User Centred Networked Health Care
A. Moen et al. (Eds.)
IOS Press, 2011

doi:10.3233/978-1-60750-806-9-839

Metadata - an International Standard for Clinical Knowledge Resources

Gunnar O KLEIN[a1]
[a] *Dept of Microbiology, Tumour and Cell Biology, Karolinska Institutet, Sweden*

Abstract. This paper describes a new European and International standard, ISO 13119 Health informatics – Clinical knowledge resources - Metadata that is intended for both health professionals and patients/citizens. This standard aims to facilitate two issues: 1) How to find relevant documents that are appropriate for the reader and situation and 2) How to ensure that the found knowledge documents have a sufficient or at least declared quality management? Example of use is provided from the European Centre for Disease Control and Prevention.

Keywords. metadata, decision support systems, clinical knowledge resource, standard, ISO 13119

1. Introduction

The internet is rapidly changing the way we access medical knowledge. Health professionals use web based knowledge sources and digital documents are provided from databases and via e-mail. Also the patients/citizens turn to the internet for advice.

The European Commission has published a set of quality criteria for health related websites [1] as one way of establishing trust in web resources. A trust-mark indicating a "minimum" level of trustworthiness requires:

- a set of quality requirements. This might be very difficult to agree on as relevant for all contexts.
- third party control by governmental bodies or professional associations of all possible documents to receive the mark.
- reliance on a self-declaration by the issuer in which case the user of the information has no real guarantee that the criteria are met even if the mark is there.
- Instead of reviewing the actual content of the medical knowledge resources, we can define processes behind their development.

Health authorities in many countries and in co-operation with the Commission have considered the possible requirement for legislation and control procedures, but generally the conclusions have been that rather than trying to ban bad quality information, one should facilitate for the citizens as well as for the health professionals to find the type of information they request where quality criteria behind a knowledge resource are easily accessible.

One feasible and important approach is to establish a set of metadata for each knowledge resource to describe the content and the procedures behind its production.

[1] Corresponding author. Gunnar Klein, 177 77 Stockholm, Sweden, E-mail: gunnar.klein@ki.se

In this paper the development of a standardised set of metadata is described. The following issues are addressed: What are the possible uses of a standardised set and What are the basic principles of the new standard in the field?

2. Materials and Methods

This study is based on the work of the European and International standards organisations during the years 2000-2010. This started with a literature review on the use metadata for various health care purposes and the general development of metadata for intersector use, the Dublin Core and various initiatives to propose metadata especially for clinical guidelines. The development of the first draft and discussions of the standardisation working groups was followed by extensive and repeated international review and comments with suggestions for improvements from many nations. The author was the project leader of the standardization project started in CEN that led to the publication of the CEN/TS 15699:2009 [2] and further enhanced in ISO to the ISO 13119 [3] also to be published as a European standard.

3. Results

3.1. The Scope of the Metadata Standard

This standard defines a number of metadata elements that describe documents containing medical knowledge, primarily digital documents provided as web resources, accessible from databases or via file transfer, but can be applicable also to paper documents, e.g. articles in the medical literature. It is based on the ISO 15836:2009 Information and documentation- Metadata – The Dublin Core metadata element set [4]. The metadata should:

- support unambiguous and international understanding of important aspects to describe a document, e.g. purpose, issuer, intended audience, legal status and scientific background
- be applicable to different kinds of digital documents e.g. recommendation from consensus of a professional group, regulation by a governmental authority, clinical trial protocol from a pharmaceutical company, scientific manuscript from a research group, advice to patients with a specific disease, review article
- be possible to present to human readers including health professionals as well as citizens/patients
- be potentially usable for automatic processing e.g. to support search engines to restrict matches to documents of a certain type or quality level

3.2. Characteristics of the Metadata Element Set

In the element descriptions below, each element has a descriptive label intended to convey a common semantic understanding of the element, as well as a unique, machine- understandable, single-word name intended to make the syntactic specification of elements simpler for encoding schemes.

Each element is optional and repeatable. Metadata elements may appear in any order. The ordering of multiple occurrences of the same element (e.g. Creator) may have a significance intended by the provider, but ordering is not guaranteed to be preserved in every system.

To promote global interoperability, a number of the element descriptions suggest a controlled vocabulary for the respective element values. The Dublin Core set assumes that different domains develop where necessary controlled vocabularies as specialisations of the content of the general purpose metadata element set and adding other metadata elements as required. This standard is such a specialisation for the medical knowledge domain.

3.3. Metadata Groups

The metadata elements are grouped purely for human navigational purposes:

- Resource form
- Intended use
- Subject and Scope
- Identification and source
- Quality control

The total number of metadata element tags in this standard is 150.

3.4. Examples of Specialisations

In some areas the standard contains an enumerated list of specialisations to be used for the content under some metadata elements.

3.4.1. Healthcare Specialization for Type

One example is for the element Type defined by Dublin Core:

Definition: Nature or genre of the content of the resource. The following Types are from the Dublin Core 2009: Text, MovingImage, StillImage, Sound, Dataset, InteractiveResource, Software, Device.

Table 1: The following terms may be used to describe Type.Text in health care:

Journal_article	Teaching_material	Known_uncertainty
Book_chapter	Computable clinical	Observational_study
Book	information model	Qualitative_study
Report	Terminological_resource	Randomised_controlled-
Abstract	Metainformation	trial
Patient_information	Case_report	Research_study
FAQ	Proposal	Review
Algorithm	Event	Systematic_review
Clinical guideline	Service_description	Structured_abstract
Policy-strategy	Product_information	Care_pathway
Information_standard	Critically_appraised_topic	

3.4.2. Example of Healthcare Specialization for Situation

This defined as: Description of the situation where the knowledge is intended to be used (HC). This can also be understood as the intended role of the knowledge resource. Healthcare specific specialisation:

- Clinical_guidance
- Self_guidance
- Supporting_software
- Research_protocol
- Background_knowledge

3.5. Overview of the Metadata Classes

Figure 1 shows an overview of the classes.

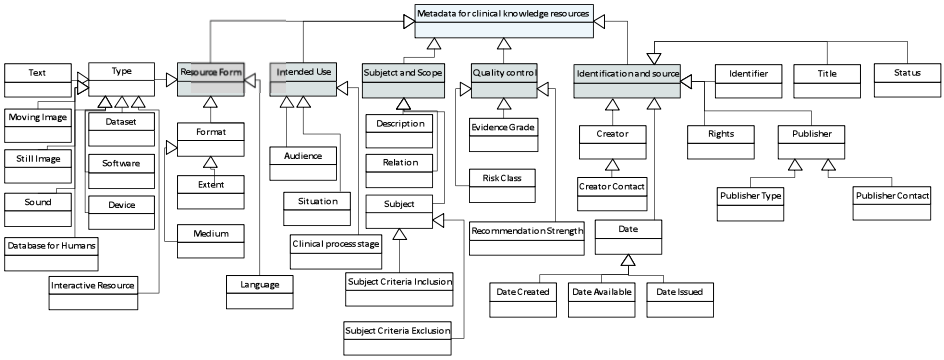

Figure 1. Overview of the Metadata elements for Clinical Knowledge Resources.

3.6. A First User of the New Standard - ECDC

The European Centre for Disease Control and Prevention, ECDC, which is a rather new European Union agency that is mainly active in the surveillance and prevention of communicable disease has an ambitious programme for knowledge management that shall serve not only its internal staff an specially commissioned experts but also the member states of the European union with their national agencies for control of communicable diseases. This organisation was looking for their own implementation of the Dublin Core metadata standard when they were approached and studied the new standard already at the draft stage. They have now implemented its use as a routine part of their work together with other strategies for knowledge retrieval in an Enterprise wide search system.

4. Discussion

A system of metadata tags, the names of the elements can have many different uses. The first uses have been within larger organisations that have a need to ensure that their documents, the most common form of a knowledge resource, can be found using automatic retrieval methods. If metadata are assigned in a consistent and well

structured way to each document, this can ensure complete retrieval of all relevant documents meeting the search profile whereas various other ways indexed or not using the core content of a document without any metadata can usually not ensure that all relevant documents are found.

The other major feature of a good metadata based system is to exclude irrelevant documents because of a much more specific search profile. With the explosive growth of various document resources, this becomes more and more important. It is something a clinician is frequently experiencing searching knowledge on various dedicated medical knowledge sites. It is of course also a common problem for the general public using general purpose search engines as Google on the general world wide web.

It should be emphasized that there is no requirement in the standard to use all of the metadata elements available. This is a set of optional elements and typically a publisher of a type of knowledge resource only uses a small subset. If required it is possible to extend the set by additional elements.

For some elements there is very detailed guidance provided where there was some good justification to propose details. In other areas the users will have to develop their own guidance documents if consistency is to be achieved.

There is also another use of metadata that is not related to retrieval but for the user of the resource to be able to understand what the resource is, its intended use, source and possibly quality control. The latter is achieved largely through reference to the Grade system for clinical guideline documents, which is also acknowledged by the WHO [5].

Acknowledgments: This study was supported by the European Union Network of Excellence: Semantic Mining. During the first years of this work the author was co-operating with Dr Anders Thurin, Göteborg, Sweden. His important contributions are gratefully acknowledged together with many other experts of CEN/TC 251/WG 2 and ISO/TC 215/WG 3.

References

[1] The European Commission, COM(2002) 667, eEurope 2002: Quality Criteria for Health related Websites.

[2] CEN/TS 15699:2009. Health informatics – Clinical knowledge resources – Metadata. The European Committee for Standardization, Brussels, (2009).

[3] ISO/DIS 13119:2011. Health informatics – Clinical knowledge resources – Metadata. International Organization for Standardization, Geneva, (2011).

[4] ISO 15836:2009 Information and documentation- Metadata – The Dublin Core metadata element set. International Organization for Standardization, Geneva, (2009).

[5] H.J. Schünemann, A. Fretheim, A.D. Oxman. Improving the use of research evidence in guideline development: 9. Grading evidence and recommendations. *Health Research Policy and Systems* **4** (2006), 21.