

Development and Application of Subject Knowledge Environment (SKE) in Chinese Academy of Sciences

中国科学院专业领域知识环境

Jinxia Huang

Wen Song

Yi Liu

(National Science Library, Chinese Academy of Sciences)



About SKE

1. The need for SKE from the researchers in Chinese Academy of Sciences

- An effective tool on knowledge organization / knowledge discovery
- A domain-focused platform on knowledge management / knowledge sharing

2. Related work

Literature review

Some researches on Virtual Research Environment (VRE), Semantic Web, Scientific Workflow and Knowledge Organization System (KOS) started fairly early, during the 2008 and early 2009.

Review on Virtual Research Environments

Virtual research environments (VRE) are core to e-Research and help researchers within and across specific disciplines to manage the increasingly complex range of tasks involved in carrying out their research. Current studies on VRE pay attention on these points: better understanding on the scientific workflow, integration/interoperability with widely national or domain e-infrastructure, implementing the interface into the digital library framework, embedding the software of equipments that researchers need in the scientific workflow, satisfying the special needs on communications and collaborations.

Review on Semantic Web

Many research groups are currently maintaining their group portals using Semantic Web technologies. The researches on Semantic Web from 2007-2008 were focused on semantic annotation, semantic inference, semantic search, and semantic web services. The study showed that the research of semantic web in the Virtual Knowledge Environment focused on targeted research, and the comprehensive applications were few, and the practical application focused on small-scale applications.

Review on Knowledge Organization System (KOS)

The roles of KOS in tagging, browsing and searching are obviously. The researches on KOS were focused on using classification and glossary, metadata, ontology and semantic relationships to construct KOS. Currently KOS is developed from Natural language to Controlled language. KOS is becoming data resources, and is also becoming linked data. Ontology is recognized as the top level of KOS and the knowledge based system. With the development of the semantic web, the research and the application on KOS will obtain a new peak.

Survey

Some surveys on KOSs and ontologies, contents and services in CAS institutes' websites, and Social Network Services (SNS) were also made.

Survey on KOSs and ontologies

In order to know whether the KOSs and ontologies existed can be reused in the development of SKE, two surveys were made: one is on KOS (different types of resource ontologies and domain ontologies, or ontologies developed by the classifications), the other is on the Semantic Webs in which ontologies are the critical parts. The results showed that it was very difficult to reuse the existed ontologies in SKE because of the different goals and the different granularity of the ontologies.

Survey on CAS institutes' websites

In order to model the classes and properties in SKE ontology, 20 institutes' website of CAS were surveyed on the contents and the services which include the system structure, information and services in institutional websites. More than 2,000 terms were collected, such as people, project, communication, publication, and some of them have high-frequency. These high-frequency terms were used as the classes in SKE ontology. Based on the survey results from the services in these websites, some functions on SKE were also developed.

Survey on Social Network Services (SNSs)

To build the user-interested service webs and reduce the cost on popularizing the web serving (such as SKE), libraries may get some illuminations from social networking services (SNS). We surveyed on some popular SNS websites which have users more than 1 million, such as DouBan, Kaixin, QQ, Facebook, Flickr, YouTube. Some characteristics on the webpage designing, system functions, attracting the users were reviewed. Based on this review, some SNSs were applied on the web building of SKE.

3. Functionality of SKE

- (1) SKE provides the services on the dissemination of scientific information and academic products through supporting academic groups to construct the community in which the scientists communicate with each other.
- (2) SKE supports the integration of domain knowledge information which is based on the application of mash-up and ingest tools and supports the collaboration of librarians and the end-users on selecting, collecting, organizing and utilizing domain knowledge information.
- (3) SKE supports the librarians to provide the real-time and high-valued knowledge services which are facing to the fundamental research areas, grand projects and important academic groups.
- (4) SKE supports the researchers to organize and manage the knowledge resources basing on their needs and preferences and basing on the help from professional librarians.
- (5) SKE supports the knowledge mining and knowledge discovery basing on one SKE or the integration of more than one SKE.

As a semantic tool, SKE is driven by ontology, which integrates and organizes all kinds of information needed in the research life-cycle (Figure 1)

The research life-cycle may include some key points, such as applying for projects, literature review & indexing, identification of collaboration, project management, scientific workflow, communication, dissemination & publication. In each point, scientists may need different kind of information. For example, when a scientist applies for a project, she/he needs the information about the funding opportunities; when a scientist wants to identify collaborations, she/he needs the information about the researchers or organizations in the same research field. Based on these different kinds of information, SKE ontology was generated to support the linked data of the information mentioned above and support the users to develop SKEs with different application aims, such as institute SKE, project SKE.

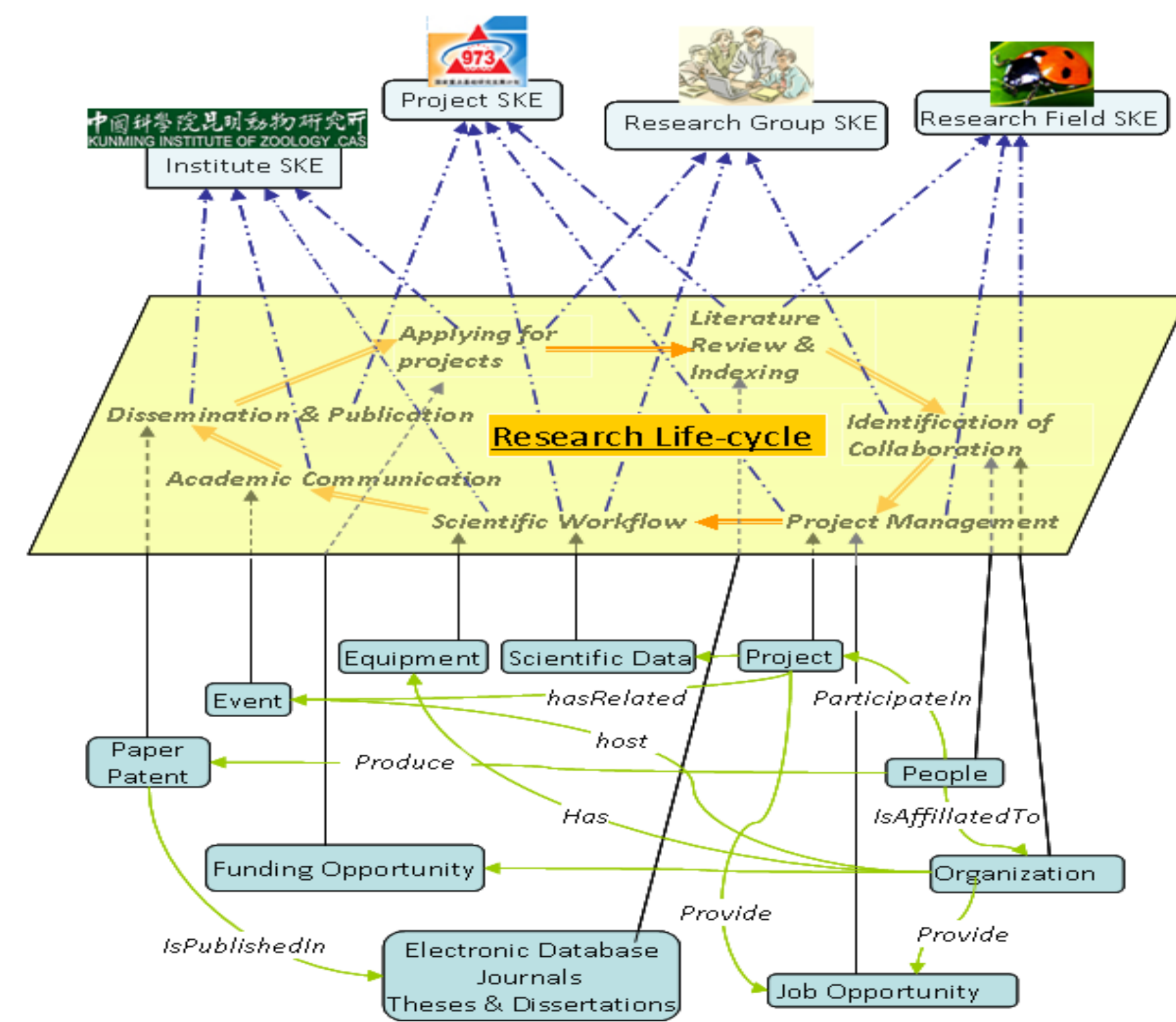


Figure 1 Context and contents in SKE

SKE overall architecture

The component "user application" provides the access for the users and third-systems, the ontology-based editing template for the users to release the information which is structure-organized and stored in SKE, the open AIP interface for the remote applications (software agents) to access the contents in the knowledge base, and the function on mashup for SKEs to dynamical integrate the contents from the third systems.

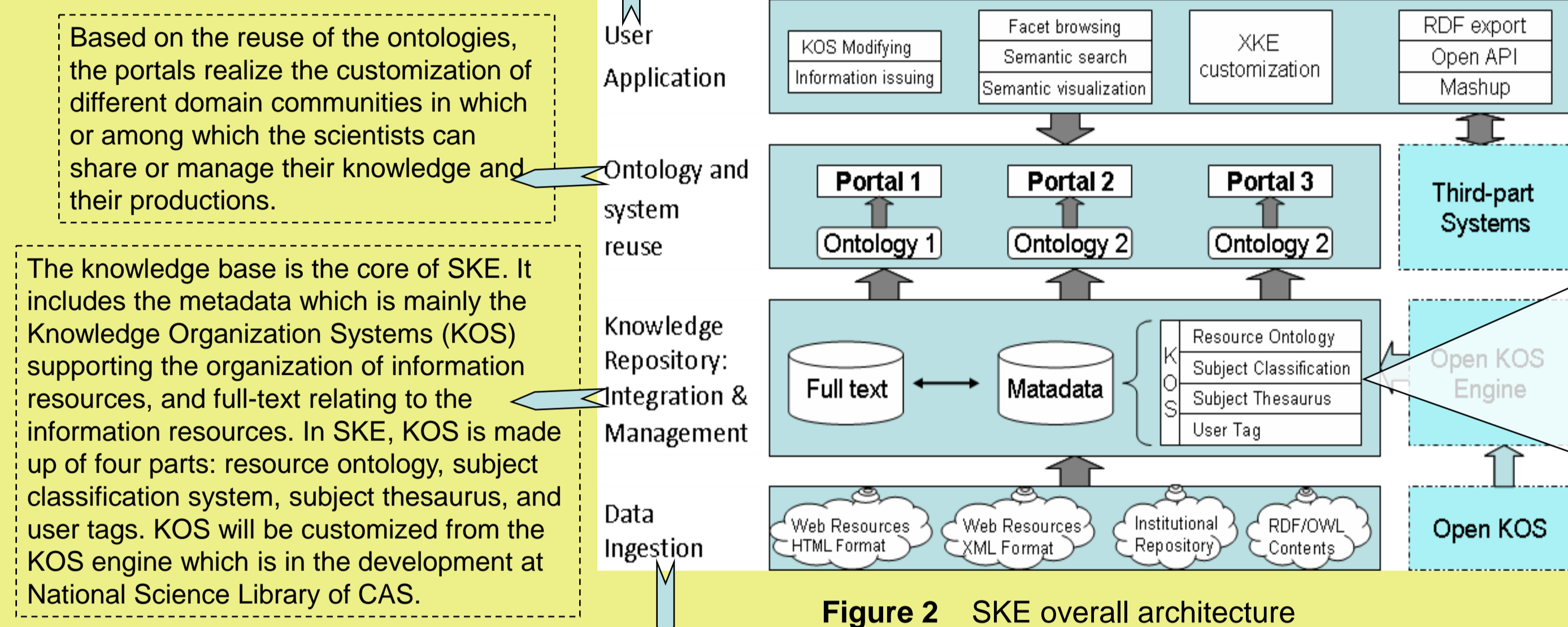


Figure 2 SKE overall architecture

The main function on the component "Data Ingestion" is to facilitate the collection, ingestion and encapsulation of the data from Institutional Repository, websites of top institutes and other information systems automatically. It provides different templates and wrappers to handle some common data formats, such as HTML, XML, and to integrate these data into SKE.

SKE System development based on Vitro

1 Vitro' localization and system development

- 1) Rebuilding the webpage
 - 2) Translating into Chinese
 - 3) Developing the SKE application
- mechanism including four types of user, three types of object, three types of permission to support the customization of SKE

2 Development on new functions

- 1) Developing the class trees for the users easily adding the instances
- 2) Developing the browsing- road map
- 3) Developing the portal customization
- 4) Providing the cross-search into other systems such as NSL Crosssearch System
- 5) Providing the functions on uploading or downloading files in an instance
- 6) Adding some reasoning rules
- 7) Adding RSS (Really Simple Syndication)

Main functions and services provided by SKE (SBKE as an example, Figure 3)

- 1) Providing the user to browse and search the science & technology information which is shown in the different types, and browse them in other systems from SKE.
- 2) Providing the relationship between different information, providing the browse-road map, providing the derived results.
- 3) Developing an open platform for the librarians to collect and integrate the information.
- 4) Providing the academic communities for researcher to communicate their knowledge and share their research productions.
- 5) Developing the individual SKEs becomes easy for the users.
- 6) Providing some "SNS".

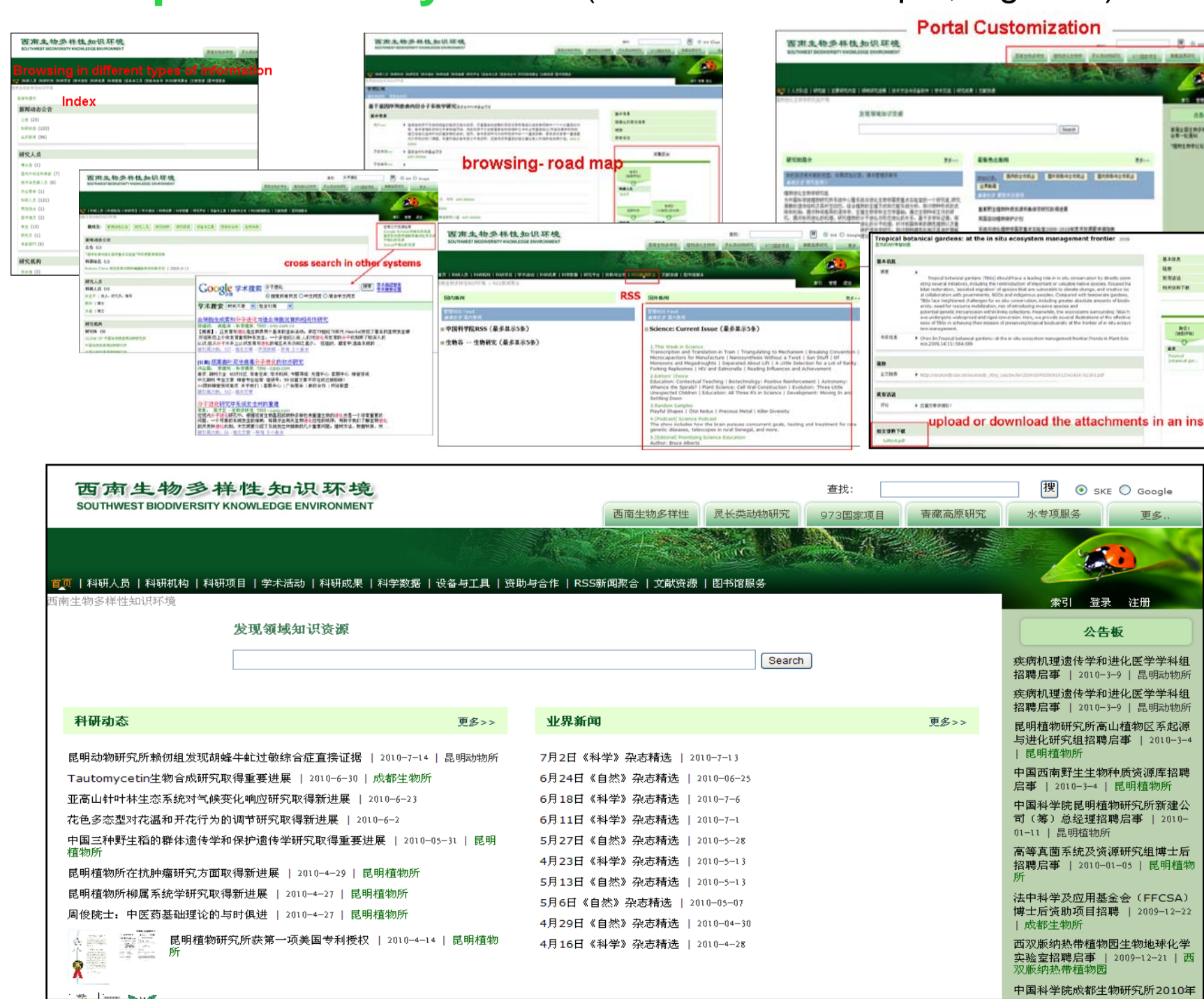
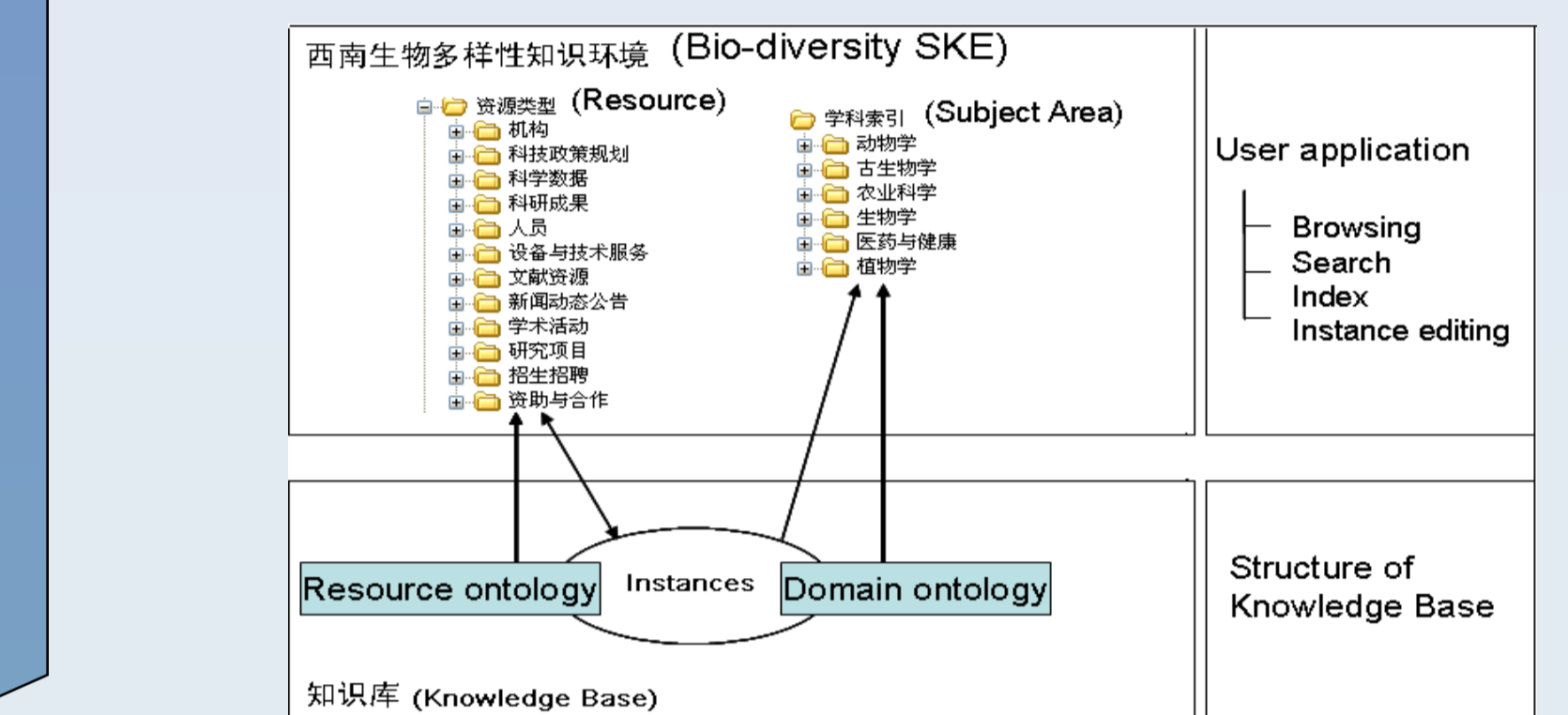
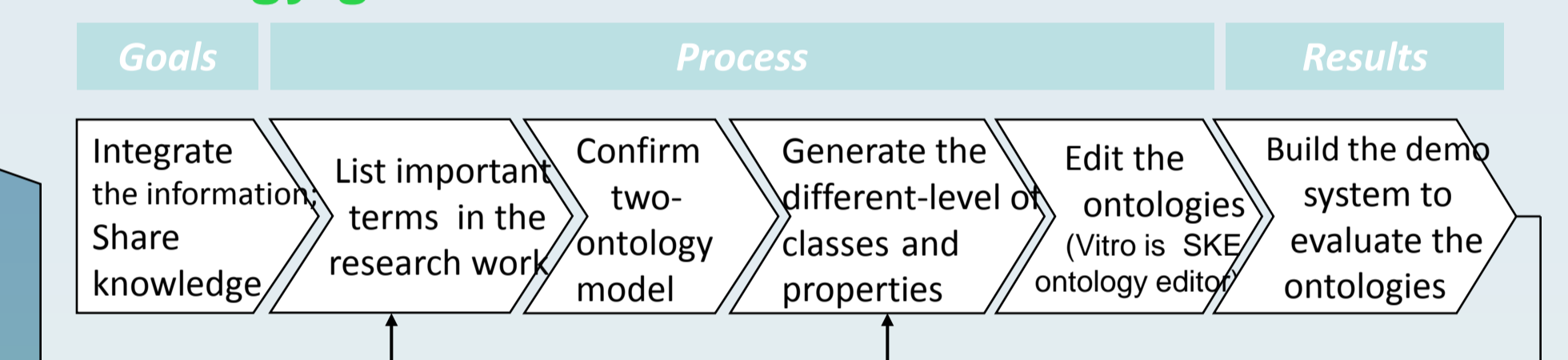


Figure 3 Homepage of Bio-diversity Knowledge Environment (SBKE)

Ontology-based KOS model



Ontology generation



Application of SKE

- Institute's SKE which aims to support the knowledge sharing in an academic organization
Biomedical and Health Knowledge Environment (<http://health.ske.las.ac.cn>)
Qingzang Plateau Knowledge Environment (<http://ske.las.ac.cn/qingzang/>)
- Project's SKE which aims to support the information communication in a project group
Special Water Project Knowledge Environment (<http://ske.las.ac.cn/water/>)
- Research group's SKE which aims to support the information/knowledge communication in a research group
Plant Evolutionary Biology Knowledge Environment (<http://ske.las.ac.cn/botany>)
- Domain knowledge environments which aims to support the knowledge communication and dissemination in a special subject area
Primates Knowledge Environment (<http://sbke.las.ac.cn/monkey/>)



Acknowledgements

This work is funded by Chinese Academy of Sciences. We thank every member in SKE team besides Jinxia Huang, Wen song, Yi Liu and Zheng Liu: Rui Yang, Yijie Tang, Zijian Zhou, Shinan Zhang, and Li Jing.

SKE in CAS <http://ske.las.ac.cn>