ICIC Express Letters
Part B: Applications
Volume **8**, Number **1**, January **2017**

ICIC International ©2017 ISSN 2185-2766
pp. **85–91**

RESEARCH AND DEVELOPMENT OF SCIENTIFIC AND TECHNICAL INFORMATION SHARING AND SERVICE PLATFORM FOR MASS INNOVATION

Jing Zhao[1], Yonghong Cheng[2] and Yi Huang[1]

[1]Institute of Scientific and Technical Information of China
No. 15, Fuxing Road, Beijing 100038, P. R. China
zhaojing@istic.ac.cn

[2]China National Institute of Standardization
No. 4, Zhichun Road, Beijing 100191, P. R. China

Received June 2016; accepted September 2016

Abstract. *In 2015, with the supportive policies and measures of "maker space" and the development of Internet+ action plan, mass entrepreneurship and innovation is boosted in China. However, in the information industry, there are some urgent needs to build a unified information services platform to support innovation activities. Thus, we proposed a platform design concept that really simulates the whole process of a maker with creative solutions, but with no capital, resources, condition nor technology to start an entrepreneurship, and then developed the scientific and technical information sharing and service platform for mass innovation, covering the whole entrepreneur process as well as the needs for technical resources. Then we constructed the platform, and through application in chemical field, this platform fulfils its design concept to enable a maker with excellent idea to complete the whole innovation process with the platform.*
Keywords: Mass innovation, Scientific and technical information, Platform development

1. **Introduction.** From lack of information in the past to the current information explosion, lots of useful information is buried under the mass and disorder information database, which makes intelligence institutes play an important role in information collection, processing, analysis, and service for the public [1]. With "mass entrepreneurship and innovation" economic development strategy and the "Internet Plus" action plan, good information service should also be able to provide industry and market advice, professional technology, Internet technology, science and technology financial assistance and other services to innovative entrepreneurs. It becomes an inevitable choice to build a connection to promote economic restructuring and development, employment and entrepreneurship to create new technologies, new products and new services. Currently, there are some open innovation platforms in China, but they offer services limited to those who have ideas as well as resources and technologies [2]. Thus, we design and develop the scientific and technical information sharing and service platform for mass innovation to promote innovation and entrepreneurship both online and offline, as well as incubation and investment combined to provide low-cost innovative services for small and micro business and individual entrepreneurship [3-5]. The rest of this paper is organized as follows. Section 2 introduces idea and framework. Sections 3 describes the methods and key technologies. Sections 4 describes the platform and its feature. Application and demonstration are shown in Section 5. Finally, a brief conclusion and future work are given in Section 6.

2. **Idea and Framework.** To effectively address the problems mentioned above, we propose an idea that simulates the whole process of a maker with creative solutions, but with no capital, resources, nor technology to start an entrepreneurship, to raise funds by

crowd-funding model, obtain new products or new technology by crowdsourcing, and recruit elites to build entrepreneurial teams and broaden the market channels in accordance to "Internet Plus" action plan. The platform aims to attract talents, projects, and teams to come together to work out excellent solutions and creative evaluation. The platform will also integrate all types of data, such as the latest technical research, development and other relevant information to become a professional information distribution hub, as so to achieve the mission of leading innovation.

With this design model, there are three issues to be addressed, i.e., resources, systems and technologies. We need to integrate technologies, resources, systems in this platform to meet the needs of mass innovation in scientific and technical information industry, such as natural language processing technology, subject headings resources, and information processing systems. The framework is shown as Figure 1.

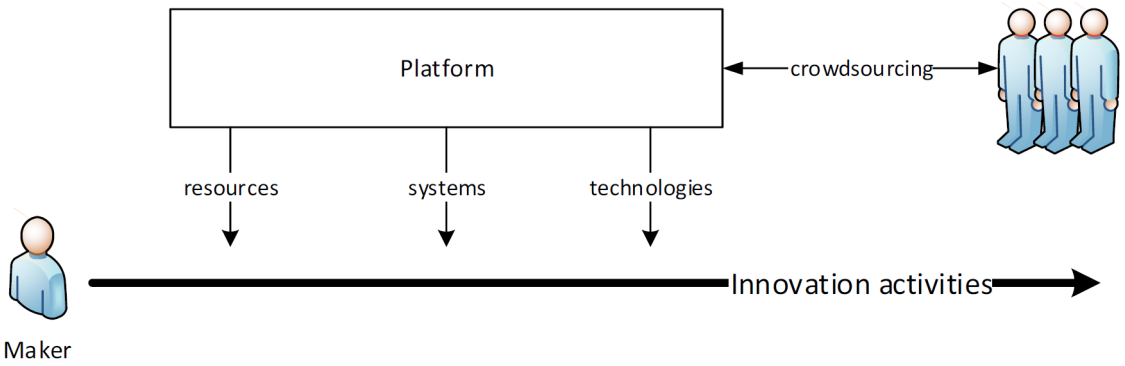

FIGURE 1. Design idea and framework

3. **Methods.** With the design idea, and based on the technologies, databases and systems built in recent years, we intend to provide tools, resources, systems and support services to public to promote innovative activities.

3.1. **Resource development and integration.** Resources are an essential part in the mass innovation process. By integrating all databases available, we provide resources to support innovative ideas and programs. The databases include "high-level scientific and technical talent database", "national research and experimental base database", "Chinese technological vocabulary system", "Chinese Thesaurus (Engineering and Technology Edition)", "China research project database", "Chinese science and technology archive", "national science and technology report database", and "Chinese patent database", allowing users to query and grasp the current status and trends related to the development of latest scientific research, patents, and other aspects of an innovative idea or product, as well as obtaining resources to carry out the innovative work.

3.2. **Systems integration.** It is essential to provide a maker with systems and tools to promote innovative activities. The systems include "information technology monitoring and analysis system", "technology innovation service platform", "knowledge organization and service platform", "information resources deep labels and key construction platform", "information research and analysis service platform", "Chinese science and technology vocabulary system service platform", "intelligent retrieval service platform", "scientific and technical information and documentation services", "top articles from outstanding S&T journals of China-F5000", "China science and technology information network", and "multi-source information fusion of scientific literature intelligence services platform". These systems can provide users with comprehensive query services, facilitate them to conduct further evaluation and application, and boost innovative ideas and methods [6-8].

3.3. **Pipeline technology.** We have integrated resources and systems, but it did not form an organic whole to meet the varied and diverse innovative ideas of a maker in the information industry. Pipeline technology is the discipline that combines knowledge from information technology and knowledge from management sciences and applies this to customized operational processes in order to complete certain tasks. With pipeline technology, a maker with innovative idea can choose the resources, systems or tools needed to form a customized technological process to complete a certain task to get satisfying results. The basic idea of pipeline technology is shown as Figure 2.

It has the following advantages:

Targeted. Users can choose a more reasonable processing unit according to their own needs, such as corpus, user dictionary, and natural language processing tool.

Scalability. Since the processing unit interface would meet the system requirements, we can easily add some existing algorithms or natural language processing technology to better satisfy our needs.

Efficiency. The whole process becomes more effective because of customization by only picking necessary processing units rather than a general process for all tasks.

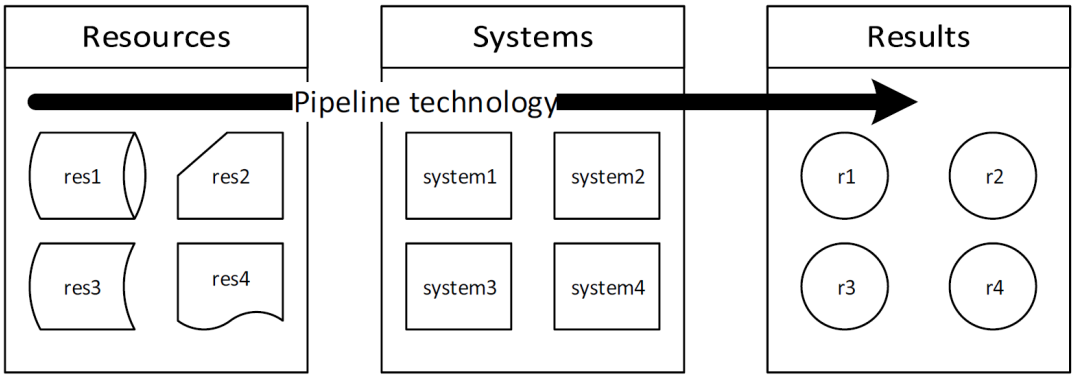

FIGURE 2. Basic idea of pipeline technology

With pipeline technology, we focus on developing a series of technical tools to achieve a resource-oriented process, and it not only can be freely combined but also be compared the pros and cons of different processing units so as to deal with the needs of pluralism and diversified problem.

A variety of resources parsing tools, including natural language processing tool, ontology evolution tool, semantic search engine, etc. are developed and provided to achieve free combination for parsing multiple formats of data resources. Natural language processing tools provide a variety of NLP tools, such as word segmentation, POS tagging, syntactic analysis, semantic analysis, semantic role labeling, and discourse analysis tools. Corpus processing tool is also provided to proofread the processing results due to some errors existing in Chinese because of the diversity of Chinese expressions. Ontology evolution tool aims to build a professional, dynamic, and interrelated knowledge structure for users in the information field. Semantic search tool uses a dual indexing mechanism: full-text indexing and semantic indexing. Ontology evolution tool and semantic search tool form a life cycle that enables the implementation of semantic annotation with domain ontology and ontology evolution with semantic annotation results. The semantic indexing does not index the content based on the whole document, but splits the document into different particle sizes (in paragraph level or sentence level), aiming to manifest the semantic content of the knowledge fragment in the form of semantic index.

4. **Platform Development.** Based on the platform design concept, the architecture of the platform consists of four layers: view layer, application layer, supporting layer and

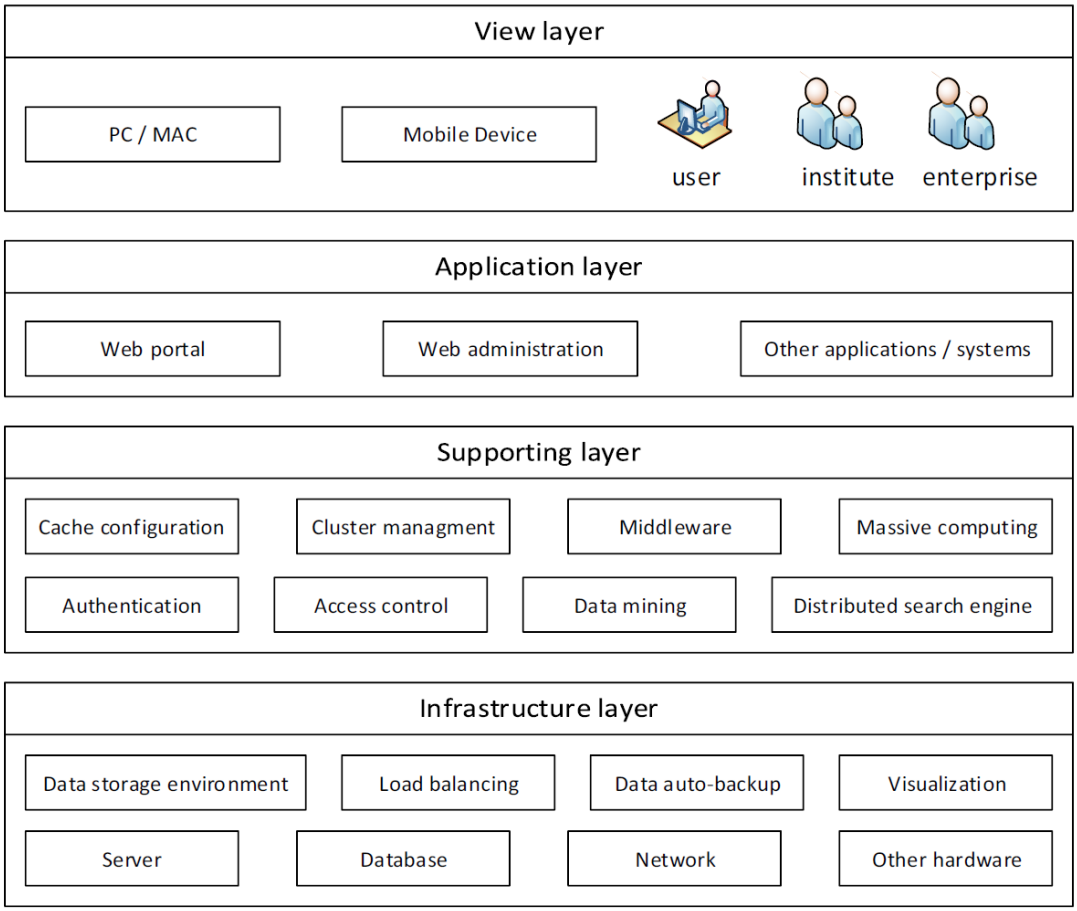

FIGURE 3. Framework of the platform

infrastructure layer. The view layer displays the information of the platform such as web portal, and web administration; the application layer controls the platform's functionality by conducting detailed processing; the support layer includes cache configuration, middleware, and authentication services; and the infrastructure layer includes the data persistence mechanisms, process management, visualization, etc.

The platform provides crowdsourcing, crowd-funding, merchandising business processes, online communication combined with offline activities, and comprehensive Internet information resources and information services. It integrates application systems, data resources and Internet resources into a single information management platform with a unified user interface presented to users to establish information retrieval channels. These services provide expert assessments or the corresponding decision-making basis for the user to do technology forecasting and technology decision-making reference.

It main features include:

1) one-stop-service: through visualization, intuitive entrepreneurial step demonstration, decomposition and show with each step corresponding solutions and tutorial programs.
2) core services: with forming a team, seeking experts, and seeking investment services to enable entrepreneurs to find the most feasible solutions.
3) customer community: information exchanges between the same team, and automatically push industrial related information to deepen mutual communication within the team and information industry.
4) new trend report: offer entrepreneurs deep industry custom reports to grasp the first-hand industry information.

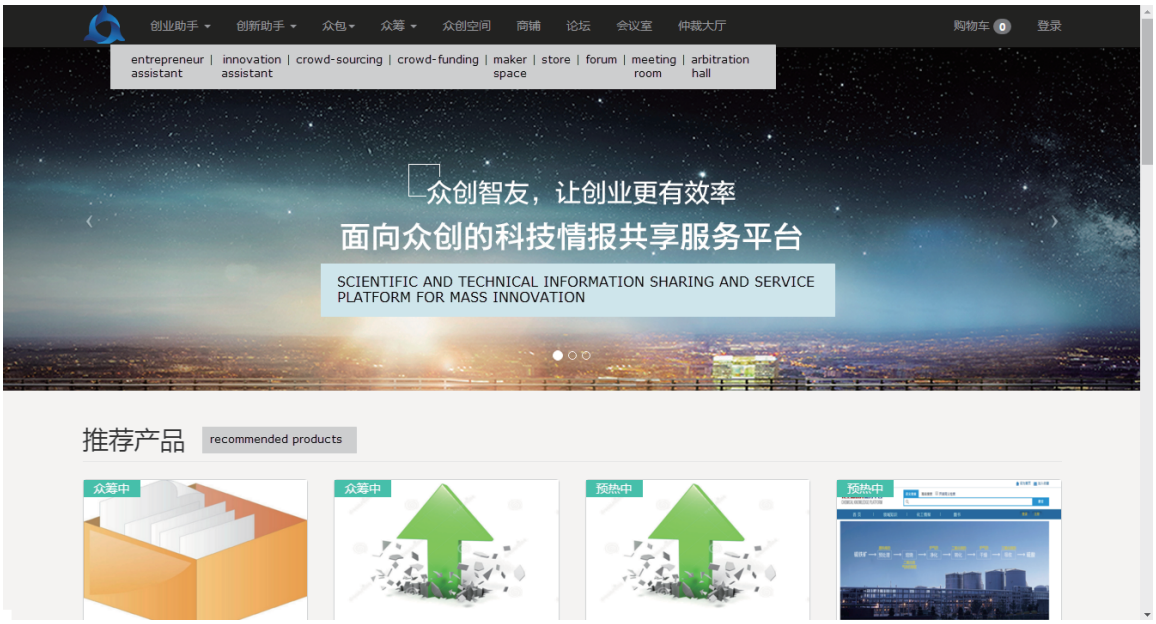

FIGURE 4. Homepage of the platform

5. **Application.** This study conducts application demonstration in chemical field. Chemical Industry Press (CIP) plans to process a batch of books to make preparation for knowledge services, and posts a task in the platform. A maker with deep understanding of knowledge service and natural language processing background notices this task, and proposes a processing model in order to deeply process the scientific and technical literature in CIP. The customized process model is constructed as follows:

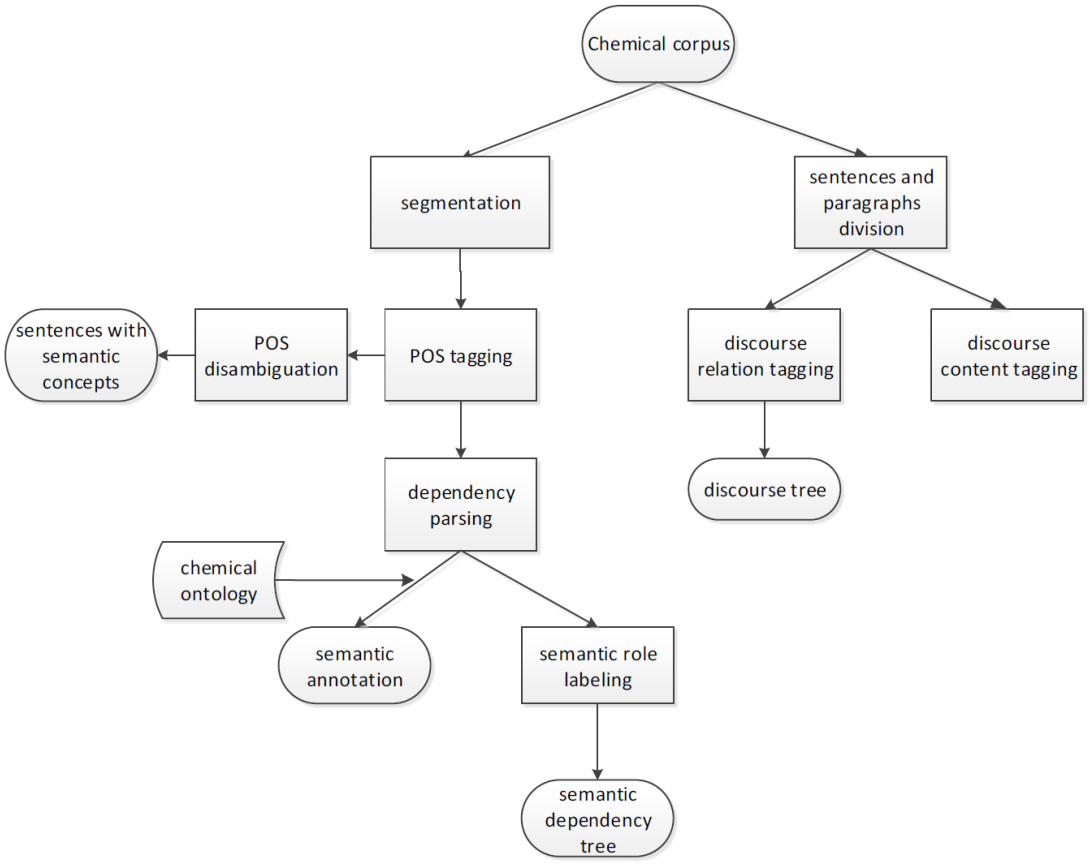

FIGURE 5. The customized process model for processing CIP data

This model includes the entire process of natural language processing and semantic annotation to achieve the desired requirements based on the nature of the task or application control requirements, methods employed, and the request output. Based on the idea and key technology above and the actual situation of CIP, the following processes are conducted with the platform: CIP launches a project to build a chemical knowledge database; a maker only with ideas bids for the project, buys electronic chemical dictionaries, books, and authorization to access the tools necessary (ontology evolution tool, semantic search engine, etc.) for constructing the knowledge database from the platform, finds workers to edit, and proofread the results in the platform by crowdsourcing, and finally delivers the database to CIP to complete the project.

FIGURE 6. Homepage of the chemical knowledge processing platform

6. **Conclusions.** In this paper, we proposed a platform design concept that really simulates the whole process of a maker with creative solutions, but with no capital, resources, condition nor technology to start an entrepreneurship, and then developed the scientific and technical information sharing and service platform for mass innovation, covering the whole entrepreneur process as well as the needs for technical resources. And through application in chemical field, this platform fulfils its design concept to enable a maker with excellent idea to complete the whole innovation process with the platform. In the future, we intend to proceed along two lines in parallel: on one hand, to integrate more resources and tools into the platform; on the other hand, to broaden its application fields other than information industry.

Acknowledgment. This work is partially supported by National Key Project of Scientific and Technical Supporting Program No. 2013BAH21B02 and China Central Fundamental Research Program No. 522014Y-3354. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] Y. Liu, X.-F. Chen, Z. F. Sui, Y. W. Hu and Q. L. Zhao, Research on semantic method of library resources' organizing, *ICIC Express Letters*, vol.5, no.4(A), pp.1011-1017, 2011.

[2] S. Scholten and U. Scholten, Platform-based innovation management: Directing external innovational efforts in platform ecosystems, *Journal of the Knowledge Economy*, vol.3, no.2, pp.164-184, 2012.

[3] Z. F. Sui, Y. Liu and Y. W. Hu, Extracting hyponymy relation between Chinese terms based on term types' commonality and sequential patterns, *ICIC Express Letters*, vol.3, no.4(B), pp.1233-1238, 2009.

[4] Z. F. Sui, K. Wei and T. Ye, Synchronously extracting instances and attributes for the concepts from the web, *International Journal of Knowledge and Language Processing*, vol.3, no.3, pp.1-17, 2012.

[5] Y. Liu, X. F. Chen, S. J. Li and R. J. Wang, A semantic analyzing method in the field of technological literature, *ICIC Express Letters*, vol.5, no.9(A), pp.3225-3230, 2011.

[6] Y. Liu, H. Q. Shi, D. J. Zheng and Y. Huang, Study on semantic annotation for professional literature, *ICIC Express Letters, Part B: Applications*, vol.5, no.5, pp.1383-1389, 2014.

[7] Z. F. Sui, Y. Liu, J. Zhao and H. Zhang, The development of an NLP-based Chinese ontology construction platform, *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, pp.190-194, 2010.

[8] J. Li, R. J. Wang and Y. Liu, Research on semantic metadata online auxiliary construction platform and key technologies, *Chinese Lexical Semantics*, Springer, Berlin Heidelberg, 2013.