Construction of NSFC Project Oriented Information Sharing System

Jianjun Li[1], Jin Wang[1,2]

[1] *Information Center, National Natural Science Foundation of China, Beijing China*
[2] *College of Information Engineering, Yangzhou University, Yangzhou, China*

Abstract

With the fast increasing of Natural Science Foundation of China (NSFC) supported projects, management of project relevant achievements such as publication, awards, books, and patents is becoming more and more important. In this paper, we make a comprehensive explanation of our NSFC project oriented information sharing system, which has been developed in the last 10 years. Some key system functions such as project based searching, achievement based searching, statistical displaying, user feedback etc. are analyzed in detail with illustrative figures and tables. We first propose our overall architecture for OA library. Then, we present detailed design from upper layer OA webpage with 3-level display hierarchy to the middle layer some key function modules. Next, bottom layer raw data acquisition module and other important system function modules like interface module, people management and security modules are explained with illustrative figures and table.

Keywords: NSFC, Project management, Information sharing system, Searching

1. Introduction

Founded on February 1986, the National Natural Science Foundation of China (NSFC) is a governmental organization which is in charge of managing national natural science projects. In the past 30 years, NSFC has been playing a very crucial role on supporting basic research, fostering talented researchers and scientists, developing international cooperation and promoting socioeconomic development etc. [1].

Until now, NSFC has accumulated very huge amount of academic and management data ranging from projects, experts, affiliations and achievements (e.g. papers, technical reports, patents, dissertations and awards etc.) and various project reports. Based on some statistical reports, there are about 1 million project application proposals, 0.3 million granted projects, 8 million participants, 3170 affiliations and 1.6 million electronic achievements inside the NSFC data repository [2].

Thus, how to manage the huge NSFC data repository and display it in a well organized manner is a very important and challenging task. As a member of Global Research Council (GRC), NSFC also has the duty to construct a NSFC project-oriented basic research knowledge library platform which is open to the nationwide researchers and the mass as an academic resource sharing system.

The "NSFC Project oriented Information Sharing System" (http://npd.nsfc.gov.cn, NPD for short) was firstly open on May 2006 [3], with main purpose to increase transparency of NSFC project management, to stimulate basic research information sharing and utilization, and to strengthen supervision and academic moral construction. Unlike the public academic resource websites nor the personal websites from various universities, NPD is not just a website with huge amount of documents like publication, person, but that it is a national level NSFC project oriented data repository with

underlying support data mining, indexing and searching, distributed file storage and cloud computing etc.

The benefits of NPD construction are obvious. To the NSFC project managers, they can supervise the progress of undergoing projects with very convenient tools from NPD system. To nationwide researchers and scientists, they can share research findings from their recent publications and projects. In this way, they can also strengthen the value of their work and increase their academic popularity. To the mass, it provides a convenient way to them so that they can know the state-of-the-art scientific results in basic research area. Besides, academic misconduct like plagiarism can be easily found, which will finally lead to a much better and benign research environment.

The rest of this paper is organized as follows. Some related work is first presented in Section 2. The primary project based searching system function is given in Section 3, followed by the achievement based searching function in Section 4. Section 5 provides additional system functions of statistical displaying and user feedback functions. Section 6 concludes this paper.

2. Related Work

Table 1 lists some similar websites to our NSFC project oriented information sharing system. We mainly studied the first two websites (namely NSF and EC), since their functions and orientation are close to our NSFC. We found that most of the contents therein are text-based material, including downloadable PDF files, pictures and few video documents.

Table 1. Some Similar Websites

Website links		Brief description
www.research.gov	[4]	The National Science Foundation (NSF) official website.
www.ec.europa.eu/research	[5]	The European Commission (EC) research and innovation official website.
www.research.google.com	[6]	Google's research and academic website.
www.research.microsoft.com	[7]	Microsoft's research and academic website.
www.research.ibm.com	[8]	IBM's research and academic website.

We made a classification of the above relevant websites from their function, contents and ways of displaying as follows.

A. Function Classification

Searching is the first and basic function for such websites. As the data is becoming more and more huge, searching module or engine is necessary and of vital importance. How to find the results for users in a fast, accurate and convenient manner is a crucial problem. The searching function can be further divided into fast searching and advanced searching. Advanced searching can help user find some specific items with constraints of year, journal name and keyword of title etc, while fast searching can help to find a range of similar items like paper title or journal etc. Users can further narrow down their specific items by using second time searching, namely searching from the first time results. All these functions can be simply implemented by clicking some tool boxes from the above website systems.

Resource classification and displaying is the second function for such systems. Based on different classification criteria, information resource can be divided into various kinds such as from different research area, geographical location, affiliation, project type etc.

From multiple dimensional searching aspect, different searching classification can be presented on the same webpage so that users can click and shift conveniently. The key point behind this technique is how to find and then build internal interconnection between such resources, and meet users' various kinds of requirements precisely and conveniently.

The third function is statistical function, where some relevant statistics will be shown in column chart or pie chart. Statistics mainly include finished project number in different research area, finished project number for different kinds of projects, different kinds of academic achievements etc. We will illustrate the statistics in the following section.

B. Content Classification

There are various kinds of website contents ranging from achievement, people, project, hot research topics, and geographical information etc. Table 2 shows relevant research contents with brief description. Due to space limit, we will not go into details.

Table 2. Some Representative Research Contents Description

Research contents	Brief description
Achievement	Publication, awards, patents, thesis and books etc.
People	Role as PI or participant, relevant project number, project member, affiliation etc.
Project	Granted project number, research area, funding, project type, duration etc.
Geographical information	Statistical number of projects, achievements etc. in different geographical area.
Hot research area	Some frequently used keyword in different research areas, which will be shown with different color and size.
Progress report	Annual report, mid-term report and final report etc., where academic achievements will be reported and uploaded.
Academic activities	Various academic conferences, seminars or meetings will be announced. Some journal special issue (SI) will be provided.

C. Displaying Classification

Usually, there are four kinds of ways to display relevant contents on our NPD website. The most commonly used one consists of various texts and pictures and some audio/video contents, which is very simple and easy to understand. The second way of displaying is tables or information lists which can usually be downloaded with csv, excel or xml formats. The third way is displaying as resource digest, which may include basic information like title, abstract, author, year, funding, duration etc. Sometimes, necessary pictures are also provides like journal cover, copy of awards etc.

The fourth way is geographical displaying. As is shown in Fig.1, IBM has its worldwide 12 research labs distributed on 6 continents. We can see detailed information by clicking on these icons.
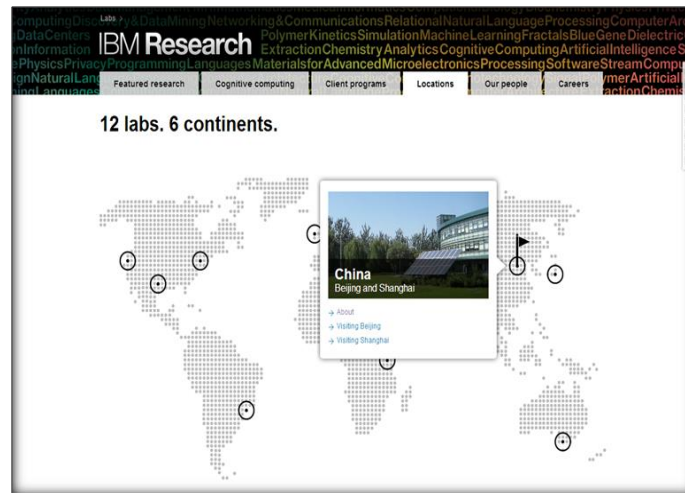
Figure 1. Illustration of Geographical Display of IBM 12 Labs

D. Our Uniqueness

Comparing with most of the similar websites in Table 1, the uniqueness of our NSFC project oriented information sharing system mainly include the following three aspects. First, we provide the searching function based on different research area, and the second time further searching function from searched results is also provided. The 8 natural science research areas are mathematics and physics, chemistry, life science, earth science, engineering and material science, information science, management science and medicine, which we will explain more next.

The second uniqueness is that the final research report is provided in the system, which mainly includes project achievement like publication, patents, awards etc. The novel ideas as well as their unique contributions are explained in detailed therein, together with the usage of project funding.

Finally, our system provides the searching function based on 14 different project types ranging from general program, key program, major program, to special fund for research on national major research instruments. The merit of this is that different researchers can target on their project focus and find the latest hot research topics from the recently awarded projects. Interested readers can refer to [9] for more information.

3. Project based Searching Function

Fig. 2 shows the homepage of our NSFC project oriented information sharing system. As can be seen from the top boxes, searching function mainly include project searching and achievement searching. Besides, some statistics analysis is also provided. From this section, both these functions will be explained.

Figure 2. NSFC Project Oriented Information Sharing System Homepage

It is worth noting that there are about 116608 finished projects and 1669928 achievements until now. For the project searching, it can be further divided into granted project searching and finished project searching. The granted projects include both recent undergoing projects and finished projects with final project report. Project searching can also be classified into research area based searching, and project type based searching.

3.1 Finished Project Searching

To some degree, finished project is more popular than granted project searching. This is because more detailed information like final project report and complete achievements will be provided therein.

Fig. 3 shows the webpage of finished project searching, where users can search their interested items such as granted project number, project name, PI name, affiliation, and research area etc.
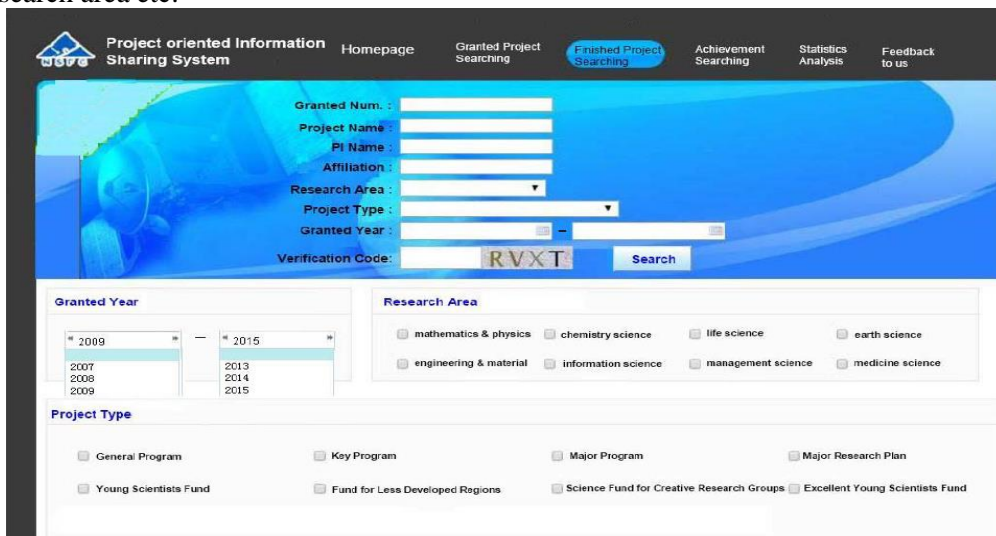

Figure 3. Finished Project Searching Webpage

Detailed research area covers 8 major natural science areas which are mathematics and physics, chemistry, life science, earth science, engineering and material science,

information science, management science and medicine. There are all together 14 project types in the year 2015, and interested readers can refer to [9] for more details.

If we continue selecting "information science" research area with project type "General Program", the following Fig. 4 will appear with search results accordingly. We found that there are 6390 finished projects which meet the search conditions. If we further narrow down the year, affiliation or PI name, more specific and accurate projects will be found. It is very fast and convenient since users only need to click the choosing box.

Taking the first project as an example, Table 3 shows the meaning of each line for the first project. For the blue color part, users can further click and get more detailed information. If users click project title, more comprehensive details of the project such as abstract, keyword, final report, as well as final achievements will be shown. If users click PI's name, authors with same name as project principle investigator will appear. And detailed information about publication, award, and books will be shown once the blue items are clicked.

Figure 4. Finished Project Searching Result Demonstration

Table 3. Explanation of First Project as an Example

Research contents	Explanation
Granted Number	69876029
Project Type	General program
Affiliation	Xidian University, China
PI Name	Zhang Hongxin
Fund	127,000 Yuan (RMB)
Achievement	Journal (1), Conference (0), Book (0), Award (0)

It is worth noting that if users click the affiliation like "Xidian University, China", the 374 relevant searching project results will appear, with 10205 academic achievements. These 374 projects can be further searched based on research area or project type, as is shown on the left upper part. The 10205 achievements can be classified as books (162), conference papers (2634), journal papers (7260) and awards (149). Detailed information can be retrieved based on users' own interests.

Figure 5. Demonstration of Affiliation based Searching Result

3.2 Awarded Project Searching

Awarded project is an extension of finished project searching, where more undergoing projects are added into the searching database. The main difference between them is that there is no detailed information about current ongoing projects such as final project report or academic achievements etc.

The benefit of this searching function is that researchers can get the first-hand information about newly awarded project information, such as project title, PI name and affiliation etc. It can stimulate academic communication and collaboration, so that researchers can know the value of their own work as well as the others' work.

Once the awarded project finishes, there is one index label so that users can click it to see detailed information of this project. Since this searching function is similar to the finished project searching, we will not explain too much here.

3.3 Research Area and Project Type based Searching

Finally, if users click "Research area based searching" or "Project type based searching" in Fig. 2, the searching results are almost the same as Fig. 3 when users choose corresponding "Research area" or "Project type". The important thing is that internal interconnection should be carefully established and organized.

The advantage of such kind of system design is that users can easily find interested items from multiple places. And they can simply click to choose the items, and then narrow down their targets via second time searching and so on. This is also the designing principle of good website with huge amount of data.

4. Achievement based Searching Function

In parallel with project based searching, achievement based searching is another important searching function where users can easily find their interested achievements such as journal or conference publication, award, and patent etc.
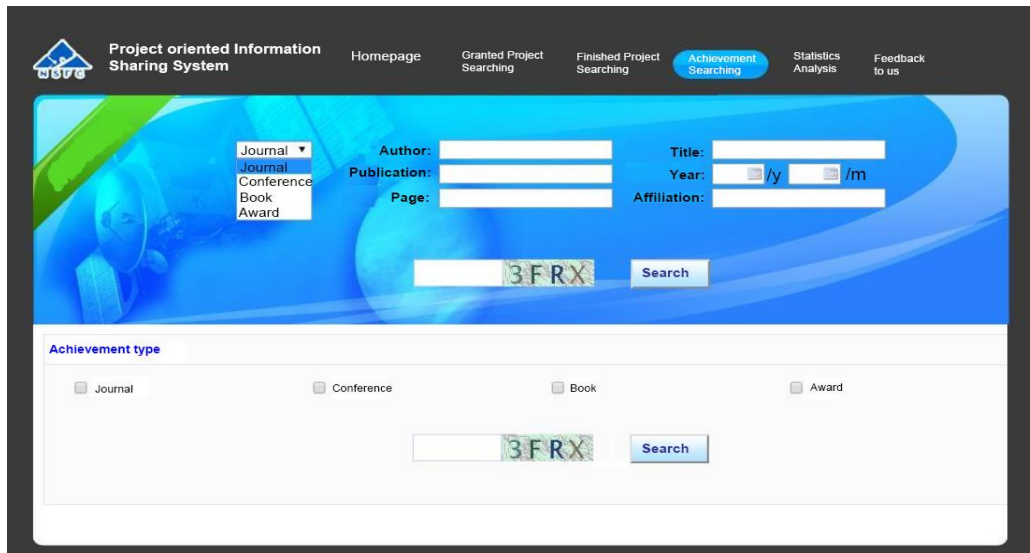
Fiure 6. Achievement based Searching Webpage

Fig.6 gives the webpage of our achievement based searching, where some important searching items include author name, paper title, publication journal/conference name, published data (y/m/), page number and affiliation etc.

Figure 7. Demonstration of Achievement based Searching Result

If users input author name like "Jin Wang" as an example, 127 searching results in Fig. 7 will appear. Interested users can further click the paper title to see detailed information. If this achievement is from finished project, bottom line relevant project can be clicked to re-direct users to project information webpage. We can also find that there are 1288475 journal papers, 324030 conference papers, 28178 books and 29245 awards therein.

It is worth noting that interested users can download some of the project oriented achievements by referring to [10, 11].

5. Statistical Display Function

Even though the statistical data is relatively few, it is very important to make comparison between different metrics, such as research area, project type, achievement type and year etc. Such statistics can be dynamically adjusted based on users' demands.

Fig.8 gives the webpage of statistical display function of system, when users choose year from 2006 to 2013 and click the "year based" statistical item. We can see the awarded project number in different year, which can be dynamically adjusted. Besides, if users click left part items with specific number, detailed information will be displayed accordingly.

Similarly, if users are interested about research area based statistics or project type based statistics, they can find such statistical data very conveniently and then make comparison. Due to space limitation, we will not go into further details.

Figure 8. Demonstration of Achievement based Searching Result

Finally, simple user feedback function is also provided so that the engineers can improve system performance. As is shown in Fig. 9, users can easily submit their suggestions, ideas or complaints on the system if they have any question or concern about NSFC system platform, project achievements etc.

Figure 9. Demonstration of Achievement based Searching Result

6. Further Research Topics

Similar to the geographical displaying function in Fig. 1, our NSFC project oriented information sharing system can provide such function. Fig. 10(a) shows the awarded projects distribution among different areas/provinces in China. We can see that the darker the area is, more projects are awarded. We can further provide the geographical distribution of different types of projects if needed.

Hot word searching function can also be available, as is shown in Fig. 10(b). Some most frequently searched hot words are displayed with different size based on searching frequency. The more times they are searched, the bigger the size will be. It provides readers a very intuitive way to know current hot research areas and key words in their different areas.

(a) Geographical display function   (b) Hot word searching function

Figure 10. Demonstration of Achievement based Searching Result

7. Conclusions and Future Work

With the fast increasing of NSFC supported national level projects, project relevant achievements such as journal/conference publication, books, thesis and patents are growing even faster. To well organized and manage such huge amount of projects and achievements, the NSFC project oriented information sharing system is constructed, where there are many challenging research problems. In this paper, the project based searching and achievement based searching as main functions are well explained with illustrative figures, where project based searching can be further divided into finished project, awarded project searching. And the finished project searching can get narrowed down to research area based and project type based searching etc. Besides, the statistical displaying function is also provided.

In the near future, we plan to extend this work in the following ways. First, we plan to add more PI authors' information like their collaborators, and other projects etc. to make them know by more people. Then, more academic achievements can be provided such as their programming source code, rough research ideas to make contribution to the research society. Finally, system performance will be improved step by step with support from cloud computing and advanced big data techniques.

References

[1] http://www.nsfc.gov.cn/publish/portal1/ Retrieved 1 December, **(2015)**.
[2] http://www.nsfc.gov.cn/publish/portal1/tab157/ Retrieved December, **(2015)**.
[3] http://npd.nsfc.gov.cn Retrieved December, **(2015)**.
[4] www.research.gov Retrieved December, **(2015)**.
[5] www.ec.europa.eu/research Retrieved December, **(2015)**.
[6] www.research.google.com Retrieved December, **(2015)**.
[7] www.research.microsoft.com Retrieved December, **(2015)**.
[8] www.research.ibm.com Retrieved December, **(2015)**.
[9] http://www.nsfc.gov.cn/publish/portal1/tab285/ Retrieved December, **(2015)**.
[10] http://or.nsfc.gov.cn/ Retrieved December, **(2015)**.
[11] J. Li, M. Zhang, D. Li, W. Zhang and J. Wang, "Construction Scheme of NSFC Open Access Library", International Journal of Security and Its Applications, vol. 9, no. 8, **(2015)**.

Authors

Jianjun Li. He received the B.S. degree from Beijing Forestry University in 1989. Now, he is the director of the Information Center of National Natural Science Foundation of China (NSFC). His main research interest includes NSFC information system management ranging from system architecture design, construction, operation, database maintenance, and security etc.

Jin Wang. He received the B.S. and M.S. degree from Nanjing University of Posts and Telecommunications, China in 2002 and 2005, respectively. He received Ph.D. degree from Kyung Hee University Korea in 2010. Now, he is a professor in the College of Information Engineering, Yangzhou University. His research interests mainly include routing protocol and algorithm design, performance evaluation and optimization for wireless ad hoc and sensor networks. He is a member of the IEEE and ACM.