Studies of Chinese Digital Health Resources Metadata Proposal (CDHR-m)

Zhu Dong Ping
The aim of the DC

About Metadata

Background

The current situation

[2]
A new type of information-structured data, is defined commonly as “Data about data”. Dublin Core Metadata Element Set (DC) was initiated in 1995 by the OCLC (USA Online Catalog Library Center) and NCSA (the United States Center for Supercomputing Applications).
The aim of DC[^2]

The aim of DC is to search and acquire the needed information rapidly through description, collection and organization of the information resources by using simple identifications.
Current situation \textsuperscript{[3]}

Study on the organization of Internet information has been a research focus in the field of information management. During recent 10 years, DC-m has been rapidly developed and applied in many fields including medical information of many digital resource organizations.
Medical-m Proposals based on DC
However, up till now, there has been no report on standard metadata scheme profile for Chinese health resources. [10]
The problems and challenge

Purpose and Objectives

What we want
Problems and Challenge

(1) Rapid development of Internet and the computer technology.

(2) Growth of the digitalized Chinese biomedical health resources.
Due to the fact that the information on the Internet is dynamic and disordered, searching and utilization of Chinese digital health resources has become difficult.
What we want

The aim of this research is to establish a set of Chinese medical metadata application schema based on DC for the use of organizing, searching and utilizing both the present Chinese digitalized medical information and those from Internet.
What we want

It is named CDHR-m (Chinese Digital Health Resources Metadata), which can be fitted to the characteristics of contemporary medical information and be applied to network and digital medical information resources.
To define DC as a standard metadata scheme for CDHR-m

To establish the contents of CDHS-m coding schemes

To select and determine the project of description object for CDHR-m

To analyze the factors of contents description objects for CDHR-m
Selecting DC as a standard metadata scheme for Chinese health resources

The first reason (for selecting DC)

(1) DC-m has been repeatedly revised, and improved to become an appropriate and mature international metadata standard.

(2) It has been widely used in many fields in the past ten years.
The first reason (for selecting DC)

(3) It has to be adaptive to the regulation, standardization, internationalization and the principle of interoperability.\textsuperscript{[11]}
The 2nd reason:

There are a number of foreign-based medical meta-data DC projects \(^{[12]}\) which are sophisticated enough to be used as a model for Chinese digital health resources regarding to consultation, reference, reusing and nesting when creating CDHR-m.
The 3rd reason:

The present model: "Chinese journals metadata to describe the rules of the recorded" has just been widely used in China, and it is based on the DC metadata standard. Therefore, there is a ready-made sample DC-based local metadata to provide CDHS-m of the reference reuse.
Establishment of the Contents of CDHS-m Coding Schemes

CDHR-m based on DC-metadata

- **Advantages**
  - International standards
  - Highly condensed summary

- **Disadvantages**
  - Not "exhaustive"
  - Not "matching directly"
Using 15 core elements of DC metadata merely to interpret and recapitulate highly specialized resources is obviously inadequate for the description of medical literature.
To make The CDHS-m perfect:

▲ Follow the framework of DC provided by the standard coding system.

▲ Use the present foreign several medical metadata programs, such as MCM, CISMeF, EBM metadata and NLM-metadata as reference.

▲ Quote the medical MeSH Vocabulary as encoding scheme in the "Subject" element of the subject coding system.
Consult the “EBMC” and the “EBMS” (which is used to describe the evidence-based programs medical coding system), CLC ("Chinese Library Classification 4th edition")

Refer to HL7 (Health Information transmission standard).
Selection and Determination of Description Object of CDHR-m

(1) Medical journals are selected and defined as the metadata resources.

- one reason:

Most of the medical core papers are published in these selected journals.
The 2nd reason:

These are the main medical information resources owing to the fact that they publish most recent and reliable medical information.\[15\]
Many electronic medical journals and books (in both full text and abstract) are already on the Internet.

The 4th reason:

Most of the medical journals are well written and in good format.
Medical journals are regarded to be available representation in all types of medical resources.
CDHR-m major description object - issue medical journals

CDHR-m minor description object - the other types: Medical multimedia, network resources, electronic medical records, medical conferences, various types of agreements, tele-medicine, and medical imaging, etc.
Analysis of Factors in CDHR-m
Contents Description Objects
Evidence-based medicine (EBM) is the integration of the best research evidence with clinical expertise and patient values. An important process of EBM practice is to find significant evidence supported by clinical studies (both primary studies and secondary studies such as systematic reviews) published in clinical journals or databases.\textsuperscript{[16]}
Four elements should be considered when we put the CDHR-m project as the main content of evidence-based medicine resources:
An example
Japan's EMB-m (evidence-based medicine metadata project), is a comprehensive search for EBM resources.

- the degree of evidence depending on research design
- clinical aspects of the study (e.g., pathogenesis, diagnosis, therapy and prognosis)
- patients (with or without diseases) interventions and outcomes.
Therefore, we should absorb experiences from Japan's EMB-m projects for our CDHR-M. The CDHR-m proposal should be consistent with the recognition of primary and secondary studies for a comprehensive EBM search tool.
The second factor:

HL7 (Health level Seven) Standard [18]

HL7 is a kind of “medical messages data interchange system”, which was originally designed for economic services, mainly for the management of medical expenses and hospital information system, but it is in the lack of the contents of the medical literature resources. [19]
In view of the HL7’s popularization and application, we should build a national HL7 that not only is under the framework of HL7 standard classification, but also perfect and modify the content of standard classification, coding and contents.
It would be the future target that HL7 standards and the medical literature resources could be shared, exchanged.
One of the interesting work is that development of CDHR-m project for clinical care and medical research to provide more flexible and multi-level medical information related to the elements are expected to eventually become the localization part of the HL7 standard.[20]
The third factor:

Medical assistant information is an important part of health resources.

[21]
such as

★ **Author information**

name, unit, address, telephone no., brief introduction — can supply personal information of author himself;

★ **The reference information**

can indicate that the scope and depth of author's thesis, as well as its citation situation;

★ **The information of Fund**

supply resource evaluation, the level of its publication, impact factor, factor, etc.) and give us the information about the scientific and technological content and quality positioning.
Outcome

Article supporting comprehensive medical assistant information can not only help readers quickly and efficiently catch the idea of content, but also play an important role in information transmission, communication, retrospective retrieval, bibliometrics and statistics.\textsuperscript{[22]}
Effect

The fund support information available helps us to know the types of support, intensity, distribution (geographical or major) and et.al, in order to provide this integrated information content to the government or the relevant departments in accordance with decision-making.
**Effect**

Setting up the Chinese Medical Professional expert’s database system is useful in terms of author’s personal information. Meanwhile the medical assistant information description will become important factor to be considered in CDHR-M.
The fourth factor

Numerous types of multimedia springing up

The remote medical
major medical network protocol
The medical home
Medical digital images
Audio and video, bio-signal graph
Electronic medical records
Other non-text form special

This is the main characteristic of medical resources digitization\cite{23}.\[23\]
Review

The application of DC metadata has a strong advantage in the description and organization of such resources. There are three medical metadata proposals in the world: MCM, CISMeF and EBM. All of them have expanded from original DC. type restrictions.
The **MCM metadata** created 35 sub-elements;

The **CISMeF metadata** has resource type 20 added, which covered all of the network resource types.

The **EBM metadata** reused all types of DC metadata description in resources elements of both MCM and CISMeF.

The **MeSH Thesaurus** is used in more than half of the selected words.
selected words.
A set of CDHR-m

<table>
<thead>
<tr>
<th>The name of metadata element</th>
<th>Element Refinement</th>
<th>Encoding Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>alternative</td>
<td></td>
</tr>
<tr>
<td>Creator</td>
<td>modeinstitution or personalcreator resume</td>
<td></td>
</tr>
<tr>
<td>Subject</td>
<td>keyword subject classification no.</td>
<td>MeSH、CLC 4ed.EBMC<em>EBMS</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td>HL7</td>
</tr>
<tr>
<td>Description</td>
<td>Content description</td>
<td>URL</td>
</tr>
<tr>
<td>-------------</td>
<td>---------------------</td>
<td>-----</td>
</tr>
<tr>
<td></td>
<td>catalog(contents)abstract:objectives,design,method &amp; material,result,conclusion 遵证医学: system reviews、meta-analysis、RCT、clinical trial HL7: patients management info. 、医嘱处理信息、临床观察信息、医疗病案管理、医疗咨询服务信息、财务管理信息; Fund class: commentary structure description: paragraph,table,graph,chart,formula,footnote,comment,links,thanks,sub-title(1-4 level)</td>
<td></td>
</tr>
<tr>
<td>Publisher</td>
<td>Name of pub.,ad.,postalcode,E-mail, mail code、tel number</td>
<td></td>
</tr>
<tr>
<td>Contributor</td>
<td>editor in charge</td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>Date createdDate validDate availableDate issuedDate modifiedDate acceptedDate copyrightDate submitted</td>
<td>ISO8601[W3CDTF] (YYYY-MM-D)</td>
</tr>
<tr>
<td>Type</td>
<td>Bibliographies（Medline）Clinical ConferenceClinical TrialComment Consensus Development ConferenceIndex of name’s addressDirectoryEditorialDegree PaperMulticenter studyReviewReview of Reported Electric CasesReview TutorialTechnical ReportHealth News(Internet)Full Text JournalsAbstractsHomepageDecision ToolsRadiographs（CT、DR、MRI、PET-CT）Ultrasound ImagesPath ImagesClinical ImagesDatabasesMedical ViewAudioVidioProceduresMeetingsTextbooksMeeting ReportsMedical ForumsMedical EducationPatient EducationExpert Advise</td>
<td>DCMIType[DCMIType]MCMType、Medline-PT</td>
</tr>
<tr>
<td>---------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Format</td>
<td>Extent medium</td>
<td>IMT（for the type of network。MIME）</td>
</tr>
<tr>
<td>Identifier</td>
<td>Reference list:name of journal,ISSN,yy-mm-dd：ISSNISBN</td>
<td>URI URL DOI SICI ISSN ISBN DOI</td>
</tr>
<tr>
<td>Source</td>
<td>URI</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td>Relation</td>
<td>isVersionOf</td>
<td>hasVersion</td>
</tr>
<tr>
<td>Coverage</td>
<td>(spatial) (temporal)</td>
<td>Period,W3CEDTF</td>
</tr>
<tr>
<td>Right</td>
<td>access Right</td>
<td></td>
</tr>
<tr>
<td>Resource Rating</td>
<td>scientific and technological paper citation database, influencing factor, keyword journal,</td>
<td></td>
</tr>
<tr>
<td>Audience</td>
<td>medical personnel, patients and others</td>
<td></td>
</tr>
<tr>
<td>Article Classification</td>
<td>article, comment, case report, review, clinical research, et al.</td>
<td></td>
</tr>
</tbody>
</table>

To be continued
The CDHR-M aims at establishing DC metadata model designing to look for a new method of organizing and searching Chinese health resources. Currently it is on the development stage and it is only a rough idea. There is still a lot to be done for its establishment and improvements.
Conclusion

More efforts need to be made in terms of the metadata element model of CDHR-M and medical experts from decision-making body of the government are required to make deep research on its work, so as to
provide a scientific basis for the development of the standard of CDHR medical metadata.


5. Appleyard R. Medical core metadata project: enhancing internet medical information retrieval with metadata encodings [EB/OL]. http://www2.sub.uni-goettingen.de/cgi-bin/ss-gfi/anzeige.pl?db=meta&nr=000103&ew=SSGF-I.


References


E-mail:zdping2002@vip.sina.com
Thank You!