We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Scientific Data Sharing Virtual Organization Patterns Based on Supply Chain

Hui Zhao[1], Jie Peng[1] and Guoqing Huo[2]
[1]Center for Resource Sharing Promotion
(Institution of Scientific and Technical Information of China)
[2]School of Management
(Graduated University of Chinese Academy of Science)
China

1. Introduction

Since 1990s, as technique and social progress more quick, the world economy has experienced significant change. People ask for the varieties and specifications of products according to their own needs of work and life. So enterprises should become highly flexible and response rapidly to fit with constantly changing conditions. Therefore, a new organization mode appeared. Under this mode, the modern businesses make knowledge, technique, capital, materials, market and management resources in everywhere together using information technology. This organizational structure and management pattern was known as Virtual Organization (VO).

Sharing of scientific data is an important contemporary scientific and technical infrastructure construction activity. This activity focuses on how scientific data resources, especially those supported by public research funding to be obtained, are organized, large-scale produced, processed and preserved within the maximum extent. The goal of this activity is to provide useful data services for researchers conveniently, rapidly and efficiently. Many countries and regions, such as USA, English, Germany, China, and so on, are mobilizing the relevant institutions and resources to build the scientific data infrastructure for the whole society. One of the most important organization types is virtual organization, which is a way of structuring and managing goal-oriented activities.

This thesis introduces the concepts, development, applications, research status of virtual organizations, analysed the Scientific Data Sharing Virtual Organization (SDSVO) operation problems. Then the work prompted that supply chain management should be led into SDSVO to solve above problems. In the aspects of champion selection, members choose and IT system design, it would be best for SDSVO managers to supply chain integration. Three theoretical basis includes resources dependences, transaction cost minimization and Gametheory must be considered while supply chain integration.

The paper analysed the scientific data supply chain includes data creators, data centres, data service providers and data users based on three theory, then defined analytical frameworks, then chose two cases, which are German National Science Data Infrastructure (GNSDI) and China National Scientific Data Sharing Program (CNSDSP), to discuss operation mechanism construction, such as champion selection, members composition, IT system development,

etc. In the end the similarities and differences between the two modes were compared and discussed.

2. Literature review

The chapter research is involved in virtual organization, supply chain management research.

2.1 Virtual organization

The VO stresses that cooperation partners should make full use of integrated resources to maximize values and profits in competitive circumstances. VO has some charming characters, such as:

- Boundary crossing.
- Complementary Core Competences/the Pooling of Resources.
- Sharing of Knowledge.
- Geographical dispersion.
- Changing participants.
- Participant equality.
- Electronic communication.(above, Wendy Jansen 1999)
- Alternation in the constitution, members can be in or out as context.(Travica 1999)
- Switching assignment with requirement. (Mowshowitz 1999)
- Virtualization. Through the network technology (especially e-commerce) to coordinate and organize activities (Venkatraman and Henderson 1996)

Based on the above characteristics, scholars have done many works on organization pattern, operation mechanism, member management, etc.

In respect of organizational mode, Lethbridge(2001) put forward six patterns of VO from the view of organization communication form and content, which are the virtual interface, the star alliance, market alliance, common alliance, values coalition and parallel alliance. From organizational structure and function, Guo (2004) proposed three VO modes including cellular mode, star mode and federal mode. The cellular structure mode organizes the internal business units as the core, the stars structure model is applicable to grasp the enterprise, federal vertical supply chain structure model is similar to the coordination commission.

About organization management, Zhu&Peng (2003) summarized three patterns that are technical cooperation mode, contract cooperation mode and integration mode based on manufacture-learning-research cooperation. As co-operations became gradually close, technical collaboration would turn to contract cooperation then integration mode at last. And during the VO's Operation period, incentive mechanism (interest distribution mechanism, dynamic mechanism), management and supervision mechanism (internal supervision mechanism, and external supervision mechanism, operation mode) and update mechanisms should be resolved. Hong, et al (2004) proposed five factors must be considered in R&D VOs, which were coupling between team and projects, each member's resources conditions and associativity, internal communication and coordination mechanism in VO, explicit knowledge system type (recessive/explicit), R&D project innovation levels (mutation/gradual), etc. Xie (2007) put forward three VO patterns in the field of e-government to share information standing by the view of information system architecture.

The patterns are concentrated sharing model (solve data on centralism processing), P2P sharing model (solve department data sharing), and services mode (based on SOA application, solving collaboration problems).

In the side of relationship of member management, Qi Sun (2006) discussed partnership organization structure and work order, partner selection problem, and partnership performance evaluation. He deemed that the host of partnership should better belong to the third party because he or she has no conflicts with VO members. The host's responsibility is to undertake planning, prepare for all the partnership workshops, guide partnership management teams, the relevant training and the implementation process. First of all, partner selection evaluation index should be built.

But the differences of organization mode study are too big and hard to use in practice. And these organizations form in common is that they organize the members according to business relationship along supply chain. So **VOs can be divided into two basic types on the basis of data sharing supply chain. These are vertical VOs along the longitudinal business chain, and horizontal VOs on joint transverse business chain**. And in respect of VOs research, operation mechanism and membership management is concentration.

2.2 Scientific data sharing management and VOs

Scientific data refers primitiveness, basic data which are generated in the procedure of humanity understanding and transformation the world, and data products and related information which produced according to different requirements. The data can come from two main sources: 1) Observation and measurement. 2) Calculation. (Niinimaki1, et al. 2004) The data sources are diversiform. The format each is not identical. Unified standards are lacked in management and storage the data. And the storage is very scattered. (Zhang, 2008) So the use of scientific data is very inconvenience. At the same time, the mass data stored in government funding agencies that are bound by confidential treaty are difficult for the public to access.

Scientific data is not only the scientific research base, the indispensable source for information, but also the traction of modern science. A good database may bring out a new discipline, or re-activate an old discipline to enter a new stage of development. (Lu, et al. 2003) With the scientific data strategic role strengthening, a new science research form has taken shape. That is the strategy of scientific research project driving scientific data accumulation gradually turns into the strategy of scientific data driving project development. In 2000's, some government departments, national institutions and public welfare institutions went into action to promote scientific data sharing.

International organizations. In early 2002, the International Committee on Data for Science and Technology (CODATA) world congress took "data synthesis and data interoperability problems, scientific data applications and sharing legal problems, the information economics of scientific data services and ethics in data acquisition and application process" as the theme of the congress. CODATA and the International Council for Science (ICSU), etc. also constructed the world weather monitoring network, the global resources information database, international disaster information network and international Marine resources system, and so on, using Internet technique. These databases formed a global scale scientific data exchange and sharing service system initially. (CODATA handbook, Honglie Sun & Chuang Liu, 2003)

The United States. The non-profit scientific data "full open" policy is a national strategy in the United States. Since 1990, national science data centres group has established including 9

data centres. In 1995, "G1obal Data and Information System (GCDIS)" project has launched. The project further integrated nine data centres information system to one system (EOSDIS). Its aim is to provide more useful format assembly data to users to fulfill existing datasets potential. (Shi &Cui, 2002; Chuang Liu & Zhengxing Wang, 2002) National Institution of Health (NIH) of USA made data sharing policies about scientific research data produced by NIH supported project to archive and share science datasets to public. (NIH, 2004)

Britain. In 2002, The UK government put forward e-Science plan. One of the key research contents is scientific data sharing based on grid (Chen, 2004). Jung & Yeom (2010) proposed a dynamic bridge for e-Science Grid, implementing Web 2.0 service in order to share experimental data effectively. Sànchez-Artigas & García-Lópezwe (2010) explored the possibility of using the P2P paradigm for data-intensive e-science applications on the Grid. Using that method achieved fast access to the huge and widely distributed amounts of data and proposed e-Science Grid to overcome the scalability barriers in e-science communities.

Germany. In 2003, German Science Foundation launched a project which aim is to make science data open access. First, the action limited in geoscience. The project reached the desired objectives. In 2006, the project executive, the German National Science and Technology Library, extend the scope of data to other areas, such as medical, chemical, crystal structure and grey-literature and other types of content. They have established branches to manage the registration of data sets. (Brase, 2008) The virtual organization became bigger and stronger than ever.

China. In 2003, the State Ministry of Science and Technique of China started the scientific data sharing program. The program has established a leading group, a co-ordination group, experts group, the working group and offices whose responsibility is project design, implementation, oversight, project management, and so on. Program involves meteorological, seismic, hydrological, marine, agriculture, basic sciences, medicine and health care, earth science and other fields. Nine major data centers and three data networks were the beginning of the program.

In conclusion, countries of scientific researching actively are practising in the scientific data sharing field. And virtual organization mode is adopted in above activities. But in the research content, technical scheme discussed mostly and management solutions are lack. In practice, SDSVOs faced many barriers of mechanism, coordination and cooperation such as non-cooperation, copyright infringement, long-term running etc. So management research need to develop and discussion further.

This article discussed SDSVOs management in the view of supply chain management and did the case study. Hope to provide new thoughts to scientific data sharing.

3. Data sharing supply chain and driving factors in SDSVOs

According to the resources dependence and core competence theory, an enterprise's core resources and capacity are limited. In order to maintain long-term competitive advantage, enterprise will highlight the key and develop competitive parts. The result will inevitably lead to the enterprise resource allocation unbalance. Thus enterprise should obtain resources from the external. In this case, the enterprise has to establish cooperative relations with other enterprises to get a stable resources supply (Yajun Liu & Guoxu Chen, 2008). The internal factors driving this co-operation are resource integration theory, transaction cost theory and game theory. These theories are also suitable for scientific data sharing activity.

This paper first analysis scientific data supply chain. Then discuss resource superiority, cost changes and strategy chosen on each link of scientific data sharing.

3.1 Supply chain in scientific data sharing

La Londe and Masters (1994) proposed that a supply chain is a set of firms that pass materials forward. Normally, several independent firms are involved in manufacturing a product and placing it in the hands of the end user in a supply chain—raw material and component producers, product assemblers, wholesalers, retailer merchants and transportation companies are all members of a supply chain. Ma (2000) give a more comprehensive definition in the book of "supply chain management": Around a core businesses, the supply chain controls the flow of information, logistics, capital, procures, raw materials as beginning, then produces intermediate products and final products, final distributes products to consumers through distribution network. This chain forms suppliers, manufacturers, distributors, retailers, even until the end-users to a functional and overall network of chain structure. Supply chain management theory is also suitable for information industry (Rong Wang & Ji, 2008). The principle of supply chain can be used in research scientific data sharing problems.

Since a long time, scientific data in science and technique activities is a subsidiary which is preserved in the hands of individuals. Other people could obtain it through the publication of scientific papers and other opened documents.

By now, with the development of information technique activities, more and more data are collected in science and technique activities, especially in large-scale, high-tech research projects funded by public research funding. (DANS, 2010) These data are unusual wealth which has high research value. Some data are unique. So requirements of processing, preserving data professionally and providing data service created. Therefore, the scientific data have become one of the important outputs in science and technique activities today.

The scientific data process has multiple steps including creating, processing, storage, distribution and service. There are professional organizations on each link. All organizations on scientific data procedure form a supply chain. Ideally, researchers acquire data through experiment, observation activities first. Then they archive data to data centres. Data centres release datasets to public. Meanwhile, researchers can analyse data, then publish research results in publications. In brief, scientific data supply chain has four links, namely suitable data producers, data centre, data services and data user. Each link has its responsibilities. (See also figure 1)

- *Data creator*: may be scientists, or government-owned scientific research institutions, or scientific research projects supported by public funds group. Their responsibility: data harvesting, data production. Costs are from ten thousand dollars to millions, even more than billions. Sometimes costs can't even assessment;
- *Data centre*: or information centre, etc. Its main task is data storage, quality assurance, making metadata, things like that. Costs of production is usually 1% of data creating;
- *Data services*: or information supplier, library, etc. Responsible for providing directory, retrieval results, such as DOI; Costs is only hundreds of dollars;
- *Data user*: using data sets, giving suggestions and opinions to data centres.

3.2 Resources need by scientific data sharing VO

Enterprise resources foundation theory is put forward by Wernerfelt officially in 1984 through a marked publication named A Resource-Based View of the Firm. (Wernerfelt 1984)

Wernerfelt said that resources and products like the pros and cons of a coin, most of finished product needs some resources input and service., enterprise's main mission is to create and grasp resource advantage. Barney (1991) thought that differences among enterprises come from each enterprise strategic resources, the core competence of enterprises relies on value, scarce, not imitate resources.
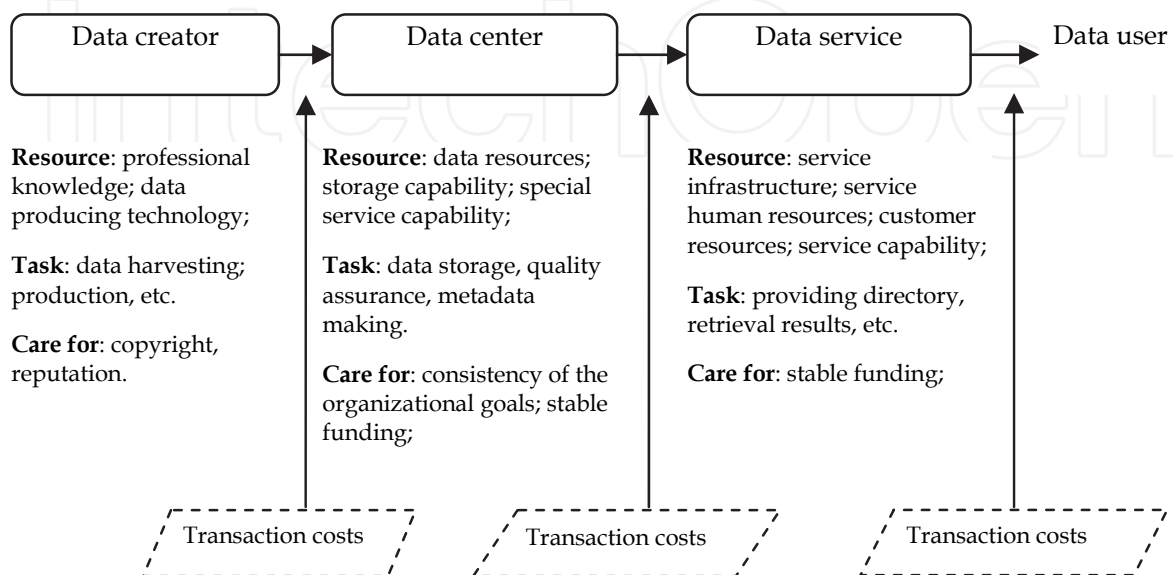

Fig. 1. Scientific data supply chain

In scientific data sharing field, a large number of scientific data resources by observation, detection and investigative is a kind of one-time-gained scarce resources. Efficient utilization of these resources is the basic guarantee for social development. Owing these resources make the organization to have some kind of congenital, special strength in competitive, but not enough. (Lievesley, 1998) The advantages of researchers (data producers) are data obtaining professional knowledge and skills, but they have no data storage equipment and professional knowledge. One of the biggest problems in scientific data sharing field is that data cannot be long-term preservation and identification, but these are data centre advantage and expertise. Data centre has specialized storage equipment and data management skills, can complete the tasks of data resources long-term preservation and identify. But the data centre is shortage of public information service ability and relative human resources, and is good at special subject services. Data service, information service centre, or library for example, is expert in public service. These information institutions are universal channels for the individuals and groups in society. They have rich experiences in interdisciplinary information providing. Otherwise, they can standardize the collecting resources and make resources available by worldwide researchers. And most services provided by information supplier are available to users on the internet. The combination of data producers, data centres and information service providers can maximize the benefit of scientific data sharing.

3.3 Transaction cost analysis in SDSVOs

Transaction cost economics is first proposed by Coase R. H. and Williamson O. E. developed the theory. Transaction cost theory researches economic activity execution efficiency. Williamson (1998) think that, two kinds of factors influence the transaction cost:

one is the principal transactions to humanity hypothesis, another is trading factors including trade frequency, uncertainty, the asset specificity, market environment, etc.

SDSVO mode can reduce the following transaction costs. (1) Break the constraints of human nature hypothesis principal transactions. SDSVOs restrain bounded rationality and opportunism from organization management level, forcing data resource protection turn to data resources opening or service. (2) Reduce the trading uncertainty among data sharing chain. When data resources aren't effective shared, data server and data users are not sure whether transaction was completed in the future. Therefore trade agreement can't be signed. After SDSVOs are established, participators clearly know the transaction rules and limits and data users can quickly find accurate achived data. Thus trade contracts can be signed easily. (3) Lower the assets specificity, such as special equipment of data processing and storage, personnel, etc. (4) Establish a more diverse, transparent trading environment. SDSVO decrease the monopoly of data suppliers by the methods including selecting service members freely through regional advantages, or resources advantage, or service capability. The market offers a variety of options, to reduce the transaction cost.

3.4 Game theory analysis in scientific data VOs

Game theory is the science about people chosen strategy. A complete game generally includes five aspects: participants, game information, game strategy or behavior choice, game order, game earnings. In scientific data sharing field, participants includes data producer, data centre, services provider and data user. All participants disseminate information, such as data requirement, data storage requirement, service capabilities, etc. in the market. Based on profit condition, the data owner decides whether to provide data to others, or private use. If the owner can get more interests from data open than data privacy, he will choose open strategy. If he gets less interest, he will use data in private only.

The barrier of science data sharing is that all data owners want to maintain their own advantages in resources and get other's benefits. One side is data owners want to get more interests, the other side is that interests are higher while the cooperative is disposable, the data owner may choose the strategy not to provide data resource and destroy cooperative basis. And while SDSVO was built, it can change the Nash equilibrium to some extent. First, the SDSVO members need to sign an agreement to join the federation. The agreement can ensure data sharing execution legally. The members who do not abide the agreement shall bear legal risk. Second, SDSVOs can provide funds or other preferential conditions (such as get complementary resources, reduce risk and sharing the cost, etc.) to members. These values can change participants' pay vector which change negative vector into positive one. That makes sharing revenue exceed the not-sharing. Third, during VOs sustaining, there are multiple games instead of disposal game between participants. That can promote parties to take share data resource action.

3.5 Factors comparison between two patterns of scientific data sharing

In section 2.1, we transform SDSVO mode into two kinds. One kind is the vertical VO mode (VVO), and the other is the horizontal VO mode (HVO).

VVO conjoin two or more institutions along supply chain together. The consortium provides services to users as a whole. There are five benefits of this mode: 1) Make scattered data resource centralized storage., That can reduce storage cost, improve storage efficiency; 2) Integrated services can reduce trading opportunities finding cost; 3) Specialized storage and

service reduce maintenance and release cost of data producers; 4) Enhance data producer influences, in favour of copyright protection, perhaps have economic gains; 5) Gear up the business professional development of data centre and data services supplier.

HVO is a federation of similar business organizations. The federation have an union catalogue. The user through joint catalogue search resources, but the league according to regional and industry divide customer. Each member provides services to a particular customer group. There are three main benefits of this mode: 1) Integrated information service system to reduce the user's information finding cost; 2) Union catalogue and services reduced resources construction costs and improve efficiency of resource use and services. 3) Improve development efficiency and depth based on resources integration, reduce resources purchasing cost.

The difference between two modes based on three theories shows in table 1.

	Vertical VO mode (VVO)	**Horizontal VO mode (HVO)**
Resource foundation	Mainly rely on resources integration along supply chain;	Mainly rely on services integration on service link;
Transaction cost	Centralized storage can reduce storage costs; Integrated services can reduce trade opportunity finding costs; Specialized storage and service reduce maintenance and release cost of data producers	Integrated information service system reduces the user's information finding cost; Union catalogue and services reduced resources construction costs and improve efficiency of resource use and services.
Game theory	Enhance data producer influences, Copyright protection, had better have economic gains; Promote the business Professional development of data centre and data services supplier.	Improve development efficiency and depth based on resources integration, reduce resources purchasing cost.

Table 1. Comparison of VVO and HVO based on three theories

4. Two cases of SDSVOs

Operation mechanism and membership management of VOS are concentration. As scientific data is a kind of information product, its information management system need to be considered. (Xueping Liu, 2007) Next, the thesis choose two national scientific data sharing projects, Germany's national scientific data infrastructure and China National Scientific Data Sharing Program, to discuss on organization structure, the leader of alliance selection, partnership construction and IT system design, etc.

4.1 Project background
- **Germany's national scientific data infrastructure (GNSDI)**

On 22 October 2003, a group of leading research institutions and research funding institutions published the 'Berlin Declaration on Access to Knowledge in the Sciences and Humanities' (Berlin Declaration, 2003) in order to "promote the Internet as a functional

instrument for a global scientific knowledge base and human reflection and to specify measures which research policy makers, research institutions, funding agencies, libraries, archives and museums need to consider."

As a consequence, the German Science Foundation adopted access to data as part of their policy in their 'Recommendations for Good Scientific Practice'. To make this policy effective, scientists themselves need to be convinced that preparing their data for online publication is a worthwhile effort. It would be an incentive to the author, if publishing of data became a citeable publication, which would add to his reputation and ranking among his peers. In 2005, National Science and technique library (TIB) officially became DOI (Digital Object Identify) registered organization of European Union. GNSDI project enters substantive construction phase.

- **China National Scientific Data Sharing Program (CNSDSP)**

In 2003, the State Ministry of Science and Technique (MOST) started the scientific data sharing program. The program has established leading group, a co-ordination group, a experts group, the working group, and offices whose responsibility is project design, implementation, oversight, project management, and so on. Program involves meteorological, seismic, hydrological, marine, agriculture, basic sciences, medicine and health care, earth science and other fields. Nine major data centres and three data networks were the beginning of the program.

By December 2005, the program had integrated and transformed 864 databases, more than one million data form, which involved in about one-third of public welfare, basic scientific data types in whole country.

4.2 Organization structure

Two projects took a different way of organization and structure. The organizational structure of GNSDI has a core institution--TIB, and rely on data supply chain to alliance each link to form a scientific data sharing infrastructure. That is a kind of VVO. The program participants of CNSDSP are mainly data centres. It has a virtual leader group and working group to fulfil top-level design and project implementation inspection. Every data centre paralleled. That is a kind of HVO. The detailed information is below.

4.2.1 GNSDI

GNSDI have three layers of architecture. The first layer is TIB who led DOI registration of scientific data, provided unified retrieval services. The second layer is composed by several Data centres whose responsibility is scientific Data storage, quality assurance and management. The third layer is all scientific data acquisition and production institutions and individuals. The connection among the three layers is through DOI system and the URL system.

In this project, according to the core standards and information system model proposed by International DOI Foundation (IDF), TIB is responsible for scientific data DOI registration and resolution as DOI regional agents of the European Union. TIB have built relationships with earth science, climate, oceans, satellite data centres by the DOI registration. Data centres are in charge of metadata of data sets abiding by DOI standard and submitting metadata to DOI register database. Data centres assure data sets and metadata quality and long-term permanent preservation conditions for registered data sets. TIB is responsible for

data sets DOI registered and resolution services. At the same time, TIB established relationship between data sets metadata and the literature resources preserved in TIB that integrated data sets services and literature services at one stop. That function is very useful and welcome by researchers.

In 2006, TIB associated with Thieme Publishing Group launched a project. The task of project is to register research data DOI while researchers publish their paper on SYNLETT and SYNTHESIS. All registered data set of this project is stored in Fachinformationszentrum Karlsruhe and be open access.

The organizational chart is as follows (See also Figure 2). We can see that TIB is the core of alliance. It is a bridge between users and data centres. It also link with publishers and users.
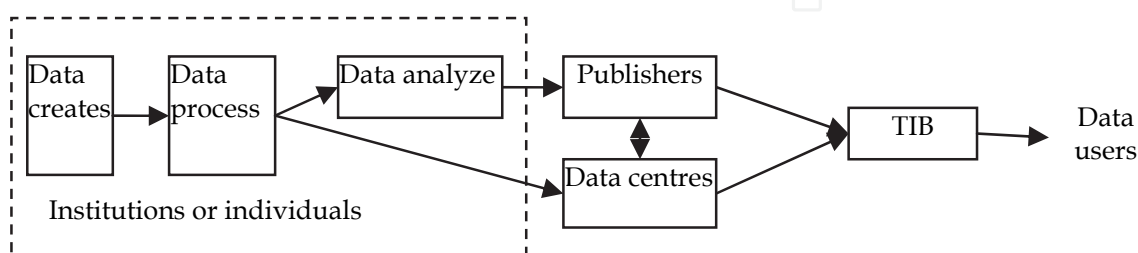
Fig. 2. GNSDI organizational chart

- **Why TIB is the core of VVO?**

There is a core enterprise in VVO mode, which responsible for sales/service. Why TIB became the core institution? Because data centres think that libraries have extensive user group. If TIB provides data sets directory retrieval, data sets can be used more widely. So data centres are very happy to cooperate with TIB to jointly promote the scientific data sharing. Their union is a kind of spontaneous organization and coordination, for the common goals and aspirations.

For making scientific data open access, they have established a virtual organization based on a supply chain with German Science Foundation's support. The core of virtual organization is the National Science and Technique of Germany Library (TIB) who plays the role of data publisher. Others are data author, data centre and data user.(see figure 2)

- **Relationship between TIB and data centres**

TIB signed a contract with each data centre. The contracts states that data centre responsible for solving the data sets copyright conflict. Data centre also is responsible for data quality control, metadata processing, as well as in the contract period to ensure that the data can be access. Data centres pay fees as €1 per data sets when registering DOI because TIB provide extra services to data centres. The services include directory retrieval, related literature recommendation, DOI register and solution, etc. This is a kind of market operation mechanism.

4.2.2 CNSDSP

CNSDSP has two layers of architecture. The upper layer is a virtual leader group and working group to fulfil top-level design and project implementation inspection. The leader group was composed by administrative leaders and executive directors come from MOST, Chinese Academy of Sciences (CAS), Ministry of Education, ministry of agriculture, Natural

Science Foundation of China (NSFC), and so on. The workgroup includes directors of business departments and fields experts. The second layer was project groups composed by data centres. Every group has a core data centre and other data centres delivered data to the core centre. So every group is parallel. Each data centre is responsible for project sponsor and contract. There is no legal restraint between each other inside project group. So this organization mode is a kind of HVO.

- **The role of virtual leading group and workgroup**

Virtual leading group's main function is to play a part in administrative leader and coordination. It is the central of the whole program. It fulfilled top-level design and project implementation inspection. The standard development project and one-stop web portal development are two projects in the program. The undertaking unit is responsible for developing and running IT system and have no contract with data transferring institution. They don't coordinate and communicate officially, but through leading group and workgroup instead, even if there is obvious upstream and downstream relationship between participants of program. Every participant signed a contract with science and technique administration. This pattern reflects the planned coordination mechanism, not market coordination mechanism.

So whether the whole program implement success depends on virtual leadership group management ability and level.

- **Program participants relationship**

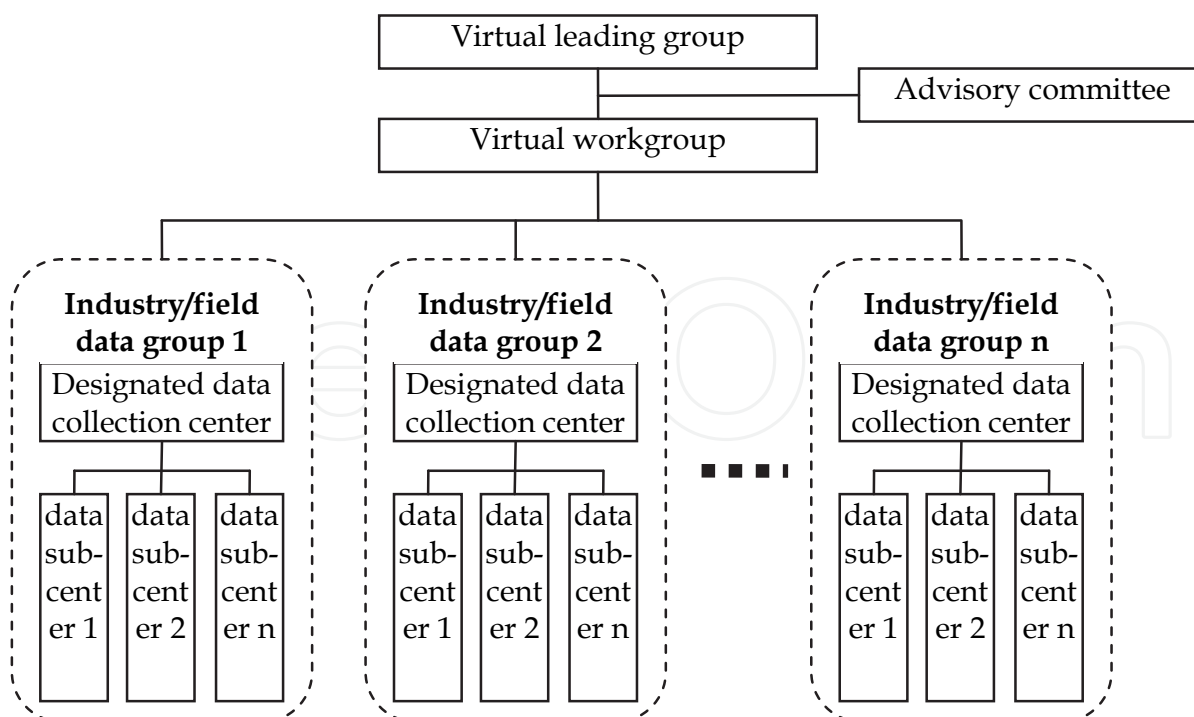The program organization chart is as follow figure 3.

Fig. 3. CNSDSP organizational chart

In figure 3, we know that every industry / field data group is a project group. Project members have technical relationship with each other and no contract relationship with each other. All management responsibility and risk belong to virtual leading group and workgroup.

But the relationship between HVO members also is based on agreement. Just CNSDSP has some particularity.

4.3 IT system design
4.3.1 GNSDI IT system

GNSDI project use DOI standards to develop a unique identifiers register, release and service system. Data centre through the system registered their data set DOI into TIB system and IDF system. Using DOI, registered data sets can be easily located to storage URL and metadata. That is a kind of convenient service for data sets users. The system architecture is as figure 4.
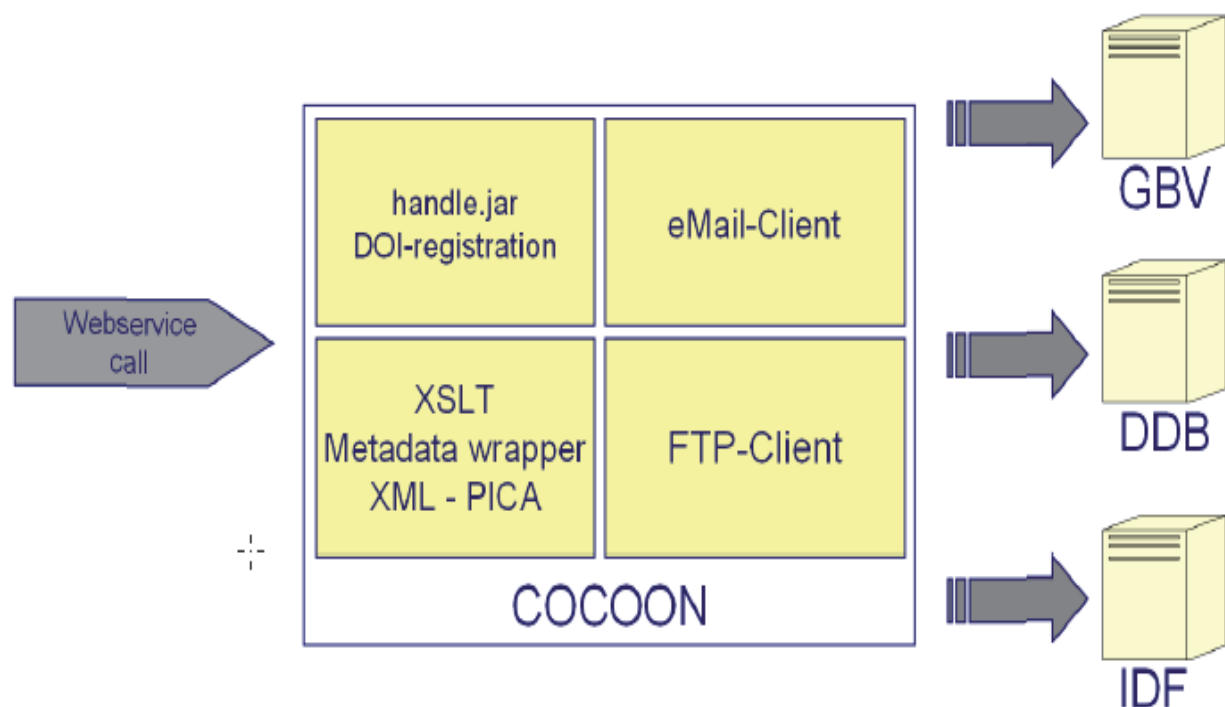

Fig. 4. GNSDI IT system architecture diagram (Schindler, 2005)

The system manages two kinds of data set. One kind is the citable core data set (cited in the literature), the other is a core data set while important, but in the literature not cited. The core data set can upgrade to the citable core data set. The system is a web service system to complete data provider and TIB information interaction.

Data providers submit 4 kinds information to the system:
1.   Register DOI information for core data sets (citable and non-citable);
2.   Upgrade non-citable core data sets to citable core data sets;
3.   If the citable core data set metadata change, create a new record;

4. If the data sets URL change, modify URN register database and IDF resolution database.

TIB have also developed some compatible middleware, such as assist registration plug-ins to decrease integration cost and work load.

However, data centre also make their IT system to suitable for various applications. For example, PANGEA (one data centre signed agreement with TIB) IT system support four data standards for various application including Open Geospatial Consortium (OGC) standard, ISO9115, Dublin Core and science and technique dataset DOI.

This shows that data centre has strong desire to recommend their data sets to public to let more people use these data. This desire and TIB's needs are matching. Therefore, both sides signed cooperation agreements to promote data sharing substantive. This is one important factor to achieve Nash Equilibrium.

4.3.2 CNSDSP IT system

CNSDSP built at least two levels of information systems. The top level is a portal website (http://www.escience.gov.cn/) which function is releasing all datasets metadata uploading by data centres. Users can search datasets by category and keyword. The items of metadata include resource name, resource ID, keyword, resource describe, data centre name, contact information, update date, resource category, etc. When the user click "details" button, he will obtain detail metadata which is stored in data centre's database if the data is open access. Otherwise, he will enter the data collection centre website to get permission to access the data.

CNSDSP IT system is simply for data information sharing, not to provide other information linking service for data centres. And the portal website collected part metadata from data centre. The system of data centre is brand new which is separated from their existing datasets resources service system.

4.4 Two cases comparison

In 2006, TIB extended the scope of registration to other areas, such as medical, chemical, and other like crystal structure and gray literature. They have established branches to manage the registration of datasets. The virtual organization became bigger and stronger than ever. By October 2007, TIB has registered 475,000 datasets, 12,500 scientific movies, 6302 case studies, 342 technical reports, as well as learning objects 112.

By September 2009, CNSDDP have integrated sharable data resources more than 140TB, exceeding more than 3,000 systemization datasets, attracting more than 1.6 million registered users, the download data more than 430TB, have provided scientific data support for more than 1,500 national level projects, such as the manned space flight project, national Marine rights and Qinghai-Tibet railway construction, etc.

Compare the above two cases as following table2.

Though two cases have so many differences, but cooperation member selection for both is similar. After investigating TIB and nine data centres of CNSDSP, four first level index and eleven second level index are identified. The index and their meanings show in table 3 as following.

Competence basis index reflect business capabilities and resources advantages. Information environment basis index show the cooperation desire. Cooperative basis and efficiency can preliminary evaluate cooperation quality.

Feature	GNSDI	CNSDSP
Project implementation environment	Data centre business development is relatively mature, have burning desire and motivation to further expand the operations scale and service channel. Data centres are active.	Data center construction just get started, and data sharing demand is very strong, so data centres were asked to grow rapidly. Data centres is passive to work
Organization structure	VVO Along data sharing supply chain	HVO parallel
the leader of an alliance	TIB (Selected reasons: more massive user base, user influence, Management, planning, implementation capabilities, integration capability, etc)	Virtual leading group (committee) (role: coordinating parties, planning, looking for funding, etc.)
Participants	Data Centres，Library	Data centres only
Relationship between participants	Should sign agreement, library provide extra service for data centres, data centres pay service fee	Have a virtul leading group who is program sponsor; No legel restrain between each data centre
IT system	Library: A datasets DOI register system combined with literature service system;Data centres: integrated datasets service system facing to various application	Have a vitrual datasets metadata integration portal; Data centres: separate datasets sharing system;
Mechanism design	DOI can anounce copyrigh. Cooperation can achieve every participants organization goal and get their due interests	The state financial capital is the important factor to attract cooperation, and the project participants improve their own ability and capability.
accomplishments	All of the participants expanded their business and services. Scientific data sharing virtual organizations develop healthly. The success mode can be extended to regions and countries where data services market is relatively mature.	People become more familiar with the data sharing function and significance. The standardization and regulation level of data centres improved, and the total amount of valuable data resources increases. Scientific data industry has developed effectively.
Opportunities and Threats	1) Is this organization mode applicable to other countries and regions; 2) Scientific data set of long-term preservation and addressable still un-solved fundamentally. 3) How to get long-term operational funds for virtual organization.	1) How scientific data sharing virtual organization steadily develop and long-term sustain? 2) Change coorperation mode from government instructions to the participants voluntary cooperation. 3) Extensions scientific data sharing service chain to deepen service contents and improve service quality.

Table 2. Comparison of two cases

First level	Second level	Meaning of index
Competence basis	The working information system level	If a member have had data processing platform, compatibility should be considered
	Data basis	Data resource scale, type and quality
	Researchers	Support staff structures, scale, etc. for software and hardware
Information environment based	Support extent by leader	Whether there is a combination of desire. If no, the institution can't be a candidate.
	Target harmony degree	If the goal gap between a member and virtual organization is too big, the cooperation cannot reach agreement.
	Business saturation	If a member's business is saturated, the virtual organization's work will be unable to complete.
Cooperative basis	Cooperation experience and skills	Ever have similar cooperation with other organizations
	Cooperation creditworthiness	The cooperation with other institutions whether smoothly
Cooperation efficiency	Built-up time	
	Built-up cost	
	Cultural compatibility	If cultural compatibility, the cooperation easy achieve success

Table 3. Scientific data sharing virtual organization member selection index

5. Conclusion

In this paper, the driving factors of SDSVO based on supply chain were discussed first. In brief, scientific data supply chain has four links, namely suitable data producers, data centre, data services and data user. Creators work includes data harvesting and data production. Data centres tasks are data storage, quality assurance, making metadata, and so on. Data service responsibility is providing directory, retrieval results. Data users use data and give suggestions and opinions to creators, data centres and services. Every link has its own advantage resources and capabilities. For example, data centres have integrated data resource, storage capabilities. Data creators have data production professional knowledge and they can collect data, but they can't preserve data permanent. Thus, data centres can cooperate with data creators. Data centres have more data resources. At the same time, data creators get more storage space and don't worry about the storage device maintain. Their cooperation can decrease both cost—collecting data cost for data centres and storage cost for creators. And so forth, data sharing supply chain form. At the SDSVO operation stage, mechanism based on Gametheory should be considered. Data creators care for copyright and reputation, data centres care for organization goals and profit. If the mechanism can satisfy all the demand, SDSVOs can run fluently. The difference between VVO and HVO based on three theories are discussed following. Then keys of case study are further explicated. That is organizational structure, leader of SDVOs, partnership and IT system.

Cases study shows that GNSDI organizational structure is a kind of VVO. TIB is the core. It integrates various datasets or other forms data resources, provide DOI register and

resolution service, and relative literature retrieval service. Data centres provide datasets metadata to TIB and pay fee for its service. Interests constraints based on agreement. The mechanism is fit for mature science data sharing environment.

Meanwhile, CNSDSP organizational structure is a kind of HVO. There is no core institution, but a virtual leading group. Data centres are participants. They transferred datasets metadata to web portal system on which user can search metadata by catalogue or keywords. Participants shared metadata according to project contract which signed with project sponsor—scientific and technical administration department. This mode is built while data sharing industry is still not mature, need government support and promote.

In IT system developing, system function design should match organizational goal and responsibilities. Integrating platform had better provide more useful functions and tools to improve datasets metadata harvesting efficiency. If platform develop functions which can improve datasets usage and influence, it will welcome.

When the leader of alliance was selected, there are different index. The chairman of VVO should have more massive user base, user influence, Management, planning, implementation capabilities, integration capability, etc. While the committee of HVO forms, optional conditions include: coordination capability, planning, looking for funding, and so on.

The member selection should consider capabilities and resources advantages, cooperation desire, cooperation quality, etc. totally eleven second level index.

For each case, there are some suggestions. GNSDI should look for long-term stable funding to datasets permanent preservation and addressable. CNSDSP partnership should change state-directed to agreement between participants each other, attract information service such as library to join the alliance to extend data sharing service contents and quality.

Although some conclusion were got in this paper, there are many further research should be done. The future work includes: the member selection index empirical study, profit distribution quantitative analysis, and the design of incentive mechanism, etc. These researches can provide more guidance for practice.

6. Acknowledgement

7. References

Mowshowitz. (1997). On the Theory of Virtual Organization, *Systems Research and Behavioral Science*, Vol.14, No.6, (Nov/Dec, 1997), pp. 373–384, ISSN 1099- 1743

Barney. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, Vol.17, No.1, pp. 99-120, ISSN 0149-2063

Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, www.zim.mpg.de/openaccess-berlin/berlin_declaration.pdf

Brase. (2008). The German National Library of Science and Technique as a DOI-Registration Agency for Scientific Conten，*China Science&Technology Resources Review*. Vol. 40, No.1, pp.37-40, ISSN 1674-1544

Byrne. (1993). The Virtual Corporation. *Business Week*, February 8, pp. 98-102, ISSN 0007-7135

Chen. (2004) E-Science plan administration and the status quo. *Quanqiu Keji Jingji Liaowang*. No.9, pp.6-8, ISSN 1009-8623

Codata online handbook, *codata websites*, Mar. 31, 2010, Available from: http://www.codata.org/ handbook/ index.html

DANS. (2010). *Preparing data for sharing*. Pallas Publications, Retrieved from: http://www.google.com/books?hl=zhCN&lr=&id=sqrjSpzEMkoC&oi=fnd&pg=PR7&dq=Working+Group+on+the+Preservation+and+Sharing+of+Statistical+Material:+Information+for+Data+Producers&ots=pmAHnUnmqg&sig=0JCsiaVyPbHBVXPyRwMMU0FFItw#v=onepage&q&f=false

Final NIH Statement On Sharing Research Data, Available from http://grants.nih.gov/grants/guide/ notice-files/ not-od-03-032.html

Finholt. (2002). Collaboratories. *Annual Review of Information Science and Technology*.Vol.36, No.1, pp.73-107, ISSN 1550-8382

Gray&Igbaria. (1996). The Virtual Society, ORMS Today, December, pp44-48

Henderson&Venkatraman. (1992) Strategic Alignment: A model for organisational transformation through information technology, *Transforming Organizations*, Thomas A. Kochan and Michael Useem,(Ed.), 50, Oxford University Press, ISBN13: 978-0-19-506504-6, New York, USA

Hong, et al. (2004) Virtual R&D Organization basic mode. *Research and Development Management*. Vol.25, No.1, pp. 43-48, ISSN 1000-2995

Klump&Wächter. (2009). *OPEN ACCESS TO DATA AND THE 'BERLIN DECLARATION'*, the STD-DOI Consortium. Available from www.codata.org/04conf/papers/ Klump-paper.pdf,

Jung & Yeom. (2010). A Dynamic Bridge for Data Sharing on e-Science Grid Implementing Web 2.0 Service , In: *Future Application and Middleware Technology on e-Science*, O. – H. Byeon et al. (eds.), pp.73-83, Springer Science+Business Media, ISBN 978-1-4419-1718-8, LLC 2010, USA

Katzy, B. R. (1998). Design and Implementation of Virtual Organisations. *Thirty-First Annual Hawaii International Conference on System Sciences*, Vol. 4, pp. 44-48. ISBN 0-8186-8242-6, Kohala Coast, HI, USA, January 06- 09,1998

La Londe, Bernard J. and James M. Masters (1994), Emerging Logistics Strategies: Blueprints for the Next Century, *International Journal of Physical Distribution and Logistics Management*, Vol. 24, No. 7, pp. 35-47, ISSN 0960-0035

Lethbridge, N. (2001). An I-Based Taxonomy of Virtual Organisations and the Implications for Effective Management. *Informing Science*, Vol. 4, No.1, pp.17-24, ISSN 1521-4772

Lievesley. (1998). The value of Data Dissemination to Both Data Users and Producers. In: *Information Dissemination and Accesss in Russia and Eastern Europe*, R Walker and M. F. Taylor(Eds.) pp.65-70, IOS press, ISBN 9051994206, Amsterdam, Netherlands

Bing Liu. (2002). Network organization system analysis and our country enterprise organization system strategic adjustment，*Inquiry Into Economic Issues*, No.5, pp.46-50, ISSN 1006-2912

Chuang Liu & Zhengxing Wang. The Experience of U.S. Global Change Data and Information Sharing and Its Indication To China，*Advance in Earth Sciences*, Vol. 17, No.2, pp. 151-157, ISSN 1001-81662002

Xueping Liu. (2007). Information products supply chain system conception desing. *Market modernization*. No.2, pp.11, ISSN 1006-3102

Yajun Liu, Guoxu Chen. (2008). The re-recognition on resource-based theory. *Science and Technology Management Research*. Vol.28, No.11, pp. 167-170, ISSN 1000-7695

Lu, et al. (2007). Investigation on scientif ic data sharing situation and its analysis. *Earthquake*. Vol.27, No.3, pp.125-130, ISSN 1000-3274

Ma. (2000). Supply Chain Management, *China Machine Press*, pp. 24, ISBN：711107978, Beijing, China

Mowshowitz, Abbe, (1994). Virtual Organization: A Vision of Management in the Information Age, *The Information Society*, Vol.10, No.4, pp. 267-288, ISSN 1087-6537

NIH. (2004) Other Data Sharing Documents and Resources. *NIH grants policy*, Available from:http://grants.nih.gov/grants/policy/data_sharing/ data_sharing_resources.htm

Niinimaki1, etal. (2004). Using Virtual Organizations Membership System with EDG's Grid Security and Database Access, *Proceedings of Database and Expert Systems Applications, 15th International Workshop on (DEXA'04)*, ISBN: 0-7695-2195-9, Zaragoza, Spain, August 30-September 03, 2004

Palmer J. W. and Speier，C. (1998). Teams: Virtualness and Media Choice, *Proceedings of Thirty-First Annual Hawaii International Conference on System Sciences,* Vol.4, pp.0131, ISBN: 0-8186-8242-6, Kohala Coast, HI, USA, January 06- 09,1998

Roland Klüber. (1998). A Framework for Virtual Organization. *Organizational Virtualness*, Pascal Sieber and Joachim Griese (Ed.), pp.93-106, *Proceedings of the VoNet - Workshop*, ISBN 3-9521463-2-3, Bern, Switzerland, April 27 -29, 1998

Sànchez-Artigas & García-Lópezwe. (2010). eSciGrid: A P2P-based e-science Grid for scalable and efficient data sharing. *Future Generation Computer Systems*, Vol.26, No.5, May 2010, pp.704-719. ISSN 0167-739X

Schindler, et al. (2005). Webservices infrastructure for the registration of scientific primary data. *Research and Advanced Technology for Digital Libraries*. No. 3652, pp.128-138, Available from: hdl:10013/epic.23912 or http://hdl.handle.net/10013/epic.23912

Scientific Data Shar ing Project，Available from http://www.amadata.net.cn/gxgc.aspx

Honglie Sun, Chuang Liu. (2003). International science and technique data development research frontier, *Advance in Earth Sciences*, Vol.18, No.3, pp.329-323, ISSN 1001-8166

Shi & Cui. (translate). (2002). NASA earth and spatial science data availability and usability evaluation. *Land And Resources Information*, NO.12, pp.12-17, ISSN 1674-3709

Qi Sun. (2006). Project Partner Relations Based On The Virtual Organization, Degree thesis repositories of CNKI, Mar. 30, 2010, Available from: http://cdmd.cnki.com.cn/Article/ CDMD-10056-2007043578.htm

Travica, B. (1999). ICAAAC Model of Virtual Organization. Proceedings of AMCIS 1999, Available from http://aisel.aisnet.org/amcis1999/8

Rong Wang & Ji. (2008). Research on Strategic Causes and Structural Characteristics of Information Goods Supply Chain. *Shanghai Management Science*. No. 1, pp.23-26, ISSN 1005-9679

Wendy Jansen et al. (1999). Electronic Commerce and Virtual Organizations. Available from imwww.fee.uva.nl

Wernerfelt, (1984). A Resource-based View of the Firm. *Strategic Management Journal*, Vol. 5, No. 2, pp. 171-180, ISSN 1056-4926

Williamson, Oliver E. 1998. Transaction Cost Economics: How It Works, Where It Is Headed, *De Economist*, Vol.146, No.1, pp. 23-58, Available from http://www.springerlink.com/content/ tg682662l836g406/

Xie, etal. (2007). E-government development trends and information sharing mode，*China Information Times*，No.5, pp.22-26, ISSN 1671-3370

Zhu&Peng. (2003). Production-study-research cooperation innovation network organization mode and the operation mechanism research. *Soft Science*. Vol.17, No.4, pp. 49-52, ISSN 1001-8409

Zhang. (2003). Scientific data "solo combat" needs change in the view of science data sharing, *High Technology And Industrialization*, No.3, pp.110-111, ISSN 1006-222X

Supply Chain Management - New Perspectives

Edited by Prof. Sanda Renko

Over the past few decades the rapid spread of information and knowledge, the increasing expectations of customers and stakeholders, intensified competition, and searching for superior performance and low costs at the same time have made supply chain a critical management area. Since supply chain is the network of organizations that are involved in moving materials, documents and information through on their journey from initial suppliers to final customers, it encompasses a number of key flows: physical flow of materials, flows of information, and tangible and intangible resources which enable supply chain members to operate effectively. This book gives an up-to-date view of supply chain, emphasizing current trends and developments in the area of supply chain management.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

INTECH
open science | open minds