

Cultural Knowledge Co-Creation on Social Networking Paradigm

Virach Sornlertlamvanich

National Electronics and Computer Technology Center
Pathumthani, Thailand
virach.sornlertlamvanich@nectec.or.th

Thatsanee Charoenporn

Digital Media Informatics Laboratory
National Electronics and Computer Technology Center
Pathumthani, Thailand
thatsanee.charoenporn@nectec.or.th

Abstract— Since 2006, by the Ministry of Culture, there was an effort in creating cultural portal. Each provincial cultural office was assigned to survey and report any cultural practice according to a designed template. The technological potential of each office was so different and most of them had to relied on the local system developer and service provider. As a result, the collected contents cannot fulfill the requirement of documentation standard and service level of media presentation. In contrast, cultural knowledge is redefined to an agreed structure for interoperability and computability. Moreover, the contents must reflect the up-to-date daily practice and thoroughly be covered by any individuals. We reuse the existing contents by recovering any typos and format distortion. The social networking system is prepared to allow individual participation to co-create the contents under an authorized supervision. As a result, more than 80 percents of the original contents have been recovered. Each provincial office is able to strategically plan and co-create their own contents to establish the total cultural knowledge.

Keywords— component; cultural knowledge; digital cultural communication; co-creation; social networking

I. INTRODUCTION

Under the collaboration between Ministry of Culture and Ministry of Science and Technology by National Electronics and Computer Technology Center (NECTEC), a collection of cultural knowledge has been embodied since 2010. We did not start the work from scratch. Actually the work has been done since 2006 by forming a set of servers individually operated by each province. Each province has to take care of their own contents about their responsible area. The initiative has been carried out for the purpose of creating a reference site of the local cultural knowledge. The distributed system is an aim to decentralize the management and to maintain the uniqueness of each specific area. However, there is a trade-off between the independent design and cost of maintenance that covers the service operation, interoperability and integrity. There are currently 77 provinces in Thailand, and each province is allocated an office for provincial cultural center. With the approach of the above-mentioned distributed system, it is too costly to maintain the service and the standard for data interchange.

The newly designed platform-based approach for digital cultural communication has been introduced. It is to build a co-creative relationship between the cultural institution and the community by using new media to produce audience-focused cultural interactive experience [1]. First, we

collected the existing provincial cultural knowledge and convert them to conform to a standardized set of metadata. This is to prepare the cultural knowledge for an open data schema and interoperability. The metadata is defined to follow the Dublin Core Metadata Element Set with some additional elements to fulfill the requirement information during the recording process. Second, we assign representatives from each province and train them to be a core cultural content development team for community co-creation. The contributed content needs to be approved by the core team before public visibility. Third, the cultural knowledge will be put on service to the audience of such scholars, who may be interested in the cultural practice, or business developers who may benefit from attaching the cultural knowledge to their products, or tourists who may seek for cultural tourism.

These cultural media assets will be linked and annotated by a governed conceptual scheme such as Asian WordNet [2]. The semantic annotated and linked data will be serviced as a fine-grained cultural knowledge for higher level of applications. The new media for recording the cultural knowledge is in the form of audio, text, video, animation, image incorporated with GPS data for visualization on the map.

II. CULTURAL KNOWLEDGE PORTAL

The efforts in archiving cultural knowledge can be found in several regions. Mainly the purpose is to store and disseminate knowledge in forms of text, image, video, audio, and animation. Therefore, several powerful software utilities such as search engine, pattern recognition, summarization, question and answering as an efficient human interface are developed. Multilingual supports can also be found in the non-English portals.

The Internet Archive is a non-profit that was founded in 1996 to build an Internet library. Its purposes include offering permanent access for researchers, historians, scholars, people with disabilities, and the general public to historical collections that exist in digital format. The Internet Archive includes texts, audio, moving images, audios and software as well as archived web pages for education in the collections, and provides specialized services for adaptive reading and information access for the blind and other persons with disabilities [3].

Europeana project was launched in 2008, with the goal of making Europe's cultural and scientific heritage accessible to the public. The project is funded by the European

Commission. It is based in the National Library of the Netherlands, the Koninklijke Bibliotheek. Application Programming Interfaces (APIs), virtual exhibitions and better search capabilities are provided to improve the accessibility. The archive contains texts, images, videos, and sounds. Currently, it aims to support the accessibility in European languages [4].

In order to showcase the cultural heritage and natural resources of Taiwan to the world as well as complement the existing Chinese portal (<http://digitalarchives.tw>) which demonstrates TELDAP's (Taiwan e-Learning and Digital Archives Program) accomplishments, Taiwan has created a new "Digital Taiwan - Culture & Nature" English portal. It provides the global community with an in-depth and convenient way of appreciating Taiwan's culture and digitization achievements. As the contents of archives increase, so the functions of the portal will expand to meet the public's expectations. In addition, the portal will make Taiwan's academic and digital archives activities more visible to the world community. The program consists of Taiwan e-Learning & Digital Archives Portal; Union Catalog of Digital Archives, Taiwan; Digital Taiwan Culture & Nature; Exhibition of Cyber Island, Taiwan; Learning Object Repository in Taiwan; e-Learning Standards; e-Learning Quality Certification Center. The learning object repository is one of the anticipating showcases [5].

Many other similar efforts can be found in Info-plosion project 2005-2011 (new IT infrastructure for the information-explosion era) which aims to establish the fundamental technologies to deal with the information explosion phenomena [6]; the World Digital Library (WDL) which makes cultural contents available on the Internet in multilingual format, significant primary materials from countries and cultures around the world [7].

Digitized Thailand gives precedence to the technology for contents creation and software utilities development [8]. Resulting from the efficiency in archiving the digital contents, we are anticipating a service platform for digital cultural knowledge dissemination. The service facilitates the content distribution and supporting utilities that makes appropriate interlink between the related cultural knowledge. Cultural knowledge is created in the co-creation manner. It is open to public but supervised by the authorized group of cultural knowledge curators. The provincial cultural offices are in charge of content preparation. They may form a team to provide a guideline for cultural knowledge collection and encourage the communities to participate in appealing their particular practice. The qualified and appropriate contents will be approved by the curators to open for public view. To do this, an online social networking system has been developed to accumulate the country wide cultural knowledge. Any individuals can be a member to upload and tell their stories.

To provide a scope for collection, the target subjects of cultural contents are categorized into four main categories namely, cultural person or organization, cultural artifact, way of life, and cultural site. Each subject contains a certain number of categories to elaborate the grouping of the contents. One record may belong to more than one category.

Followings show the list of categories to assign to the subject of the contents.

- Cultural Person / Organization
 - Artist
 - Scholar
 - Religious Monument
 - Writer/Author
 - Society/Association/Foundation
 - Cultural Network
 - Cultural Unit
- Cultural Artifact
 - Archaeological Finds
 - Artwork
 - Visual Art
 - Book/Press
 - Audiovisual Media
 - Utensil
 - Costume
- Way of life
 - Ethnic
 - Religion and Belief
 - Tradition and Rite
 - Language and Literature
 - Local Wisdom
 - Performing Art and Music
- Cultural Site
 - Archaeological Site
 - Historical Park
 - Historical Site
 - Architecture
 - Religious Place
 - Museum
 - Library
 - Archive
 - Monument
 - Theater
 - Tourism Location

III. DIGITAL CULTURAL COMMUNICATION

To build a co-creative relationship between the cultural institution and the community, we provide an open social networking system that the authorized provincial curators can work with the individuals in the community to create the cultural knowledge. Major new media such as text, image, audio, video and animation are allowed to upload and be a part of story telling. The community is encouraged to participate in content co-creation to produce the audience-focused cultural interactive experience. The digital cultural communication will therefore be formed up on the linkage between the cultural institute and the community.

Figure 3 shows the corresponding activities of how the institute interacts with the community and the audience. The institute initiates the task by collecting the cultural contents from the existing or newly generated resources. The structure of the contents has to conform to the defined standard. The contents are made available to the audience through a particular set of viewpoints namely, full-text or subject search, and location-based access. Contents are also necessarily rewritten to serve the interests of the audience in

terms of whatever for studying, sightseeing, produce or service finding, etc. The institute and the audience will then be bound to the community through the generated contents. The interrelation among the institute, the audience and the community will sustainably form the digital cultural communication.

Followings are three steps we executed to draw the existing contents to the standardized cultural knowledge service.

Step 1: Cultural knowledge curation

It is the first step to bring in the existing contents. The data will be clean up to follow an open standard guideline. This is to make the contents conformed to the requirement from others in terms of interoperability. The metadata is designed to follow the Dublin Core Metadata Initiative (DCMI) guideline [9].

Step 2: Community co-creation

An online social networking system is provided for the provincial cultural office to work with the community to co-create the cultural contents. The audience-focused provincial intellect can extract the interest of both the community and the visitors.

Step 3: Cultural knowledge service

The collection of cultural contents will be analyzed to create the linked data between the related contents to produce a set of cultural knowledge. As a result, the cultural knowledge platform for application service development will be developed.

A. Cultural Knowledge Curation

The existing cultural data has been collected and cleaned up to conform to the designated standard metadata. The absent data are supposed to be revised and augmented by the experts from the Ministry of Culture. A few tens of thousands of records have been collected but most of them are captured in a coarse-grained image. Narratives and images are revised by a group of trained expert to create a seed of standardized annotated cultural knowledge base. Some new records have been added together with animation, video, panoramic photograph, etc. New technique in capturing the cultural image is aggressively introduced to create value added and gain more interest from the audience.

The standardized annotated cultural knowledge base is presented through a set of viewing utilities to the audience. Filter according to the location and province is prepared for customizing to page for each province. This is to allow the unique presentation of each province. The administration of each province will be responsible to its content correctness and coverage. Actually, the attractive presentation and narrative are required to attract the audience.

B. Community Co-Creation

Social networking system is introduced to invite the participation from the communities. The institution representatives are actively encouraged to create their own community. The results from the community co-creation will keep the content maintained and clean up to compete with each other.

It is significantly that the provided framework can encourage the data accumulation and fulfill the needs from the audience. Community co-creation also collects the feedbacks of actual requirement that can encourage the improvement of the quality of the content. Institution plays an important role in mediating between community and the audience. As a result, the multiple types of content are generated on the designated standard. The annotated metadata can be used as a guideline for higher level of data manipulation such as semantic annotation, cross language and link analysis.

C. Cultural Knowledge Service

Cultural knowledge is created based on the defined standard to maintain the interoperability and openness. Figure 4 shows the XML exchange format. Based on the Dublin Core Metadata Element Set, Version 1.1 [9], additional metadata has been proposed to elaborate the annotation scheme, especially for extended computational usages such as, ‘dt.tag’ is for adding tag set, and ‘dt.price’, ‘dt.operation.daytime’, ‘dt.suggestion’ are for particular site service information.

The culture knowledge service retrieves the target contents via the full-text or subject search, and location-based accessibility as shown in Figure 1.

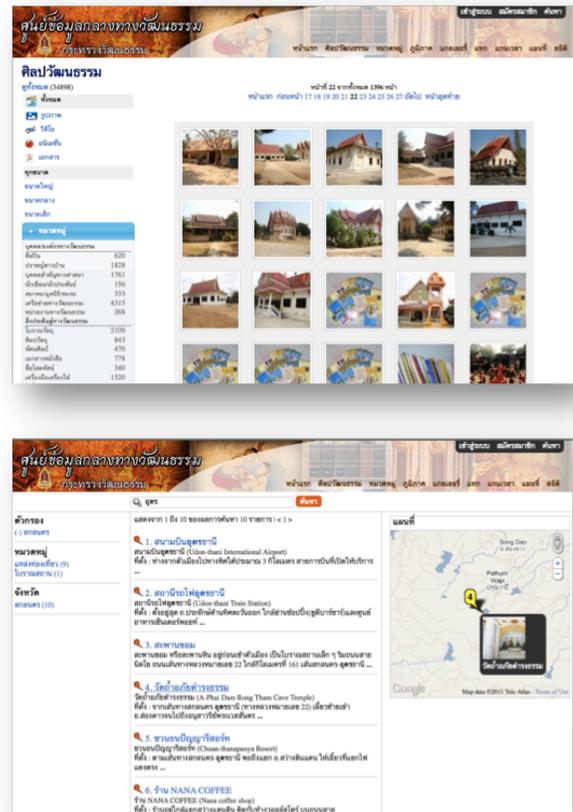


Figure 1. Search result of full-text and subject search

Figure 2 shows a sample of the retrieved content. The metadata is defined to structure the content that provides multiple images, location map, title, description, tag, subject, and so on. Following the XML exchange format as shown in Figure 4, various applications can therefore be developed to provide the services related to the cultural knowledge.

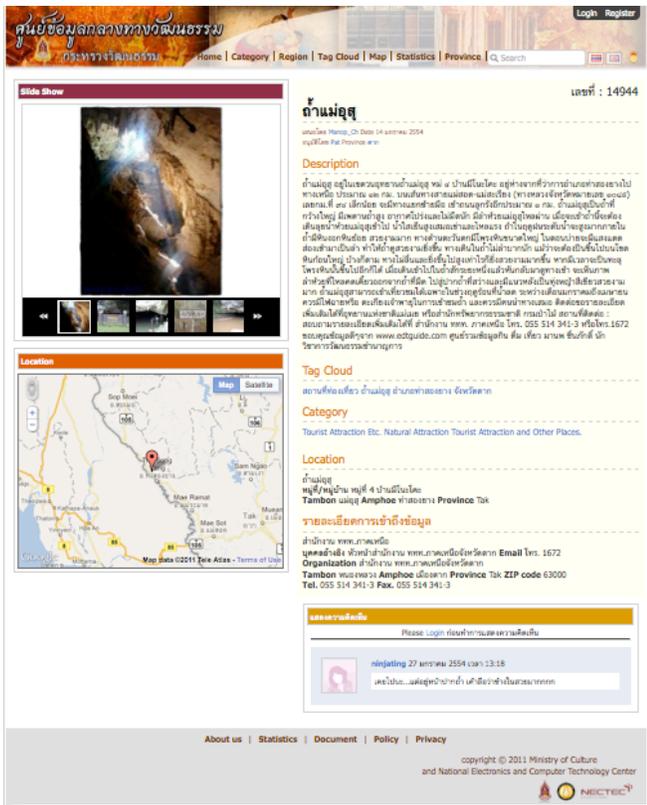


Figure 2. A sample of retrieved content

IV. ACCUMMULATION

In November 2010, the portal is started from the existing records crawled and converted from the dispersed database. Each provincial cultural office is assigned to revise and improve their records by targeting at 1,000 records for each province. Table 1 shows the monthly growth of the records.

TABLE I. MONTHLY GROWTH OF NUMBER OF RECORDS

Month-Year	Approved	Pending	Total
Nov. 2010	9,966	600	10,566
Dec. 2010	234	44	278
Jan. 2011	2,805	180	2,985
Feb. 2011	13,546	458	14,004
Mar. 2011	4,304	57	4,361
Apr. 2011	3,428	36	3,464
May. 2011	28,692	1,029	29,721
Jun. 2011	17,909	807	18,716
1-25 Jul. 2011	2,968	3,460	6,428
Total	83,852	6,671	90,523

Starting from the existing data, the amount of records in the first month is quite high. Nonetheless, the growth in the following months can be kept constantly in a high pace. This shows the high contributions from the offices that are institutionalized under the ministry. Currently, 53 from the total of 76 provinces can fulfill the target of 1,000 records per province. The data is collected from social network participation and mainly from the institutionalized counterparts therefore the data can constantly be collected. But, we found some crucial problems getting from the audience feedbacks in terms of the usefulness.

- Information for management against information for public use. Some records are not within the interest of public use. Information for internal management should be separated from this cultural knowledge portal.
- Title naming. There is a trend to describe the record in the title rather than giving the topic. Short and
- Description. The description is too deep in academic detail rather than attractive description to draw the audience's attention.
- Photo and other media. More attractive photos or other media are required. The total relation between the photos, description, title and tag are required.

These are the major topics that need to be improved in utilizing the social network in data collection. This is one of the reasons why the approval procedure is introduced. Every submitted record is reviewed for appropriate expression and correct information. The pending records need revision before releasing for public viewing.

V. CONCLUSION

Digital cultural communication is framework that we introduced to create the relation between institute, community and audience. Provincial cultural office pays an important role to attract community to participate in cultural content co-creation based on the interactive requirement from the audience. The collected cultural contents are connected to develop linked data for the persistent cultural knowledge resource providing. The cultural knowledge is defined on an open standard to maintain the interoperability and openness. The future service will perform as a platform for cultural knowledge providing and service application development. Linked data and the tools to create the linked data are the expecting result from the existing resources. The cultural based service applications are also in our range of future development.

ACKNOWLEDGMENT

The project is a part of Digitized Thailand program, dedicated to cultural content collection and cultural knowledge creation. The project is subsidized by the Ministry of Culture, Thailand under the agreement of the technology support from the National Electronics and Computer Technology Center (NECTEC), Ministry of Science and Technology.

REFERENCES

- [1] A. Russo and J. Watkins, "Digital Cultural Communication: Audience and Remediation" in *Theorizing Digital Cultural Heritage* eds. F. Cameron and S. Kenderdine, Cambridge, Mass., MIT Press, 2005.
- [2] V. Sornlertlamvanich, T. Charoenporn, K. Robkop, C. Mokarat, and H. Isahara, "Review on Development of Asian WordNet" in *JAPIO 2009 Year Book*, Japan Patent Information Organization, Tokyo, Japan, 2009.
- [3] The Internet Archive, <http://www.archive.org/>
- [4] Europeana Project, <http://europeana.eu/>
- [5] Taiwan e-Learning and Digital Archives Program (TELDAP), <http://www.teldap.tw/en/>
- [6] New IT Infrastructure for the Information-Explosion Era (Infoplosion), <http://www.infoplosion.nii.ac.jp/info-plosion/ctr.php/m/IndexEng/>
- [7] World Digital Library, <http://www.wdl.org/en/>
- [8] Digitized Thailand, <http://www.digitized-thailand.org/>
- [9] Dublin Core Metadata Initiative (DCMI), Dublin Core Metadata Element Set, Version 1.1, <http://dublincore.org/documents/dces/>

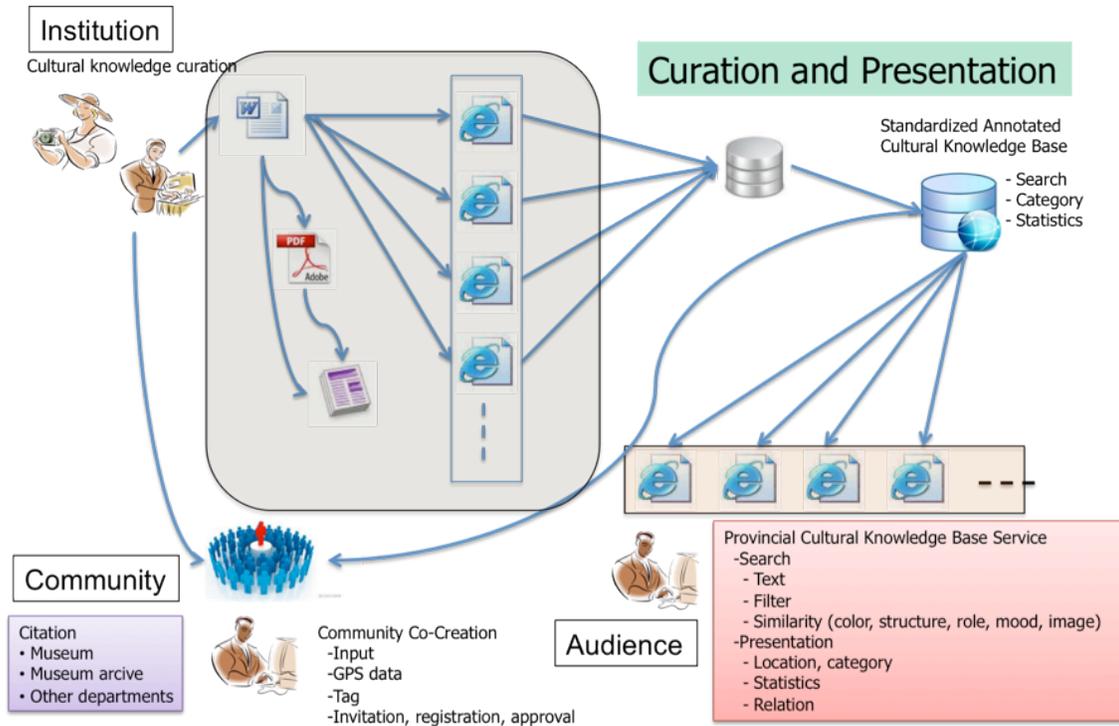


Figure 3. Community Co-Creation Cultural Knowledge Base

```

<collection>
  <record>
    <dc.title>title of the cultural record</dc.title>
    <dt.tag>tag name to determine more possible keywords, there
can be multiple tags</dt.tag>

    <dc.description>content to describe the nature or the
characteristics about the title
  </dc.description>

  <dc.subject>predefined cultural category, there can be
multiple subjects</dc.subject>

  <dc.relation.resource>relating other data files, there can
be multiple resources
  <dc.type>type of the related file, it can be a file of
photo, video, document image, animation</dc.type>
  <dc.source>position of the related file, it can be URI of
the file</dc.source>
  <dc.format.extent type="encoding">encoding type of the
related file, it can be photo (jpg, png, gif, bmp), video
(flv, avi, 3gp, wmv, mov), document image (pdf), animation
(swf)</dc.format.extent>
  <dc.format.extent type="resolution">resolution of the
image</dc.format.extent>
  <dc.format.extent type="size">file size
</dc.format.extent>
  <dc.date.taken>date of image taken</dc.date.taken>
  </dc.relation.resource>

  <dc.coverage.name>address of titled object</dc.coverage.name>
  <dc.coverage.no>address number</dc.coverage.no>
  <dc.coverage.moo>number of division</dc.coverage.moo>
  <dc.coverage.soi>name of lane</dc.coverage.soi>
  <dc.coverage.road>name of road</dc.coverage.road>
  <dc.coverage.tambon>name of region</dc.coverage.tambon>
  <dc.coverage.amphoe>name of district</dc.coverage.amphoe>
  <dc.coverage.province>name of province</dc.coverage.province>
  <dc.coverage.postcode>postal number</dc.coverage.postcode>
  <dc.coverage.telephone>tel number</dc.coverage.telephone>
  <dc.coverage.fax>facsimile number</dc.coverage.fax>
  <dc.coverage.email>email address</dc.coverage.email>
  <dc.coverage.website>URL of the website</dc.coverage.website>
  <dc.coverage.zone>zone of geo position</dc.coverage.zone>
  <dc.coverage.easting>pos of easting</dc.coverage.easting>
  <dc.coverage.northing>pos of northing</dc.coverage.northing>
  <dc.coverage.latitude>latitude</dc.coverage.latitude>
  <dc.coverage.longitude>longitude</dc.coverage.longitude>

  <dc.creator type="author">
    <dc.creator.name>name of the author</dc.creator.name>
    <dc.creator.email>author email address</dc.creator.email>
    <dc.creator.telephone>author tel no.</dc.creator.telephone>
  </dc.creator>

  <dc.creator type="inputter">
    <dc.creator.name>name of the data inputter</dc.creator.name>
  </dc.creator>

  <dc.creator type="approver">
    <dc.creator.name>name of the data approver</dc.creator.name>
  </dc.creator>

  <dc.publisher>name of publisher</dc.publisher>

  <dc.contributor.name>contributor name</dc.contributor.name>
  <dc.contributor.organization>organization name
  </dc.contributor.organization>
  <dc.contributor.no>address number</dc.contributor.no>
  <dc.contributor.moo>number of division</dc.contributor.moo>
  <dc.contributor.soi>name of lane</dc.contributor.soi>
  <dc.contributor.road>name of road</dc.contributor.road>
  <dc.contributor.tambon>name of region</dc.contributor.tambon>
  <dc.contributor.amphoe>district name</dc.contributor.amphoe>
  <dc.contributor.province>province</dc.contributor.province>
  <dc.contributor.postcode>postal no.</dc.contributor.postcode>
  <dc.contributor.telephone>tel no.</dc.contributor.telephone>
  <dc.contributor.fax>facsimile number</dc.contributor.fax>
  <dc.contributor.email>email address</dc.contributor.email>
  <dc.contributor.website>website URL</dc.contributor.website>

  <dc.rights>right on the content</dc.rights>

  <dc.date.last_update>last update date</dc.date.last_update>
  <dc.date.created>date of creation</dc.date.created>
  <dc.date.submitted>date of submission</dc.date.submitted>
  <dc.date.approved>date of approval</dc.date.approved>
  <dc.date.contributed>contribution date</dc.date.contributed>

  <dc.identifier>identifier of the content</dc.identifier>
  <dc.langugae>language of description</dc.langugae>
  <dt.price>entrace fee of the site</dt.price>
  <dt.operation.daytime>operate day-time</dt.operation.daytime>
  <dt.suggestion>suggestion and opinion</dt.suggestion>
  <dt.documentnumber>document number</dt.documentnumber>
</record>
</collection>

```

Figure 4. XML exchange format