Kuwait Geographic Information Metadata Research

Yongping ZHAO and Jamal Al-AKROKA, Kuwait

Key words: GIS, ISO/TC 211, Kuwait, Metadata, Profile

SUMMARY

Geographic information Metadata is one of the important GIS related standards for geospatial information sharing. ISO/TC 211 has been developing the standard since 1995 and more than 400 metadata elements involved in the approved standard by ISO in 2003. Actually, there are only limited metadata elements may be used for GIS in typical countries.

Kuwait Geographic Information Center (KGIC) at Kuwait Institute for Scientific Research (KISR) is one of the top research centers on geographical related information research and management in Kuwait. Several national GIS related projects have been finished in the last more than 10 years. Anyway, due to the GIS projects were implemented via different software platforms and lack of referenced GIS standards in the implementation, much work still need to be done in order to share and use the plentiful geographic and environmental information resources.

KGIC aims to establish an open GIS center to server KISR and Kuwait on Geomatics information sharing such as GIS, RS, and photogrammetry datasets. Based on the data control and data assurance requirement in KIEIN project which funded by UNDP and KISR, a GIS Metadata Draft is suggested to make. This metadata will be originally developed for KIEIN project and later extended to the national level. As the result, Kuwait Metadata Profile (KMP) based on ISO/TC 211 19115 has been developed in KGIC. This profile accepted ISO/TC 211 metadata architecture and used UML structure to define the relationship of metadata entities and elements. All the needed metadata packages, entities, and the elements which can be used for projects oriented management and sharing have been listed in the document by elements dictionary. The other elements which may be used for other countries situation have been diminished from the draft. A metadata software implementation method based on the KMP is also discussed in the paper.

All the KMP research and development procedures are guided by ISO/TC 211 standards and demands. This is to make sure all the geographic information managed by KMP can be shared in KGIC and the outreach in future. Clearly this is a more efficient metadata standard than ISO/TC 211 19115 for the GIS related information management in Kuwait.

Kuwait Geographic Information Metadata Research

Yongping ZHAO and Jamal Al-AKROKA, Kuwait

1. GIS IN KUWAIT

GIS has been widely used in city planning, resources management, and scientific research due to its efficient analysis and visualization capability. Since 1990's, more than 10 GIS related projects have been launched in Kuwait. Kuwait Geographic Information Center (KGIC) at Kuwait Institute for Scientific Research (KISR) has been involved in the GIS projects development and research. Among them, Kuwait Environment Information System Phase I (KEIS I, 1994-1996) & Phase 2 (KEIS II, 1998-2000) (Abdul Nabi, 1997) and Kuwait Soil Survey (1995-1999) (Gerard G., Peter K., Samira O., et al., 2005; KISR 1999) were two master projects in helping to establish Kuwait environmental GIS datasets and architecture.

By the successful implementation of the two GIS projects, Kuwait environmental related data sources which developed in history have been converted to digital formats and stored in GIS database. Anyway, due to the limited international GIS related standards and the weak on the cooperation between GIS projects, the GIS datasets were integrated with different GIS platforms. This resulted in much reedit work and burden on GIS datasets sharing and usage today.

However, due to the fast development of the huge GIS application market and the increasing demands on GIS datasets application, many GIS related new projects were created recently to establish new GIS databases in Kuwait, such as the Development, Analysis and Integration of Hydrogeological Dataset; Bubiyan Island Environmental GIS Project; and the Kuwait Integrated Environmental Information Environment Network (KIEIN) project, etc. According to the experiences accumulated in the previous projects and in order to efficiently manage GIS datasets and sharing them in future, referenced GIS standards are needed in the projects design and implementation.

Standard development is a time and money consuming work. It usually needs many years from ideas development until to get a standard approval. Kuwait has no referenced GIS standards in traditional. Clearly, it is hard to develop GIS series standards overnight. An alter method is to analyze and accept referenced international GIS standards and develop the related profiles. This is a popular method in internationality for standardization cooperation research and standards making.

Since the set up of KGIC in 1990's, GIS has been recognized as one of the key topics in the center. It helps to solve the scattered scientific database issue and integrated them to centralized Geodatabase in supporting 5 divisions related to Food Resources & Marine Sciences, Environment & Urban Development, Petroleum Research & Studies Center, Water Resources, and Techno-Economics research in KISR. KGIC is interested in implementing GIS projects on the national level. Therefore, KGIC aims to establish an open GIS center to server KISR and Kuwait on Geomatics information sharing such as GIS, RS, and

photogrammetry datasets. Based on the data control and data assurance requirement in KIEIN project which funded by UNDP and KISR, a GIS Metadata Draft is suggested to make. This metadata will be originally developed for KIEIN project and later extended to the national level.

2. CONTENT OF KUWAIT METADATA

2.1 ISO/TC 211 Metadata Standard Analysis

Geographic Information/Geomatics standard has been recognized as one of the key research fields in GIS world (Zhao 1999). There are several international organizations which have been involved in GIS standardization research, such as International Organization for Standardization (ISO) Technical Committee for Geographic Information/Geomatics (ISO/TC 211), Federal Geographic Data Committee (FGDC), European Committee for Standardization on Geographic Information (CEN/TC 287), Open GIS Consortium (OGC), and the other organizations. Among them, ISO/TC 211 is the leading organization with international influence. Many countries in the world are following ISO/TC 211 standards and creating their own profiles (Ota, 2001).

Among GIS standards developed by ISO/TC 211, 19115 Metadata Standard is the popular one which has been approved as an international Metadata standard by ISO in 2003. In 19115, more than 400 metadata elements are involved in describing GIS datasets (ISO 19115, 2003). It is a fully developed international Metadata Standard and spanned for about 10 years until it approved.

ISO Metadata Standard has been used in several countries as the national profiles (Ota, 2001). Actually, there are only limited metadata elements may be used for GIS in typical countries. Based on the demands of KGIC information management and KIEIN project, ISO/TC 211 Metadata Standard is selected as the referenced document in making the GIS Metadata Draft. By the revising of ISO/TC 211 19115 Metadata Standard, Kuwait Metadata Profile (KMP) has been developed by the support of KGIC and KIEIN project. This profile accepted ISO/TC 211 metadata architecture and used UML structure defining the relationship of metadata entities and elements. All the needed metadata packages, entities, and the elements which can be used for projects oriented management and sharing have been listed in the document by elements dictionary. The other elements which may be used for other countries situation have been diminished from the draft.

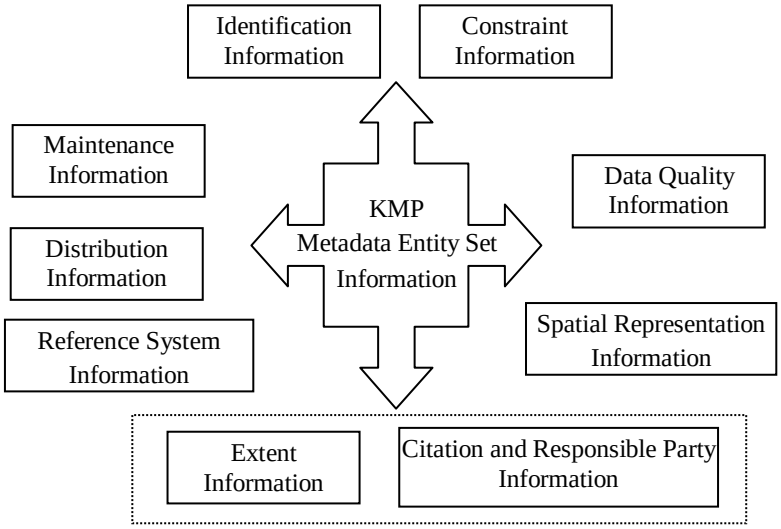

Fig. 1 KMP Metadata Entity Frame

2.2 KMP Content

Kuwait Metadata Profile (KMP) is organized with the same structure as ISO/TC 211 (ISO 19115:2003). According to the KMP application requirement, the first step of the draft will be used for projects management in KIEIN. Therefore several entities in ISO/TC 211 were diminished in KMP. Those deleted entities include MD_ContentInformation, MD_PortrayalCatalogue, MD_MetadataExtentInformation, MD_ApplicationSchemaInformation.

As the result, the following entities are selected in KMP (Fig. 1). KMP Metadata Entity Set Information which named in the profile with MD_Metadata entity is an aggregate of MD_Identification, MD_Constraints, DQ_DataQuality, MD_MaintenanceInformation, MD_SpatialRepresentation, MD_ReferenceSystem, and MD_Distribution. The Citation and Responsible Party Information and the Extent Information are two referenced entities which are mostly used by the other entities. The more detail classes and elements included by each of them are as follows:

Identification information contains information to uniquely identify the data; Constraint information contains information concerning the restrictions placed on data; Data quality information contains a general assessment of the quality of the dataset; Maintenance information contains information about the scope and frequency of updating data; Spatial representation information contains information concerning the mechanisms used to represent spatial information in a dataset; Reference system information contains the description of the spatial and temporal reference system(s) used in a dataset; Distribution information contains information about the distributor of, and options for obtaining, a resource.

Extent information is an aggregate of the metadata elements that describe the spatial and temporal extent of the referring entity. Citation and responsible party information provides a standardized method (CI_Citation) for citing a resource (dataset, feature, source, publication, etc.), as well as information about the party responsible (CI_ResponsibleParty) for a resource. The CI_ResponsibleParty datatype contains the identity of person(s), and/or position, and/or organization(s) associated with the resource. All of them could be referenced by other metadata entities whenever they existed.

The other classes and elements included in KMP are also revised in each entity. More details could be referenced to the content of the KMP metadata.

3. KMP CORE ELEMENTS SCHEMA DESIGN

Most of the elements included in KMP are seldom to be used in KIEIN project. So the XML Schema designs for the KMP are divided into two stages. The first step is to implement the XML Schema for KMP core elements. The next stage will be left for future design.

In ISO 19139, the Metadata Schema implementation is based on the other several ISO standards which gave more details definition for conformance rules (ISO 19139, 2003). The application of domain restriction types are also defined from those standards. It is hard for KMP to accept all of the requirements at the initial design.

KMP core elements are defined in Table 1. Those elements came from MD_Metadata, MD_Identification/MD_DataIdentification, MD_Constraints, MD_DataQuality, and MD_Distribution. These are elements used to define metadata content, dataset content, dataset access, dataset quality, dataset acquirement, and contacts information, etc. The elements could give a simple but useful description of GIS dataset and the related information. These elements could match most of datasets application requirements.

Table 1 KMP Core metadata elements

No.	Name	Definition	Condition
0	**Metadata cataloguing information** (MD_Metadata)	root entity which defines metadata about a resource or resources (datasets and/or dataset series)	M
1	**Dataset title** (MD_Metadata > MD_DataIdentification.citation > CI_Citation.title)	name by which the cited resource (dataset) is known	M
2	**Dataset Alternative title** (MD_Metadata > MD_DataIdentification.citation > CI_Citation.alternateTitle)	Short name or other language name by which the cited information is known	O
3	**Edition** (MD_Metadata > MD_DataIdentification.citation > CI_Citation.edition)	version of the cited resource (the dataset)	O
4	**Series Name** (MD_Metadata > MD_DataIdentification.citation > CI_Citation.series > CI_Series.name)	name of the series, or aggregate dataset, of which the dataset is a part	O
5	**Issue Identification** (MD_Metadata > MD_DataIdentification.citation > CI_Citation.series > CI_Series.issueIdentification)	information identifying the issue of the series	O
6	**Dataset reference date** (MD_Metadata > MD_DataIdentification.citation > CI_Citation.date)	reference date for the cited resource. (date and time when the dataset was published or otherwise made available for release)	M

7	**Dataset responsible party** (MD_Metadata> MD_DataIdentification.pointOfContact >CI_ResponsibleParty)	identification of, and means of communication with, person(s) and organization(s) associated with the resource(s)	M
8	**Geographic location of the dataset (by four coordinates or by geographic identifier)** (MD_Metadata>MD_DataIdentification.extent> EX_Extent> EX_GeographicExtent> EX_GeographicBoundingBox or EX_GeographicDescription)	Horizontal and vertical extent covered by the dataset	C
9	**Dataset language** (MD_Metadata > MD_DataIdentification.language)	language(s) used within the dataset	M
10	**Dataset character set** (MD_Metadata > MD_DataIdentification.characterSet)	full name of the character coding standard used for the dataset	C
11	**Abstract describing the dataset** (MD_Metadata > MD_DataIdentification.abstract)	brief narrative summary of the content of the resource(s)	M
12	**Purpose** (MD_Metadata > MD_DataIdentification .purpose)	summary of the intentions with which the resource(s) was developed	O
13	**Progress Code** (MD_Metadata > MD_DataIdentification.status)	status of the resource(s)	O
14	**Dataset topic category** (MD_Metadata > MD_DataIdentification.topicCategory)	main theme(s) of the dataset	M
15	**Keywords information** (MD_Metadata > MD_Identification > MD_Keywords.keyword)	commonly used word(s) or formalised word(s) or phrase(s) used to describe the subject	M
16	**Keywords Type information** (MD_Metadata > MD_Identification > MD_Keywords.type)	subject matter used to group similar keywords	O
17	**Keywords Thesaurus Name** (MD_Metadata>MD_Identification> MD_Keywords.thesaurusName)	name of the formally registered thesaurus or a similar authoritative source of keywords	O
18	**Access Constraints** (accessConstraints) (MD_Metadata > MD_LegalConstraints.accessConstraints)	access constraints applied to assure the protection of privacy or intellectual property, and any special restrictions or limitations on obtaining the resource or metadata	
19	**Use Constraints** (MD_Metadata > MD_LegalConstraints.useConstraints)	constraints applied to assure the protection of privacy or intellectual property, and any special restrictions or limitations or warnings on using the resource or metadata	
20	**Lineage** (MD_Metadata > DQ_DataQuality.lineage > LI_Lineage)	information about the events or source data used in constructing the data specified by the scope or lack of knowledge about lineage	O
21	**On-line resource** (O) (MD_Metadata > MD_Distribution > MD_DigitalTransferOption.onLine > CI_OnlineResource)	information about on-line sources from which the dataset, specification, or community profile name and extended metadata elements can be obtained	
22	**Spatial representation type** (MD_Metadata > MD_DataIdentification.spatialRepresentationType)	method used to represent geographic information in the dataset	O
23	**Reference system** (MD_Metadata > MD_ReferenceSystem)	information about the reference system	O
24	**Distribution format name** (MD_Metadata > MD_Distribution > MD_Format.name)	name of the data transfer format(s)	O
25	**Distribution format version** (MD_Metadata > MD_Distribution > MD_Format.version)	version of the format (date, number, etc.)	O
26	**Spatial resolution of the dataset** (O) (MD_Metadata > MD_DataIdentification.spatialResolution > MD_Resolution.equivalentScale or MD_Resolution.distance)	level of detail expressed as the scale of a comparable hardcopy map or chart/ ground sample distance	O
27	**Metadata parent identifier** (MD_Metadata.parentIdentifier)	file identifier of the metadata to which this metadata is a subset (child)	O
28	**Metadata file identifier** (MD_Metadata.fileIdentifier)	unique identifier for this metadata file	O
29	**Metadata standard name** **(**MD_Metadata.metadataStandardName)	name of the metadata standard used	O
30	**Metadata standard version** (MD_Metadata.metadataStandardVersion)	version (profile) of the metadata standard used	O

31	**Metadata language** (MD_Metadata.language)	language used for documenting metadata	C
32	**Metadata character set** (MD_Metadata.characterSet)	full name of the character coding standard used for the metadata set	C
33	**Metadata point of contact** (MD_Metadata.contact > CI_ResponsibleParty)	party responsible for the metadata information	M
34	**Metadata date stamp** (MD_Metadata.dateStamp)	date that the metadata was created	M

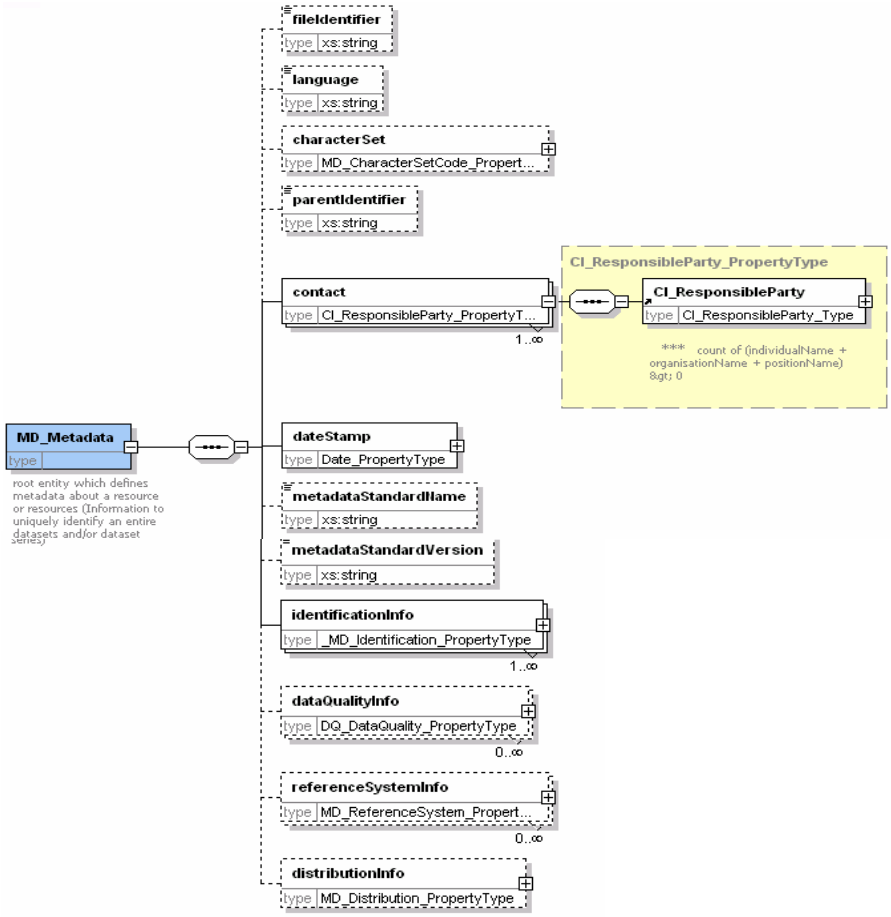

Fig. 2 Schema example designs for MD_Metadata

By the defining of KMP core elements, a XML Schema implementation of it is designed (Fig. 2). The Schema started from the root element MD_Metadata. Both Simpletype and ComplexType XML types are used to define the Schema. Self reference method is used to call global elements. Most of the Domains and data types are defined from XML clauses directly.

Figure 3 is the example XML Schema of how to implement special case for complicate entity. MD_DataIdentification is the special case of MD_Identification.

4. KMP METADATA SCHEMA APPLICATION TREND

GIS standards generally refer to information technology standards and/or spatial data standards. A GIS standard may result from the adoption or adaptation of an information technology standard for GIS applications. In standards development, a common practice is to first consider the adoption or adaptation of an existing standard. Developing a standard is usually a last resort, since the time for developing and approving standards is a lengthy process.

```
<xs:element name="MD_DataIdentification" type="MD_DataIdentification_Type" substitutionGroup="_MD_Identification">
  <xs:annotation>
    <xs:documentation>
***  characterSet: documented if ISO 10646-1 is not used
***  {MD_Metadata.hierarchyLevel = "dataset" implies count (extent.geographicElement.EX_GeographicBoundingBox) + count
(extent.geographicElement.EX_GeographicDescription) &gt;=1}
***  {MD_Metadata.hierarchyLevel notEqual "dataset" implies topicCategory is not mandatory}
  </xs:documentation>
  </xs:annotation>
</xs:element>
<xs:complexType name="MD_DataIdentification_Type">
  <xs:annotation>
    <xs:documentation/>
  </xs:annotation>
  <xs:complexContent>
    <xs:extension base="smXML:_MD_Identification_Type">
      <xs:sequence>
        <xs:element name="language" type="xs:string" maxOccurs="unbounded"/>
        <xs:element name="characterSet" type="MD_CharacterSetCode_PropertyType" minOccurs="0" maxOccurs="unbounded"/>
        <xs:element name="topicCategory" type="MD_TopicCategoryCode_PropertyType" maxOccurs="unbounded"/>
        <xs:element name="extent" type="EX_Extent_PropertyType" minOccurs="0" maxOccurs="unbounded"/>
        <xs:element name="spatialResolution" type="MD_Resolution_PropertyType" minOccurs="0" maxOccurs="unbounded"/>
        <xs:element name="spatialRepresentationType" type="MD_SpatialRepresentationTypeCode_PropertyType" minOccurs="0" maxOccurs="unbounded"/>
        <xs:element name="environmentDescription" type="xs:string" minOccurs="0"/>
        <xs:element name="supplementalInformation" type="xs:string" minOccurs="0"/>
      </xs:sequence>
    </xs:extension>
  </xs:complexContent>
</xs:complexType>
<xs:complexType name="MD_DataIdentification_PropertyType">
  <xs:sequence>
    <xs:choice>
      <xs:element ref="smXML:MD_DataIdentification"/>
      <xs:element ref="scXML:Reference"/>
    </xs:choice>
  </xs:sequence>
</xs:complexType>
```

Fig. 3 Schema example code for MD_Dataidentification entity

KIEIN adapted ISO/TC 211 Metadata Standard and developed it to KMP. It is a right decision at the first step. Anyway, some of the content elements which included in KMP may be changed after the datasets description test finished.

KGIC has established a WebGIS site. Most of historical Kuwait GIS datasets have been involved in the server (Fig. 4). GIS Metadata server and Datasets sever technology has been used in the implementation of the KGIC WebGIS clients/servers architecture. All users could search GIS Metadata sets through internet and access them via the Website with interactive GIS interface. If the needed GIS datasets are found by users, further consideration or order requirement could be discussed by the contact to KGIC members.

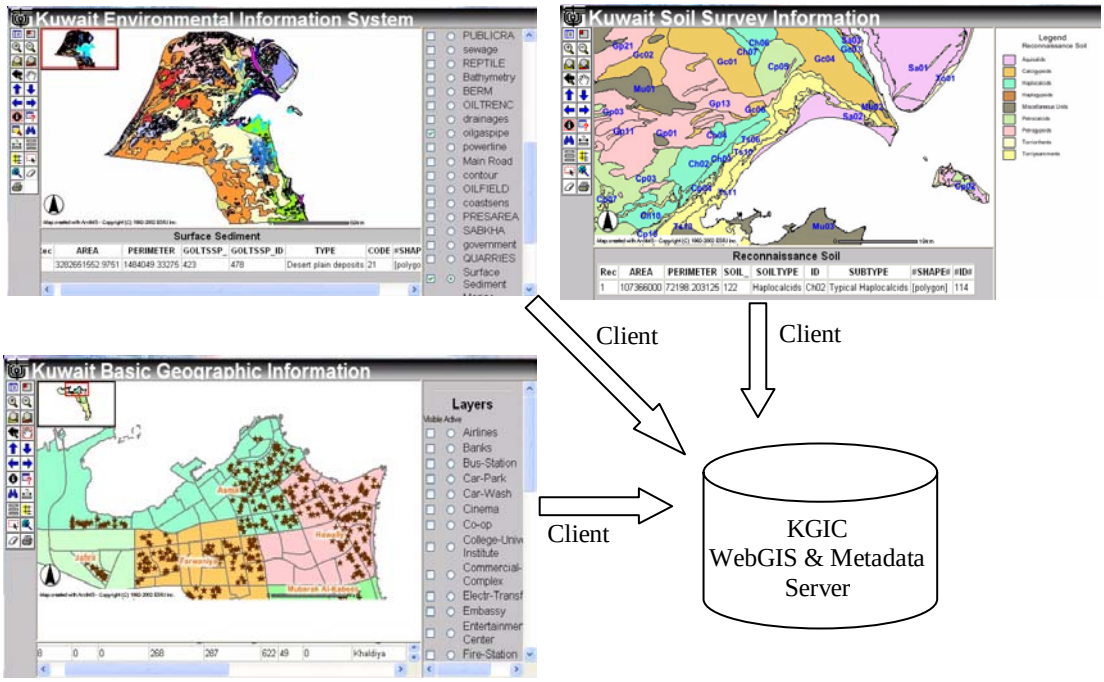

Fig. 4. KGIC WebGIS and Metadata Clients and Server Architecture

The KMP schema also could be used to update ISO/TC 211 Metadata schema which implemented by ESRI in ArcGIS. Due to ArcGIS is one of the KISR main GIS software, most of the new GIS projects in KISR will use it as the GIS platforms. After the KMP Schema integrated to ArcGIS platform, it will supply a better chance for KMP test and implementation rather than only with KIEIN datasets.

Geographic information standardization research in Kuwait is just started. Based on the widely application needs such as E-Government, Kuwait National Environmental Information Network (KNEIN), and several national GIS related projects, Metadata standard should be urgent developed in Kuwait. This will be very helpful for Kuwait geographic information to be benefited to the national information interests. However, more cooperation and effort should be done before it could be accepted as the national metadata standard draft.

REFERENCES

Abdul Nabi Al-Ghadban, 1997, Towards a GIS-Based environmental information system for Kuwait, GIS Conference'97, Qatar.

Gerard Grealish, Peter King, Samira Omar, Waleed Roy, 2005, Geographic information system and database for the soil survey for the State of Kuwait – design and outputs, Kuwait J. Sci. Eng., 31, (1), pp. 135-148.

KISR 1999, Soil Survey for the State of Kuwait – Volume II Reconnaissance Survey, AACM International, Adelaide, Australia, 1-270.

ISO 19115:2003, Geographic Information–Metadata, 2003/05/01.

ISO 19139:2004, Draft technical specification 19139 Geographic Information–Metadata – XML schema implementation, 2004/06/30.
Ota, M., 2001, Japan Metadata Profile (JMP) for geographic information clearinghouses, International Conference on Dublin core and Metadata Applications 2001, 2001/10/22-26, Tokyo.
Zhao Yongping, Guo Jingjun, Chen Aijun, 1999, Application Model to Build Metadata Sharing Architecture, Journal of Tsinghua University, 39 (12), 90-93.

BIOGRAPHICAL NOTES

Yongping Zhao

Dr. Zhao is a GIS Specialist at the Systems Development Department of National Scientific and Technical Information Center in Kuwait Institute for Scientific Research. He obtained his Ph.D. degree in Cartography and Remote Sensing from the Institute of Remote Sensing and GIS at Peking University, P.R.China. He was a postdoctoral fellow at University of Waterloo, Canada, an Assistant Research Scientist at George Mason University, USA, and an Associate Professor at Tsinghua University P.R.China. He is specialized in remote sensing, GIS, and photogrammetry theories and technologies. His research areas include digital photogrammetry, image process, GIS standards, environment, and geospatial analysis.

Jamal Al-AKROKA

Mr. Al-AKROKA is the Department Manager at the Systems Development Department of National Scientific and Technical Information Center in Kuwait Institute for Scientific Research. He obtained his Master degree in Computer Engineering from George Washington University, USA. He is specialized in data analysis, database design, information system implementation and management.

CONTACTS

Dr. Yongping ZHAO
Kuwait Institute for Scientific Research
SDD/NSTIC
P.O. Box 24885,
13109 Safat
Kuwait,
Tel. +965 483 6100 Ext.5278
Fax: +965 481 5175
Email:yzhao@kisr.edu.kw
Website:www.kisr.edu.kw