of utility systems for compiled data is an important subject. It will support and expand utilization of the nuclear data. The amount of data becomes larger year by year, and the structure of data will be more complex. In recent years, so many remarkable developments of computer technologies, infrastructures and software allow us to extend our activities on academic research. For example, we can easily access to the foreign computational resources through the network and download or upload a large amount of data with a help of the computer facilities. These developments will go on ever after, and we will handle the larger amount of data, which is not capable for us to do now. The development of the graphical user interface (GUI) is also one of the key-technology to achieve the human-oriented user interfaces. Recently, for this purpose, many groups try to develop the retrieval system with new architectures [4].

In this paper, a new type of nuclear data retrieval system is reported, and in which NRDF (Nuclear Reaction Data File), a kind of CPND compilation of Japan, is applied as an application example. The data storage and retrieval system of NRDF was experimentally constructed on a mainframe at Hokkaido University Computing Center in 1980's [5]. However, this system would be out of date, since users are restricted to the members who have the users' accounts on the mainframe, and the user interface and indication of retrieved data are based on a useless CUI. Therefore, for NRDF, the development of CPND retrieval system on a new architecture is also one of the urgent subjects. With the foregoing backgrounds, we have developed the other data retrieval systems: Microsoft Windows application (WinNRDF) [6] and WWW search server [7]. In addition, to get benefits from recent computer and network technologies, we adopt the IntelligentPad architecture [8] as the framework of the present system. IntelligentPad has many useful features for handling multimedia, media-oriented system construction, graphical user interface, and so on.

We call the present system CONTIP, which is an abbreviation of "Creative, Cooperative and Cultural Objects for Nuclear data and Tools on IntelligentPad". The prototype of the system was built by one of the authors (M.C.) on a UNIX workstation in 1995 [9]. For the practical use, development of the system is now pursued [10]. The purposes of this paper are i) presentation of the feature of CONTIP, ii) discussion for the functions and the elements of CONTIP, from the viewpoint of the effective utilization of nuclear data. In Section 2, NRDF compilation is briefly summarized. Section 3 gives the feature of the IntelligentPad architecture. The system design and major function of CONTIP are shown in Section 4. Based on the present function of our system, the discussion of ideal systems for the effective use of nuclear data is given in Section 5. Section 6 gives a summary with future extensive issues.

2. NRDF – A charged particle nuclear reaction data compilation –

For the purpose of international exchange of experimental nuclear reaction data, International Atomic Energy Agency (IAEA) manages EXFOR compilation [11]. EXFOR includes not only CPND but also neutron nuclear reaction data, and it is used as a de facto standard. Historically, EXFOR was designed to accumulate mainly neutron nuclear reaction data. Therefore, the data description rule of EXFOR was not necessarily suitable for CPND.

From the above background, The Japan Charged Particle Reaction data Group (JCPRG) designed NRDF (Nuclear Reaction Data File), as a study of database, to accumulate scientific information of CPND from experiments of nuclear reactions. Using the own data description rule in NRDF, which is different from EXFOR, the contents of NRDF are taken from published papers mainly associated with nuclear experimental physics. The detail of the data structure of NRDF is shown in Ref. [5,12].

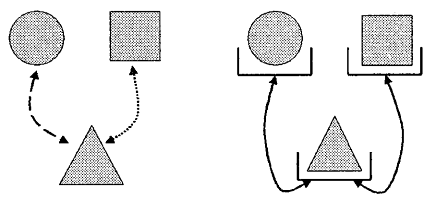
Although NRDF and EXFOR use independent data description rule, both are relevant. IAEA coordinates several nuclear reaction data centers (NRDC) in the world [13]. Each center has regional responsibility for the compilation. Almost all centers adopt the data description rule of EXFOR, hence they directly contribute as EXFOR data compiler. In contrast, JCPRG contributes as one of the members of the centers by supplying transformed data of NRDF to EXFOR [14].

JCPRG has compiled CPND on NRDF for more than 20 years. NRDF consists of over 20,000 of data mainly produced in Japanese experimental facilities [5]. The compilation policy of NRDF is based on the consideration of the following backgrounds of CPND. Due to the large variety of particles and physical quantities, CPND has a large number of reaction types in contrast to the neutron nuclear reaction data. Furthermore, measured reaction types are evolved through the development of the experimental techniques. In addition, motivation of experiment and aspects of consideration are changed along with the progress of research, thus it is not proper to define the scope of interest in advance, i.e., it is not obvious which information should be encoded in the data file. Therefore, NRDF contains the maximum information about experimental processes, conditions, devices and methods used in making data, to be encoded from the papers.

3. IntelligentPad - Meme media architecture-

The IntelligentPad architecture has been proposed and developed at Hokkaido University since 1987 [8]. The development was started as making somewhat object-oriented system construction toolkit, which is based on the graphical user interface (GUI). A pad in the IntelligentPad can be treated as an object of GUI with a view of "real paper pad", and each pad has functions such as data control programs, input/output devices between other pads, and so on. On this environment, programming of tools can be carried out by intuitively pasting and peeling off operation with pads. In addition, through the research and the development of the IntelligentPad, fundamental prospect and directivity of this architecture is being clarified: IntelligentPad is not only a specific software package, but also the fundamental environment architecture to support effective utilization of academic resources. This concept is called "Meme media architecture" [15]. Now we briefly summarize the concepts and features of the IntelligentPad architecture in this section.

Generally, current computers treat various types of intellectual resources. In Fig.1-(a), a circle, a triangle, and a square display schematic images of different type of resources. It is important to integrate these resources where linkages between any pair of different types of resources are given. Suppose that we have n different types of resources, we require $O(n^2)$ types of linkages. Moreover, future additions of new types of intellectual resources cause the same kind of difficulty [15].



(a) object-oriented architectures (b) media-based architectures

Fig.1 Object-oriented architectures and media-based architectures.

This problem can be solved by using media-based architecture [15]. As shown in Fig. 1-(b), the media-based architecture is designed as the restricted (constrained) object-oriented architecture: i.) A container of media, which schematically corresponds to a tray of each resource in Fig. 1-(b), and the content of resource are separated. ii.) The interface and the logical structure of container media are standardized. Thus, each intellectual resource is treated as a media object consists of its container and its functional contents.

The IntelligentPad architecture is the one of the media-based architecture. The pad is a kind of container of a media object. Each pad has an arbitrary number of jacks called slots, and pad has a single pin-plug connecting to one of the slots of another pad (Fig. 2). The slot holds the content of the media object as the slot value. Such pad architecture and the standard linkage facility are provided by the kernel of the IntelligentPad systems. Then, neither users nor developers have to manage them. As shown in Fig. 3, when pad B is pasted on another pad A, the pad B becomes a sub-pad of pad A, and pad A becomes the master pad. The linkage operation to the other pads is done by using the slot of master pad A. To compose pads, we only paste a pad on the other pad, with making a slot connection, just like putting the pin-plug into the jack.

When two systems share the same class libraries, pads can be easily transported from one to another through the network or by using off-line media, in the same manner as text and image files. The transport of a composite pad is achieved by only sending the pad ID's and the current status of the slots.

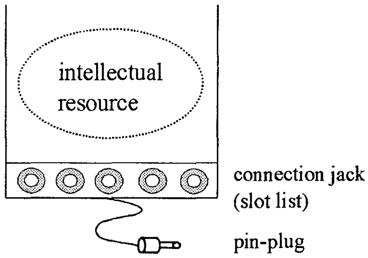


Fig.2 The logical structure of a pad as a media object.

The IntelligentPad system also provides a very powerful framework to integrate DBMS into its environment. To communicate with external objects, we only need to define each proxy pad (Fig. 4). The proxy pad for an external object provides both a view as a metaphor of programmed object and an interface of program communicating with the external object. Once the proxy pad is provided, any database system is treated as a pad media, and it can be easily integrated with other tools in the IntelligentPad architecture. Here, we explain an essential structure of database proxy pad. As shown in Fig.4, it has slots list including the query slot and the result slot. The query slot receives an SQL query from other pad, which specifies the query statement for the specific database, and the data proxy pad sends the query to the database system. When the searched result is sent back from the database, the result slot holds the list of results. Other tools can easily use the query result as the slot value. Therefore, the concept of proxy pads allows us to treat the external computational resources as a pad media.

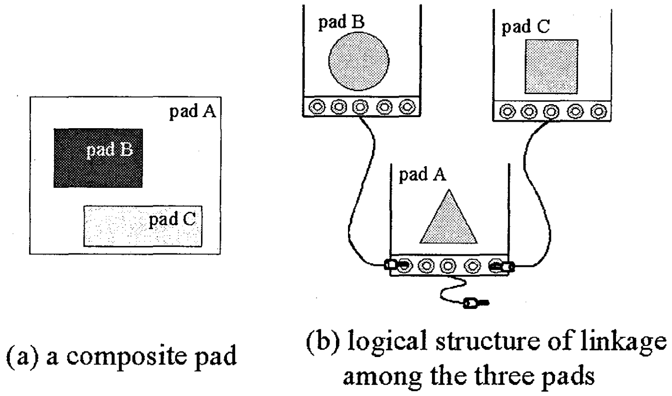


Fig.3 The application linkage in the IntelligentPad architecture.

As shown in Ref. [15], Meme media is the environment where we can share and re-use the resources of all over the world connected to the network. In order to develop and enlighten this architecture, more than 60

Japanese and other foreign companies organize the IntelligentPad Consortium (IPC)[16]. Furthermore, in the future, the kernel system of the IntelligentPad is planed to distribute freely to all users. Recently, the IntelligentPad system is available as some commercial software packages [17,18]. These products allow us to use the IntelligentPad on personal computers.

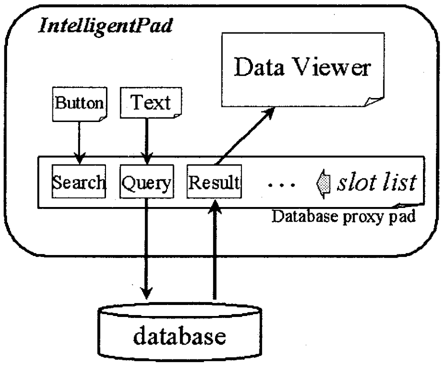


Fig.4 Schematic image of database proxy pad as pad as an interface with external object.

4. CONTIP

4.1. System design

The recent trend of the computer environment can be expressed as following key words: "Network", "Graphical", "Interactive", "Reuse of resources", and "Human oriented". The fundamental concept of our data retrieval system is based on them. We design CONTIP on the IntelligentPad architecture, which supports GUI with interactive operations, communicating through the network, and re-using resources.

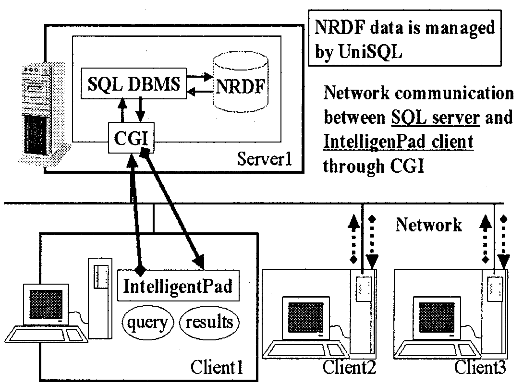


Fig.5 Network-based system.

Figure 5 shows a basic image of network communications of this system. Using a database management system (DBMS) based on SQL, we construct the NRDF data management server on a UNIX workstation. UniSQL [19] is adopted for the present system. Using a common gateway interface (CGI) on this server, the network communications between the NRDF server and clients are achieved.

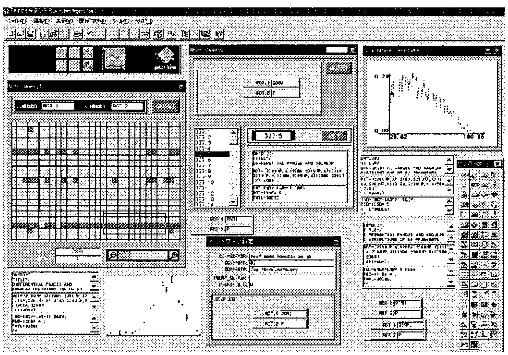


Fig.6 Overview of the client system.

We have constructed this CONTIP client on Microsoft Windows95, 98/NT-based IntelligentPad [17]. Figure 6 shows the overall appearance of CONTIP for NRDF data retrieval system. On this system, each utility component is constructed as a composition of primitive pads. The IntelligentPad software package, what we adopt, provides many primitive pads, e.g. a text browsing pad, a button pad, a slider pad and so on [17]. To create a new primitive pad, we use Microsoft Visual C++ [20] as a programming language. We develop several specific primitive pads for CONTIP. For example, the database proxy pad is used for the connection to the database server as shown in previous section. Composing the database proxy pad and other primitive pads, we construct "Data navigation pad" and "Data retrieval pad". The functions of these composite pads are given in the following.

4.2. Reference example and functions of the trial system of CONTIP

We show a reference example to discuss features of the present system. In this example, we suppose to retrieve the data of the ^{238}U(p, X) X reaction. The procedure is shown in the following subsections step by step. Here, "^{238}U" is "Uranium 238", which is the target nuclei, and "p" is the proton, which is an incident charged particle. "X" means "all possibilities within the database".

4.2.1. Data navigation.

As a first step, we show the function of the data navigation pad. Using this pad, existence of the ^{238}U(p,X) is navigated. Figure 7 shows each step of the search: i) Set the network address of SQL server. ii) Specify the two attributes of the NRDF database, in this case, target and projectile. iii) Execute the search with a click the "QUERY" button on the pad.

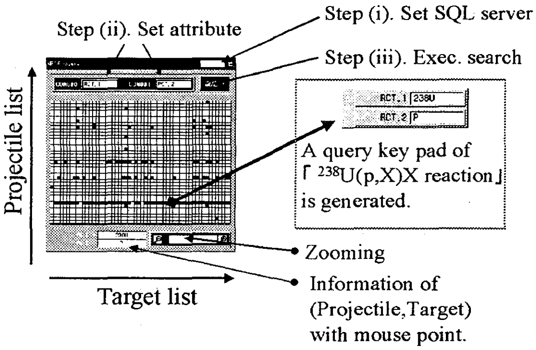


Fig.7 Data navigation npad.

For beginners, who don't know about the specific commands of DBMS, the utility system should navigate the contents of databases with ease. To connect with the database server, the data navigation pad must produce an SQL statement with the condition of the query. It is achieved by introducing a pad to generate the SQL statement. The generated SQL statement, which corresponds to the step ii), is sent to the database proxy pad, and the database proxy pad receives the results. Further, the received results are sent to a visualization component in the data navigation pad.

The result is displayed as a grid in the middle region of the pad. The status of the NRDF database is mapped in the 2-dimensional *x*-*y* plane, where *x*- and *y*- axes correspond to targets and projectiles, respectively. Dark colored points show the NRDF database has some reaction data for the target (*x*), and projectile (*y*). In this way, we find that NRDF has some data about the ^{238}U(p,X)X reaction. Using click and drag operations, a query key pad, which is used for the query of this reaction data, is generated.

4.2.2. Data retrieval.

After the query key pad of the ^{238}U(p,X)X reaction is obtained, we retrieve the data from database using the data retrieval pad (See Fig.8). The processes of the data retrieval are shown as follows: i) Set the network address of the database in the same way as the data navigation pad. ii) Drag the query key pad into the data retrieval pad, next, execute the search using the "QUERY" button. Then, the data set of the reaction is listed in the lower left part of the pad. iii) Select a dataset number from the list and get the data using "Get" button. Thus, the data pad, which specifies the retrieved data, is generated in the lower right part of the data retrieval pad. Consequently, we obtain a data pad for the ^{238}U(p,X)X reaction. Repeating step iii), other data pads of the reaction are also retrieved.

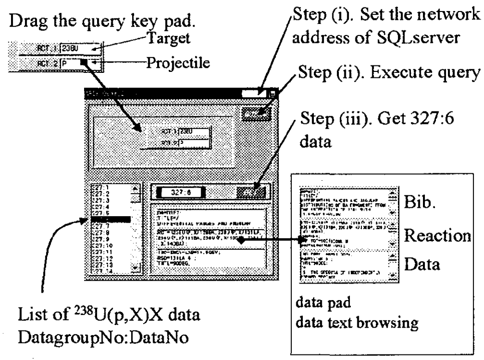


Fig. 8 Data retrieval pad.

In the same manner as we did the data navigation pad, the SQL statement is generated by a pad, and the data retrieval is done by the database proxy pad. The data pad holds the NRDF formatted code, and the contents are displayed by using the text-browsing pad. In this example, each data in the list of the retrieved data for the ^{238}U(p,X)X reaction is labeled by data file ID and data table ID. These labels can be customized by other points of view within the attributes of the database, e.g., reaction types, incident energies, and so on.

4.2.3. Interactive data visualization and comparison.

The data pad is also used for visualization as shown in Fig.7. Once we drag the data pad into the 2D data plot pad, we can see the graphical representation of the data. When the 2D data plot pad receives the numerical data from the data pad through the slot connection, a programmed function in the data plot pad visualizes it.

Data comparison is also achieved by just a drag and drop operation. 2D data plot pads can be compared by transparent superposition on the graph base pad. On this pad, selection and rejec-

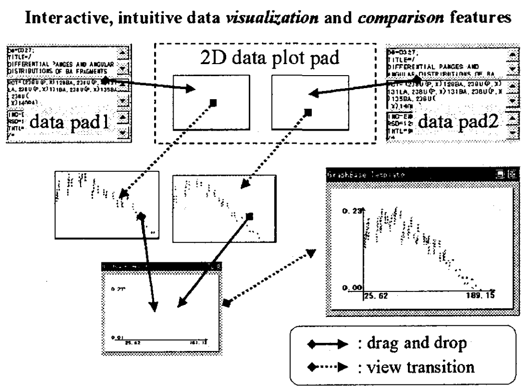


Fig. 9 Interactive data visualization and comparison.

tion can be done by pasting and peeling off operations.

5. The aim of reaction database towards the effective use

In the previous section, we show the feature of CONTIP using a reference example. Major features of this system are: i) 2-dimensional intuitive data navigation, ii) interactive data visualization and comparison. Although CONTIP is currently applied to NRDF, the framework of CONTIP can be generalized to use the other nuclear database.

For effective utilization of nuclear data, seamless linkages between measured experimental data and its application should be important. Considering these linkages, it is essential to link accumulation, evaluation and circulation on the same system. Furthermore, there are interdependences among them: evaluation of accumulated data, circulation of evaluated data, and re-accumulation evaluated data. Therefore, we should consider constructing the framework so as to achieve this continuous cycle. In addition, we should consider integrating the different databases which have different compilation policy or evaluation policy. Using the integrated database, we can retrieve and utilize the various resources concerned with nuclear data.

With the above background and motivation, we discuss the aim of the framework of CONTIP for nuclear reaction data with future computer facilities.

5.1. Re-production, Re-edit, Re-use

On the IntelligentPad, composite pads can be decomposed to more primitive pads. As shown in Fig. 10, composition of the data pad with the text-browsing pad or 2D plot pad gives each specific function. Resources are encapsulated in the primitive pads, and linkages among them are standardized. Thus, we can customize the function of pad by adding other functional pads, and/or by replacing to other implemented pads.

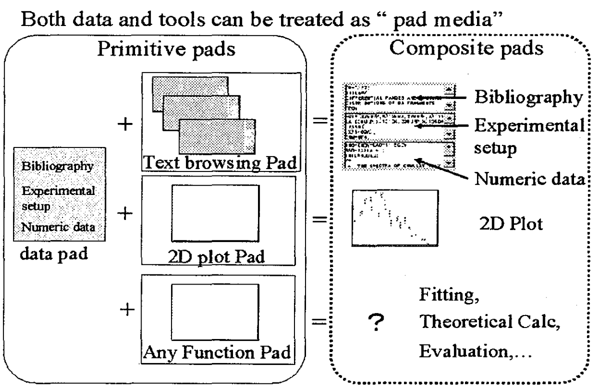


Fig.10 Synthetic features of pads.

For the further development, construction of utility pads, which, e.g., interpolates experimental data, or compares the data with theoretical calculations, supports to lead us to a kind of interactive analysis. Such a synthetic feature of the IntelligentPad would be achieved by many users' cooperative activities of "Re-produce, -edit, -use" of both nuclear data and useful tools.

5.2. Circulation, Coordination and Evolution

As discussed in Section 3, the IntelligentPad architecture supports cooperative fundamentals through the network. Once we construct a server to distribute and circulate not only data but also tools, it already gives a basis of the circular system of nuclear data as integrated intellectual resources. Figure 11 shows a schematic figure for the circular system of nuclear data and tools as the pad media by using PIAZZA [21], which is a software technology to enable to circulate pads through the network. Using PIAZZA, many users can retrieve many tools and data as the pad media. Ultimately, although any resources are scattered all over the world, we can circulate them as pad media on the IntelligentPad architecture (Fig.12).

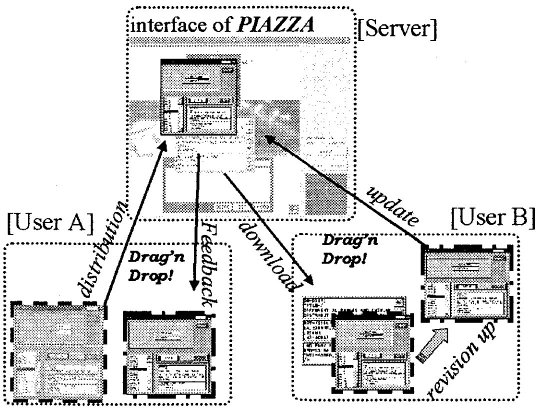


Fig.11 PIAZZA: circulation and coordination are supported by using these architectures.

5.3. Knowledge discovery

The amount of nuclear reaction data is getting huge and increasing day by day. In addition, variety of the data will also be more complex. Hence, as the amount of data increases, it will be more difficult to get essential information from huge databases. The data navigation pad of CONTIP is one of our answers for this issue. The data navigation pad enables us to investigate the correlation among the attributes of experimental data. Generally, experimental data has many attributes, hence the utility, which supports to show the correlation among them, would be helpful. Consequently, development of these kind of utility pads realizes easier navigation to needed data, and would support a new discovery of knowledge concerned with nuclear data.

6. Summary and future extension

Development of CONTIP is reported where NRDF is applied as an example. The current trial system has mainly two features: i) interactive data visualization and comparison and ii) 2D intuitive data navigation. Although the present system is constructed only for NRDF, the framework of the system can be generalized for utilization of the other nuclear data resources. As discussed in Section 5, we attempt to extract the important elements for effective utilization of nuclear data. Consequently, we conclude the system should be provided with the following three features: i) Re-production, Re-edit, and Re-use, ii) Circulation, Coordination, and Evolution, iii) Knowledge discovery.

We have developed the system using the IntelligentPad software package [17]. The IntelligentPad is not only a programming package but also media architecture to share the knowledge of our research activities. Although, now we use the Microsoft Visual C++ [20] to develop primitive pads, there are several projects [21] to develop the IntelligentPad software packages on the other programming languages. Ultimately, once the standardization of application programming interfaces among other programming languages achieved, this architecture proceeds to be independent of the programming languages.

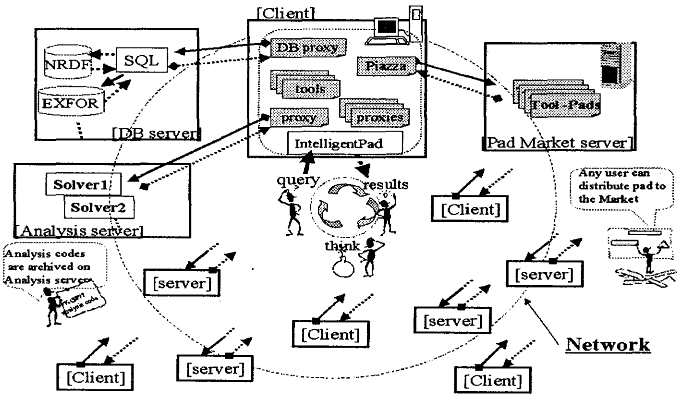


Fig.12 A schematic image of the ideal environment for nuclear data.

Although it is easy to "distribute" the nuclear data and tools programmed by the other competing software architectures, e.g., Java and Active-X control, through the networks, it is still impossible to "re-edit" the objects on the client and also impossible for Java to "re-circulate" them from the clients. The IntelligentPad is currently the only one to have ability to "re-edit" the objects as pads and to "re-circulate" them through the network [22]. Therefore, the concept of the IntelligentPad seems to be one of the prospects for effective utilization of nuclear data. We will improve the utilities of CONTIP, which will be used by a huge number of users and which would be supported to share and to evolve the knowledge of all users.

As the future extensions, inclusion of EXFOR [11] to CONTIP provides more comprehensive retrieval system for the nuclear reaction data. We are planning to distribute the system to researchers and to blush-up the trial system with feedback. Furthermore, through our development, we aim at giving the framework of the ideal environment for nuclear data as discussed in Section 5.

Acknowledgement

The authors thank Professor Y. Tanaka for helpful discussions on integrating the nuclear data utility system into Meme media architecture.

References

[1] For example, the list of major compilation is shown on the web page, http://www-nds.iaea.org/.

[2] M. Sasaki, T. Yamanaka and H. Yokobori, Accelerator conceptual design and needs of nuclear data for boron neutron capture therapy, JAERI–Conf 99-002 (1998) 25.

[3] S. Kubono, Development in nuclear astrophysics, JAERI–Conf 99-002 (1998) 31.

[4] For example, see the Chapter 4 of, Conference proceedings Vol.59 Nuclear data for science and Technology, SIF, Bologna, 19-24 May 1997.

[5] M. Togashi and H. Tanaka, An information system for charged particle nuclear reaction data, Journal of Information Science 4 (1982) 213.

[6] S. Aoyama, et al., Development of utility system of charged particle nuclear reaction data on unified interface, JAERI-Conf 99-002,222(1999).

[7] NRDF web page: http://nucl.sci.hokudai.ac.jp/~nrdf/.

[8] Y. Tanaka, IntelligentPad as Meme media and its application to multimedia databases, Information and Software Technology 38(3) (1996) 201.

[9] M Chiba, "An IntelligentPad system for the reuse of nuclear reaction data", Conference proceedings Vol.59: Nuclear data for science and Technology, SIF, Bologna (1997) 1057.

[10] Y. Ohbayasi, et al, Development of Charged Particle Nuclear Reaction Data Retrieval System on IntelligentPad, JAERI-Conf 99-002 (1999) 228.

[11] http://www-nds.iaea.org/exfor/.

[12] K. Katô, Charged particle nuclear reaction database NRDF – Present status and its usage –, Genshikaku Kenkyu 39 (1995, in Japanese) 63.

[13] For example, see INDC (NDS)-401. this document is available via http://www-nds.iaea.org/.

[14] M.Chiba, T. Katayama and H. Tanaka, A database translator of nuclear reaction data for international data exchange, Journal of Information Science 12 (1986) 153.

[15] Y. Tanaka, Meme Media and a World-Wide Meme Pool, The fourth ACM International Multimedia Conference, November, Boston (1996) 175.

[16] IntelligentPad Consortium, http://www.pads.or.jp/.

[17] Microsoft Windows version, http:/www.fujitsu.co.jp/hypertext/softinfo/product/use/ipad/.

[18] Apple Macintosh version, http://www.hitachi-sk.co.jp/Products/IntelligentPad/HomePage.html.

[19] UniSQL web page (in Japanese), http://unisql.www.nttdata.co.jp/.

[20] http://www.microsoft.com/.

[21] PIAZZA project is developed under the initiative of IntelligentPad Consortium.

[22]Y. Tanaka, private communication,