

Evaluation and Design Issues of Nordic DC Metadata Creation Tool

Preben Hansen

SICS – Swedish Institute of computer Science
Box 1264, SE-164 29 Kista, Sweden
preben@sics.se

Abstract

This paper presents results from the analysis of data collected during a 3-year user-oriented longitudinal and empirical on-line evaluation of the use of the Dublin Core metadata creation tool within the Nordic DC Metadata Project. The paper is concerned with how humans create metadata. In particular, the paper explores different categories of requirements. The evaluation is part of an on-going design process of the metadata creation tool and inform on valuable issues to be considered in the future design process.

Keywords: *Dublin Core, Metadata tool, Evaluation, HCI*

1 Introduction

The World Wide Web has changed the creation, distribution, storage and retrieval and presentation of information. This means that there may be difficulties for the end-user to search, browse and navigate for relevant information. Metadata is one way to overcome the management problems of large sets of both heterogeneous and homogeneous sets of distributed digital information collections.

The issue of metadata have become more relevant and necessary not only to the library environment, but recently also within other domains such as commercial settings and business, which show a great interest in adopting, integrating and developing different metadata tools, schemas and frameworks for both general and specific purposes.

In order to be able to provide useful metadata we need useful formats and tools as well as user guidelines, which make the metadata provision easy and effective as possible.

Dublin Core (DC) Metadata Initiative [1] was an early effort to overcome the problem with cataloguing, indexing and retrieval of web-based documents. The proposed set of 15 elements has emerged from an international effort of consensus building manifested through a series of workshops.

2 Metadata tools and Evaluation

There are continuous developments of new metadata management tools and systems for handling metadata processes. Even though there is a large amount of research and knowledge both within the field of library and information science regarding the traditional way of indexing by domain experts, and within the field of computer science and information retrieval (IR), there are very little research done regarding the evaluation of users creating metadata in real situations and for real life use purposes. Fidel [2] explores more in depth the distinction between automatically generated and human-generated indexing of information. Within the field of Human-Computer Interaction (HCI), methods and techniques are used for analysing and evaluating both the design of tools and systems as well as of human factors involved in the interaction of these tools and systems.

In order to provide a better way to collect, access and create metadata as a mean to more effective searching on the Internet, it is very important how such tools are designed and on what knowledge these tools are built on. Issues such as domains, users and their knowledge levels, design issues regarding the functionality of the tool, user interface, as well as the support such a tool can provide, are very important and will affect the usability [3].

There are few studies exploring and investigating the tools that enables user to create metadata for their own use as well as for general use. However, Marshall [4] presents an analysis of ethnographic data gathered when creating metadata for a mixed physical-digital collections of visual resources. She concludes that an ethnographic approach to the understanding of creation, and subsequently the use of metadata is an important way of prescribing the limits of metadata. She also says that an analysis of human-created metadata will reveal knowledge regarding describing single information units as well as collections for public use.

Our goal with this paper is to present a user-oriented empirical and longitudinal evaluation of a metadata creation tool designed within the Nordic

DC Metadata Project [5]. In this particular study, we want to investigate the following questions: What type of users do we support with our template? What problems do users have with the template functionality and with the User Guidelines? In what way could the data collected inform the design process?

3 Nordic Metadata Project 1 and 2

The first Nordic metadata project I and II [5] started in 1996 and ended in early 2001 and was one of the first international DC projects in a growing number of similar projects [1]. The aim of the project was to create a Nordic metadata production, indexing and retrieval environment. The system was intended to be primarily for production purposes. However, in an early stage of the project, we also decided that we needed rather extensive support for the users and User Guidelines were also provided.

In order to develop the design of the tools, we wanted to investigate how the template worked during the task of creating DC metadata for different purposes (private, business, library etc). One of the main tasks within our project was to evaluate the template and the guidelines provided for the users.

3.1 Metadata Template and User Guidelines

The Nordic Metadata Project designed one short and one long version of the Metadata Template [6] containing the 15 elements and boxes to be filled out with content. Guidelines for inserting metadata were provided in two ways: as short “pop-up” menus in connection to the DC element field and as a stand-alone User Guide [7]. The guidelines included among other things: an introduction, how to use our metadata creation tool, a description of the standard set of DC Elements and syntax.

4 Methodology and study set-up

The goal was to collect data for analysis of real-life experience and usage of our Nordic DC Metadata creation tool and user guidelines. At the end of the template, we invited all users to participate in the questionnaire. All data collected was optional and anonymous. The questionnaire and data collection was managed electronically via a cgi-enhanced e-mail function.

Our approach was to gather both quantitatively and qualitatively in order to support each other in the requirement phase. The questionnaire contained of 3 parts: information regarding the users occupation and knowledge levels; secondly, 5 questions along a 5-point likert scale (see app. A), and finally, the third part contained 7 fields for open comments (5

of them belonging to the questions). The Nordic DC Metadata creation tool was evaluated between 1997 and 2000. During this period, 109 answers and 171 comments were gathered.

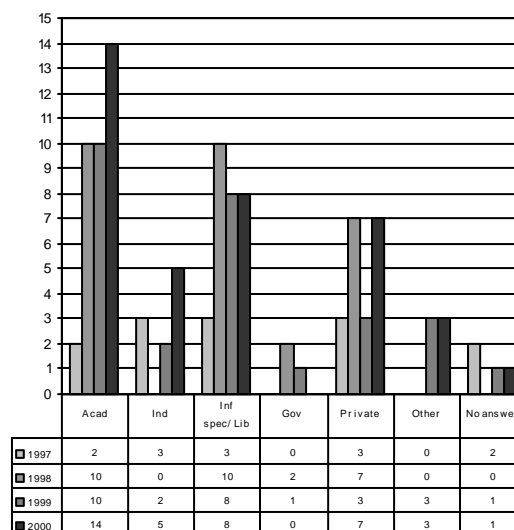
5 Findings

In the first part of this section we present the results concerning the background questions, in chapter 5.1, we present results from the 5-point Likert-scale and in chapter 5.2, we results from the written comments related to the questions.

Occupation.

Table 1 shows that 1/3 of the participants came from an academic setting. The three largest groups of participants came from the academic, library, and private sectors and correspond to 78% (33%-27%-18%) of all users. The smallest groups represented in our study were the governmental (3%) and Others (5%). Six participants (5%) did not submit complete questionnaires.

Table 1. Number of users and distribution over time/user category (N=109).

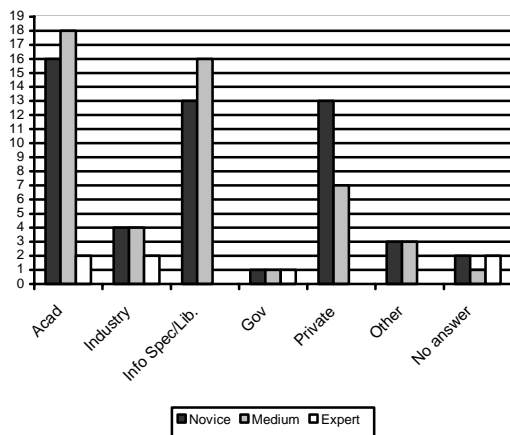


User knowledge.

This variable involves self-stated levels of perceived DC knowledge along 3 sub-categories: novice, medium and expert users (however, it is a well known fact that this is a subjective and thus problematic way of measuring).

Figure 1 shows that the groups of participants that had a high score of novice and medium knowledge levels, were academics (31% and 36%) and information specialists/librarians (25% and 32%), while users from the private sector (25%) had a high score for novices.

Figure 1: Distribution of numbers of users and knowledge levels during 1997-2000 (N=109)



5.1 Statistical analysis of questions

It must be noted that the answers related to the group of expert users (table 2) are too low in order to make any significant judgments. Instead, they should be seen as trends and indications

Table 2: Distribution of the numbers of answers for each of the 5 questions along a 5-point Likert-scale including mean values.

Question	1	2	3	4	5	Mean	Total
Question 1	Novice	5	4	7	14	12	42
	Medium	4	1	9	13	17	44
	Expert	1	1	0	3	0	5
	Total	10	6	16	30	29	91
Question 2	Novice	3	3	4	18	15	43
	Medium	3	1	4	13	23	44
	Expert	2	0	0	1	2	5
	Total	8	4	8	32	40	92
Question 3	Novice	1	7	13	12	10	43
	Medium	0	1	3	16	21	41
	Expert	0	0	0	1	3	4
	Total	1	8	16	29	34	88
Question 4	Novice	1	6	11	10	13	41
	Medium	2	3	6	19	13	43
	Expert	1	0	0	0	3	4
	Total	4	9	17	29	29	88
Question 5	Novice	0	4	9	13	14	40
	Medium	0	0	5	19	15	39
	Expert	0	0	0	0	3	3
	Total	0	4	14	32	32	82

Legend: 1=not satisfied; 3=uncertain/could not decide; 5=satisfied

Q1 focused on the *functionality* of the template and we observed a high score of satisfaction among users with novice and medium knowledge levels. However, a high degree (19%) of the novice user was also not/almost not satisfied (ratings 1 and 2) of all novice users. This question also has the lowest mean value of 3,68.

Concerning the variable of the *overall design* of the template (Q2), users with medium knowledge level showed a high mean value (4.18). On the expert level, we found a relatively high score for not satisfied with the overall design (3.20). An overall mean value for this variable was 4.00.

Q3 was concerned with the *understanding and semantics* of the DC elements in the template. Within the mean value section, we observe that there is a growing curve regarding the relationship between understanding and semantics of DC elements and user knowledge levels (novice to expert). The group of novice users had a relatively high understanding (51% on rating 4 and 5), but also a rather high degree not understanding the meaning (18% on rating 1 and 2 of all novice users). Users with medium knowledge had a high rating of 90% (rating 4 and 5) of all users with medium knowledge.

Question 4 asked if the users were satisfied with the *information provided in the User Guide*. 56% (rating 4 and 5) of all novice users did find the guidelines useful. An even higher degree of satisfaction was found among the users with medium knowledge levels (74% based on rating 4 and 5).

Question 5 investigated if the users were satisfied with *how the template and user guidelines interacted*. Medium knowledge users showed a high mean value of 4.26 and novices a mean value of 3.92.

Finally, if we assume that rating 3 would stand for uncertainty, we find that we have a high degree of uncertainty concerning novice users in question 3 (30%), question 4 (27%) and in question 5 (23%). The high scores in these questions might be the result of not having a clear opinion on the issue.

5.2 Analysis of written comments

In connection to the 5-Likert scale-based questions, additional field for written and open comments were made available for the users. Furthermore, 2 two stand-alone questions were added. A total of 171 comments were gathered.

The distribution of the comments (table 3) shows a high number of comments regarding the

functionality, and semantic issues of the template. Furthermore, there was a high number of comments made in the final comment field which concerned “other issues” to be considered and was mainly concerned with general requirements suggested by the users.

Table 3: Number of comments 1997-2000 in relation to total numbers of questions.

Question No.	1997-2000	
	Total	%
Q1: Functionality	38	22%
Q2: Overall design	12	7%
Q3: Semantic issues	28	16%
Q4: Guidelines and support	25	15%
Q5: Overall satisfaction	15	9%
Q6: Additional requirements	25	15%
Q7: Other comments	28*	19%
	N=171	

* New item from 1998-2000

In the analysis of single comment fields, we clustered the comments into different categories grouped into 3 columns: *category*, *numbers of comments* and *sub-topics* within the categories. Only categories with high rate are included in this section due to limited space.

Question 1 was concerned with the functionality of the Nordic DC Metadata creation tool and comments asked for was if the users had any problems with the template. 38 comments were submitted (22% of all comments). A large number of the comments were concerned with DC elements and the understanding of the DC Element field of Subject: classification (10 out of 38). Furthermore, the users asked for more information and examples on how to apply DC metadata to their own documents (9). Users also had general problems with the functionality of the Template (7), such as problems with copy & paste or using other web-browsers than Netscape and IE.

Question 2 asked about if there were any unclear parts of the overall design of the DC metadata template. A total of 12 comments were submitted (7% of all comments). Unclear aspects of the template, as reported by the user, were the DC Subject element field (7) and that the template was long and tedious to work with and therefore caused uncertainty among the users (5).

In the following question (*Q3*), we investigated if the users had any problems understanding any specific element in the template. 28 comments were made (16% of all comments made). These categories were usually without any content other than pointing out the object of concern and therefore only numbers are presented. The users had problems with the following DC elements:

Creator (5), Relation (4), Source (4), and Subject: Keyword (4).

Comments related to *Question 4* was concerned with if there was any further support that the users needed regarding the User Guidelines or if they were satisfied. A total of 25 comments were made (15% of all comments made). The issue that attracted most comments pointed out that users were satisfied with the information and instructions that the user guidelines provided (14). However, other comments pointed out that there should be instructions for specific domains (3), longer descriptions on the template page itself (2), or that the instructions and guidelines should be created in other more manageable formats such as .pdf or .ps.

The final comment related to the 5-Likert scale-based questions (*Question 5*) investigated if the users required other kinds of support when creating DC metadata. 15 comments were made (9% of all comments). A very high score was found regarding the interaction (11) in which the users pointed out that they found the support good and that no further support was needed. Furthermore, 4 comments pointed out that they encountered errors performing their metadata creation tasks. This may be due to the experimental stage in which the NM tool was in time to time.

Comment 6 asked the users if there were other facilities or requirements that they would like to see added to the DC metadata creation tool. 25 comments were made (15%). 8 comments pointed out that the users were satisfied with the template and 4 comments pointed out a need to be able to add new authorities to the list of controlled vocabulary for subject, such as ABN (Australian Bibliographic Network).

The final comment (*comment 7*) was concerned with any other general issue to be added. 28 comments were submitted (19%). The opinion that the template and tools are important work was highlighted (18 comments). Other comments pointed out the relation between DC metadata and search engines; that the process of filling out metadata was cumbersome; and finally, it was mentioned that the template was not stable.

6 Summary and Conclusion

The value of the evaluation and analysis of both qualitative and quantitative data is twofold:

- the evaluation may inform on the performance of the tool itself and consequently inform the design and redesign phases of the tool; and
- the evaluation may give valuable insight regarding requirements on the tool when

incorporated into a larger information access environment, e.g. within a Digital Library framework, it is important to see the tool in this context and user environment.

The methods may be used in an on-going and iterative evaluation and design process and variables may be checked and changes may be done iteratively.

Our findings may be viewed and used as a list of requirements or indicators for important issues to be considered. Regarding the design process of our metadata creation tool, a major redesign was made in 1998. During 1998 to 2000, requirements extracted from the study have been suggested to the design team and minor changes were made iteratively. In the context of understanding the development of our tool, the responses we got and the redesign process, we must remember that we were aware of that the users would come from different domains and that the user would have different levels of knowledge. However, we may report on some aspects from the findings:

A very important aspect of the study and the result is that the users did have different knowledge levels. Novices were both satisfied and not satisfied with the functionality of the template. However, one of the design criteria was to keep our tool on a general level in that it may support most of our users. In order to support different levels of user knowledge, we made two versions of the user guidelines (one "light" version and one more extensive and downloadable version). Furthermore, the users also came from very different domains with different knowledge about metadata.

The interaction model (functional and semantic) between different parts of a tool must be satisfied. Concerning the comments, users had problems with the semantics and especially understanding of the DC Element field of Subject: classification. The distinction between different elements must also be further enhanced in the instructions and the descriptions of the template and guidelines. Users with medium knowledge had a high degree of satisfaction regarding the interaction between template and guidelines. The problem with semantics is well known to the DC community and may be solved with more detailed explanations.

Tools must be user friendly in order to be used. We observed that there were a group of users pointing out that the Template were tedious and cumbersome and that this caused uncertainty among the users. This is an issue that has been taken into consideration, but the project preferred to keep a level in which we could attract as many

users as possible. However, we might anticipate that there will be a learning curve that will allow us to reduce some levels of information. This must be an issue in a future design phases.

The Guidelines or manuals must address the importance of different domains, user groups and individuals. We observed a growing level of understanding regarding the meaning of DC elements based on level of DC knowledge. This will have implication for the usefulness of the template. This implies that we may need to have differentiated guidelines. User were in general satisfied with the information and instructions that the user guidelines provided, but there was a need for more detailed examples on how to apply DC metadata to their own documents and domains.

References

- [1] Dublin Core Metadata Initiative: <http://dublincore.org/>
- [2] Fidel, R. (1994) User-centered indexing. JASIS, 45, pp. 572-576.
- [3] Nielsen, J. and Mack, R. L. (eds). (1994). Usability Inspection Methods. Wiley: New York.
- [4] Marshall, C. (1998). Making Metadata: a study of metadata creation for a mixed physical-digital collection. In Proceedings of the third ACM Conference on Digital libraries, 1998, Pages 162 – 171.
- [5] The NM Project: <http://www.lib.helsinki.fi/meta/>
- [6] Dublin Core metadata Template: http://www.ub2.lu.se/metadata/DC_creator.html
- [7] Nordic Metadata Project: User Guidelines: http://www.sics.se/~preben/DC/DC_guide.html

Appendix A.

Question 1: What do you think of the *functionality* that the Nordic DC Metadata creation tool provide?

Question 2: What did you think of the overall *design* of the DC Metadata template?

Question 3: Did you understand the *meaning* of the elements?

Question 4: Did the User Guidelines provided enough *support* to complete your DC metadata description task?

Question 5: Are you satisfied with how the template and user guidelines interacted?