

eContent Programme reUSE digital master files of printed publications

(eContent Project No. 11173)

Deliverable D No. 3.2 EVALUATION REPORT



Date of delivery: March 2006

Responsible author: Alenka Kavčič-Čolić

Author's organisation: National and University Library, Slovenia, (NUK)

Version: 1.0

**Contributors: Alenka Kavčič-Čolić (NUK), Mateja Šmid (Faculty for Civil and
Geodetic Engineering of the University of Ljubljana, Ljubljana, (LJU)), Darko
Majcenović (NUK), Matjaž Depolli (NUK),**

in cooperation with

**Špela Razpotnik (NUK) and Tobias Steinke (German National Library, Frankfurt,
(DDB))**

Executive summary

The main objective of the reUSE project was to establish digital repositories in demonstrator partners' institutions from Austria (the ALO repository jointly used by the University Libraries of Graz, Innsbruck and Linz), Germany (EDOC, repository of the Humboldt University Library, Berlin) and Estonia (DIGAR, repository of the National Library), and to fill them with digital master files of printed materials. These repositories were designed according to the OAIS Reference Model standard and RLG/OCLC trusted digital repository guidelines. All partners succeeded in achieving the set goals. However, due to different national roles of the partners' libraries and their different infrastructures, the results varied from library to library. In order to get optimal results, assure that the goals of the project were achieved, and advise the demonstrator partners in their efforts, three partners from Germany (National Library) and Slovenia (National and University Library, Faculty of Civil and Geodetic Engineering of the University of Ljubljana) had to evaluate the demonstrators' digital repositories. This report presents the findings of all the surveys undertaken during the project evaluation period.

The evaluation partners implemented an overall user-centered evaluation approach, which was focused on the external and internal organizational and political environment as concerns the implementation of long-term preservation policies and strategies; the technical characteristics and functioning of the repositories; and the services for the users, including their impact on and benefit to the local environment. These three topics were evaluated from different aspects. Since at the moment of deciding on the evaluation methods there was no methodology which would encompass an integrated interdisciplinary approach, we had to produce our own methodology, combining different methodologies and using parts of them for the purpose of evaluation. This new methodology combines two approaches: multi-attribute evaluation, which is based on different standards and methodologies; and a SWOT analysis of the three repositories.

The implementation of the project evaluation work package WP3 was scheduled for the period from July 2004 to December 2005. Due to the late completion of the demonstrators' repositories, the evaluation was extended to March 2006 and for the same reason the planned evaluation of the value-added services was excluded.

The main methods of gathering data were the questionnaires, interviews and site visits. We also clarified the information gathered on reUSE meetings. On 10th January and 1st December 2005 two evaluation meetings were organized in Ljubljana; the first one in order to define the evaluation methods and combine them with the White Paper survey approach, the second one to disseminate the most important findings of the surveys undertaken and to clarify the partners' missing information.

Initially the idea of the project was very well conceived. The main approach was to focus on the development of the digital repositories on RLG/OCLC concepts of trusted digital repositories and to base it on the reference model for an OAIS (ISO 14721/2003). These two starting points helped all partners to achieve the reUSE project goals.

The survey on the **organizational aspects** of the demonstrators showed a mature external environment in their countries, concerning political awareness of the importance of reUSE goals. From the point of view of internal environment the development of digital repositories has got a very strong support from the top management in all demonstrators' institutions. In all of them the library is in close cooperation with other departments, which are of key importance for the development and maintenance of the digital repositories. They have written policies for ingest, storage and long-term preservation processes within their repositories. The digital repositories are included in the mission statements of their

organizations. Most mandatory responsibilities of OAIS archives are fulfilled by all demonstrators' digital repositories.

Concerning the decision on metadata types, all reUSE demonstrators use the basic descriptive metadata set in Dublin Core 1.1, which is mapped in an XML encoding standard. The preservation metadata is collected according to the OAIS information model.

With all repositories, we can speak of financial viability. After the conclusion of the reUSE project all demonstrators' institutions plan to assign a part of the operational budget for the functioning and maintenance of the digital repository. They expect a necessary increase of approx. 5% to 20% in the investment sources.

Concerning organizational viability, the demonstrators had given special attention to building teams of permanent staff from different fields that could do a log in digital preservation. During the reUSE project the expertise regarding repository development and long-term preservation practice has increased considerably.

The evaluation of the **technical aspects** of the repositories showed them to vary in the design and in ways of achieving the same goal. They use different kinds of software – mainly open source and self developed software. All examined repositories can be upgraded with support for new file types and workflows. Service quality is also quite heavily dependent on the standards observed and used in the design of the repository. All partners followed basic guidelines of the OAIS reference model as much as possible and the recommendations in the RLG/OCLC Report on Trusted Digital Repositories.

All demonstrators took measures to assure the reliability of their repositories. They all have written policies for disaster readiness, response and recovery. Recovery includes a rebuild of the entire system from backup server or location – of which all partners have at least two. They all do daily backups and use systems to minimize the effect of failure on the entire service. Documentation and training is also provided for future reference in case of updating or maintaining the repositories, as well as instructions for content providers. Staff responsible for the repositories is trained additionally for the specific needs of the repository though that training is often non-formal and specific to the field of expertise of personnel working on the project.

Management of the content varies between partners; they use different takes of ingest process, supported file formats and metadata. Repository contents should also have various services attached to it and among these search and presentation of search results are paramount to the end-users. All partners allow for both simple and complex search within roughly similar data fields, but their presentations vary. On top of these there should also be some value-added services, though these were not available at the time of this evaluation and were not evaluated. Even so there are available services such as print-on-demand and CD delivery, multiple language support and indexation by external search engines.

The evaluation of the **repositories from the users' perspective** was very important since it gave us very important information on the functioning and use of the repositories. The users are an important element, which influences the vitality and success of the digital repository. There were two phases of survey in which questionnaires for users and non-users were distributed. They were asked about their opinions on repositories' content, recognition, performance and efficiency, error tolerance, and their personal and subjective perception of them.

Due to rather big differences between the repositories it was impossible to compare their user services, instead, we rather searched for interesting qualities and specifications of each repository and outlined them as advantages, which might be useful for different developers and implementers.

Many different opinions were expressed by the respondents in their comments, which should serve as a guide for implementers to improve their systems.

Our findings from both evaluation periods show that it is very important what purpose the repository serves and who accesses it. For example, while in digital repositories run by university libraries more recent scientific and professional publications should be available,

the national library should run repositories with more heterogeneous publications for different kinds of purposes. The fact that it is the users, who predetermine the collection should be highly considered. Implementers should understand and take into account who and why predetermines what and how! The key challenge then, is to know the target group and their primary needs.

Based on the survey responses, we suggested some recommendations, which might improve the usability of the repositories.

Additional purpose of the evaluation work package was to get repository's **cost-benefit** information. However, for this kind of information we did not have enough time, since the profit of these kind of services could be estimated on a long-term basis and not in such a short term. Consequently, our analysis was directed towards the expectations of the content providers and immediate benefits to content providers and librarians, who offer the service.

The survey showed that the main motivation of the majority of content provider respondents for their contents to reUSE repository was the access and dissemination of their contributing printed documents. The access to the publications for disabled people is a very important motivation factor of an average of 50% of respondents. Long-term preservation is a very important task, of which most of the respondents are aware. The content provider respondents do not see reUSE repositories as a way of achieving strategic goals, but rather to provide access to their contents. However, the possibility of increasing the interest in their organizations as well as increasing the dissemination of their activities and work was very highly assessed. They are less convinced that they would be able to increase the recognition of their organization among their stake-holders or attract new sponsors through reUSE repositories.

Most of the content provider respondents think that the benefits for their stake-holders would be more important than its costs. They expect neither any monetary advantages nor money saving. This means that some of the content providers do see the benefits in this collaboration to be more important than its costs. Broader access and dissemination as well as long-term preservation are very good reasons for collaboration.

The delivery system and negotiations between the partners' demonstrators and the content providers are functioning very well and do not require much additional time and effort from the content providers.

According to the data received from the librarians-operators of the repositories, it seems that the processing of electronic documents requires less time than the processing of printed materials. Where the metadata schema is not so complex as with the printed materials the processing could take only 4 minutes. Negotiations and document delivery require additional work and time.

At the end, considering all gathered information during the survey, a **SWOT analysis** was done describing the strengths, weaknesses, opportunities and threats for each individual repository.

Table of contents

Executive summary	2
Table of contents	5
Table of figures	8
Table of tables	11
Preface	12
INTRODUCTION / by Alenka Kavčič-Čolić	13
1 EVALUATION OF ORGANIZATIONAL ASPECTS OF reUSE DEMONSTRATORS' REPOSITORIES / by Alenka Kavčič-Čolić	16
1.1 BACKGROUND	16
1.2 EVALUATION METHODOLOGY	17
1.3 EVALUATION OF DIGITAL REPOSITORIES ACCORDING TO PRE-DEFINED ATTRIBUTES	18
1.3.1 POLICY	18
1.3.1.1 National policies related to digital repositories	18
1.3.1.2 Organizational policies concerning digital repositories	20
1.3.1.2.1 Organizational mission statements	20
1.3.1.2.2 Organizational structure	20
1.3.1.2.3 Digital preservation policy and strategy	21
1.3.2 PERCEPTION OF TOP MANAGEMENT	23
1.3.3 ORGANIZATIONAL CULTURE	24
1.3.4 METADATA	24
1.3.4.1 Preservation metadata	25
1.3.5 BUDGET PLANNING OF THE REPOSITORY	27
1.3.6 STAFF INVOLVED IN THE OPERATION AND ACCESS TO DIGITAL REPOSITORIES	28
1.3.7 CONTENT PROVIDERS	30
1.4 REVIEW OF ALL FINDINGS IN THE EVALUATION OF ORGANIZATIONAL ASPECTS	32
2 TECHNICAL ASPECTS EVALUATION / By Matjaž Depolli and Darko Majcenovič	34
2.1 BACKGROUND	34
2.2 EVALUATION METHODOLOGY	36

2.3 EVALUATION OF DIGITAL REPOSITORIES ACCORDING TO PRE-DEFINED ATTRIBUTES	37
2.3.1 QUALITY ISSUES	37
2.3.1.1 Functionality.....	37
2.3.1.1.1 Repository design	37
2.3.1.1.2 Logs and statistics.....	38
The EDOC repository's usage statistics for the time period from 1 st to 31 st October 2005:.....	38
The DIGAR repository's usage statistics for the time period from 26 th September to 25 th October 2005:	39
The ALO repository's usage statistics for the time period of August 2005	40
2.3.1.1.3 Security	41
2.3.1.2 Reliability	42
2.3.1.3 Efficiency	43
2.3.1.4 Maintainability	43
2.3.1.5 Flexibility	45
2.3.2 TECHNICAL SPECIFICATIONS	46
2.3.3 STANDARDS	46
2.3.4 SOFTWARE	47
2.3.5 REPOSITORY SYSTEMS & ADMINISTRATION.....	48
2.3.5.1 Collection access by user types	48
2.3.5.2 User administration	49
2.3.5.3 Content submission administration	49
2.3.6 CONTENT MANAGEMENT	50
2.3.6.1 Document formats	50
2.3.6.2 Metadata.....	51
2.3.6.3 Content import and export	52
2.3.6.4 Updating and indexing	53
2.3.7 SERVICES	53
2.3.7.1 Search capability	53
2.3.7.2 Presentation of search results	54
2.3.7.3 Value-added functions.....	54
2.4 REVIEW OF ALL FINDINGS IN THE EVALUATION OF TECHNICAL ASPECTS	54
3 EVALUATION FROM THE USERS' PERSPECTIVE / by Mateja Šmid	61
3.1 BACKGROUND.....	61
3.2 EVALUATION METHODOLOGY	61
3.2.1 METHODOLOGY OF DATA ACQUISITION	61
3.2.1.1 Target group	61
3.2.1.2 Sample.....	62
3.2.1.3 Type of inquiry	61
3.2.1.4 Period of inquiry.....	63
3.2.1.4 Procedure	63
3.2.2 METHODOLOGY OF DATA ANALYSIS	63

3.3 WORK DONE	64
3.4. END USER EVALUATION	66
3.4.1 DEMOGRAPHIC AND BASIC DATA	66
3.4.2 CONTENT	69
3.4.3 RECOGNITION	74
3.4.4 PERFORMANCE AND EFFICIENCY	78
3.4.5 PERSONAL AND SUBJECTIVE PERCEPTION	85
3.4.6 ERROR TOLERANCE AND SUITABILITY FOR INDIVIDUALIZATION	97
3.4.7 ADVANTAGES OF TRUSTED DIGITAL REPOSITORIES	102
3.5. NON-USER EVALUATION	103
3.6. REVIEW OF ALL FINDINGS IN THE EVALUATION OF THE REPOSITORIES FROM USERS' PERSPECTIVE	108
4 COST-BENEFIT ASPECTS / by Alenka Kavčič-Čolić	110
4.1 BACKGROUND	110
4.2 EVALUATION METHODOLOGY	110
4.3 CONTENT PROVIDERS SURVEY	112
4.3.1 CONTENT PROVIDERS' INTEREST IN INCREASING THE ACCESS TO THEIR PUBLICATIONS	112
4.3.2. LONG-TERM PRESERVATION	118
4.3.3. STRATEGIC REASONS FOR COOPERATION WITH REUSE DEMONSTRATORS	118
4.4 LIBRARIANS IN CHARGE OF DIGITAL REPOSITORIES	130
4.5 REVIEW OF THE FINDINGS OF THE COST BENEFIT ISSUES' EVALUATION	130
4.5.1 FINDINGS OF THE SURVEY ON CONTENT PROVIDERS	130
4.5.2 FINDINGS OF THE SURVEY ON THE LIBRARIANS	131
5 SWOT ANALYSIS / by Alenka Kavčič-Čolić, Darko Majcenović, Mateja Šmid 	132
6 REFERENCES	141
ANNEXES	
Annex 1: Used abbreviations	144
Annex 2: Glossary	146
Annex 3: Distributed questionnaires	151

Table of figures

Figure 1: <i>Evaluation framework</i>	14
Figure 2: Total number of documents held by reUSE repositories.....	31
Figure 3: <i>Comparison of visits, sites and accesses to EDOC repository during one month survey phase</i>	39
Figure 4: <i>Comparison of the total number of queries and the number of sessions in DIGAR's one month survey phase</i>	40
Figure 5: <i>Total number of visits in ALO's one-month survey phase</i>	41
Figure 6: <i>An example of the end-user questionnaire with special fonts designed according to Remark Office OMR software.</i>	62
Figure 7: <i>Marking OMR fields which will later be scanned in each questionnaire.</i>	64
Figure 8: <i>The number of returned end-user and non-user questionnaires from ALO, DIGAR and EDOC in June 2005</i>	65
Figure 9: <i>The number of returned end-user and non-user questionnaires from ALO, DIGAR and EDOC in October 2005</i>	66
Figure 10 and Figure 11: <i>The age of end-users in all three libraries for both evaluation phases is presented</i>	66
Figure 12: <i>Detailed basic data for the end-user survey in June 2005</i>	67
Figure 13: <i>Detailed basic data for the end-user survey in October 2005</i>	67
Figure 14: <i>The format preferences from the ranked list of searching criteria for ALO and DIGAR repositories for the 1st evaluation phase (June 2005)</i>	68
Figure 15: <i>The format preferences from the ranked list of searching criteria for ALO, DIGAR and EDOC for the 2nd evaluation phase (October 2005)</i>	68
Figure 16: <i>The diagram presents the ranked list of searching criteria for the 1st evaluation phase (June 2005)</i>	69
Figure 17: <i>The ranked list of searching criteria for the 2nd evaluation phase (October 2005)</i>	69
Figure 18: <i>Comparison of results from: Content among ALO and DIGAR end-users in the 1st evaluation period (June 2005)</i>	72
Figure 19: <i>Comparison of results from: Content among ALO and DIGAR end-users in the 2nd evaluation period (October 2005)</i>	73
Figure 20: <i>Correlation (*) between long procedure before finding and opening a certain document and usage among ALO end-users (June 2005)</i>	74
Figure 21: <i>Correlation (*) between long procedure before finding and opening a certain document and experience among ALO end-users (June 2005)</i>	74
Figure 22: <i>Correlation (*) between long procedure before finding and opening a certain document and year of birth among ALO end-users (June 2005)</i>	75
Figure 23: <i>Comparison of results from part: Recognition among ALO and DIGAR end-users in the 1st evaluation period (June 2005)</i>	76
Figure 24: <i>Comparison of results from part: Recognition among ALO, DIGAR and EDOC end-users in the 2nd evaluation period (October in 2005)</i>	77
Figure 25: <i>Comparison of results from part: Performance and efficiency (part I) among ALO, DIGAR and EDOC end-users in the 1st evaluation period (June 2005)</i>	79
Figure 26: <i>Comparison of results from part: Performance and efficiency (part I) among ALO, DIGAR and EDOC end-users in the 2nd evaluation period (October 2005)</i>	80
Figure 27: <i>Comparison of results from part: Performance and efficiency (part II) among ALO, DIGAR and EDOC end-users in the 1st evaluation period (June 2005)</i>	81
Figure 28: <i>Comparison of results from part: Performance and efficiency (part II) among ALO, DIGAR and EDOC end-users in the 2nd evaluation period (October 2005)</i>	82
Figure 29: <i>The amount of time needed to get the requested item (1st evaluation period, June 2005)</i>	83
Figure 30: <i>The amount of time needed to get the requested item (2nd evaluation period, October 2005)</i>	83

Figure 31: <i>The amount of time needed to find the proper information (1st evaluation period, June 2005)</i>	84
Figure 32: <i>The amount of time needed to find the proper information (2nd evaluation period, October 2005)</i>	84
Figure 33: <i>Comparison of results from part: Personal and subjective perception (part I) among ALO and DIGAR end-users in the 1st evaluation period (June 2005)</i>	86
Figure 34: <i>Comparison of results from part: Personal and subjective perception (part I) among ALO and DIGAR end-users in the 2nd evaluation period (October 2005)</i>	87
Figure 35: <i>Comparison of results from part: Personal and subjective perception (part II) among ALO and DIGAR end-users in the 1st evaluation period (June 2005)</i>	88
Figure 36: <i>Comparison of results from part: Personal and subjective perception (part II) among ALO and DIGAR end-users in the 2nd evaluation period (October 2005)</i>	89
Figure 37: <i>Comparison of results in part Personal and subjective perception (part III) among ALO and DIGAR end-users in the 1st evaluation period (June 2005)</i>	90
Figure 38: <i>Comparison of results from part: Personal and subjective perception (part III) among ALO and DIGAR end-users in the 2nd evaluation period (October 2005)</i>	91
Figure 39: <i>Correlation (*) between use of e-documents and age from the 1st evaluation period (June 2005)</i>	92
Figure 40: <i>Correlation (*) between preferences of e-documents and age from the 1st evaluation period (June 2005)</i>	93
Figure 41: <i>Correlation (*) between future use of e-documents and age from the 1st evaluation period (June 2005)</i>	93
Figure 42: <i>Interface of the ALO repository</i>	94
Figure 43: <i>The interface of the DIGAR repository</i>	95
Figure 44: <i>The searching engine at DIGAR digital repository</i>	95
Figure 45: <i>Interface of EDOC digital repository</i>	96
Figure 46: <i>Comparison of results from part: Error tolerance and suitability for individualization (part I) among ALO and DIGAR end-users in the 1st evaluation period (June 2005)</i>	98
Figure 47: <i>Comparison of results from part: Error tolerance and suitability for individualization (part I) among ALO and DIGAR end-users in the 2nd evaluation period (October 2005)</i>	99
Figure 48: <i>Comparison of results from part: Error tolerance and suitability for individualization (part II) among ALO and DIGAR end-users in the 1st evaluation period (June 2005)</i>	100
Figure 49: <i>Comparison of results from part: Error tolerance and suitability for individualization (part II) among ALO and DIGAR end-users in the 2nd evaluation period (October 2005)</i>	101
Figure 50 and Figure 51: <i>While the most preferred advantage for ALO end-users is document availability, answers show that among DIGAR end-users the most important advantage is full-text</i>	102
Figure 52, Figure 53 and Figure 54: <i>The preferred advantages for each repository obtained in the 2nd evaluation phase (October 2005)</i>	103
Figure 55: <i>Correlation (*) between age and future use of the digital repository in the June evaluation</i>	104
Figure 56: <i>Correlation (*) between age and future use of the digital repository in the October evaluation</i>	104
Figure 57: <i>Correlation (*) between repositories, age and use by repositories in the October evaluation</i>	104
Figure 58: <i>Detailed basic data for the non-user survey in June 2005</i>	105
Figure 59: <i>Detailed basic data for the non-user survey in October 2005</i>	106
Figure 60: <i>The future use of the digital repositories among non-users from the 1st evaluation period (June 2005)</i>	107
Figure 61: <i>The future use of the digital repositories among non-users from the 2nd evaluation period (October 2005)</i>	107
Figure 62: <i>Ratio between the total number of content providers and received responses</i> ..	111
Figure 63: <i>Type and number of content providers that responded to the questionnaire</i>	111

Figure 64: Main motivation for collaborating in the reUSE digital repository related to the access and dissemination of publications	114
Figure 65: Easier and faster access to their publications as one of the motivation factors of the content providers for contributing to the reUSE repository.....	115
Figure 66: Possibility for disabled people to access to content providers' publications as one of the motivations for contribution to the reUSE repository.....	116
Figure 67: Easier and broader distribution of content providers' printed documents as one of the motivation factors for their contribution to the reUSE repository.....	117
Figure 68: Increased public use of content providers' publications as one of the motivation factors for their contribution to the reUSE repository	117
Figure 69: Long-term preservation of content providers' electronic publications as one of the motivation factors for their contribution to the reUSE repository.....	118
Figure 70: Main motivation factor for collaborating with the reUSE digital repository related strategic issues.....	120
Figure 71: Increased public interest in content providers' organization as one of the motivation factors for their contribution to the reUSE repository.....	121
Figure 72: Increased dissemination of content providers' activities and work as one of the motivation factors for their contribution to the reUSE repository.....	122
Figure 73: Increased recognition of content providers' organization among their stake-holders as one of the motivation factor for their contribution to the reUSE repository.....	123
Figure 74: Attraction of new sponsors as one of the motivation factors of the content providers for contributing to the reUSE repository	124
Figure 75: Content providers respondents' expectation concerning the benefits of the digital repository for their stake-holders.....	125
Figure 76: Expectations concerning monetary advantages by collaborating in the reUSE project.....	126
Figure 77: Opinion of the content providers respondents on the expenses for the repository.	127
Figure 78: Content provider respondents' expectations concerning saving money with the reUSE project.....	128
Figure 79: Content provider respondents' expectations regarding investing more time in producing the electronic format of printed publications.....	129
Figure 80: Content providers respondents' expectations concerning negotiations and transfer of electronic versions.....	129

List of tables

Table 1: <i>Evaluation attributes related to organizational aspects of reUSE repositories</i>	17
Table 2: <i>Written policies regarding demonstrators' repositories</i>	21
Table 3: <i>Metadata schema used by reUSE demonstrators</i>	24
Table 4: <i>Representation information and Preservation information according to OAIS Information Model</i>	26
Table 5: <i>Methods used by demonstrators' institutions to increase the level of staff expertise concerning digital preservation</i>	29
Table 6: <i>Number of content providers by reUSE demonstrators</i>	30
Table 7: <i>Review of all findings in the evaluation of organizational aspects</i>	32
Table 8: <i>Technical aspects evaluation attributes and related technical aspects</i>	36
Table 9: <i>Basic information on reUSE demonstrators' repositories</i>	38
Table 10: <i>Staff involved in the operation and maintenance of reUSE demonstrators' repositories</i>	44
Table 11: <i>Software in reUSE demonstrators' repositories</i>	47
Table 12: <i>reUSE demonstrators' submission formats</i>	50
Table 13: <i>reUSE demonstrators' storage formats</i>	51
Table 14: <i>reUSE demonstrators' dissemination formats</i>	51
Table 15: <i>Searchable metadata fields</i>	53
Table 16: <i>Search types</i>	53
Table 17: <i>Ability to browse search results</i>	54
Table 18: <i>Summary of partners' repositories evaluation findings with the description of basic properties of each repository</i>	56
Table 19: <i>The list of contents which end-users would like to find in the digital repository - data from the 1st evaluation period (June 2005)</i>	70
Table 20: <i>The list of contents, which end-users would like to find in the digital repository - data from the 2nd evaluation period (October 2005)</i>	71
Table 21: <i>Ranged advantages among ALO and DIGAR end-users in the 1st evaluation period (June 2005)</i>	102
Table 22: <i>Ranged advantages among ALO, DIGAR and EDOC end users in the 2nd evaluation period (October 2005)</i>	102
Table 23: <i>Access and dissemination aspects that produced interest among the content providers in cooperation with reUSE repositories</i>	112
Table 24: <i>Strategic aspects quoted as reason for cooperation with reUSE repositories</i>	119

PREFACE

The reUSE project was started on a very firm foundation, which was developed in a ten-year long research and development in the field of digital archiving and long-term preservation of the access to digital publications all over the world. We need also mention the endeavours of the NEDLIB project partners, which influenced the affirmation of the reference model for an Open Archival Information System (OAIS), which years later became an important standard (ISO 14721/2003) for building digital repositories, or the RLG and OCLC efforts and activities, which for the first time defined the concept of trusted digital repositories. There are also many successful and failed experiences of many national and university libraries, which were used in the development of reUSE digital repositories.

However, the experience of reUSE is unique in many aspects, i.e.:

- For the first time, we are trying to implement the new concept of trusted digital repositories to digital deposit institutions from different systems, approaches and designed communities;
- For the first time we are saving the digital master files of printed documents and providing them with new forms of digital access;
- We are developing new forms of collaboration with the content providers;
- For the first time we are implementing an evaluation of digital repositories with a holistic approach.

The experience of the project was very valuable not only for the demonstrators but also for the evaluators of the project. We were able to follow very closely the experiences of the demonstrators and, in order to evaluate their repositories we had to review and analyse the experiences of other libraries.

Bust most importantly, we started a new methodology for the evaluation of digital repositories, which could be used for the assessment of other repositories. This Evaluation Report is an evidence of quite different experiences, which could guide those libraries that are still deciding on or are just developing their own digital repositories. Looking back, we sometimes think that it would have been much easier for us if we had known what we know now, when we started developing our repositories. We hope that the findings in this Report are useful to the rest of the library community.

Alenka Kavčič-Čolić
reUSE Evaluation Work Package Coordinator

INTRODUCTION

By Alenka Kavčič-Čolić

The main objective of the reUSE project was to establish trusted digital repositories in demonstrator partners' institutions and to fill them with the digital master files which have been the basis for printed publications. In this regard two main goals have been reached: the repositories will guarantee long-term availability of the digital content, which is nowadays lost, and the digital content, which is distributed over thousands of public sector organizations will be directly accessible via the Internet and via electronic library catalogues.

Such digital repositories were set up at the University Library of Innsbruck (UBI), University Library of Graz (UBG), University Library of Linz (I3S3), University Library of Berlin (UBER) and the National Library of Estonia (NLE). At the beginning of the project, the Austrian National Library (ONB) also joined the demonstrators, but unfortunately was not able to complete the repository for the evaluation phase.

Another reUSE important project goal was to set up an evaluation framework for reviewing the results of the project and to create clear and objective data for further exploitation of the model on a European level. Three organizations were involved in the evaluation activities: the Faculty of Civil and Geodetic Engineering of the University of Ljubljana (LJU), the National Library of Germany (DDB) and the National and University Library of Slovenia (NUK), which was also the coordinator of this work package.

Setting up the evaluation framework was the most important phase in the evaluation process. It defined the overall methodological approach as well as methodology for the evaluation of the digital repositories.

Our assumption was that the overall aim of the evaluation was to make user-centred repositories, which would be at the same time most efficient in technical and organizational regards. Following this assumption we defined the aspects, which could mostly influence the development, operation and maintenance of the trusted digital repositories, and on which the evaluation should be focused, i.e.: (See **Figure 1**)

- internal and external organizational environments together with long-term preservation planning policy and strategy (dealt in project work package WP1),
- the digital repositories developed by the demonstrators (dealt in project work package WP1) from their technical and content points of view,
- the services, including value-added services to the users (dealt in project work packages WP2 and WP4) and their impact on and benefit to the local environment.

These three topics have been evaluated from different aspects.

Our purpose was to gather all the existing experiences in the field of digital repository building. But soon we found out that we were entering a completely new field and that there was no methodology that could be used for the project. To use the existing methodology would be easier for the evaluation process, but at the moment of making a comparative analysis or compiling different results, this approach could fail, because every methodology depends on a pre-defined context. Furthermore, the existing methodology is mainly centred on specialized research topics. Therefore, we had to decide whether to define our own methodology which would be based on an interdisciplinary approach, or use various types of methodology independently.

We needed an objective and an overall approach, which would provide us with the whole picture of the repositories and help us to get a holistic overview of their operation. Therefore, we decided to develop a new methodology integrating different types of methodology, using

various parts for the purpose of evaluation in order to build a new and more comprehensive evaluation framework, for which we prepared a table of attributes.

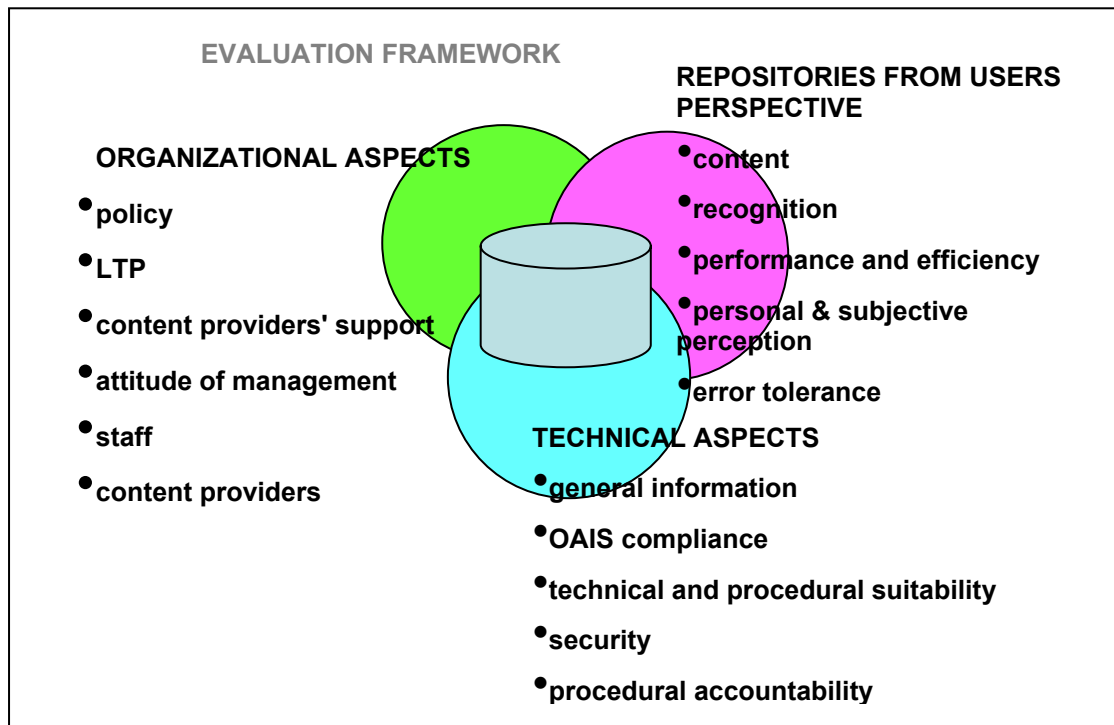


Figure 1: Evaluation framework

This new methodology combined two approaches:

- Multi-attribute evaluation, based on different standards and types of methodology; and
- SWOT analysis of the three repositories.

For the multi-attribute evaluation we consulted different standards and types of methodology and extracted the relevant attributes for the purpose of the evaluation. In this regard, we used DINI methodology for the definition of evaluation attributes and criteria of the Organizational aspects. For the evaluation of the technical aspect of the repository and evaluation of services we used the experiences of the European 5th Framework Program project SciX, which produced an extensible assessment and evaluation report. The coordinator of the project was the Faculty of Civil and Geodetic Engineering of the University of Ljubljana, which is one of the evaluators in this project. We consulted several standards related to usability and software quality as well, i.e. ISO/IEC 9126 (internal SW quality), ISO 9241-11 (usability), ISO 13407 I (human centered design processes for interactive systems), ISO 12207: (Software lifecycle processes), and ISO 14721 (OASIS).

The implementation of the project evaluation work package WP3 was scheduled for the period from July 2004 to December 2005. However, due to the late completion of the repositories and a very unsuitable summer time for evaluation, as the number of library users drastically decreased, we had to postpone the evaluation from July 2005 to September 2005. Consequently, we had to extend the report delivery deadline to March 2006 instead of December 2005.

For the same reason the project value-added services work package WP4 was postponed and there were no available results yet to provide us with data for the evaluation. Thus, the evaluation was focused on different aspects of the trusted digital repositories and their services excluding value-added services.

The main methods of gathering data were questionnaires, interviews and on-site visits. We also used the opportunity of reUSE meetings for clarifying the gathered data. On 10th January and 1st December 2005 two evaluation meetings were organized in Ljubljana, the first one in order to define the evaluation methods and combine them with the White Paper survey approach, the second one to disseminate the most important findings of the surveys undertaken and to clarify the partners' missing information. After the meeting, the first draft was produced and sent to DDB for their review.

The results obtained are described in the following chapters. We started with the organizational aspects, which are followed by the technical aspects of the repository and the evaluation of the repository from the users' perspective. This is followed by a qualitative cost-benefit survey from the point of view of the content providers and librarian operators of the digital repositories, and finally by a SWOT analysis.

1 EVALUATION OF ORGANIZATIONAL ASPECTS OF reUSE DEMONSTRATORS' REPOSITORIES

By Alenka Kavčič-Čolić

1.1 BACKGROUND

According to the definition by RLG and OCLC *a trusted digital repository is the one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future.*¹ (p.13) The word "trust" refers to a number of high-level organizational and curator responsibilities that the digital archival organization has to assume. The RLG/OCLC report define the responsibilities of such repositories, which relate to their administrative responsibility, organizational viability, financial sustainability, technological and procedural suitability, system security and procedural accountability (pp. 21-23). A trusted digital repository should fulfil the following conditions (p.13):

- it has to accept the responsibility for the long-term maintenance of digital resources on behalf of its depositors and for the benefit of current and future users;
- it must have an organizational system that supports not only long-term viability of the repository, but also the digital information for which it is responsible;
- it has to demonstrate fiscal responsibility and sustainability;
- the design of its system should be in accordance with commonly accepted conventions and standards to ensure the on-going management, access, and security of materials deposited within it;
- it should establish types of methodology for system evaluation that meet community expectations of trustworthiness;
- it must be dependable to carry out its long-term responsibilities to depositors and users openly and explicitly;
- it must have policies, practices, and performance that can be audited and measured.

reUSE demonstrators agreed to develop trustworthy digital repositories since the reUSE Digital Repositories White Paper stated that "All reUSE demonstrator organizations plan to certify their repository as a trusted digital repository as soon as such a certification scheme is available." (Digital Repositories White Paper, p. 13).

The achievement of these goals depends on the organizational internal and external environment. One of the purposes of the evaluation process of the reUSE repositories was to find out to what extent these organizations fulfil all the conditions specified by RLG/OCLC for setting up trusted digital repositories and to what extent the goals defined in the reUSE White Paper were achieved.

¹ Trusted Digital Repositories : Attributes and Responsibilities. An RLG-OCLC Report. - RLG : Mountain View, CA. - May 2002. (<http://www.rlg.org/en/pdfs/repositories.pdf>)

1.2 EVALUATION METHODOLOGY

The development, operation, and maintenance of a trusted digital repository are based upon a favourable organizational policy and long-term strategy, its management, financing and available resources. In the Evaluation Framework we defined the most important attributes, which we evaluated in the evaluation phase. They could be divided in organizational and technical aspects. While the technical aspect will be discussed in a separate chapter, the organizational aspect describes the environment, in which the digital repository is operating (See **Table 1**). Every evaluation of digital repositories should consider the context in which the repositories were developed and their main functions in it. The context tells us a lot about the goals and mission of the repository, the interests of the stake-holders in its maintenance. It also tells us whether the organizational strategy and organizational culture are favourable to the development of such a repository and whether the organization is prepared to undertake such an important task. If, for instance, the management is not interested in including the methods of future finance of the repository in the strategic plan, or does not consider its role in the community important, we can hardly speak of its viability. It is also necessary to consider the national legal framework concerning electronic archiving, which could represent an obstacle or a promoter in the implementation of its roles.

Table 1: Evaluation attributes related to organizational aspects of reUSE repositories

Attributes	Evaluation topics	Detailed topics
1. Policy	Existence of a public policy	Policy regarding: <ul style="list-style-type: none"> • standards • operation of the repository • content • management • rights and obligation of repository operators • rights and obligation of authors and/or editors • services • long-term preservation • metadata
2. Content providers' support	Guidelines for content providers	<ul style="list-style-type: none"> • formats • design • data transfer • metadata
3. Legal aspects	Clearly defined all copyright holders and copyright arrangement with them	<ul style="list-style-type: none"> • legal foundation for archiving • type of rights • rights owners
4. Staff	Number of staff members dealing with the repository and their professional profiles. Influence on decision-making	<ul style="list-style-type: none"> • number • professional profile • affiliation • organizational level/situation (chart)

In the development of our evaluation framework of organizational matters we used different sources (DINI, RLG/OCLC, ISO standards) to find out which are the topics that could help us describe the organizational context. In this regard, we decided to focus on the following:

- Existence of national and organizational policies concerning digital repositories – legal framework, organizational strategy, budget planning etc.,
- Rights management of digital material,

- Preservation planning,
- Staff involved in the operation and access to the digital repository,
- Level of knowledge concerning digital preservation,
- Funding of the repository planning,
- OAIS mandatory responsibilities fulfilled by the repositories,
- Metadata types,
- Organizational culture,
- Perception of top management attitude.

For each of the listed aspects we tried to get information through a questionnaire or an interview. It was important to gather the data from different levels in the organization. It is possible that the understanding and perception of the top management of the core organizations of the mission of the digital repository differs from the perception and understanding of the librarians or information specialists in charge of the operation of the repository.

That is why there was a detailed questionnaire prepared for the staff in charge of the repository. Their answers provided us with contextual data on the practical level of the repository and its users. Being aware that the top management will hardly take more than five minutes to answer our questionnaire, we tried to interview them by phone and evaluate their responses.

The continuation includes the information we got for every aspect regarding the organizational context.

1.3 EVALUATION OF DIGITAL REPOSITORIES ACCORDING TO PRE-DEFINED ATTRIBUTES

1.3.1 POLICY

1.3.1.1 National policies related to digital repositories

The country political framework, which could affect the activities of the organization changes slowly and depends on many factors that are not under the organizational control. In the field of electronic publications' archiving we have to distinguish two kinds of institutions, deposit and non-deposit ones. reUSE partners, according to their functions in their local environment, intend to become electronic deposit institutions. This gives national libraries an advantage, since their activities are clearly defined in the national legislation, especially in Legal Deposit Law. However, in reUSE demonstrators' countries the Legal Deposit legislation still does not address on-line publications or digital master files of printed publications, and consequently, their acquisition depends on voluntary basis.

The University Library of Innsbruck (UBI) is a deposit library for the Austrian region of Tyrol and the University Library of Graz (UBG) is the deposit library of the Austrian region of Styria. As university libraries they collect academic electronic production and as deposit libraries also collect publications from the government and the public sector.

The partner I3S3, which is actually the Information sciences Institute, closely cooperates with the University library of the Linz University. As a regional deposit institution this library has a similar role as other two Austrian university libraries.

UBER maintains a thematic repository that is focused on the university. Their content providers are members of Humboldt University.

In the last ten years European countries have become aware of the urgent need of digital preservation. UNESCO initiatives have contributed a lot with their Charter for the Preservation of Digital Heritage. Documents and guidelines were produced in order to disseminate the importance of digital preservation to all European policy makers and relevant Ministries.

In Europe, several national and international projects, focused on preservation aspects, have recently been completed (NEDLIB, DELOS). There are many national and international initiatives also related to long-term preservation of digital materials. In Germany a national digital preservation forum NESTOR was constituted, which is designed as a network of several institutions that are involved in digital preservation. Another initiative is KOPAL, German acronym for Cooperative development of a long-term digital information archive, a project funded by the German Ministry for Education and Research. The KOPAL project's main goal is to solve problems in digital preservation.

In Austria there are some initiatives for building a network similar to NESTOR, but we do not know whether they have been successful.

In Estonia the national policies concerning digital preservation are accessible on the web. In Germany this will be possible with the NESTOR project.

The awareness of the importance of long-term preservation issues is present in most European national and university libraries as well. The recent survey on the long-term preservation of digital documents at European libraries conducted by the ARGE Group² (in continuation referred as ARGE Survey) has shown that major European, particularly national libraries, consider the long-term preservation of digital documents a very important (76%) or one of the most important tasks (23%). At present approximately 38% of the questioned libraries already maintain a repository of digital documents. The same percentage plans to establish digital repositories in 2006 in order to ensure a long-term preserving of the digital documents. The main reason for this is that they are aware of the important role libraries have in this field. 91% of responding libraries declared that long-term preservation of digital documents should be an intrinsic library task.

What is the situation in reUSE demonstrators' countries? According to the ARGE survey mentioned above, in Austria 3 out of 12 libraries surveyed already maintain a repository of digital documents and another 5 of them plan to establish one very soon. In Estonia unfortunately there were only two libraries surveyed and we suppose that one of them was the national library, which established a digital repository during the reUSE project. Germany has by far the highest percentage of existing repositories of digital documents. 30 out of 46 libraries declared that they already have a digital repository and 8 of them plan to establish one in the near future. We see the reason for this situation in a high government awareness and its preparedness to support national policies and programmes in this field.

² Krimbacher, Monika, Neuhauser, Michael, Vogl, Martina: Survey on the long-term preservation of digital documents in European libraries 2005. - Innsbruck : ARGE Kulturconsulting - Medienkonzepte - Marktforschung ; Innsbruck : reUSE Project, 2005.

If we compare the national policies of Austria, Estonia and Germany we can say that Germany's is the most favourable to digital preservation. Estonia is doing a lot in this field and is not far behind. Such level of awareness is present at institutional and governmental levels. The last is very important, since it could be favourable for promoting national programmes and financing of electronic deposit institutions at a national level.

1.3.1.2 Organizational policies concerning digital repositories

Before continuing with this section, it is necessary to point out that although there are 5 demonstrators in the project, three of them are cooperating in the development and maintenance of one repository. This is the case of UBI, UBG and I3S3, which contribute to building their reUSE collection under the ALO repository. UBI is the main deposit institution, while the content providers and different policy arrangements regarding the access to and the building of the digital collection differ in each partners' organization. Since we are trying to analyze organizational aspects, we have to focus on partners' organizations. In the next sections, when talking about the repositories themselves we are focused on three repositories, i.e. **ALO** (UBI, UBG, I3S3), **EDOC** (UBER) and **DIGAR** (NLE).

1.3.1.2.1 Organizational mission statements

The RLG/OCLC report says that the organizations that choose to become trusted digital repositories have to establish themselves in ways that demonstrate their viability. This way "their mission statements will reflect a commitment to the long-term retention, management of, and access to digital cultural assets on behalf of depositors and users."³

As the organizational level is concerned, UBI, I3S3, NLE and UBER have written mission statements of the organization and the repository. In NLE and UBER they are web accessible.

UBI, NLE and UBER provide a written statement about the repository to the designated community.

UBG and NLE have included the goals, operation and maintenance of the reUSE repository in the organizational strategy, while UBI plans to do so.

All demonstrator partners have written mission statements of the organization and the repository.

1.3.1.2.2 Organizational structure

The assessment and description of existing and planned digital repositories in the mentioned ARGE survey show that in most European libraries "long-term preservation of digital documents is accomplished cooperatively by several departments (46%). In 40% of the libraries, an existing department was charged with this task, whereas 14% created an entirely new department dedicated to the long-term preservation of digital documents" (p.75). A very similar situation can be found in reUSE demonstrator partners.

In NLE, the development and maintenance of the repository are performed by the collaborative effort of the Information Systems Department, Collection Development Department and Marketing Manager. The whole initiative is supervised by the Director of

³ RLG/OCLC Trusted digital repositories : report p. 17.

Technology Services and approved by the Director General. This is important for securing the continuity and sustainability of the repository. As the national library, NLE is responsible for long-term preservation of electronic publications.

At UBI a department has been designed for development and maintenance of the digital repository. This operates in close cooperation with the university library. The other two university libraries in Graz and Linz do have digital repositories, but they are not trusted digital repositories. Nevertheless, they contribute digital contents to the joint ALO repository in Innsbruck.

At UBER the departments in charge of the repository are the Electronic Publishing Group of the Computer and Media Services and the University Library with permanent and project staff dedicated to the continuous further development of the collection EDOC, ensuring this way the high quality and sustainability of the service. The university library of UBER is responsible for the deposit and long-term preservation of doctoral dissertations and scientific publications of UBER and now also for electronic publications in the public sector (reports, yearbooks and studies from non-profit organizations in the area of science, culture and history). UBER digital repository has been certified by DINI.

In all organizations, the digital repository is developed by the infrastructure and service departments, which are far from the decision making level when it comes to universities. They closely cooperate with the library staff. At NLE the organizational structure is more flat and top management is more embedded in digital repository activities. However, these facts have no influence on the operating of the repository. The core ALO repository is physically located At UBI, and its collections are built and maintained jointly by UBI, UBG and I3S3.

1.3.1.2.3 Digital preservation policy and strategy

The ARGE survey (p. 58) showed that European libraries, which already have a digital repository are much more likely to have related written guidelines. This is the case in 20% of all questioned libraries.

NLE, UBI and UBER provide written policies for the ingest, storage and long-term preservation processes within their repositories. At I3S3 there are written policies for the ingest and storage processes only. These policies are web accessible at NLE and UBER, while at UBI only the ingest process policy is web accessible. Written policies for the access to the repositories are available at NLE and UBER only, and they are web accessible. (See **Table 2**)

Table 2: *Written policies regarding demonstrators' repositories*

Do you provide written policies for the:	Yes, web accessible	Yes, but not web accessible	No
Ingest process	NLE, UBER, UBI	I3S3	
Storage process	NLE, UBER	UBI, I3S3	
Long-term preservation process	NLE, UBER	UBI	I3S3
Access to the repository	NLE, UBER		I3S3, UBI

Most of the mentioned policies did not exist or were not available to the public during the White Paper's production, which means that this is an important output of the reUSE project in the development of digital trusted repositories.

At NLE and UBER there is a well defined preservation planning policy. At NLE all preservation activities are based on half-year while at UBER only on quarterly work plans. They review their written policies, NLE on an annual and UBER on a biannual basis. The preservation processes are regularly audited in order to assure their quality. At NLE this is done every 6 months, while at UBER every year. At UBI there is no specific digital preservation planning policy yet. At I3S3 there is no preservation planning at all.

According to the RLG/OCLC report, a trusted digital repository should be in compliance with the reference model for an OAIS. Most mandatory responsibilities of OAIS archives as stated in the RLG/OCLC report and reUSE White Paper⁴ are fulfilled by the three repositories. These are (White paper, p. 82):

- They negotiate for and accept appropriate information from information producers
- They obtain sufficient control over the information provided for the level needed to ensure long-term preservation
- They determine, either by itself or in conjunction with other parties, which communities should become the Designated Community and, therefore, should be able to understand the information provided
- They ensure that the information to be preserved is independently understandable to the Designated Community
- They follow documented policies and procedures, which ensure that the information is preserved against all reasonable contingencies, and which enable the information to be disseminated as authenticated copies of the original, or as traceable to the original (fulfilled by NLE and UBER only)
- They make the preserved information available to the Designated Community

As seen above, the demonstrators declared to follow the OAIS reference model and according to the RLG/OCLC report fulfil most of the trusted digital repositories obligations, i.e.:

- They accept responsibility for the long-term maintenance of digital resources on behalf of their depositors and for the benefit of current and future users
- They have an organizational system that supports not only long-term viability of the repository, but also the digital information for which it is responsible
- They demonstrate fiscal responsibility and sustainability (NLE and UBI only)
- They design its system(s) in accordance with commonly accepted conventions and standards to ensure the ongoing management, access, and security of materials deposited within it
- They can depend upon to carry out their long-term responsibilities to depositors and users openly and explicitly (NLE and UBER only)
- They have policies, practices, and performance that can be audited and measured.

None of them establishes types of methodology for system evaluation that meet community expectations of trustworthiness.

No repository has obtained a certification as a trusted digital repository yet, but most of them plan to do so. EDOC (UBER) holds a DINI certificate, the primary objective of which is "to improve interoperability and cooperation between German higher education institutions that run digital repositories; and to provide an instrument for the repository operators that could be used to raise the visibility, recognition, and importance of the digital repository within the

⁴ RLG Trusted Digital Repository, p. 55: Appendix C: Operational Responsibilities Checklist and reUSE Digital Repositories White Paper, p. 82.

university. The DINI certificate distinguishes the repository from common institutional web servers and assures potential users and authors of digital documents that a certain level of quality in repository operation is warranted. In addition, DINI sees its certificate as an instrument to support the Open Access concept. It can be viewed as a “soft certificate,” where the coaching idea prevails, and works on the basis of self-disclosure by the repositories.”⁵ NLE on the other hand plans to certify it in the future.

All demonstrators have written policies for the ingest, storage and long-term preservation processes within their repositories. Most of them have written policies for the access to the repositories. These policies were produced during the reUSE project.

At NLE and UBER there is a well defined preservation planning policy, which is periodically reviewed. They also have periodical work plans of preservation activities.

Most mandatory responsibilities of OAIS archives are fulfilled by all demonstrators' digital repositories. Neither of them establishes types of methodology for system evaluation that meet community expectations of trustworthiness.

UBER is the only digital repository that has been certified by DINI.

Concerning ALO (UBI), we recommend it to pay more attention to the issue of digital preservation by defining a more clear planning policy in order to follow the statements approved in the White Paper and implement a trusted digital repository.

1.3.2 PERCEPTION OF TOP MANAGEMENT

Two different questionnaires were used to survey the attitude of the top managers: one was prepared for the top managers and another for the librarians in charge of the repository. We thought that if the answers came directly from the top managers as well as from the operational level we would be able to get the real picture of the decision management's attitude and therefore of the future of digital repositories.

The main criteria for the interview of the top managers was their influence on the strategic planning of the organization, where the repository was located. Some directors with influence on the organizational strategy and policy making processes were contacted by phone, others were delivered a questionnaire. This way we interviewed directors of university libraries and a director of information and technology.

Their responses were very positive concerning reUSE digital repositories. All of them, with the exception of I3S3, consider a digital repository for master files of printed materials very important for their organization. The manager at I3S3 believes that it could enrich some services but that it is not so important for their organization.

NLE, UBG and UBI top managers completely agree that their organizations have an important role in their countries in collecting electronic documents.

NLE, UBI, UBG and I3S3 top managers completely agree that the digital repositories containing electronic documents, which are valuable for the public and stake-holder should be preserved for the future, and that keeping the master files of printed documents can help them to keep these documents for the future.

⁵ Dobratz, Susanne & Schoger, Astrid : "Digital Repository Certification: A Report from Germany". In: RLG DigiNews: Issue index: Oct 15, URL: 2005,http://www.rlg.org/en/page.php?Page_ID=20793&Printable=1&Article_ID=1779

NLE, I3S3 and UBG top managers completely agree that a digital version of printed materials can provide their more effective and wider use, and they are completely aware that the reUSE project is important for setting up such a digital repository and that it is an opportunity for acquiring international experience and knowledge in the field of electronic archiving. UBI top manager only partly agrees with these two statements.

The positive attitude of the top management was felt also by the librarians-operators of the repository. At NLE and UBI they were very satisfied with the top managers and they said that they were very supportive of the reUSE project.

Top managers of demonstrators' main institutions are very supportive of reUSE idea.

1.3.3 ORGANIZATIONAL CULTURE

The development of the trusted digital repository has introduced new processes in the organizational work-flows as well as new services and responsibilities. The increased number of digital publications has to be catalogued, saved and maintained differently from printed publications. We expected that it could influence the overall organization of demonstrators' libraries. However, it seems it was too early for this kind of analysis. UBI and I3S3 stated that the new digital services had not influenced staff dissatisfaction or increased their complaints while NLE claimed that they did not have information on this issue.

It is still too early to get a real picture of the changes in the organizational culture as a consequence of the introduction of new processes.

1.3.4 METADATA

In reUSE White paper the descriptive, structural and administrative metadata were defined according to their function and use in digital repository systems. In this report we shall concentrate on the different metadata actually used by the reUSE demonstrators.

The **descriptive metadata** depends on the nature of the digital materials kept in the repository. As shown in **Table 3** all the partners use the basic metadata set in Dublin Core 1.1. format which enables the interoperability with other systems. For these purposes NLE only uses the interoperability protocol Z39.50. Following the White Paper all demonstration partners use an XML encoding standard for mapping Dublin Core metadata.

Table 3: Metadata schema used by reUSE demonstrators

Metadata schema	NLE	UBI	UBER	I3S3
MPEG21			planned	
Creative Commons Metadata			X	
Dublin Core 1.1	X	X	X	X
Dublin Core qualified	X		X	
MAB			X	
METS	X	X	X	X
METSRights.xsd	X			
MIX or Z39.87	X			
Self developed			X	

All reUSE demonstrators use the basic descriptive metadata set in Dublin Core 1.1. As far as structural metadata is concerned they use XML encoding standards for mapping Dublin Core metadata. NLE, UBI, and I3S3 use METS while UBER uses a self-developed standard based on XML.

1.3.4.1 Preservation metadata

The preservation metadata includes all types of metadata and information that are needed for managing the digital objects in the future and preserving long-term access to their contents. The preservation metadata in this report follows the OAIS Information Model and the research for the reUSE Digital Repositories White paper. The basic concept in OAIS are the information packages. In the preservation context we speak of an archival information package, in which we distinguish four types of information:

- content information, which consists of the digital content and representation information,
- preservation information, which consists of reference information, context information, provenance information and fixity information,
- packaging information,
- descriptive information, which is located outside the archival information package.

The descriptive information provides the descriptive metadata that were already mentioned above. In continuation we focussed on the metadata related to representation information and preservation information.

Although most of the preservation metadata elements are included in the representation information data on HW and OS descriptions, most of this information is automatically extracted and is not collected by reUSE demonstrators. The rest of the metadata is listed in **Table 4**.

Table 4: Representation information and Preservation information according to OAIS Information Model

PRESERVATION METADATA (according to OAIS)	NLE (DIGAR)	UBI (ALO)	UBER (EDOC)
Representation information			
software descriptions	X	X	
descriptions of standards, data types or mapping rules		X	X
other:	X		data formats
Preservation information			
• Reference information			
○ System-generated internal identifier	X	X	X
○ persistent identifier	X		X
• Type of persistent identifiers			
○ URN based on NBN (national bibliography number)		X	X
○ URN only	X		
• How the persistent identifiers are assigned			
○ by institution itself	X	X	
○ centralized, by a national agency			X
• Context information (describes the relationships between content information objects)			
• info about subject-based collections	X		
• info about versions of the same content in alternative formats	X		X
• info about parent-child relations	X	X	X
• pointers to related content	X		
• Provenance information (documents the history of the content information)			
• creation/change history	X	X	X
• refreshment/migration history	X		
• Fixity information (validates the authenticity and integrity of the content information)			
• checksums	X		
• digital signatures/watermarks			X

As we see in Table 4 all the categories of information are covered by reUSE demonstrators. As far as reference information is concerned, reUSE partners decided to and do use URN persistent identifier (White Paper, p. 20). They are assigned by the repository institution or as in the case of UBER, there is a centralized system for its assignment. The context information is much better documented by NLE than the rest of the partners. As far as provenance information is concerned the creation/change history is documented by all partners. NLE has decided to use checksums as fixity information, while UBER uses digital signatures.

As far as formats go, the White Paper mentioned that all reUSE demonstrators were in the process of evaluating Adobe PDF/A as the best suitable format for the preservation of textual documents. An XML-based format was recommended as the second best format. The choice of appropriate formats still depends on the representation information, which varies from one document to another. All demonstrators took care in choosing the appropriate formats for preservation.

In their repositories all demonstrators collected preservation metadata according to the OAIS information model. Most of the metadata concerning preservation and preservation information for every digital document are kept in the repository.

1.3.5 BUDGET PLANNING OF THE REPOSITORY

The costs of preserving digital materials need to be considered in light of the relative benefits. There are many cost models aiming to define the cost of digital preservation per unit of digital object⁶. They depend on the technology development, the collection volume, repository architecture, operating system etc. However, as the digital repository's collection grows an annual increase in the costs of its maintenance should be considered in the budget planning.

Financial sustainability is very important for the process of long-term preservation. It is also one of the preconditions for the establishment of a trusted digital repository according to the RLG/OCLC report. Concerning budget planning issues, reUSE demonstrators have different approaches.

In European libraries 32% responding libraries which were surveyed by ARGE (ARGE survey p. 78) declared that they use the library budget as the sole source of finance for the digital repository. 42% of them use the library budget as one of several sources for financing the digital repository. 15% of the digital repositories are financed mainly or exclusively from a third party funds.

To the question of how demonstrators' repositories will be funded in the future, NLE, UBI and UBER top managers replied that a part of the institutions' operational budget will be assigned for this purpose, while I3S3 expects to get financial resources from outside of the organization. UBER expects to get additional funding introducing services' fees and to get financial resources from outside of the organization, i.e. they could be provided by EU and German Research Foundation (DFG).

Only NLE foresees a special budget for long-term preservation. This is not surprising as NLE top management supports the development of the digital repository and is aware of the importance of these activities. At UBG they seem to be supportive but are at the same time not prepared to assign additional financial resources for the maintenance of these repositories. NLE and UBI estimate that the financial resources of their repositories will be sufficient to guarantee the preservation of digital objects over longer periods of time. I3S3 on the other hand, considers these financial resources insufficient for this purpose. Additional investments for the maintaining of digital repositories will be needed after the conclusion of

⁶ A Workshop on *Cost Models for preserving digital assets* organized by DCC/DCP took part in the British Library on 26 July 2005. The report and some of the papers are available at URL: <http://www.dpconline.org/graphics/events/050726workshop.html>.

the reUSE project. NLE estimates approximately 5,0% of increase, UBI approximately 20% of increase and I3S3 approximately 10% of increase in their investments.

We asked the top managers whether they would be prepared to support the maintenance and the development of the repository after the conclusion of the reUSE project, and received different answers:

- The NLE Director of Technology Services would be prepared to support it financially, with additional staff and with additional information technology and software.
- The Director of UBI is prepared to support it with additional staff.
- The Director of UBG is not at all prepared to support it. We suppose this is due to their planned fund raising for this purpose.
- The Director of I3S3 is prepared to support it financially.

NLE has included the costs of the digital repository in the budget plan while UBI and I3S3 plan to do so.

After the conclusion of the reUSE project all demonstrators' main institutions will assign a part of the operational budget for the functioning and maintenance of the digital repository. In addition, NLE foresees a special budget for long-term preservation as they have included it in their budget plan, while UBI and UBER expect to get external sources of funding. The estimations of the increase in the investment sources differ from demonstrator to demonstrator. NLE calculates approx. 5%, UBI approx. 20%, and I3S3 10% of increase in their investments.

1.3.6 STAFF INVOLVED IN THE OPERATION AND ACCESS TO DIGITAL REPOSITORIES

Experienced professionals are very important in the creation and development of organizational knowledge. If they are trained to implement all processes in the digital preservation, they can ensure organizational viability and contribute to the trustworthiness of the organization.

The demonstrators NLE, UBI, I3S3 and UBER have given special attention to building teams of permanent staff from different fields that could keep a log in digital preservation. It is quite difficult to define the exact FTE (full-time equivalent) of the staff working with reUSE repositories, since they are involved in many other tasks at the same time.

In the NLE during the reUSE project time the digital preservation team included members from two departments: the Information Systems Department and the Collection Development Department. The Digital Preservation Council was working as an administrative group debating periodically (2-3 times a year) the preservation issues of the national library. NLE had to build a completely new system from scratch and for that reason engaged four part-time (0,5 FTE) specialists were engaged for the development of the digital repository. It makes 2 FTE per year. After that during the project time NLE had 1 part-time (0,5 FTE) professional librarian for cataloguing and file processing purposes (working under the Collection Development Dept.) and one part-time librarian for communication with content providers.

After the reUSE project is completed one part-time systems librarian (Information Systems Dept.), one part-time computer programmer (Information Systems Dept.) and one part-time systems engineer (Information Systems Dept.) will be needed for the maintenance of the system, which altogether takes approximately 10 hours a week. All the work will be done by

the Information Systems Department, and afterwards NLE plans to hire additional staff for the purposes of further communication with content providers, file processing and cataloguing. (part-time staff under the Collection Development Department).

At UBER the document and publication servers of the Humboldt University are under joint operation of the Computer and Media Services and the University Library. At the University Library the submission of digital documents for server distribution takes place. They employ 7 students with technical profile for 20 hrs per week and 6 additional permanent staff members, altogether 13 people in charge of the digital repository.

At UBI the system administrators are employed by the Department, while the infrastructure is provided by the University central service team. They have one professional librarian for metadata recording and consulting, and one system programmer for special tasks as well as project managers. The demonstrators did not have to increase their staff member, except UBI approx. 1 person/month.

At I3S3 they had to increase their staff members by approximately 4 persons/month in order to provide reUSE services.

The data regarding the number of staff members is very relative, since there are situations that cannot be foreseen in advance, i.e. increased number of content providers or users which requires additional support.

During the process of developing the digital repositories the demonstrators' level of knowledge increased (from intermediate to expert) concerning digital preservation in their organization. The methods used are shown in **Table 5**:

Table 5: Methods used by demonstrators' institutions to increase the level of staff expertise concerning digital preservation

Methods used to increase expertise	NLE	UBI	I3S3	UBER
regular technology watch	X			X
internal studies by institution members	X			X
training by commercial institutions				
external studies by research institutions	X			
training by software vendors		X		
Hiring consultants	X			
international workshops	X		X	X
hiring staff with expertise	X	X		X

The demonstrators NLE, UBI, I3S3 and UBER gave special attention to building teams of permanent staff from different fields that could keep a log in digital preservation. During the reUSE project the expertise regarding repository development and long-term preservation practice has considerably increased.

1.3.7 CONTENT PROVIDERS

All demonstrators were very successful in attracting different content providers. At the beginning the selection criteria were broadly defined, focusing on public digital master file producers from the public sector (non-profit organizations) and the academic field (high education institutions). Since the existing Legal Deposit Laws do not address the digital master files of printed documents all reUSE collections are built on a voluntary deposit of the content providers. In this regard, all demonstrators provided written submission agreements for content producers. They produced individual agreements for every object of a certain producer (NLE and UBER), collective agreements for all objects of a certain producer (NLE, UBI and I3S3) and collective agreements for all objects of a certain group of producers (UBER).

The cooperation between the digital repositories and the content providers/depositors is very strong, especially in the academic institutions, where a more rigid quality control regime with deposit guidelines including document templates and specifying file formats is enabled. The demonstrators in Austria and Estonia collected contents from very different types of organizations as for instance: universities, public enterprises, associations, government bodies, municipalities, political parties, church organizations, cultural institutions etc. All partners defined their designated communities and their needs.

The number of producers varies from demonstrator to demonstrator and their type depends on the demonstrator's basic activities. At the universities there are more content providers from the academic community. The total number of content providers is shown in **Table 6**.

Table 6: Number of content providers by reUSE demonstrators

reUSE demonstrator	Number of content providers (as per November 2005)
UBI	27
UBG	178
I3S3	45
UBER	19 (groups of content providers)
NLE	35

At UBER there are several types of content providers. One of them is represented by individual authors, with whom they signed separate author-agreements. The other types of content providers are: workshops, conferences and journal publishers, science institutions, and faculties.

At present, the number of documents provided by the content providers also differs from demonstrator to demonstrator. **Figure 2** shows great difference between UBG, UBER and NLE reUSE collections. However, big progress has been made in the last few months, since some of the demonstrators, like NLE, built their repository from scratch and reached an impressive number of 414 documents.

In order to analyse the benefit aspects of the project we carried out a survey on the content providers, which is presented in a separate chapter.

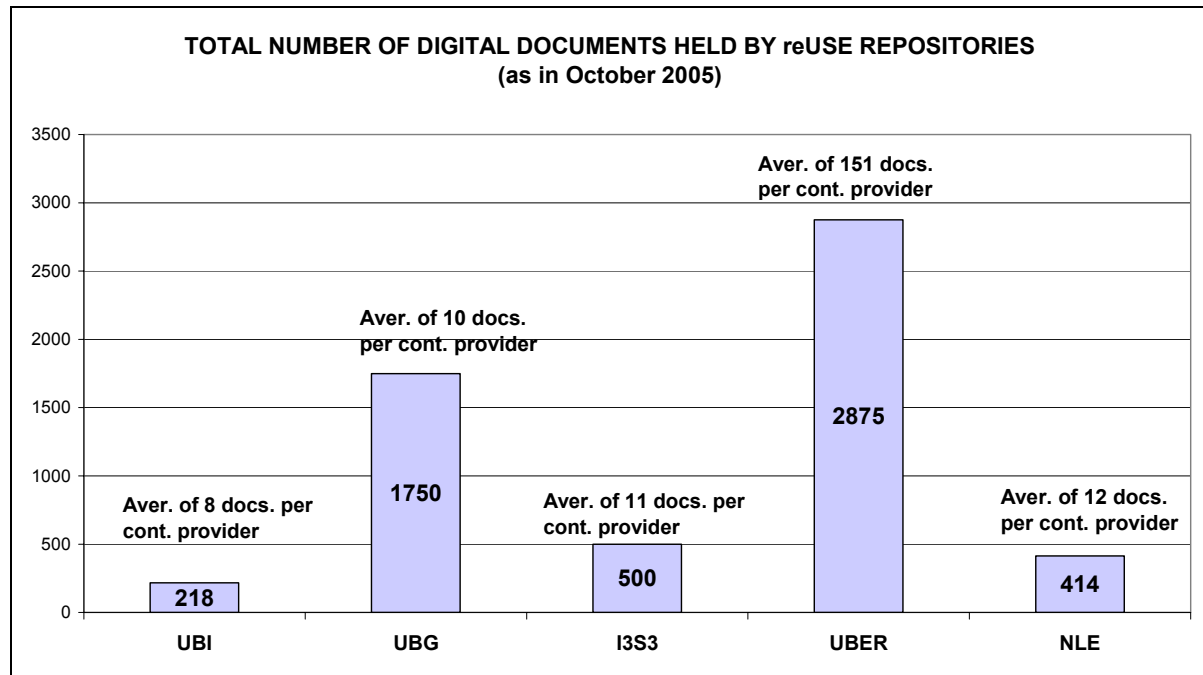


Figure 2: Total number of documents held by reUSE repositories

NLE and UBER provide training for data producers about the ingest process of their repository, while NLE provides training for users about the access process to their repository.

All demonstrators were very successful in attracting different content providers and collecting their publications.

Their collections are build upon very broad, but poorly defined selection criteria, since their publications come from very different public and private organizations.

The process of document submission is based on different types of agreement between content producers and reUSE repositories. The cooperation between them is very strong.

NLE and UBER provide training for data producers about the ingest process, while NLE provides training for users about the access process to their repository.

1.4 REVIEW OF ALL FINDINGS IN THE EVALUATION OF ORGANIZATIONAL ASPECTS

Table 7: Review of all findings in the evaluation of organizational aspects

National policy	<ul style="list-style-type: none"> • In the demonstrators' countries there are different national policies or national initiatives regarding long-term preservation of digital documents. • There is a high level of awareness at the government level which is very important for the promotion of national programmes and financing of digital repositories.
Organizational mission statement	<ul style="list-style-type: none"> • All demonstrators have written mission statements of the organization and the repository.
Organizational structure	<ul style="list-style-type: none"> • In all organizations the digital repository is developed by the infrastructure and services departments, which are far from the decision making level when it comes to universities. They closely cooperate with the library staff.
Digital preservation policy	<ul style="list-style-type: none"> • All demonstrators have written policies for the ingest, storage and long-term preservation processes within their repositories. Most of them have written policies for access to the repositories. These policies were produced during the reUSE project. • At NLE and UBER there is a well defined preservation planning policy which is periodically reviewed. They also have periodical work plans of preservation activities. • Most mandatory responsibilities of OAIS archives are fulfilled by all demonstrators' digital repositories. Neither of them establishes types of methodology for system evaluation that meet community expectations of trustworthiness. • UBER is the only digital repository that has been certified by DINI. As far as ALO (UBI) is concerned in order to follow the statements approved in the White Paper and implement a trusted digital repository, we recommend they pay more attention to the issue of digital preservation by defining a more clear planning policy.
Perception of top management	<ul style="list-style-type: none"> • Top managers of demonstrators' main institutions are very supportive of the reUSE idea.
Organizational culture	<ul style="list-style-type: none"> • It is still too early to get a real picture of the changes in the organizational culture as a consequence of the introduction of new processes.
Metadata	<ul style="list-style-type: none"> • All reUSE demonstrators use the basic descriptive metadata set in Dublin Core 1.1. As far as structural metadata are concerned they use XML encoding standards for mapping Dublin Core metadata. NLE, UBI, and I3S3 use METS, while UBER uses a self developed standard based on XML. • All demonstrators' repositories collect preservation metadata according to the OAIS information model. Most of the metadata concerning preservation and preservation information for every digital document are kept in the repository.

Financial sustainability	<ul style="list-style-type: none"> • After the conclusion of the reUSE project all demonstrators' institutions will assign a part of the operational budget to the functioning and maintenance of the digital repository. In addition, NLE foresees a special budget for long-term preservation as they have included it in their budget plan, while UBI and UBER expect to get external sources of funding. • The estimations of the increase in the investment sources differ from demonstrator to demonstrator. At NLE they calculate approx. 5% at UBI approximately 20% and at I3S3 10% of increase.
Organizational viability	<ul style="list-style-type: none"> • The demonstrators NLE, UBI and UBER gave special attention to building teams of permanent staff from different fields that could keep a log in digital preservation. • During the reUSE project the expertise regarding repository development and long-term preservation practice has considerably increased.
Content providers' relations	<ul style="list-style-type: none"> • All demonstrators were very successful in attracting different content providers and collecting their publications. • Their collections are build upon very broad but poorly defined selection criteria, since their publications are coming from very different public and private organizations. • The process of document submission is based on different types of agreement between content producers and reUSE repositories. The cooperation between them is very strong. • NLE and UBER provide training for data producers about the ingest process, while NLE provides training for users about the access process to their repository.

2 TECHNICAL ASPECTS EVALUATION

By Darko Majcenović and Matjaž Depolli

2.1 BACKGROUND

Digital repositories may have different architecture but there are certain standards that they have to follow in order to provide quality service to the users and secure a long-term preservation of their collections. In the evaluation process we analysed partners' repositories and discovered whether they follow adopted standards and fulfilled user expectations according to pre-defined evaluation criteria. Special emphasis was on long-term preservation, which is very important for the deposit institutions of digital collections.

Three demonstrators building reUSE repositories were evaluated:

- The National Library of Estonia (DIGAR) - <http://digar.nlib.ee/>
- Austrian Literature Online (ALO) - <http://www.literature.at>
- Humboldt University of Berlin (EDOC) - <http://edoc.hu-berlin.de>



The National Library of Estonia (NLE) is setting up a trusted digital repository (DIGAR) to collect, preserve and allow access to publications from public sector publishers. Digital material from NGO⁷s and private publishers is accepted if they follow the archive's submission guidelines and open access policy. The creation and ongoing maintenance of the repository are the collaborative effort of various NLE departments. From a technical perspective, the National Library of Estonia DIGAR repository system is being built upon the Fedora Digital Repository Management System⁸. The repository is accessible via an online interface. Also,

relevant records will be linked to the OPAC and other bibliographical databases. In implementing this repository the National Library of Estonia uses its extensive experience from a previous pilot-project called ARES⁹. From 1999 to 2001 digital publications of four scientific publishers were collected providing valuable insight into the publishers' attitude and workflow in electronic publishing¹⁰.

The Austrian Literature Online (ALO) digital repository was set up in March 2002. Originally designed for digital documents it also contains electronic documents in different formats (XML, PDF, RTF,..). There are currently more than 6.500 books, journals, and manuscripts online. ALO is maintained by a working group of Austrian libraries.

⁷ NGO - Non Governmental Organization.

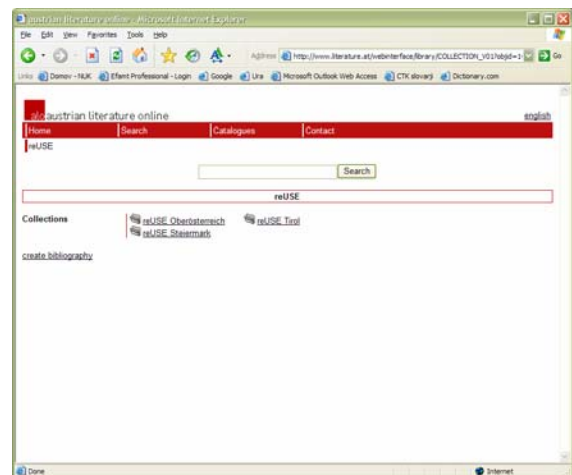
⁸ The Mellon FEDORA (Flexible Extensible Digital Object and Repository Architecture) Project, University of Virginia and Cornell University, <http://www.fedora.info/>.

⁹ ARES - Artiklite Elektrooniline Süsteem / Electronic System of Articles. ARES links in online catalogue <http://helios.nlib.ee>

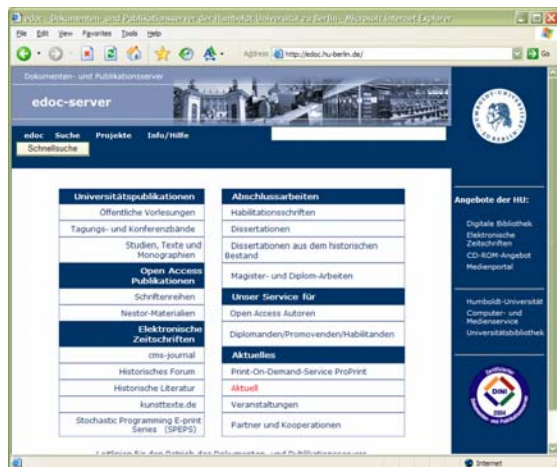
¹⁰ White Paper on Digital Repositories (reUSE Digital Master Files of Printed Material!), http://www2.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf.

The university libraries from Innsbruck (UBI) and Graz (UBG) and the University of Linz department i3s3, which is responsible for technical development. The ALO system follows the standards, the workflow within ALO is designed in a way that while documents are made available on a special website of the digital repository as well as via the local and national electronic library catalogues (ALEPH) in Austria.

A direct link is included in the MAB2¹¹ record, which allows the user to access an electronic document of the digital repository.



ALO is designed as an open-source package available for free. The University Library of Graz is a partner of the ALO consortium and one of its main users¹².



The Humboldt University of Berlin (UBER) repository EDOC¹³ is a thematic institutional repository at Humboldt University of Berlin (UBER), which incorporates scientific publications of the Humboldt University and of the cooperating partners. Repository support is carried out by the joint Electronic Publishing Group of the Computer and Media Services and the University Library, with permanent and temporary project staff in order to ensure sustainability of the service as well as continuous further development. There are already some 2.000 items in the repository, catalogued and prepared for long-term

preservation. Publications, stored within the EDOC server, use XML as preservation document format. Digital signatures and time stamps ensure the integrity of the materials over time, while open access is granted via an on-line gateway. EDOC has already set up a professional print-on-demand service for distributing bound copies of the electronic documents¹⁴.

The Austrian National Library (ONB) was interested in participating in the project by establishing a digital repository as part of its legal mandate and responsibility as a national library. But the new ONB repository could not be completed within the timeframe of technical evaluation so it was not possible to evaluate it. Targeted data suppliers are public sector institutions and commercial publishers. Digital material from private sector is accepted to a limited extent. The ONB expects thousands of objects during the next years. The preservation of digital master files is part of a range of activities at the ONB concerning the long term preservation of off-line and on-line media. Technical infrastructure and support is part of the institutional infrastructure supervised by the department for information technology

¹¹ MAB2 – Maschinelles Austauschformat für Bibliotheken [Automated Library Exchange Format], http://www.ddb.de/professionell/mab_e.htm.

¹² White Paper on Digital Repositories (reUSE Digital Master Files of Printed Material!), http://www2.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf.

¹³ Document and Publication Server of Humboldt University Berlin, <http://edoc.hu-berlin.de>.

¹⁴ White Paper on Digital Repositories (reUSE Digital Master Files of Printed Material!), http://www2.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf.

services. ONB's digital repository (commercial software) hosts diverse digital archives and connects to other activities like digitalization projects of the ONB as well as to the library system. A completely new and enhanced version of the software will be implemented during the summer of 2005 and will have gone on-line by the end of 2005¹⁵.

In the evaluation process we analysed partners' repositories and found out whether they followed adopted standards and fulfilled user expectations according to pre-defined evaluation criteria. Special emphasis was put on long-term preservation, which is very important for the deposit institutions of digital collections. We evaluated three repositories:

- The National Library of Estonia (NLE), who is setting up a trusted digital repository (DIGAR) to collect, preserve and allow access to publications from public sector publishers.
- The Austrian Literature Online (ALO) digital repository that was been set up in March 2002. ALO is maintained by a working group of Austrian libraries: The university libraries from Innsbruck (UBI) and Graz (UBG) and the University of Linz department i3s3, which is responsible for the technical development.
- The Humboldt University of Berlin (UBER) repository EDOC, which is a thematic institutional repository at Humboldt University of Berlin (UBER), and it incorporates scientific publications of the Humboldt University and cooperating partners.

2.2 EVALUATION METHODOLOGY

In the project evaluation framework we defined the attributes for the evaluation of the technical part of the repositories. They relate to different aspects regarding quality, technical specifications, repository and system administration, content management and services. **Table 8** lists the evaluation attributes and related evaluation topics. For more details see the Evaluation framework table.

Table 8: Technical aspects evaluation attributes and related technical aspects

Attributes	Evaluation topics	Detailed topics
Quality issues	Functionality	<ul style="list-style-type: none"> • Logs and Statistics • Security
	Reliability of the repository	<ul style="list-style-type: none"> • Data protection • Data recovery
	Efficiency	Guidelines for efficiency measurements
	Maintainability	Application maintenance
	Flexibility	<ul style="list-style-type: none"> • Possibility of format extension • System upgrades
Technical specifications	Standards	Used standards: <ul style="list-style-type: none"> • For repository • For client support
	Software	Software license used
Repository & System Administration	Access to collections by user types	Types of access rights given to users of the repository
	User administration	<ul style="list-style-type: none"> • Limit access • Registration, authentication & password administration
	Content submission administration	<ul style="list-style-type: none"> • Multiple collections within the same instance of the system • Submission stages • Submission support
Content management	Document / object formats, import-export	File submission management
	Metadata	

¹⁵ White Paper on Digital Repositories (reUSE Digital Master Files of Printed Material!), http://www2.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf.

	Updating and Indexing	Support for real time updating / indexing of the repository
Services	Search capability	<ul style="list-style-type: none"> • Full-text search of selected publications: • Search all descriptive metadata • Search selected metadata fields • Search functions and additional search functions
	Presentation of search results	Ways of browsing search results
	Advanced services	User interface
	Value-added functions	<ul style="list-style-type: none"> • W3C Content Accessibility • Content oriented value-added functions • Additional options

The listed attributes focused on the technical part of the repository. However some interrelated to the organizational and user points of view. For gathering information on technical aspects of the repositories a questionnaire was prepared, which had been distributed to the developers of the three repositories. In addition, the developers of the repositories were asked to provide usage logs in order to see the practical use of the repositories.

The evaluation methodology is based on different approaches, including the experience of the European project SciX, which produced an extensible assessment and evaluation report. The coordinator of the project was the Faculty of Civil and Geodetic Engineering of the University of Ljubljana, which is one of the evaluators in this project. We consulted several standards related to usability and software quality as well, i.e. ISO/IEC 9126 (internal SW quality), ISO 9241-11 (usability), ISO 13407 I (human centered design processes for interactive systems), ISO 12207 (Software lifecycle processes), and ISO 14721 (OAIS). We extracted the relevant attributes for the purpose of the evaluation for our multi-attribute methodology.

2.3 EVALUATION OF THE DIGITAL REPOSITORIES ACCORDING TO PRE-DEFINED ATTRIBUTES

2.3.1 QUALITY ISSUES

2.3.1.1 FUNCTIONALITY

Functionality as described by Bevan¹⁶ is the capability of the repository to provide functions which meet stated and implied needs when it is used under specified conditions. It could be analyzed via the repository design, logs and statistics, as well as security aspects.

2.3.1.1.1 Repository design

Design of the repository is a very important aspect which, should be closely examined because of its important impact on the future development of the repository and accessibility of the data stored within it. Average workload explains the conditions the repository normally operates in and its capabilities under that regime. We received very diverse answers to this question suggesting that average workload is very tightly connected to the software used and the hardware it is run on. We also received quite diverse answers as to how many objects each repository can hold – values vary wildly from 100.000 up to unlimited. These are very tightly connected to the software used in repository design. **Table 9** shows the basic information on reUSE demonstrators' repositories.

¹⁶ The definitions are based on the standard for software product quality ISO/IEC 9126-1 described by Nigel Bevan: Quality in Use: Meeting user needs for quality. In: Journal of System and Software, 1999.

All partners comply with and follow the OAIS recommendations as much as possible in their given situation as was required in the project description – to fulfil the main criteria set up in the RLG/OCLC Report on Trusted Digital Repositories¹⁷.

Table 9: Basic information on reUSE demonstrators' repositories

	DIGAR	ALO	EDOC
Average workload	The total number of visitors in October 2005 was 876. They viewed 2042 different objects in the repository.	The total number of visitors in August 2005 was 32863. They accessed 8751 different sites in the repository.	The total number of visitors in October 2005 was 146014. They accessed 43615 different sites in the repository.
Number of objects the repository can hold	1 million	More than 100.000	Software limitations are unknown.
OAIS compliance	Yes	Yes	Yes, for most part.
URL	http://digar.nlib.ee/	http://www.literature.at	http://edoc.hu-berlin.de

2.3.1.1.2 Logs and statistics

In order to better understand the repository usage demonstrators were asked to provide system generated use statistics and reports. This data offers an overview of the repository's peak activity and when it occurs, number of queries as well as the number of documents. Usage and statistics provide the real-world view of repository use thus revealing areas where the repository could be further improved. Because of the differences between systems in use each partner provided a different set of statistic data. The survey focuses on logs and statistics of evaluated repositories, which can serve as comparison and illustration of trends specific to each partner's repository.

The EDOC repository's usage statistics for the time period from 1st to 31st October 2005

An expected pattern has occurred while analysing data from the EDOC repository (**Figure 3**). The one-month survey period has shown that repository use appears to diminish approximately every weekend, while remaining high throughout the week. This could be explained by the fact that EDOC is a university and that its students don't use the available resources much during weekends and holidays. This also gives an approximate picture of characteristics of the main repository users.

¹⁷ More about the OAIS reference model and the RLG/OCLC Report in White Paper on Digital Repositories (reUSE Digital Master Files of Printed Material!), http://www2.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf.

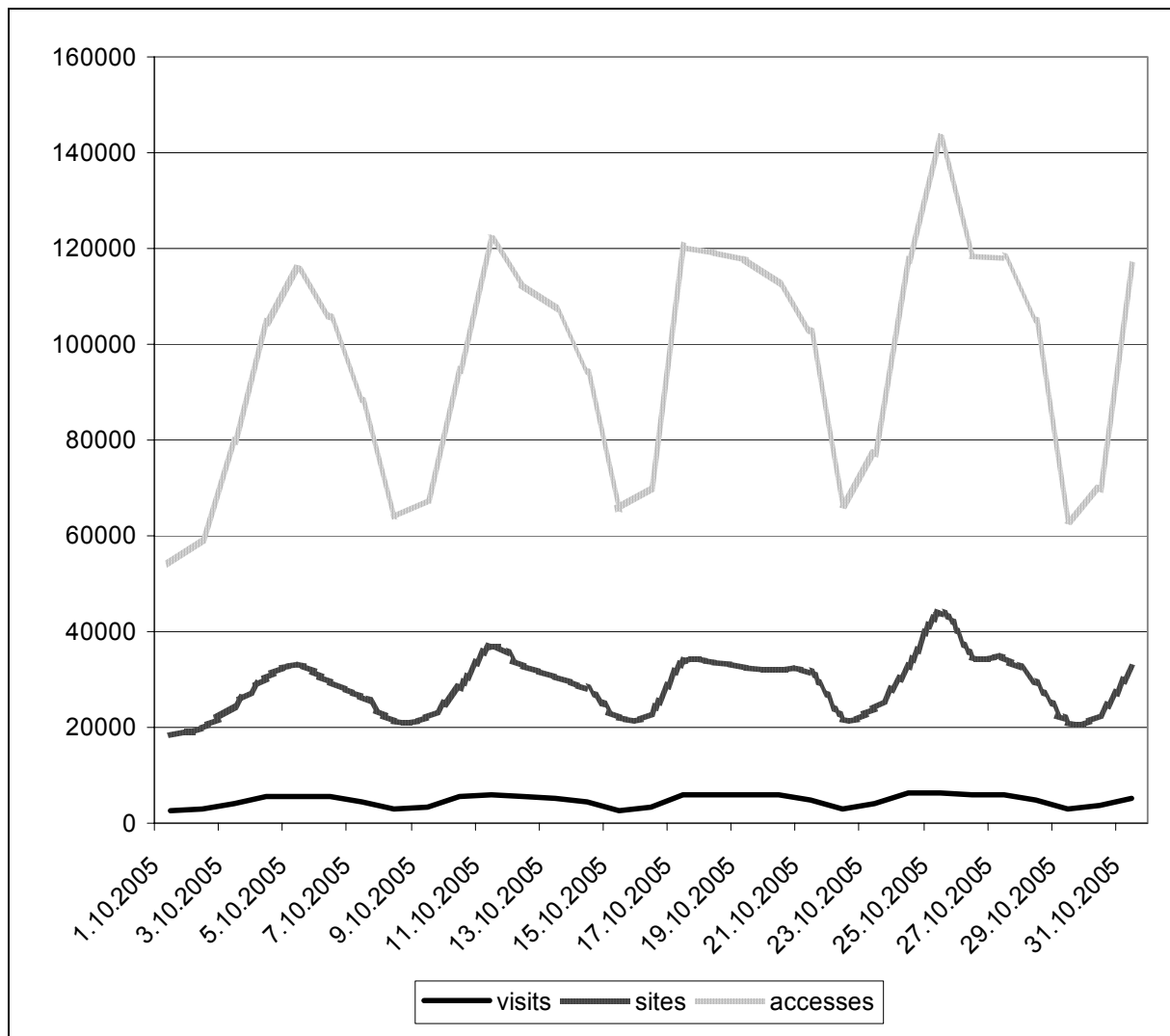


Figure 3: Comparison of visits, sites and accesses to EDOC repository during one month survey phase

66.4% of EDOC users accessed their repository via direct access (bookmarks) and only 16.9% via the web-search engines. This could prove that the repository content is valuable to its users and that they access it on purpose, storing its link in their web browser bookmarks. Less than 17% of users came to the repository via search results provided by web-search engines. The highest number of visits was on 24th October (6279) and the lowest on 1st October (2604). There was an average of 4710 visitors per day using the repository.

The DIGAR repository's usage statistics for the time period from 26th September to 25th October 2005

The same pattern as with EDOC arose in DIGAR as well (**Figure 4**). The number of repository users also diminished during weekends in DIGAR. A big number of queries at the beginning of the evaluation period was caused by the user evaluation part of survey which happened to be at the end of September. The highest number of visits was on 14th October (87) and the lowest on 8th and 22nd October (5). There was an average of 29 visitors per day using the repository.

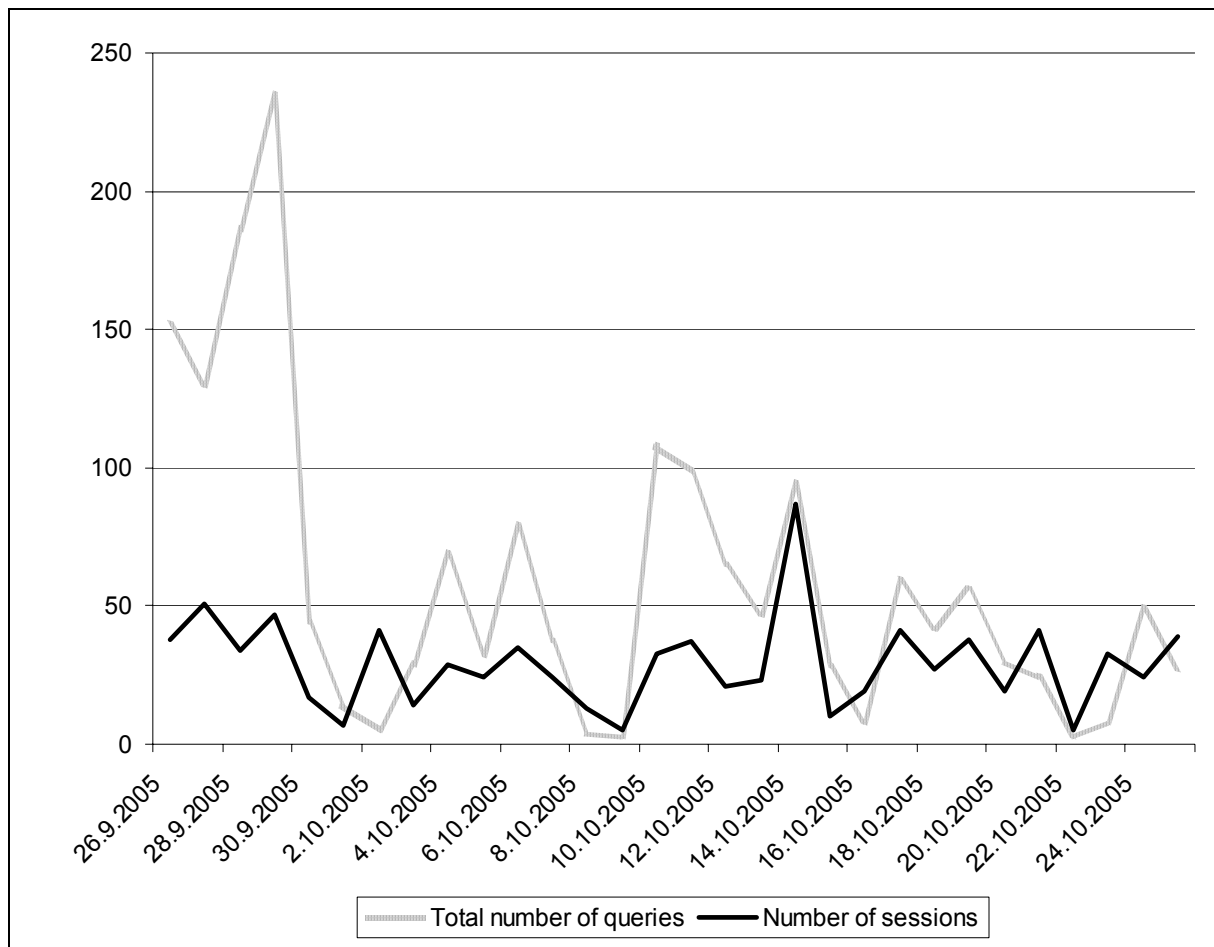


Figure 4: Comparison of the total number of queries and the number of sessions in DIGAR's one month survey phase

DIGAR offers both a simple and a complex method of searching. From the statistics provided it is evident that only 15% of users use the complex method, while 85% use the simple one. The complex method has 4 search fields: title, author, publisher and date. EDOC provides a similar method of advance search offering search fields for title, author, keyword and summary. However they could not provide statistics for the comparison.

The ALO repository's usage statistics for the time period from 1st to 31st October 2005

The number of repository users also diminished during weekends just as in DIGAR and EDOC, but remained high during the weekdays (**Figure 5**). ALOs' statistics were in raw format and while analyzing the data it became obvious that as much as 50% visits to the repository were by search-engines – mainly by Google and MSN Search. The highest number of visits was on 25th October (846) and the lowest on 30th October (14). There was an average of 267 visitors per day using the repository.

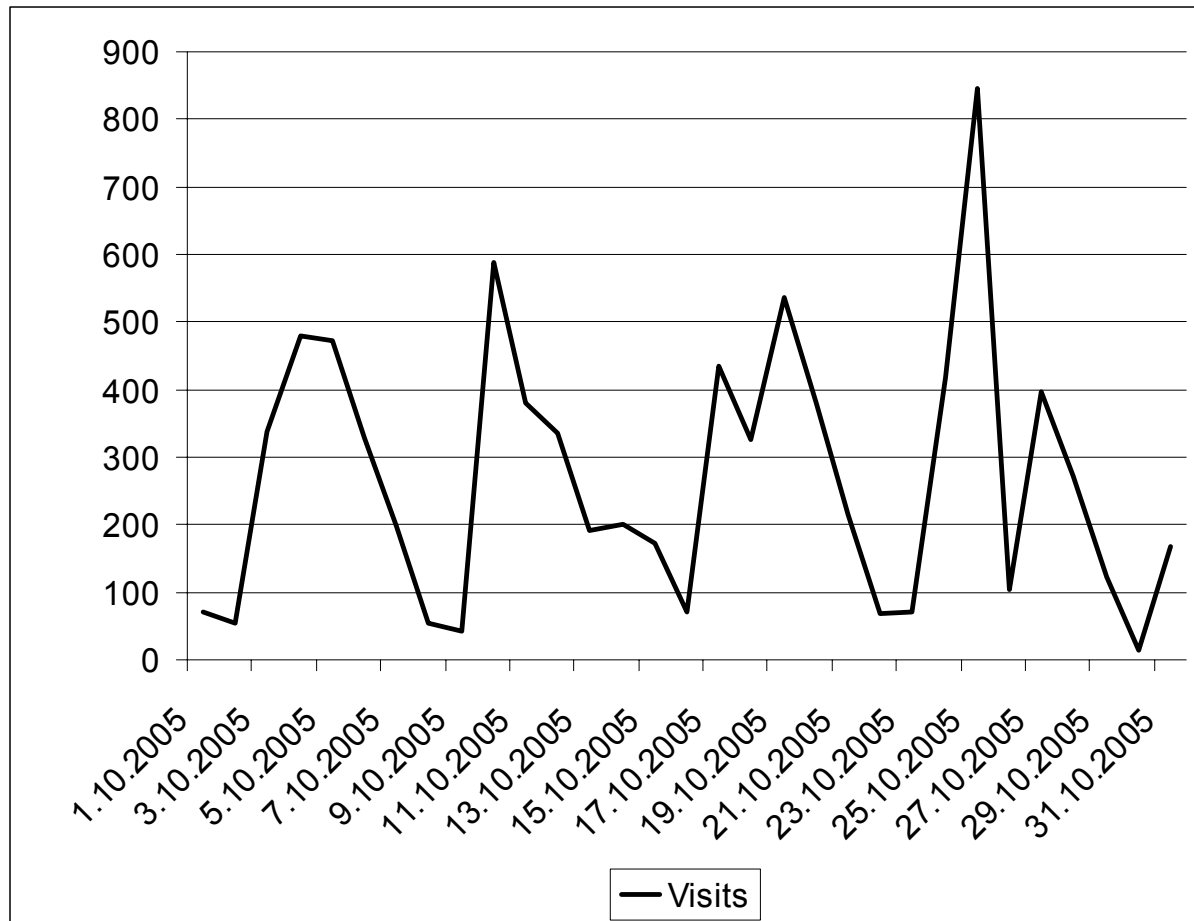


Figure 5: Total number of visits in ALO's one-month survey phase

2.3.1.1.3 Security

We set out to discover the security of the trusted repository systems in this part of the survey because an integral part of the service provided is also the capability of the repository to provide functions as stated under specified conditions. We were especially interested in the server security, database connections and restricting database access. Server Security involves restricting and limiting the actual access to the database server itself. Database Connections involve restricting access from remote locations. And Restricting Database Access involves setting up trusted IP addresses, port access security, etc. These restrictions are necessary not only for the sake of common practice but also because these are to be trusted repositories in charge of long term preservation of materials stored within them.

DIGAR limits network access to the database server and the database itself, to the local host¹⁸ users identified by passwords; while ALO does so via limitation to several IP addresses (which are determined by the operating system and not the database). EDOC uses IP addresses, log-in information and passwords to limit access to the database server. Thus, we can determine that all three demonstrators are basically using the same method.

Content providers' identification is determined by username and password in the case of DIGAR; EDOC on the other hand has no limitations, while ALO data submitters don't have direct access to the repository. EDOC limits access only to the management of the repository

¹⁸ Local host identification is actually a specialization of IP address identification.

but not to the ingest process – they provide templates that content providers must use. There are few access limitations for the end-users in all of the above mentioned repositories. DIGAR limits the end-users on object/file level so that some can only be viewed in NLE reading rooms.

2.3.1.2 RELIABILITY

Repository reliability is the capability of the repository to maintain its level of performance under specified conditions and it is therefore an aspect of its quality¹⁹. We were interested in data protection and data recovery in this section of the survey. Specific topics were fault tolerance, data recovery options, protection measures, number and frequency of backups, backup systems as well as media used and backup location.

ALO and EDOC are using the IBM Tivoli System – a systems management platform from IBM (previously Tivoli Systems, acquired by IBM in 1995 and moved into IBM's Software Group division). It includes VNC-like²⁰ capabilities and allows for remote system administration, configuration, and software installation. DIGAR is using a solution of RAID 5²¹ disk array on a separate location and a secondary backup on tapes. On all sites the primary backup storage device is hard disk drives and magnetic tapes for secondary backup storage.

EDOC uses three backup locations to preserve and protect their data while ALO and DIGAR use two. DIGAR also does periodic data duplication to another in-house server. The backup frequency of the repository in use is daily with all demonstrations.

All examined repositories have written policies for disaster readiness, response and recovery. In all cases recovery options after a critical system failure include a rebuild of the entire system from a backup server or location.

Processes that address data integrity (avoiding data-loss, detecting changes, restoring lost or corrupt data) vary among demonstrators. DIGAR relies on RAID 5 taking care of media errors and periodically verifies content file checksums. ALO uses several standard operating system routines on backup system and storage server, while EDOC uses digital signatures to verify file integrity.

¹⁹ The definition is based on the standard for software product quality ISO/IEC 9126-1 described by Nigel Bevan: Quality in Use: Meeting user needs for quality. In: Journal of System and Software, 1999

²⁰ Virtual Network Computing (VNC) is a desktop sharing system which uses the rfb (Remote Frame Buffer) protocol to remotely control another computer. It transmits the keyboard presses and mouse clicks from one computer to another relaying the screen updates back in the other direction, over a network. (<http://en.wikipedia.org/wiki/VNC>)

²¹ In computing, a redundant array of independent disks, often incorrectly known as redundant array of inexpensive disks (more commonly known as a RAID) is a system of using multiple hard drives for sharing or replicating data among the drives. Depending on the version chosen, the benefit of RAID is one or more of increased data integrity, fault-tolerance, throughput or capacity compared to single drives. A RAID 5 uses block-level striping with parity data distributed across all member disks. RAID 5 is one of the most popular RAID levels, and is frequently used in both hardware and software implementations. Virtually all storage arrays offer RAID 5. As with RAID 0, RAID 5 can be created with disks of differing sizes, but the storage space added to the array by each disk is limited to the size of the smallest disk. (http://en.wikipedia.org/wiki/Redundant_array_of_independent_disks)

2.3.1.3 EFFICIENCY

Efficiency is the repository's capability to provide the required performance, in relation to the amount of resources used, under stated conditions²². We were interested in system response time i.e. how long does it take to process a request and how long does it take it to search the database and the number of concurrent searches. Hardware and system resources involved are of big importance from the efficiency viewpoint so we collected those as well – these results are presented in chapter "2. Technical specifications / 2.2. Software."

Efficiency is a difficult parameter to determine. The average time lag between sending a request and receiving the data (client-server communication) depends on many parameters. Average search time varies between the repositories from 30 milliseconds in case of DIGAR's FEDORA repository to a few seconds in case of ALO's and EDOC's repositories.

ALO describes the efficiency of their repository as dependant on various parameters, such as the average time it takes to process a typical request for displaying a digital object - theirs is between 1-4 seconds. EDOC's repository response time is also dependant on the quantity of the data transferred and is under 10 seconds even for the transfers of big files. DIGAR describes their efficiency by specifying that FEDORA system loaded with 30 million objects is capable of keeping average response time under one second, when handling 20 concurrent users sending a request every 300 milliseconds. But that is only a prototype test and DIGAR is currently testing their repository in real world circumstances. Their repository is also designed for 300 simultaneous users in the depositor interface and 5 in management interface.

2.3.1.4 MAINTAINABILITY

Maintainability is the repository's capability to be modified²³. Modifications may include corrections, software improvements or its adaptation to environment changes, requirements and functional specifications. For this to succeed, documentation and manuals are required. Staff working with the system must also be sufficiently trained as well as have enough time to work on the repository. We were interested in the number of hours spent on repository maintenance and staff required to accomplish it.

Demonstrators provided various kinds of documentation for their developers and all types of users. DIGAR provided manuals for content managers and on-line tutorials for depositors and end-users. ALO provided a handbook for users in German language; developers are supported through SourceForge²⁴ project. EDOC offered documentation for developers, administrators, authors, and end-users.

Demonstrators provide various profiles to manage their repositories. In the White Paper on Digital Repositories the following roles are listed:

- Digital Preservation Strategy planning;
- System Administration;
- Cataloguing; and
- Library Management.

²² The definition is based on the standard for software product quality ISO/IEC 9126-1 described by Nigel Bevan: Quality in Use: Meeting user needs for quality. In: Journal of System and Software, 1999

²³ The definition is based on the standard for software product quality ISO/IEC 9126-1 described by Nigel Bevan: Quality in Use: Meeting user needs for quality. In: Journal of System and Software, 1999.

²⁴ SourceForge.net is a centralized location for software developers to control and manage open source software development, and acts as a source code repository.

The System Administration role is of special interest to the technical part of the demonstrators' evaluation. System Administration represents staff responsible for information technology of the repository system that can be integrated smoothly in the existing technology infrastructure, is easy to configure and adapt, and is well documented or comes with the necessary external support. Information technology experts contribute to ongoing technology watch for preservation purposes, while their evaluation of software and tools is important to strategic digital preservation staff for system development and refinement²⁵.

The following table (**Table 10**) describes the profiles and workload attached to them as well as the amount of training required for each profile. It is assumed that where there is no training described such as in the case of DIGAR and EDOC that staff is already trained in the area of the required expertise. The White Paper on Digital Repositories²⁶ explains that *"Staff training is a core requirement in digital preservation related tasks, for the lack of trained personnel on the market, and to keep up-to-date in the quickly changing digital environment. reUSE partners follow a variety of paths for staff development including attendance at international workshops and support in internal and private training."* This statement could explain the lack of training descriptions in the questionnaire as the training is non-formal and very specific to the field of expertise of personnel working on the project.

Table 10: Staff involved in the operation and maintenance of reUSE demonstrators' repositories

	Profile	Workload	Training needed
DIGAR	System administrator	4 hours / week	
	Technical support	0,25 person / month	
	Metadata input	0,5 person / month	
	Database administrator	6 hours / week	
ALO	Technical administrator	1 person / month	Initial training: 1-2 months; afterwards some hours / month
EDOC	Technical support	0.24 person / month	Weekly informal training
	Metadata input	0.5 person / month	

We have to take in to account here that two of the demonstrators have had previous experience and already existing repositories (ALO, EDOC) while the third (DIGAR) started building a new repository from scratch.

According to the White Paper on Digital Repositories, DIGAR *"is the only one with a stepwise organizational implementation plan, in which it accounts for a transient project phase of demonstrator development and testing. Thereby, the project is split between two teams. One team focuses on the design and implementation of the repository system and consists of four full-time equivalent (FTE) positions with an equal share of technical and library staff. The second team consisting of 2 FTE's is responsible for contacting potential data providers. Both teams are managed by a single project manager and supervised by a council of half a dozen personnel from the senior management of the National Library. This senior management support will be perpetuated once the repository moves from its initial phase into an ongoing*

²⁵ White Paper on Digital Repositories (reUSE Digital Master Files of Printed Material!), http://www2.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf.

²⁶ White Paper on Digital Repositories (reUSE Digital Master Files of Printed Material!), http://www2.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf.

service. Staff will then be lowered and split between the collection department and the department for information systems.²⁷ Since the repository is already functioning it is to be expected that the staff involved in its work will be cut down with the transition into the operating service after the end of the project. Data acquired with the technical questionnaire does not reflect this since it was collected earlier in the project timeframe.

The ALO repository is "split between technology on one side – whereby the technological infrastructure is provided by the University of Innsbruck –, and strategic and operational issues on the other. The latter team is largely an add-on to the library organization. ALO consists of library entities for the distinct university faculties, and the main university library which is the umbrella for these decentralized entities. ALO is part of the main library, has about half a dozen FTE's, and combines digitization and digital preservation."²⁸ This would explain the questionnaire results of only 1 person/month since ALO's infrastructure is already in existence and is operational.

The White Paper on Digital Repositories explains that "EDOC assumes an entirely different organizational structure. It installed a virtual 'edoc' team that is composed of staff from various departments. Additionally, the technological infrastructure is provided by the university and respective staff is not part of the core 'edoc' team. The core team includes staff from the software and multimedia departments of the Humboldt University Computer and Media Services, and service and operation departments of the University Library. Combined they contribute a workforce of about 6 FTE's. While the team members are assigned to these various departments, the 'edoc' team is an acknowledged entity within the organization and the team leader directly reports to the heads of both, the Computer and Media Services and the University Library."²⁹ This explains the questionnaire results since part of the team members belong to other departments.

2.3.1.5 FLEXIBILITY

Flexibility is the repository's capability to adapt to new conditions and requirements³⁰. Such adaptations include the ability to expand the file formats supported by the repository with new versions of file formats. The system itself should be flexible enough to allow inclusion of new workflows, new storage systems, new database back-ends and addition of new application servers.

All repositories in question can be upgraded with support for new file types. DIGAR and EDOC can upgrade their repository systems with new workflows, database systems and application server; while ALO can upgrade its repository system with a new storage system.

²⁷ White Paper on Digital Repositories (reUSE Digital Master Files of Printed Material!), http://www2.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf. p. 45

²⁸ White Paper on Digital Repositories (reUSE Digital Master Files of Printed Material!), http://www2.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf. p. 45

²⁹ White Paper on Digital Repositories (reUSE Digital Master Files of Printed Material!), http://www2.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf. p. 45

³⁰ The definitions is based on the standard for software product quality ISO/IEC 9126-1 described by Nigel Bevan: Quality in Use: Meeting user needs for quality. In: Journal of System and Software, 1999

2.3.2 TECHNICAL SPECIFICATIONS

Technical specifications are an integral part of the repository infrastructure and its ability to function according to the tasks presented to it and to fulfil the needs of its users. The quality of the service depends quite heavily on the standards used in the repository design and on the software used to accomplish it.

2.3.3 STANDARDS

Standards and commonly accepted conventions are a very important factor in assuring the long-term preservability of the data stored within the trusted repository. It has to be remembered at all times that we are dealing here with systems, which are required to function at least during next several decades all the time providing a steady and reliable service.

We examined standards used for the repository and for client support. We discovered it was important that the repository follows the OAIS specifications and relevant protocols such as OAI-PMH³¹. As for the client support the accessibility functions of the repository interfaces were very interesting. Because these repositories are for the most part accessible via a web browser we found it was important that the interface design followed the recommendations outlined by W3C³² standards and especially in the Web Content Accessibility Guidelines ver. 2.0³³. This last recommendation is important from the disabled persons' viewpoint – these repositories hold data otherwise accessible via public infrastructure and as such should be disabled-friendly as well. EDOC's web interface to the repository is compliant with Web Content Accessibility Guidelines 2.0, level 1. This means basic accordance with the standards recommendations, the basic level of accessibility.

³¹ OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) is a protocol developed by the Open Archives Initiative. It is used to harvest the metadata descriptions of the records in an archive so that services can be built using metadata from many archives. (<http://en.wikipedia.org/wiki/OAI-PMH>)

³² The World Wide Web Consortium (W3C) is an international consortium where Member organizations, a full-time staff, and the public work together to develop Web standards. W3C's mission is: To lead the World Wide Web to its full potential by developing protocols and guidelines that ensure long-term growth for the Web. (<http://www.w3.org/Consortium/>)

³³ Web Content Accessibility Guidelines ver. 2.0 (<http://www.w3.org/TR/2004/WD-WCAG20-20040730/#intro>)

2.3.4 SOFTWARE

Software used in the evaluated partners' trusted repositories is shown in **Table 11**.

Table 11: Software in reUSE demonstrators' repositories

	DIGAR	ALO	EDOC
Operating system	Debian GNU/Linux 3.1 ('sarge'), kernel 2.6.8.1 (GPL)	Linux (GPL)	Sun Solaris 2.6 (runs on: Sun Workstation 2 x 300 MHz CPUs, 1GB RAM, RAID-System)
Main repository software	Fedora 2.0 (Mozilla Public License Version 1.1)	ALO System (in-house developed) Delphi Client (in-house developed)	Self developed
Database server	MySQL 4.1.10 for Ingest Buffer database and Fedora database (GPL)	MySQL Database (GPL alike license)	Sybase (runs on: Sun Fire V880 6 x 900 Mhz Ultra SPARC-III, 36 GB)
Other software	Apache 2.0.53, PHP 4.3.10 for Depositor Management interface (Apache License, Version 2.0 (compatible with GPL)), Jakarta Tomcat 5.5.7 hosts publishers interface, object management interface and search interfaces (Apache License, Version 2.0 (compatible with GPL)), Java 1.50_02 (Sun license) ProFTP 1.2.10-10 (GPL) Bacula 1.36.1-1 for backup (GPL), Clamav 0.83-3 for virus check (GPL), Plop 2.1.0 for PDF processing (PDFLib GmbH license for 2 CPUs), Xpdf 3.00-12 for PDF processing (GPL)	Apache TomCat Java Servlet Container (GPL like license) Java SDK Sun (Sun license)	Apache web server Tomcat PHP front ends (Metaln, MetaOut/MetaSearch)

According to the recent ARGE Group³⁴ survey on the long-term preservation of digital documents at European libraries "30% of the libraries that maintain or plan to create a digital repository use or intend to use open source software for this purpose. 13% use a commercial solution, 11% opted for specifically developed software and 8% (plan to) use software that consists of various components." The survey concludes that DSpace, ePrints, Opus and MyCoRe are the most commonly used solutions. The data relevant to the reUSE project are the following findings of the survey:

³⁴ Krimbacher, Monika, Neuhauser, Michael, Vogl, Martina: Survey on the long-term preservation of digital documents in European libraries 2005. - Innsbruck : ARGE Kulturconsulting - Medienkonzepte - Marktforschung ; Innsbruck : reUSE Project, 2005.

- Austria is mostly undecided regarding which type of software to use (87,5%) and commercial software (12.5%);
- Estonia is half-way decided on the use of Open Source software (50%) with the remaining 50% of surveyed libraries undecided as to the type of software to use; and
- Germany is pro-Open Source software (50%), though they use specifically developed repository software (13.9%) and mixed licence software (13.9%) as well.

In general the survey finds national libraries more interested in commercial and specifically developed repository software, while university libraries prefer open source repository software.

All demonstrators mostly use open source software (GPL³⁵ and similar licenses). This is important for future repository development since open source software can be modified for specific needs³⁶ of the institutions by the institution itself or by contractors hired to do it for them. This also means independence of the commercial products that normally cannot be used in this way – users are forced to wait for the producer of the software to update the software and add required functions. Open source allows more flexibility yet requires more knowledgeable and trained maintainers as well.

2.3.5. REPOSITORY SYSTEM & ADMINISTRATION

2.3.5.1 Collections access by user types

Collections access is an important factor both in security and the administration viewpoints. These types of access rights given to repository users are of special interest:

- System administrator (a person who maintains the repository system),
- Content administrator (a person who administers content and process new additions to the repository),
- Content producer (a person or an institution, who contributes content to the repository),
- User (logged in to the repository and has access to all of the repository content),
- Visitor (has access to freely available and non-restricted parts of the repository).

EDOC and DIGAR provide the following user types: content administrator, content producer and end-user with full access to the repository content. ALO's only user type is end-user, since the content providers do not work directly with the system.

According to the rights given to certain types of users appropriate user interfaces are required. ALO and EDOC provide user interfaces only for their end-users while DIGAR provides different views on objects for different user types.

³⁵ The GNU General Public License (GNU GPL or simply GPL) is a free software license, originally written by Richard Stallman for the GNU project. It has since become the most popular license for free software. The GPL grants the recipients of a computer program the following rights, or "freedoms":

- the freedom to run the program, for any purpose.
- the freedom to study how the program works, and modify it. (Access to the source code is a precondition for this)
- the freedom to redistribute copies.
- the freedom to improve the program, and release the improvements to the public. (Access to the source code is a precondition for this)

³⁶ VARNER, Philip E. The Economics of Open Source Software.

http://www.cs.virginia.edu/~pev5b/writing/econ_oss/advantages.html (last accessed 2005-11-24)

Users cannot customize their profile or their user interface in any of the evaluated repositories. In certain situations this function could be beneficial – i.e. in a user account in which the users' favourite or most used documents are stored for user's easy access.

2.3.5.2 User administration

Registration authentication and password administration are important parts of the repository security and of controlling its access. We were interested in the authentication methods used, customizable user profiles and other functions available. User administration involves limiting users' access to the objects stored within the repository. This can be done at file level, object level or at collection level. Although all of the repositories provide their contents for free, DIGAR can limit access at object and file level while EDOC can limit access only at object level.

User feedback is an important source of valuable improvement tips and possible expansions of service. Demonstrators receive user comments via e-mail (and via telephone in the case of EDOC). DIGAR development team even considers the user suggestions in updates for the next development phase.

2.3.5.3 Content submission administration

In this part we set out to research submission support and submission stages. We were interested in collection management within the repository and in the type of submission parameters installed. Submission support is also a very important topic so we were interested in various notifications submitted to content contributors and in the type of services used to support the submission process.

Multiple collections within the same instance of the repository enable organizing the contents held within the repository in to logical collections. All demonstrators allow for a creation of multiple collections within the same instance of the repository system. EDOC and DIGAR can even define different submission parameters for each collection.

Contributors require at least notification of successful content submission to the repository as well as some other services. DIGAR's repository system allows submitters to view the history of their submissions that indicates the status of submissions. It also sends them a notification email upon successful ingest. ALO does not allow submitters to work directly with the repository system. EDOC's repository system notifies submitters with an email first upon the successful upload and later if the status of the document changes. Another service that all of them could implement in the future includes a list of pending submissions and approved submissions so that the contributor can easily see the progress of submission process.

All demonstrators accept the content via off-line media (such as optical discs or floppy disks) and one or more on-line methods – ftp and email attachment. None of them supports automatic updating of the content from a given URL i.e. web-harvesting³⁷. Only EDOC included anti-virus, data format and digital signature checks into their repositories ingest workflow.

³⁷ Web harvesting (also known as Web farming, Web mining and Web scraping) is the process of gathering and organizing unstructured information from pages and data on the World Wide Web. (<http://www.computerworld.com/databasetopics/businessintelligence/story/0,10801,93919,00.html>)

2.3.6 CONTENT MANAGEMENT

In this part of the survey we were interested in document formats and import / export options in place. There are many file formats in existence and to support them all is impractical. Therefore, our partners' repositories accept certain types of files, convert them for storage in other more appropriate formats, and present them to users in file formats appropriate for displaying over the internet. The repository contents must be in a file format most likely to be readable in the future – therefore well established, standardized and open file formats are preferred. The repository content must also be appropriately tagged with metadata to aid in search, organization and conversion of the content into various file formats appropriate for various situations.

2.3.6.1 Document formats

Content of reUSE repositories are largely static documents with text and images. This still means there are many possible file formats to consider when accepting the content for ingest into the repository. Each demonstrator has their own method of managing the submitted files.

According to the White Paper on Digital Repositories³⁸ ALO takes publications in any format - meaning that the format of the submitted object, of the archival object and the access object are largely the same. DIGAR also accepts largely any format provided by the data producer. They have a preferred set of file formats and they can not assume responsibility for long-term preservation of other file formats. EDOC on the other hand promotes the use of document templates when creating files in software such as LaTeX, Microsoft Word or Open Office where templates enable a largely automatic conversion to an XML based format. XML representation enables the automatic generation of a variety of other data formats, which is of special value in a trusted digital repository.

The following tables summarize the various formats used in the repositories as described in the Survey on Trusted Digital Repositories for the project reUSE conducted in October 2004³⁹. It is possible to conclude that EDOC offers the richest set of dissemination formats, followed closely by ALO. All of the reviewed demonstrators offer mainly dissemination formats appropriate for access on the internet (**Table 14**). Storage formats are largely unchanged from the submission formats as it is described in the White Paper on Digital Repositories (**Tables 12 and 13**).

Table 12: reUSE demonstrators' submission formats

DIGAR	ALO	EDOC
HTML, MS Word, QuarkXpress, Adobe PageMaker	HTML, PDF, MS Word	SGML, XML, PDF, Postscript, LaTeX, MS Word, OpenOffice.org Writer
TIFF, GIF, JPEG	TIFF, GIF, JPEG, PNG	TIFF, JPEG, GIF, MPEG, AVI
	ZIP	

³⁸ White Paper on Digital Repositories (reUSE Digital Master Files of Printed Material!), http://www2.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf.

³⁹ White Paper on Digital Repositories (reUSE Digital Master Files of Printed Material!), http://www2.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf.

Table 13: reUSE demonstrators' storage formats

DIGAR	ALO	EDOC
HTML	PDF, MS Word	ASCII, SGML, XML, HTML, PDF, Postscript, LaTeX, MS Word
TIFF, GIF, JPEG	TIFF, GIF, JPEG, PNG	TIFF, JPEG, GIF MP3, MPEG, AVI
	ZIP	

Table 14: reUSE demonstrators' dissemination formats

DIGAR	ALO	EDOC
HTML	PDF, MS Word	SGML, XML, HTML, PDF, Postscript
JPEG	TIFF, GIF, JPEG	TIFF, JPEG, GIF MP3, MPEG
	ZIP	

2.3.6.2 Metadata

Metadata are defined as data about data and can be copied, exchanged and linked, inherited among different resources and they can even be automatically transcribed between different forms of representation. Some metadata types can be created automatically and retained for future reference.

Metadata are used for a variety of different tasks, such as describing collection items, supporting retrieval, managing collection items within a digital library system, establishing their context, preserving their authenticity, protecting the integrity of archival items, enforcing a business model and rights management, facilitating interoperability and identification of the technical nature of the digital object for preservation purposes. These tasks are supported by three main types of metadata: descriptive, structural and administrative. Descriptive metadata describe a resource for discovery and identification. Structural metadata indicate the composition of items. Administrative metadata facilitate the management of a resource and include rights management metadata, technical metadata and preservation metadata⁴⁰.

Metadata schemas can be defined on different levels. The METS⁴¹ standard is a plug-in model for encoding descriptive, administrative and structural metadata while its syntax is encoded in XML Schema. XML Schemas express shared vocabularies and allow machines to carry out rules made by people. They provide a means for defining the structure, content and semantics of XML documents in more detail⁴². METS does not define metadata elements – it remains on a high structural level. Organizational metadata schemas would then comply with METS and plug-in other metadata standards with a suitable element set. All reUSE partners acknowledge and they all comply with the METS standard. Other metadata standards vary between the partners.

⁴⁰ White Paper on Digital Repositories (reUSE Digital Master Files of Printed Material!), http://www2.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf, p. 18-19.

⁴¹ METS – Metadata Encoding & Transmission Standard (Library of Congress), <http://www.loc.gov/standards/mets/>

⁴² World Wide Web Consortium: XML Schema, <http://www.w3.org/XML/Schema>

Descriptive metadata are the type most similar to traditional library cataloguing since it includes elements such as author and title. It is essential for resource discovery and identification. Best known among the relevant standards is the Dublin Core Metadata Element Set (DC)⁴³. Its main goal is interoperability and broad applicability. Since its design was intended to facilitate its operating it consist of only 15 loosely defined elements. All demonstrators use Dublin Core for descriptive metadata schemas.

Structural metadata schemas vary among demonstrators. DIGAR uses FOXML, a FEDORA extension of METS as import/export format⁴⁴. ALO uses METS. EDOC uses a complex metadata database in Sybase that exports to OAI-DC with XML and to XMetaDiss, a metadata format by the German National Library⁴⁵.

Digital preservation permeates all aspects of the repository and it can only be successful when it involves a comprehensive strategy supported by people and organization as well as technology. Preservation metadata is considered any metadata schema that is intended for supporting the management of digital resources over time. Sometimes preservation metadata merely signifies the technical descriptions of a specific digital file. This kind of metadata is often labelled as technical metadata.

DIGAR and ALO use their own in house developed long term preservation metadata schemas, while EDOC does not support long term preservation metadata schema other than conversion into XML based archive object. DIGAR based their schema on NLNZ's⁴⁶, NLA's⁴⁷ experiences and LMER schema⁴⁸.

None of the demonstrators provide automated metadata review support to automatically update metadata objects in case an error was made or more correct data is available. DIGAR and EDOC support UNICODE character set in their metadata fields, while ALO does not. This is an important feature in view of long-term preservation because of correct inclusion and support for foreign alphabets and characters.

2.3.6.3 Content import and export

DIGAR's repository enables import and export of objects comprising metadata and reference to data files in FOXML and METS formats. It also enables DC metadata harvest via OAI-PMH⁴⁹. ALO's repository allows the import and export of all objects in METS format; while EDOC's repository can export metadata via OAI-PMH, and provides a web interface for import from OPAC⁵⁰.

⁴³ Dublin Core Metadata Initiative. <http://www.dublincore.org/>

⁴⁴ FOXML is a simple XML format that directly expresses the Fedora digital object model. As of Fedora 2.0, digital objects will be stored internally in a Fedora repository in the FOXML format. In addition, FOXML can be used for ingesting and exporting objects to/from Fedora repositories. The Fedora extension of METS will continue to be supported as an ingest and export format. In upcoming releases, Fedora will also support other formats for ingest and export such as METS 1.4 and MPEG21/DIDL. <http://www.fedora.info/download/2.0/userdocs/digitalobjects/introFOXML.html#WHAT>

⁴⁵ Metadata Set of Die Deutsche Bibliothek for Online Dissertations and Post-Doctoral Theses, Including References to the Authors (XMetaPers) http://www.ddb.de/standards/pdf/ref_xmetadiss_v1-2_e.pdf

⁴⁶ National Library of New Zealand: Metadata Standards Framework – Preservation Metadata (revised, 2003), http://www.natlib.govt.nz/files/4initiatives_metaschema_revised.pdf

⁴⁷ National Library of Australia, <http://www.nla.gov.au/>

⁴⁸ LMER – Langzeitarchivierungsmetadaten für elektronische Ressourcen [Long-Term Preservation Metadata for Electronic Resources], <http://www.ddb.de/standards/lmer/>.

⁴⁹ Open Archives Initiative Protocol for Metadata Harvesting, <http://www.openarchives.org/>

⁵⁰ Online Public Access Catalog; a computerized library catalog, or the portion of the catalog available for patron use.

2.3.6.4 Updating and indexing

Updating and indexing the repository contents is a very important activity for keeping the repository contents up-to-date. DIGAR and ALO update and index their repository contents in real time, while EDOC updates and indexes after checking, conversion and approval of the workflow database.

2.3.7 SERVICES

All three demonstrators provide access to their repository via the internet, as stated in the project description. Main services that should be provided are search capability, appropriate presentation of search results and value-added services. ALO and EDOC provide print-on-demand and CD delivery as value added services in the scope of other repository parts and these services will be further extended to cover the reUSE project⁵¹.

2.3.7.1 Search capability

EDOC offers Boolean logic operators for descriptive metadata search, while ALO does too but in a limited fashion with only the AND operator and right truncation performed automatically. DIGAR allows the use of truncation, wildcards and the Boolean operator AND in their search mechanism. (Tables 15 and 16)

Table 15: Searchable metadata fields

DIGAR	ALO	EDOC
author ⁵² , title, date, subject, publisher, type	author, title, date, subject	author, title, date, keywords, date, period of time, language

Table 16: Search types

DIGAR	ALO	EDOC
Simple – search in all DC fields.	Simple	Simple
Advanced – search limited to a few predefined fields: title, author, publisher and date.	Advanced - search within four fields: title, author, date and subject	Advanced – search in four fields: title, author, keyword, summary. There are also advanced functions such as choice of language, date limitations, sorting of results and number of displayed results

Only EDOC offers full-text search capabilities (including highlighting search term) because they use Google technology.

⁵¹ White Paper on Digital Repositories (reUSE Digital Master Files of Printed Material!), http://www2.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf.

⁵² This is creator according to Dublin Core terminology.

2.3.7.2 Presentation of search results

The ability to browse search results is shown in **Table 17**.

Table 17: Ability to browse search results

DIGAR	ALO	EDOC
Provides a link to the library catalogue.	author, title, date	publication type, branch of study, institutions, author, edition, volume (e.g. journal)

2.3.7.3 Value-added functions

At the moment there is no additional support for disabled users. ALO offers multiple language support and indexation by search engines (e.g. Google, Bibliographic list). ALO plans to have a structural map and PDF support in the new version of their system. EDOC's content can be indexed by external search engines; they offer print on demand via ProPrint⁵³; and have the technology transfer e.g. bhtw Leipzig. Value-added functions are not very clearly defined by any of the partners at this moment.

2.4 REVIEW OF ALL FINDINGS IN THE EVALUATION OF TECHNICAL ASPECTS

With this evaluation we attempted to create an evaluation framework in order to examine the quality of the evaluated repositories, technical specifications, repository and system administration, content management and services. Usage logs were also examined in order to see the practical use of the repositories. We find that the repositories are quite different in comparison to each other regarding the design and in ways of achieving the same goal. They use different kinds of software – mainly open source and self-developed software. Though their security is very similar with end users and server/database access, they use different approaches for content providers.

Since a trusted repository must also be reliable, all partners ensure the reliability of their repositories. They all have written policies for disaster readiness, response and recovery. Recovery includes a rebuild of the entire system from backup server or location – of which all partners have at least two. They all do daily backups and use systems to minimize the effect of a failure on the entire service. Documentation and training are also provided for future reference in case of updating or maintaining the repositories, as well as instructions for content providers. Staff responsible for the repositories is trained additionally for the specific needs of the repository, although that training is often non-formal and specific to the field of expertise of personnel working on the project.

In order for the trusted repository to do its tasks, it also has to be expandable meaning new file types and new workflows, as well as update support and technology innovations. All examined repositories can be upgraded with support for new file types and workflows.

⁵³ ProPrint is a joint venture of the Computer and Media Service and University Library of Humboldt-University at Berlin and the State and University Library of Lower Saxony in Göttingen supported by the DFN-society. <http://dochohost.rz.hu-berlin.de/proprint/>

Service quality also heavily depends on the standards observed and used in the repository design. All partners followed basic OAIS guidelines as much as possible and the recommendations in the RLG-OCLS Report on Trusted Digital Repositories. Since these repositories are accessible over the World Wide Web it is important that partners follow relevant W3C standards to ensure maximum possible accessibility.

Content management varies among partners, they use different takes on ingest process, supported file formats and metadata (although they all support METS and DC). Repository contents should also have various services attached to it and among these search and presentation of search results are paramount to the end users. All partners allow for simple and complex search with roughly similar data fields, but their presentation varies. On top of these there should also be some value-added services provided, although these were not available at the time of this evaluation and were thus not evaluated. Even so, there are available services such as print-on-demand and CD delivery, multiple language support and indexation by external search engines.

Table 18 provides a quick summary of partners' repositories evaluation findings and describes the basic properties of each repository.

Table 18: Summary of partners' repositories evaluation findings with the description of basic properties of each repository

Evaluation topics	ALO	DIGAR	EDOC
Quality issues			
- Functionality	OAIS compliance achieved.	OAIS compliance achieved.	Close to achieving OAIS compliance.
	Can hold over 100.000 objects.	Can hold up to 1 million objects.	Software limitations are unknown.
	Peak usage occurs during weekdays; usage diminishes over weekends and holidays	Peak usage occurs during weekdays; usage diminishes over weekends and holidays	Peak usage occurs during weekdays; usage diminishes over weekends and holidays
	End-users and content providers do not have access to the database server or database.	Access to the database server and database itself limited via IP address and login information.	Access to the database server limited via IP addresses and login information. EDOC does not limit access to the ingest process because they require content to be delivered via templates.
- Reliability	IBM Tivoli System Management. Use of standard routines provided by the operating system for data verification.	A solution of RAID 5 disk array on a separate location. Periodic verification of content files checksums.	IBM Tivoli System Management. Use of digital signature for data verification purposes.
	The primary backup storage device is hard disk. Backup storage device is magnetic tape.	The primary backup storage device is hard disk. Backup storage device is magnetic tape.	The primary backup storage device is hard disk. Backup storage device is magnetic tape.
	2 backup locations.	2 backup locations and periodic data duplication to another in house server.	3 backup locations
	Has written policies for disaster readiness, response and recovery.	Has written policies for disaster readiness, response and recovery.	Has written policies for disaster readiness, response and recovery.
	In case of critical system failure rebuild of the entire system from backup.	In case of critical system failure rebuild of the entire system from backup.	In case of critical system failure rebuild of the entire system from backup.

- Efficiency	Dependant on various parameters, such as the average time to process a typical request for displaying a digital object	Yet untested in practice – testing currently in progress.	Dependant on the quantity of the data transferred.
- Maintainability	ALO provides handbook for users in German language; developers are supported via the SourceForge project.	DIGAR provides manuals for content managers and online tutorials for depositors and end-users.	EDOC offers documentation for developers, administrators, authors, and end-users.
	Profiles: Technical administrator.	Profiles: System administrator, Technical support, Metadata input, Database administrator.	Profiles: Technical support, Metadata input.
- Flexibility	Can be upgraded with support for new file types and a new storage system.	Can be upgraded with support for new file types, new workflows, a database system and application server.	Can be upgraded with support for new file types and a new storage system.

Evaluation topics	ALO	DIGAR	EDOC
Technical specifications			
- Standards	General OAIS recommendations OAI-PMH METS standard Dublin Core Metadata Element SET (DC)	General OAIS recommendations OAI-PMH METS standard Dublin Core Metadata Element SET (DC) FOXML, a FEDORA extension of METS as import/export format	Subset of OAIS recommendations OAI-PMH Web Content Accessibility Guidelines 2.0, level 1. METS standard Dublin Core Metadata Element SET (DC) A self developed standard based on XML
- Software	Open source and self developed.	Open source.	Self developed.

Evaluation topics	ALO	DIGAR	EDOC
Repository system & Administration			
- Access to collections by user types	User type: end-user.	User types: content administrator, content producer and end-user.	User types: content administrator, content producer and end-user.
	User interfaces for end-users.	Different views on objects for different user types.	User interfaces for end-users.
- User administration	None required – only user is end user.	Can limit access at object and file level	Can limit access at object level
- Content submission administration	Creation of multiple collections within the same instance of the repository system.	Creation of multiple collections within the same instance of the repository system.	Creation of multiple collections within the same instance of the repository system.
	ALO takes care of content ingest for their contributors.	Can define different submission parameters for each collection.	Can define different submission parameters for each collection.
	Does not allow submitters to work directly with the repository system.	Repository system allows for submitters to view history of their submissions and indicates the status of submissions; it also sends them a notification email upon successful ingesti.	Repository system notifies submitters with an email upon the successful upload and later if the status of the document changes.

Evaluation topics	ALO	DIGAR	EDOC
Content management			
- Document formats	Takes publications in any format.	Accepts largely any format provided by the data producer.	Promotes the use of document templates when creating files.
	The submitted object, the archival object and the access object are largely the same.	Has a preferred set of file formats.	Templates enable a largely automatic conversion to an XML based format.
	Dissemination formats appropriate for access on the internet	Dissemination formats appropriate for access on the internet	Dissemination formats appropriate for access on the internet
- Metadata	METS, Dublin Core, in-house developed long term preservation metadata schemas.	METS, Dublin Core, FOXML, in-house developed long term preservation metadata schemas, UNICODE character set in metadata fields.	METS, Dublin Core, a self-developed standard based on XML, does not support long term preservation metadata schema, UNICODE character set in metadata fields.
- Content import & export	Import and export of all objects in METS format.	FOXML, a FEDORA extension of METS as import/export format.	Can export metadata via OAI-PMH, and provides a web interface for import from OPAC.
- Updating & indexing	No automated metadata data review support.	No automated metadata data review support.	No automated metadata data review support.
	Updates and indexes the contents of their repository in real time.	Updates and indexes the contents of their repository in real time.	Updating and indexing after checking, conversion and approval of the workflow database.

Services			
- Search capability	Boolean logic operator AND for descriptive metadata search and right truncation performed automatically.	Allows use of truncation, wildcards and the Boolean operator AND in their search mechanism.	Boolean logic operators for descriptive metadata search and full-text search capabilities (including highlighting search term) because of Google technology use.
	<ul style="list-style-type: none"> - Simple - Advanced in four fields: title, author, date and subject. 	<ul style="list-style-type: none"> - Simple – search in all DC fields. - Advanced – search limited to a few predefined fields: title, author, publisher and date. 	<ul style="list-style-type: none"> - Simple - Advanced in four fields: title, author, and keyword, summary. There are also advance functions such as choice of language, date limitations, sorting of results and number of displayed results.
- Presentation of search results	author, title, date	Provides a link to the library catalogue.	publication type, branch of study, institution, author, edition, volume (e.g. journal)
- Value-added services	<ul style="list-style-type: none"> - print-on-demand and CD delivery - multiple language - support and indexation by search engines; - plans to have a structural map and PDF support in the new version of their system 		<ul style="list-style-type: none"> - print-on-demand and CD delivery; - content can be indexed by external search engines; - offers print on demand via ProPrint; - has technology transfer e.g. bhtw Leipzig.

3 EVALUATION FROM THE USERS' PERSPECTIVE

By Mateja Šmid

3.1 BACKGROUND

This report presents the results of end-user and non-user survey which was carried out in the reUSE project demonstrators' libraries of the three repositories. Along with the long-term preservation of digital publications the next important goal is definitely public access and in this regard user-centred design of all three repositories. Several repositories worldwide have shown that users are their most critical element. Repositories as such are valuable only if used. The success of the repository is often influenced by the content and the design of the interface. Many repositories are design-centred rather than user-centred.

In order to make the report easier to understand, we included the description, explanations and joint diagrams of the survey findings for all three repositories. Detailed diagrams and statistics were made separately in the excel files which enable more cross tabulations if needed in the future.

3.2 EVALUATION METHODOLOGY

User evaluation is usually carried out on target group or as a case study and is done in special test rooms on test computers. It is very complicated and expensive to carry out such an evaluation because of the specific nature of the reUSE project. One of the obstacles for the Slovenian evaluators was also the fact that evaluation in demonstrators' libraries had to be carried out in the national languages, which meant German and Estonian.

3.2.1 METHODOLOGY OF DATA ACQUISITION

3.2.1.1 Target group

The target group of this part of evaluation predominantly consist of end-users as well as some non-users. Because of the above mentioned specifications we decided to design a paper-based questionnaire and distributed it to two of the evaluated repositories: ALO and DIGAR. In this way we were also able to get some feedback from non-users, which played a key role for further dissemination. The ONB repository had not been finished yet, while the EDOC team had decided to do an on line survey. This was not a very good idea for the evaluation, since by the end of August 2005 they had gathered only 4 filled-in questionnaires out of expected 100. Because of this reason, we were not able to include them in the 1st period survey. Since the EDOC repository was set up in the last months of 1997, it has ample experience with many documents and users. Therefore, it was of great importance for the comparative analyses of all three repositories, to include them at least in the last evaluation. By the end of October 2005 they were finally able to find 33 on line filled-in questionnaires.

3.2.1.2 Sample

In order to get a stable and representative sample we required approximately 100 questionnaires of end-users and 50 questionnaires of non-users for each repository. However, this goal turned out impossible to reach. The section 'Work done' includes detailed information about the filled-in questionnaires for each repository.

3.2.1.3 Type of inquiry

We designed a special questionnaire for end-users based on ISO 9241-11:1998 (Usability standard) and ISO 13407:1999 (Human-centred design processes for interactive systems). The questionnaire consists of a demographic part and of the questionnaire, which is the main focus of the survey. The research objective was to obtain information on the following topics:

- content,
- recognition,
- performance and efficiency,
- personal and subjective perception,
- error tolerance.

There were also two open questions under content section and a special space for the end-users' comments in the questionnaire. Comments were included in the analysis as qualitative interpretation.

A part of the end-user questionnaire is presented in **Figure 6**.

reUSE
re-use about your digital library?

Dear user,
we kindly ask you to fill in the questionnaire. The information you provide in this questionnaire will help us to improve the reUSE digital repository. It is confidential and will not be shown to anyone except the project team staff.

GENERAL INFORMATION

Date _____ Country (A) (B) (C)
Year of birth _____ Gender (M) (F)
Profession _____ Field of work _____

Please print the ball with red pencil

Experience with using computers: expert (), intermediate (), novice (), none ()

Where did you learn about the repository? library (), internet (), from a friend (), other media ()

How frequently do you use the digital repository? every day (), once a week (), once a month (), few times a year (), never ()

Rank the criteria indication 1 for the most important method and 5 for the least important.

How would you like to search for information?

By author (1) (2) (3) (4) (5) (6)
By title (1) (2) (3) (4) (5) (6)
By year (1) (2) (3) (4) (5) (6)
By keywords (1) (2) (3) (4) (5) (6)
Full text searching (1) (2) (3) (4) (5) (6)
Others (1) (2) (3) (4) (5) (6)

How would you like to search result to be presented?

PDF (1) (2) (3) (4) (5) (6)
HTML (1) (2) (3) (4) (5) (6)
DOC (1) (2) (3) (4) (5) (6)
Post Script (1) (2) (3) (4) (5) (6)
Plain text (1) (2) (3) (4) (5) (6)
Not important (1) (2) (3) (4) (5) (6)

Please mark your opinion about the following statements with sign X.
Evaluation scale: 1 completely agree (1), I somewhat agree (2), I neither agree nor disagree (3), I somewhat disagree (4) I completely disagree (5)

A - PART

I'm satisfied with the documents I found in the digital repository. (1) (2) (3) (4) (5)
I think the content in the digital repository is useful. (1) (2) (3) (4) (5)
I will check for paper material as well to find all information I need. (1) (2) (3) (4) (5)

What type of the material do you expect to find in the repository ?

For what purpose will you use the content from the repository (i.e. study, research, commercial, fun, other)?

B - PART

I need to perform too many different steps to deal with a given task. (1) (2) (3) (4) (5)
The repository is well suited to the requirements of my work. (1) (2) (3) (4) (5)
In a given screen, I find all of the information I need in that situation. (1) (2) (3) (4) (5)
The repository makes it easy for me to switch between different screens. (i.e. the user has instant access from the input screen to the front page or help system) (1) (2) (3) (4) (5)
It's easy for me to move back and forth between different screens. (1) (2) (3) (4) (5)

C - PART

Help is available whenever is needed. (1) (2) (3) (4) (5)
I understand immediately what is meant by the messages displayed by the repository. (1) (2) (3) (4) (5)
The terms and concepts used in the repository are clear and unambiguous. (1) (2) (3) (4) (5)
The repository always visually marks the current entry location (e.g. by a highlight, a contrasting color, a blinking cursor, etc.). (1) (2) (3) (4) (5)
I find the collections easily accessible. (1) (2) (3) (4) (5)
The value added service (to be specified by partners) functions are useful to me. (1) (2) (3) (4) (5)

Figure 6: An example of the end-user questionnaire with special fonts designed according to Remark Office OMR software.

Since non-users are also important, a short questionnaire was designed for them as well. The outlay of both questionnaires was designed according to the requirements of the software Remark Office OMR which we used for the analyses.

The EDOC team transformed the end-users' questionnaire into an on-line form, but they kept the paper-based questionnaire for non-users.

3.2.1.4 Period of inquiry

An intensive test phase was carried out in June 2005 and from the end of September to the beginning of October 2005. It was expected that during the Summer of 2005 implementers would fill the repository with additional content which should influence the results of the 2nd evaluation period.

3.2.1.4 Procedure

Librarians of each repository were asked to help distribute the questionnaire. Every 10th visitor was asked whether they had used the repository. If the answer was negative, they were asked to fill in a short questionnaire for non-users. If the answer was positive, they were asked to fill in the questionnaire for end-users. At EDOC, where the questionnaire was on-line, they first establish a link between their news and the questionnaire. But in five weeks they received only four responses. Therefore, they activated a pop-up and received twenty-nine responses in four weeks.

With the questionnaire for non-users we were able to determine the reason of non using the digital repository.

3.2.2 METHODOLOGY OF DATA ANALYSIS

For the detailed data analysis software Remark Office OMR for statistic research was used. In order to use it properly, we first needed the required form of the questionnaire (see details above) and second, a special template with OMR fields was created. Although Remark Office OMR offers detailed statistic calculations and cross tabulations research, we rather used excel pivot function because of the more detailed diagrams. In this way data were calculated more precisely (Remark Office OMR provides calculations only for rounded up numbers and the total percentage sum is not always 100%). Data are also in the more common format (excel) and available for those who would like to make further cross tabulation analysis. For details see **Figure 7**.

The numbers of cases and percentage (frequencies) for all questions were collected and some cross tabulations (correlation between different variables) were made. Data are presented in diagrams. Comments were translated by partners and by Slovenian evaluators.

Figure 7: Marking OMR fields which will later be scanned in each questionnaire.

3.3 WORK DONE

In the first period of the evaluation which was carried out in June 2005 two digital repositories (ALO from UBI and DIGAR from NLE) were evaluated. Due to lack of data we were not able to evaluate the EDOC digital library from UBER – we got back only 4 electronically filled in end-user questionnaires. While at that time UBI's repository ALO was already running (set up in 2001), NLE had just completed their digital repository DIGAR.

Many end-users' questionnaires for different repositories were incomplete. Since it was very difficult to gather users' opinion and since data were so valuable to us, we accepted the questionnaires, which had at least some valid answers in each topic. Blank answers were marked as well and are presented in the analysis and diagrams as "missing". Some parts of the questionnaire were significantly empty; this was also taken into account and could be interpreted as end-users' indifference and lack of knowledge of certain topics.

In the evaluation of June 2005 we got back 51 end-user questionnaires and 23 non-user questionnaires from ALO; 8 of them were not valid, 1 was empty, therefore we considered 42 questionnaires as valid enough to include them in the survey. From DIGAR 14 end-user questionnaires and 134 non-user questionnaires were gathered. The DIGAR sample is a bit low so the results might be declined, but since the repository has just started operating, all information and comments were valuable. In this light it is better to consider the results from the 1st period as pre-test. **Figure 8** shows the number of returned questionnaires for all 3 repositories. The end-user group is further divided by gender.

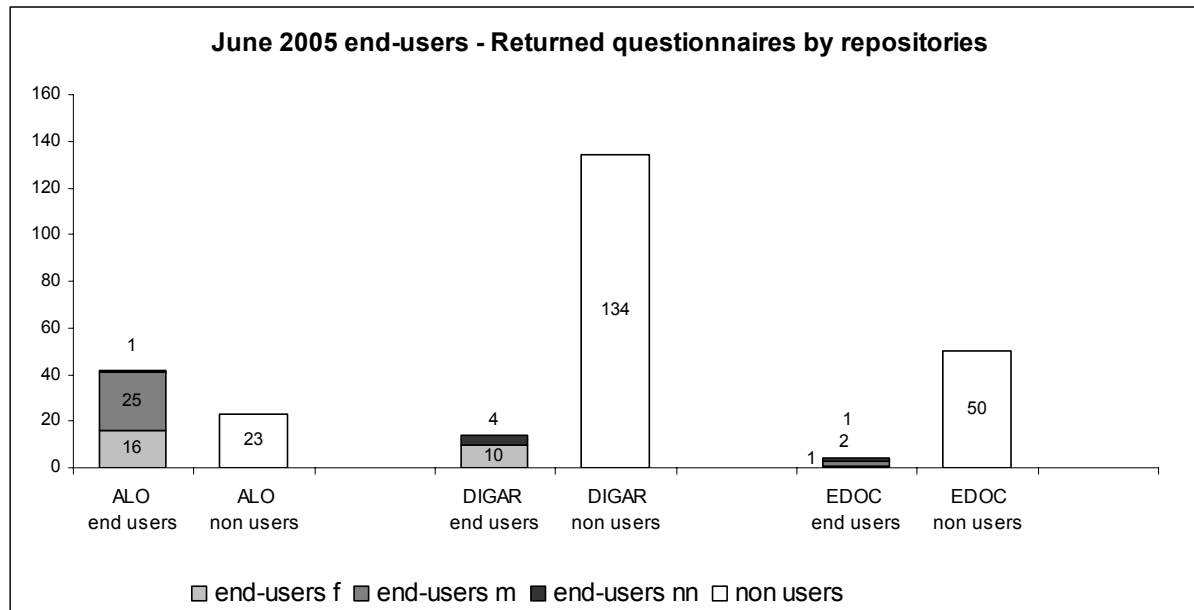


Figure 8: The number of returned end-user and non-user questionnaires from ALO, DIGAR and EDOC in June 2005

In the 2nd phase in October 2005 we got back 80 questionnaires from ALO: 19 of them were incomplete, 24 of them were empty and 37 were complete enough to be used (10 from UBI, 4 from UBG and 23 from i3s3 – Linz). From DIGAR we got 49 end-user questionnaires: 2 of them were incomplete and 47 were complete enough to be valid. Finally we got back enough questionnaires from EDOC; altogether they sent data for 33 questionnaires but 4 of them were too incomplete to be used, so we considered 29 of them as valid (See **Figure 9**).

The described situation clearly shows how difficult it was to retrieve data from on line questionnaires; paper based questionnaires and personal contact still provided more efficient feedback. Besides, paper based questionnaires were more completely filled-in than the on line ones. The ALO team distributed the questionnaires to their associate partners - University of Graz and University of Linz as well. It turned out that at the University of Graz almost no one knew the ALO repository– they only provided 4 enough complete questionnaires.

In the 1st period 23 responses to the non-user questionnaire were collected from ALO, 134 from DIGAR and 50 from EDOC. In the 2nd period we received 23 responses to the non-user questionnaire from ALO and 30 from DIGAR. In the 2nd phase in October 2005 the NLE team introduced a very innovative and interesting policy to get information from the non-users – they assembled a few non-users once a day and instead of giving them the questionnaire, they offered them a training course about the use of the digital repository. In this way they probably made the most efficient dissemination and assured many new future end-users.

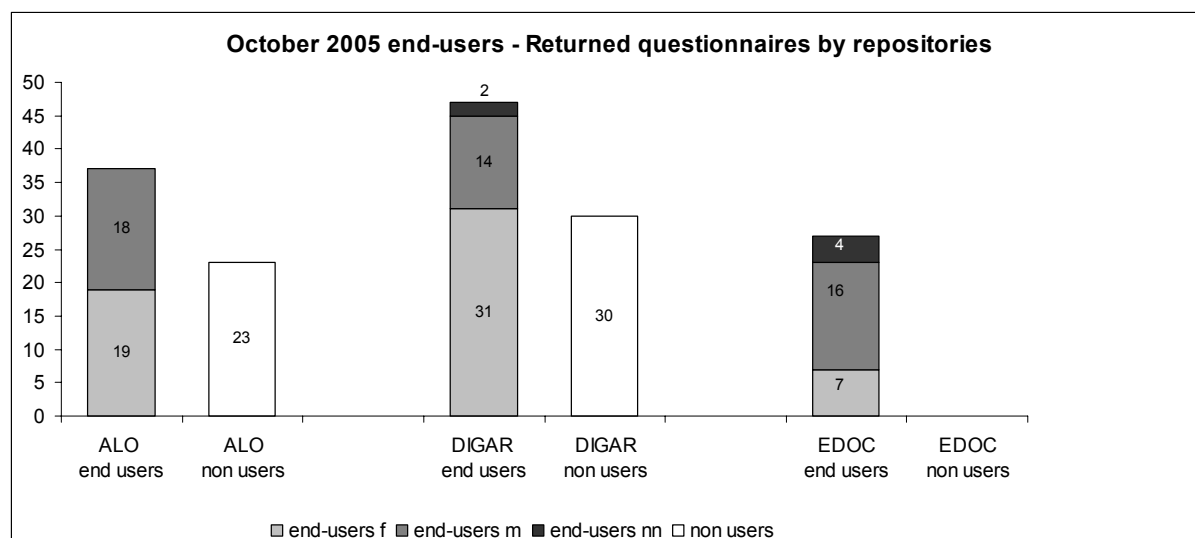


Figure 9: The number of returned end-user and non-user questionnaires from ALO, DIGAR and EDOC in October 2005

3.4 END USER EVALUATION

3.4.1 DEMOGRAPHIC AND BASIC DATA

As mentioned above, in the 1st evaluation period we were able to evaluate only ALO and DIGAR repositories. However, in the 2nd phase we managed to evaluate the EDOC repository as well. The comparison of the evaluations outlines the differences between different kinds of libraries: ALO and EDOC are university libraries while DIGAR is a national library.

Before the main questions, there was a set of demographic questions about gender, year of birth, profession and also three questions concerning experience, use and learning about a certain repository. As expected, end-users in the national library were older and of a different age in comparison to the end-users of university libraries, but surprisingly only a few end-users from the 1st evaluation phase were born in the 80s - we assumed that some people born in the 80s should have already been at university. Is there enough dissemination among the younger population or are they not aware of the advantages of digital repositories? As shown in **Figures 10 and 11** the majority of end-users were born in the 60s and 70s.

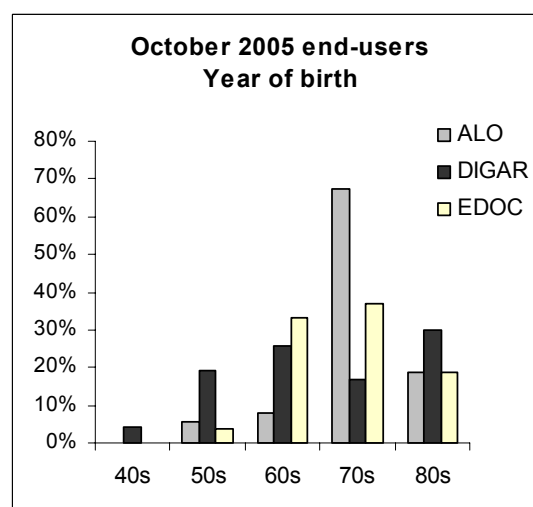
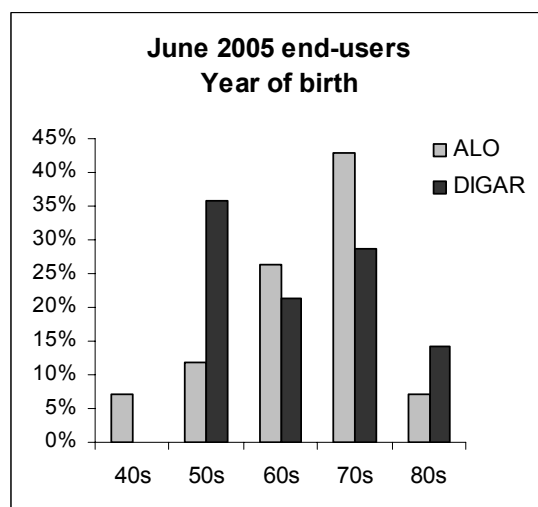


Figure 10 and Figure 11: The age of end-users in all three libraries for both evaluation phases is presented

Further analysis shows that most of the end-users have intermediate experience with computers and on average use the repository only a few times a year. The number of novices is very low. Data show that most of the DIGAR end-users learned about the repository in the library (in the 1st phase 79% and in the 2nd phase 77%) while end-users in the university library found out about the repository on the internet (EDOC 70%, ALO 41%). One of the reasons for this result is the fact that the questionnaires in DIGAR and ALO were delivered in the library, while EDOC implementers no longer have direct communication with their end-users. The question about the difficulty of the proper use of the digital repository will give the same results. Those end-users who had met with the digital repository via the web (for example in EDOC) had more difficulties accessing the reUSE digital contents than those who were trained in special courses offered by the library staff. The results also indicate the impact and the great influence national libraries have on their users. Therefore, a strong dissemination should come from their side; they are still in position to raise the awareness of a wider population and educate their end-users about proper and advance use of the digital repository. It is also interesting that many respondents from ALO (32%) claimed to have learned about the repository from a friend. More details can be found in **Figure 12** and **Figure 13**.

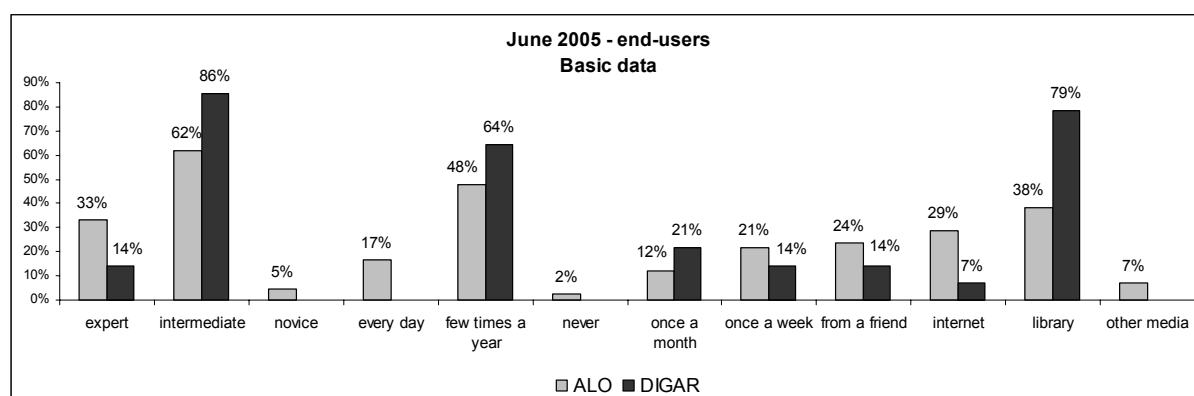


Figure 12: Detailed basic data for the end-user survey in June 2005

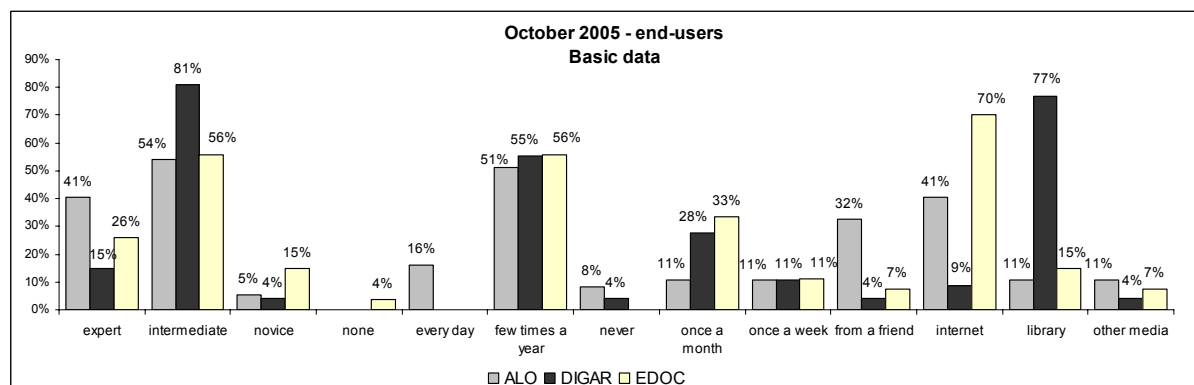


Figure 13: Detailed basic data for the end-user survey in October 2005

Unfortunately, most of the end-users didn't follow the given instructions in answering the last two questions about search criteria and format preferences. They were asked to rank the given criteria from 1 (most important) to 6 (least important) but instead marked the level of their agreement using the provided options. Therefore, we used the same methodology in these analyses as with other responses.

The most preferred format in both evaluation phases and in all three repositories is PDF. While HTML is the second most liked format in university repositories (ALO, EDOC) and in the 1st phase in DIGAR, the last evaluation phase showed that the second most preferred format for DIGAR end-users is DOC format. The end-users' dislike for PS format is evident. To some DIGAR and EDOC users format seems to be irrelevant. One of DIGAR end-users

(an architect) suggested that displayed documents should be in format A4 (always the same size). For details see **Figures 14 and 15**.

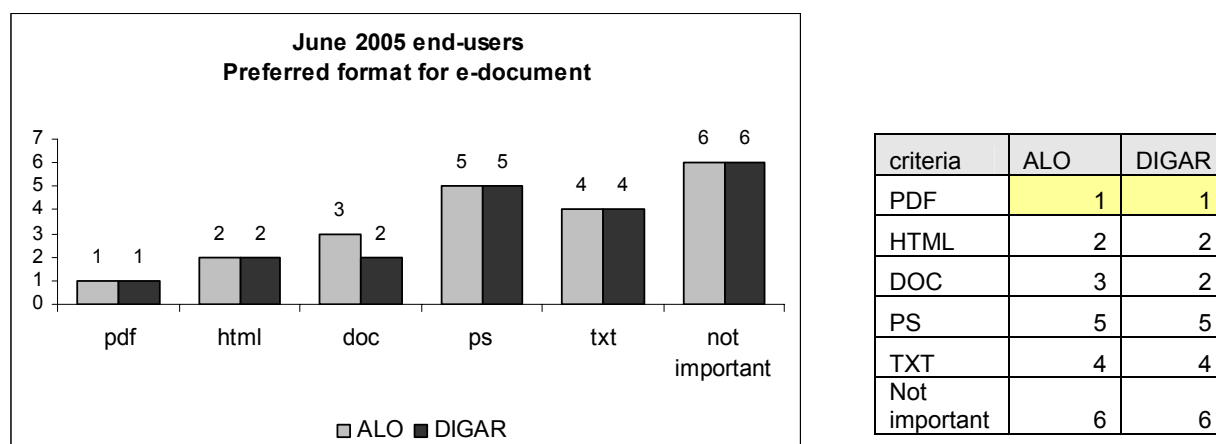


Figure 14: The format preferences from the ranked list of searching criteria for ALO and DIGAR repositories for the 1st evaluation phase (June 2005)

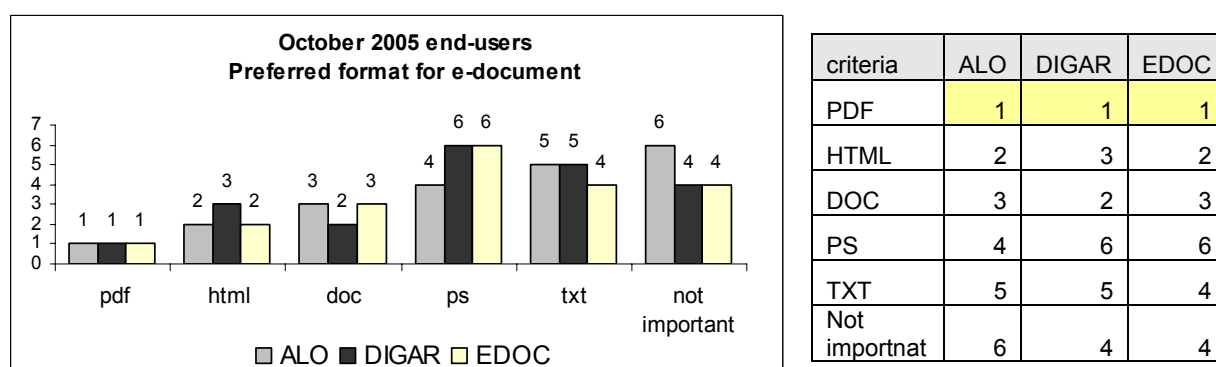


Figure 15: The format preferences from the ranked list of searching criteria for ALO, DIGAR and EDOC for the 2nd evaluation phase (October 2005)

We noticed a similar situation with search metadata preferred by end-users. In the 1st phase both, DIGAR and ALO end-users, preferred to search 'by author', as the second option they chose to search 'by title' at ALO and 'by keywords' at DIGAR. The third most preferred option to search is 'by keywords' at ALO and 'by title' at DIGAR; then followed 'full-text search' and at last 'by year' search. For details see **Figure 16**.

It was surprising that the criteria 'full-text search' reached only 4th position; either there are not many repositories that offer this useful advantage search criteria or end-users are not aware of this possibility.

However in the 2nd phase when we included EDOC's data, the situation changed as shown in **Figure 17**. We can see that users from different repositories have different preferences. While the most preferred search criteria in ALO is still 'by author', in DIGAR it is replaced 'by keywords' and in EDOC 'by title' criteria. 'Full-text search' is still not as desired and almost no one wishes to search 'by year' or by some other unspecified criteria. The reason for this might also be the fact that younger end-users born in the late 70s and 80s, who are used to search by Google, were not well represented in the survey. As we shall see in the section about personal and subjective perception, many respondents from DIGAR and EDOC expressed their wish to search 'by subject' as well.

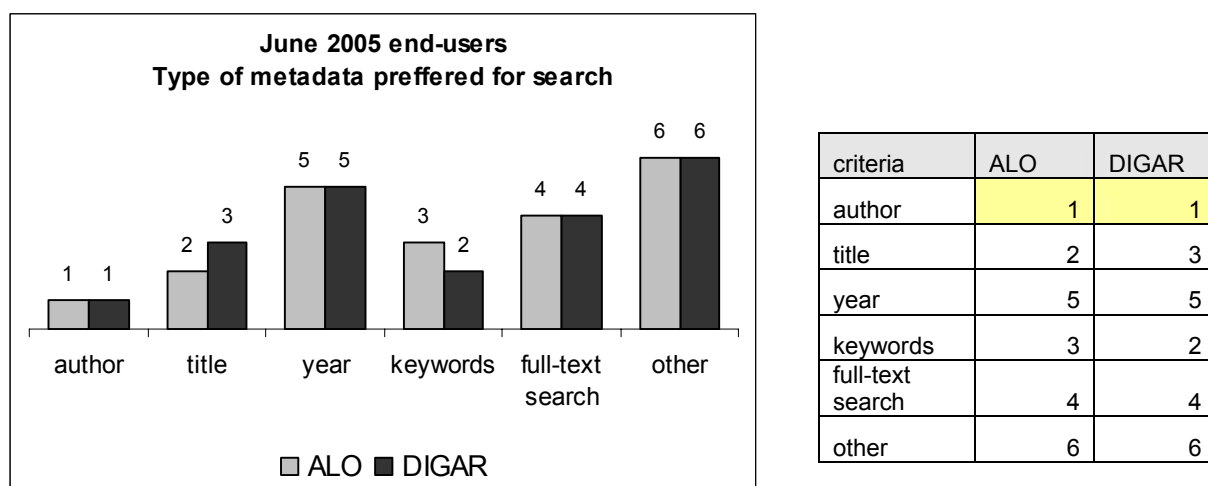


Figure 16: The diagram presents the ranked list of searching criteria for the 1st evaluation phase (June 2005)

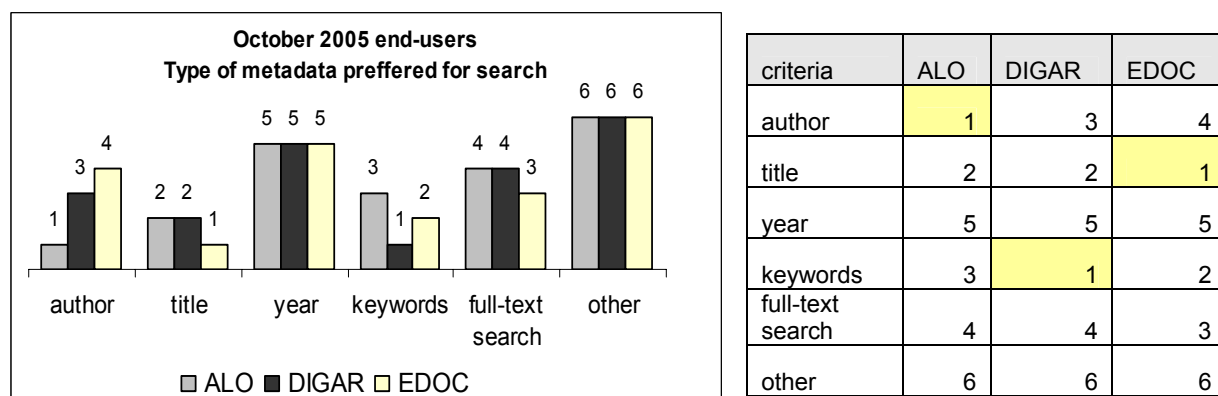


Figure 17: The ranked list of searching criteria for the 2nd evaluation phase (October 2005)

3.4.2 CONTENT

The section on digital repositories' content contained 3 close and 2 open questions. Users were asked whether they considered the content of the digital repository useful, and whether they were satisfied with it.

We were informed that there were about 3000 documents in the ALO repository. Most of the documents are different types of publications from the public sector institutions such as academic, cultural, political, clerical, and tourist organizations, social institutions, companies, communities, universities, chambers, government institutions and clubs or associations.

In June 2005 there were 71 objects⁵⁴ (containing 196 original-documents) in the DIGAR repository, but by the end of September the number of objects had already increased to 332 (containing 565 original documents). There are different kinds of publications, mainly from the public sector institutions, but also from some NGO-s. They mostly comprise annual reports, yearbooks, studies, journals, proceedings, research papers, and also some books and newsletters. Contributing organizations are mostly from the public sector i.e. State Chancellery and its subordinate agencies, the Ministries, Administrations and Boards, Inspections, Parliament, State, Audit Office, Bank of Estonia, Legal Chancellor, National Archives of Estonia, different cultural institutions and different information centres.

⁵⁴ Object - XML structure comprising metadata and content file original document - digital document (pre-print publication), sent by the publisher to the digital archive. Will be converted to archive-copy for the reason of long-term preservation and access-copy for the reason of further access (definition provided by Krista Kiisa).

The 3rd evaluated repository is the EDOC repository from Humboldt University. At the time of the user evaluation it contained nearly 2800 documents, such as dissertations, postdoctoral lecture qualifications, e-journals, digitized material (historical, handwritings), meeting and conference proceedings. But most importantly, they are dealing with scientific material from different fields (medical science, mathematics and nature science, philosophy, agriculture, history, etc...).

As shown in **Figure 18** in the 1st evaluation period DIGAR end-users were more satisfied with the content than the ALO end-users. Cross-tabulation research of the questionnaires also shows that end-users who use the digital repository a few times a year are more satisfied with digital documents than the ones who use it once a week. Why is this so? Is there enough new and fresh digital material inside the repository? Can we assume that the university library users are more critical than those from the national library?

All end-users strongly agree on the usefulness of each digital repository. The last question reveals that many end-users from both libraries (a bit more from DIGAR) will still need to search paper documents.

Comparing the results from the 2nd phase (see details in **Figure 19**) we noticed that, with the exception of a few users, the majority is quite content with the documents found in the repository. There was one end-user from EDOC, who disagreed with most of the statements and commented that there were still too few documents in the repository, and that more of them should be added. The high rate of undecided responses to such a general statement as 'satisfaction with the documents in the digital repository', was quite a surprise.

However, almost all users agreed that the digital repository content was very important. But as expected, some of them, especially from the national library (75%), claimed that they would need to check the paper material as well. It is encouraging that in ALO and EDOC, which have now been providing this kind of service for several years (ALO 5 years, EDOC 8 years) and already have a defined group of end-users, there is a certain percentage of users who find all the information they need in the digital repositories. Cross tabulation research shows that most of them are students and assistants, born in the 60s, 70s and 80s and need the digital content for their studies, research, promotion and fun.

The open question "*What type of material would you like to find in the repository?*" revealed that users were actually looking for very different content. Again there are differences between the university and national libraries, where the content should be very heterogeneous. For details see **Tables 19 and 20**. The list might be a great help for implementers for further gathering of digital content.

Table 19: The list of contents which end-users would like to find in the digital repository - data from the 1st evaluation period (June 2005)

ALO	DIGAR
books	special articles
articles	official information
catalogues	standards, economic views
(local) newspapers	e - texts of publications printed in Estonia
old and new Austrian literature	documents from various topics
scientific papers (publication)	lectures and articles
the Bible	teaching and informative material

Table 20: The list of contents, which end-users would like to find in the digital repository - data from the 2nd evaluation period (October 2005)

ALO	DIGAR	EDOC
(old) books, literature, novels	official public sector publications	books, monographs
manuscripts	articles	data bases
documentation	books, textbooks, monographs,	recent publications
scientific publications	magazines, journals	texts
journals	statistics	pictures
theses	reports, research works, studies	films
dictionaries	directives, regulations, instructions (public documents)	magazines
local newspapers	collected works	compilations
study material	general information (on education, public health, etc..)	original documents
rare documents	constructions codes	journal abstracts
	documents, reference materials	non print (incl. multimedia)
	legal acts, text of legal acts of Estonia	
	scientific literature published in Estonia	

Not surprisingly, most of the end-users use the content of the digital repositories for studies, work and scientific research. Some of them use it as well for self-education, hobbies, business and commercial purposes (DIGAR) while others use it for 'fun' (ALO) – a.k.a. hobbies. It is very clear that EDOC end-users use it almost exclusively for scientific research and investigation.

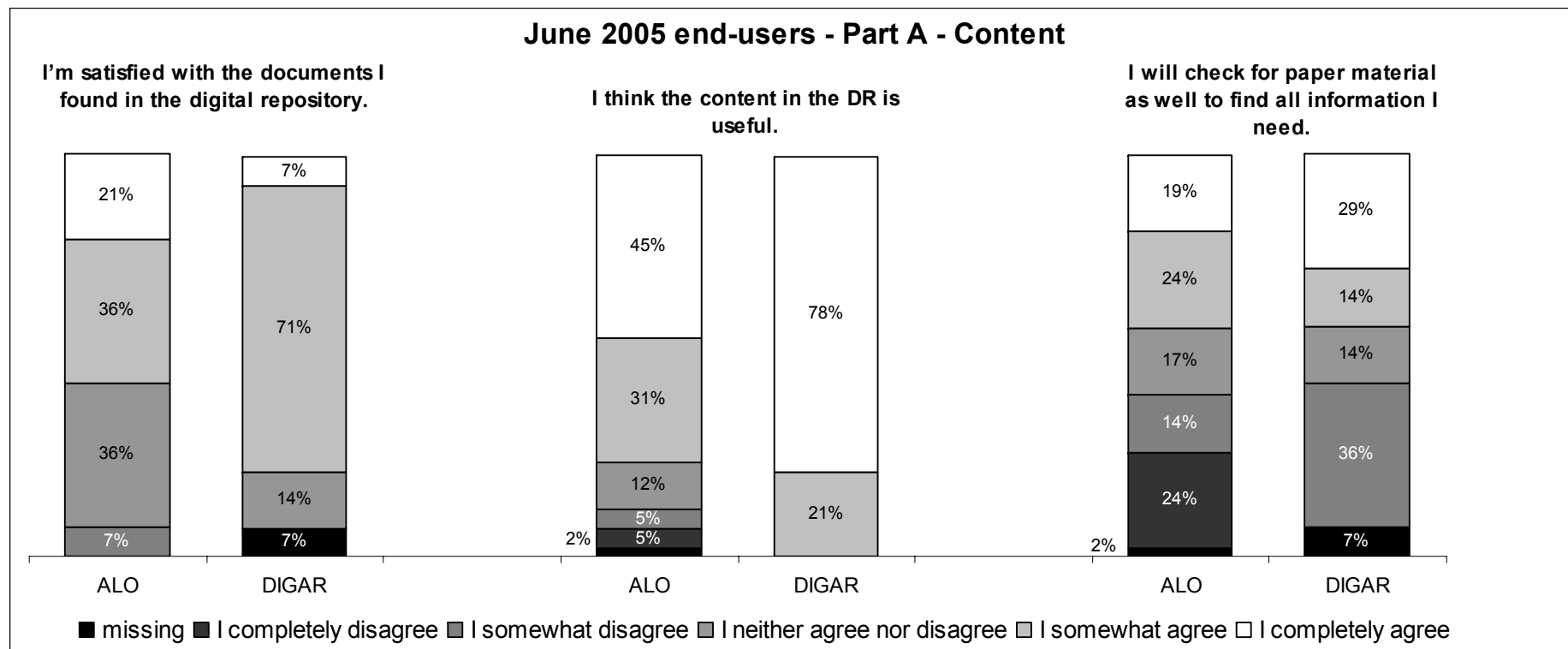


Figure 18: Comparison of results from: Content among ALO and DIGAR end-users in the 1st evaluation period (June 2005)

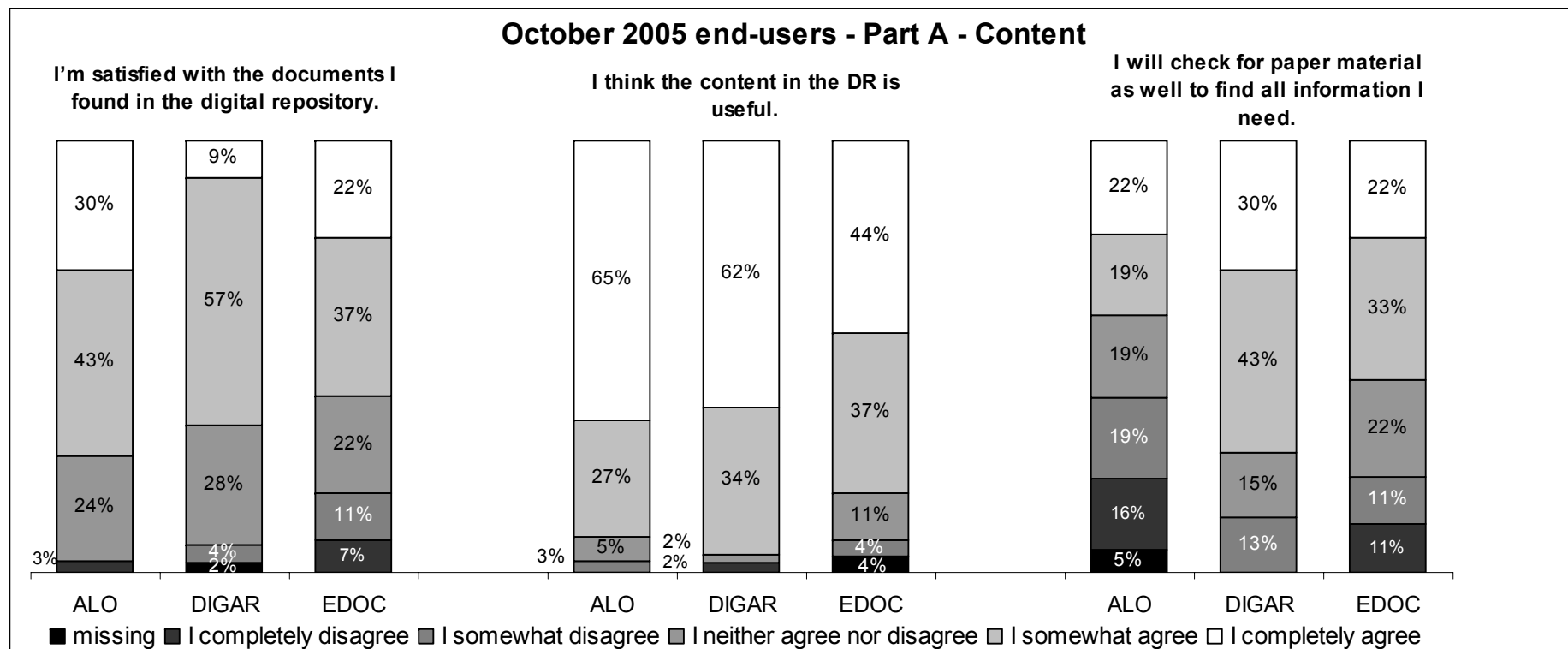


Figure 19: Comparison of results from: Content among ALO and DIGAR end-users in the 2nd evaluation period (October 2005)

3.4.3 RECOGNITION

In the next series of questions we analyzed end-users' immediate recognition of the different elements and understandability, suitability and functionality of the interface of a certain digital repository.

We had to take into consideration that in EDOC there were two end-users who didn't agree with or remained undecided about all statements from this point on. One was already mentioned above; the other one commented that there were too many questions that discouraged him; he answered most of them with a score of 5, which apologised for. There was another EDOC end-user who also chose negative answers or many times remained undecided; he commented that the organization (repository) was good but not very usual. He wished to have the possibility of sorting the results by subject for the qualification papers as well. We also assume that one or two ALO end-users used the opposite criteria in the 1st evaluation, but since they didn't comment on it, we considered only their responses.

In the 1st evaluation phase many of the ALO end-users (47%) claimed that there were too many different steps in dealing with the given task. Cross tabulation research reflects almost no significant difference according to experience, age and use habits of the individual respondents. See **Figures 20, 21, and 22** below.

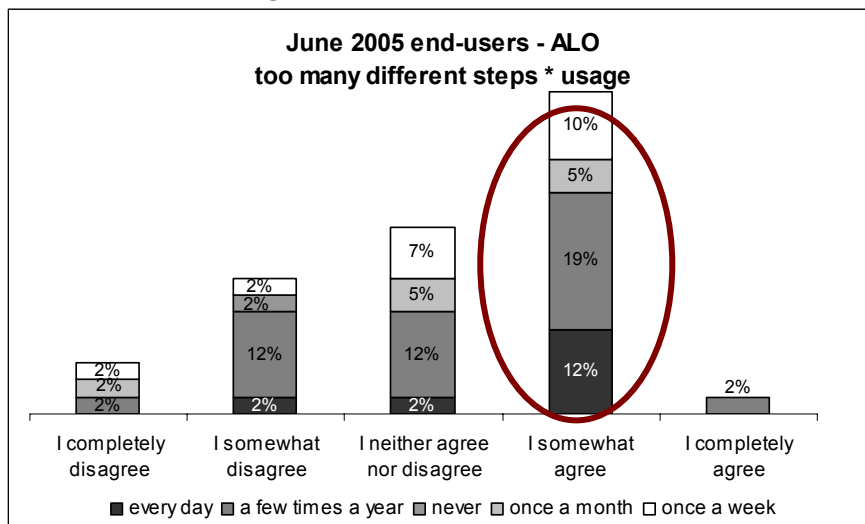


Figure 20: Correlation (*) between long procedure before finding and opening a certain document and usage among ALO end-users (June 2005)

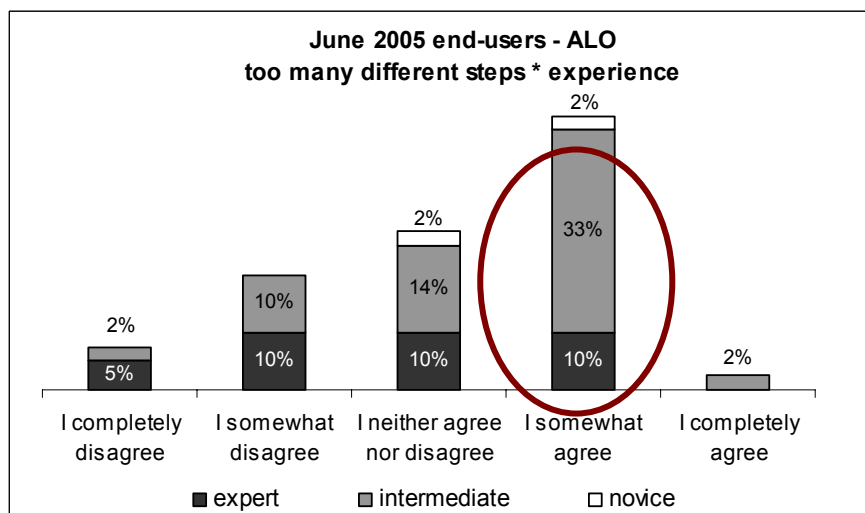


Figure 21: Correlation (*) between long procedure before finding and opening a certain document and experience among ALO end-users (June 2005)

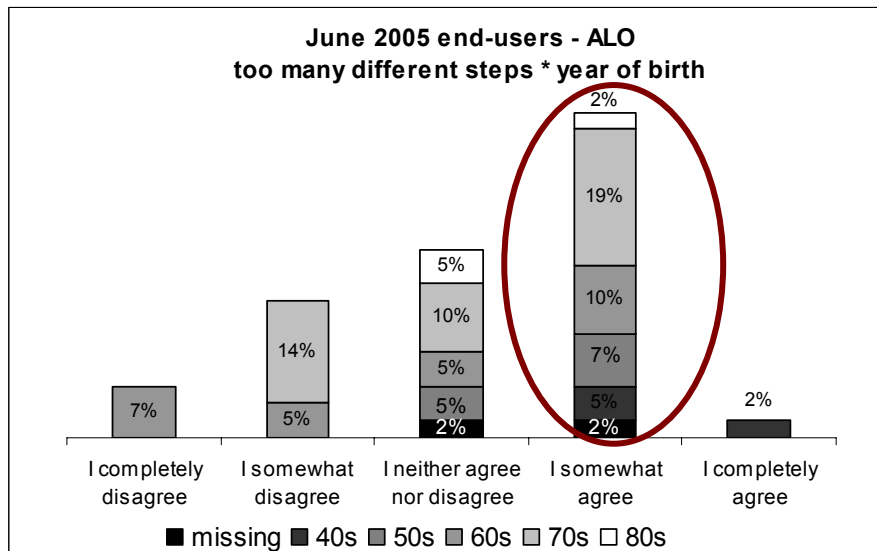


Figure 22: Correlation (*) between long procedure before finding and opening a certain document and year of birth among ALO end-users (June 2005)

The ALO results were more positive in the 2nd evaluation phase, but at this time we found out that also several DIGAR and EDOC end-users complained about the long procedure before finding and opening a certain document. A few emphasized this even in the comments, which are presented under the section: Personal and subjective perception.

Approximately half of the users agreed that the digital repository was well suited to their work. However, there are still a few exceptions in DIGAR and EDOC (already mentioned in the beginning of this section).

In the 1st evaluation period the majority of the respondents from ALO and DIGAR repositories found enough information on a given screen and it was easy for them to move back, forth and between different screens. However, there were two end-users in ALO who completely disagreed with this – they were both students, born in the 80s, who visited the repository a few times a year and one was an expert while the other's level was intermediate. Both also expressed that the repository was not well suited for their work. Unfortunately, they didn't write any comments.

In the 2nd evaluation period the results were a bit different: there were some end-users – especially from EDOC and DIGAR - who were not satisfied and obviously had problems with interface and with moving between different screens. In EDOC one claimed that the small Popup window for the bibliographic data was very bad, since forward and back buttons were missing. DIGAR end-users wished to be able to search by subject (keyword) as well and wanted a better developed search engine. They had difficulties finding the right keywords for successful search. One claimed that he could not limit his search to the obtained search results and two of them claimed that search by author was unsuccessful and limited. All details are presented in **Figures 23 and 24**.

The recognition section shows that all repositories could be improved in order to be more user-friendly, simple and efficient. At the moment, most of the end-users are able to find the information they are searching for, but are not completely satisfied with the design, structure, search engine and information on the repository interface. Implementers should try to simplify the interface in order to decrease the number of necessary "clicks and steps" while the given screen should provide only the necessary amount of information. It would be wise to enable an option for users' ability to adjust the display for the information they need.

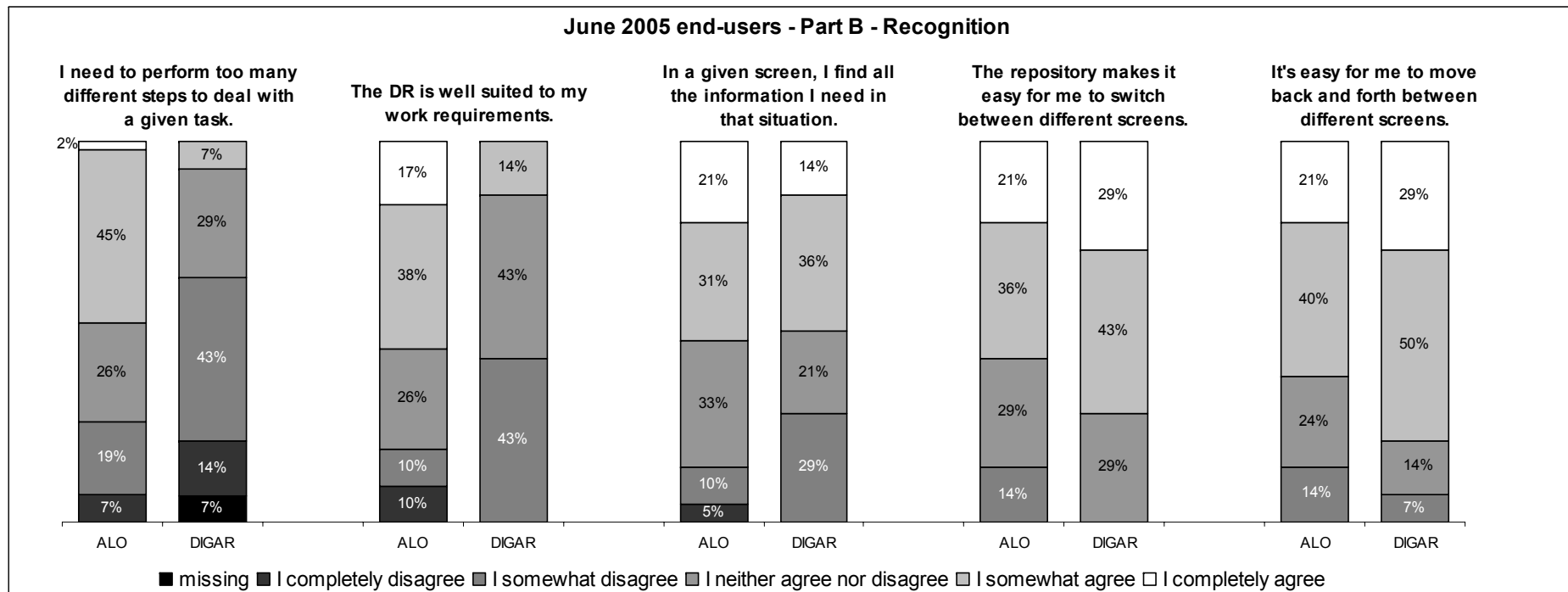


Figure 23: Comparison of results from part: Recognition among ALO and DIGAR end-users in the 1st evaluation period (June 2005)

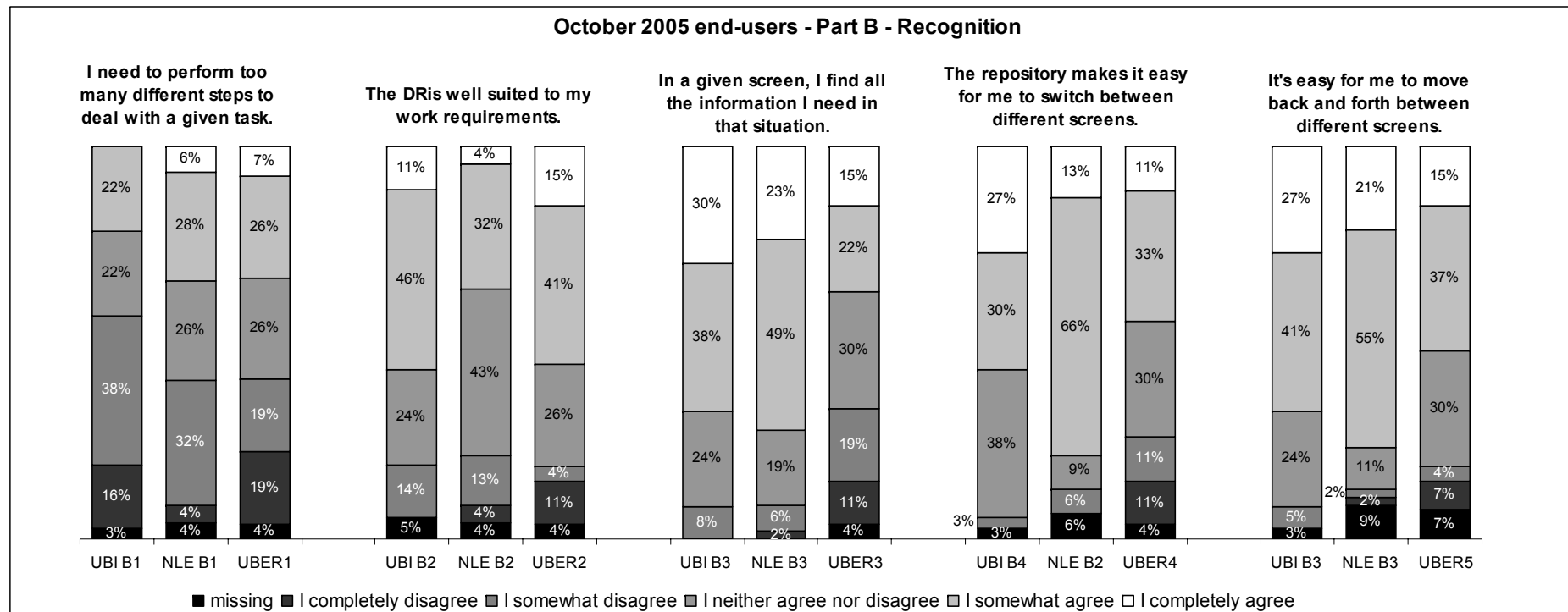


Figure 24: Comparison of results from part: Recognition among ALO, DIGAR and EDOC end-users in the 2nd evaluation period (October in 2005)

3.4.4 PERFORMANCE AND EFFICIENCY

In this section we focused on performance and efficiency. We were interested in the content accessibility, the rate of success in achieving the tasks, the digital repository productivity and whether the value-added services were useful to the users. We also wanted to find out whether the requested items opened quickly enough and the amount of time end-users averagely need to find proper information.

In both evaluation phases several end-users, especially from ALO and DIGAR, stressed that there was no help-system available; many others, mostly from EDOC, remained undecided. Technical evaluation shows that there is no help-system available in the ALO repository. The DIGAR repository provides a short help information saying that users can use a simple or a complex search system, that either small or big letters can be chosen and that truncations are possible. At http://EDOC.hu-berlin.de/e_info_en/ EDOC provides basic information and assistance on navigating the EDOC repository, viewing and printing PDF documents and searching. In the 2nd evaluation phase DIGAR end-users suggested improving the help system (the help system could be more detailed – how to search by author, subject, index, etc.) and making it available in both, Estonian and Russian.

With the exception of some end-users from EDOC (mentioned in the previous chapter), most of the respondents didn't have problems with the repository terms and concepts and understood the meaning of the on-screen message. However, many of them didn't know their exact location in the repository in any given moment – see details in **Figures 25 and 26**. It might be good for repositories to visually mark the current entry location (and hits as well) by highlights or contrasting colour.

According to the majority of end-users of all three repositories, the collections are easily accessible. But it is surprisingly that end-users (especially in ALO) expressed little interest in the usefulness of value-added service. It is possible they are not fully aware of the advantages of value-added services, but on the other hand, evaluated repositories do not offer many of them at the moment. ALO offers "digitization", but this can be more beneficial to content providers than to the end-users.

Nevertheless, some DIGAR end-users and especially majority of the EDOC end-users are aware of the usefulness of value-added services but are not really willing to pay for them. It is encouraging that DIGAR end-users are more willing to pay for the services as in the 1st evaluation phase in June (21% of users were undecided and 42% declared that they were not prepared to pay extra). DIGAR so far doesn't offer any value-added services. In the future they will make sure that the Digital archive by DIGAR content is accessible to people with disabilities and will adapt the archive to WAI guidelines, making files available to screen-readers.

EDOC already offers the following value-added services:

- different types of hosting,
- print-on-demand service (ProPrint),
- technology, software and know-how transfer

For detailed results see **Figures 27 and 28**.

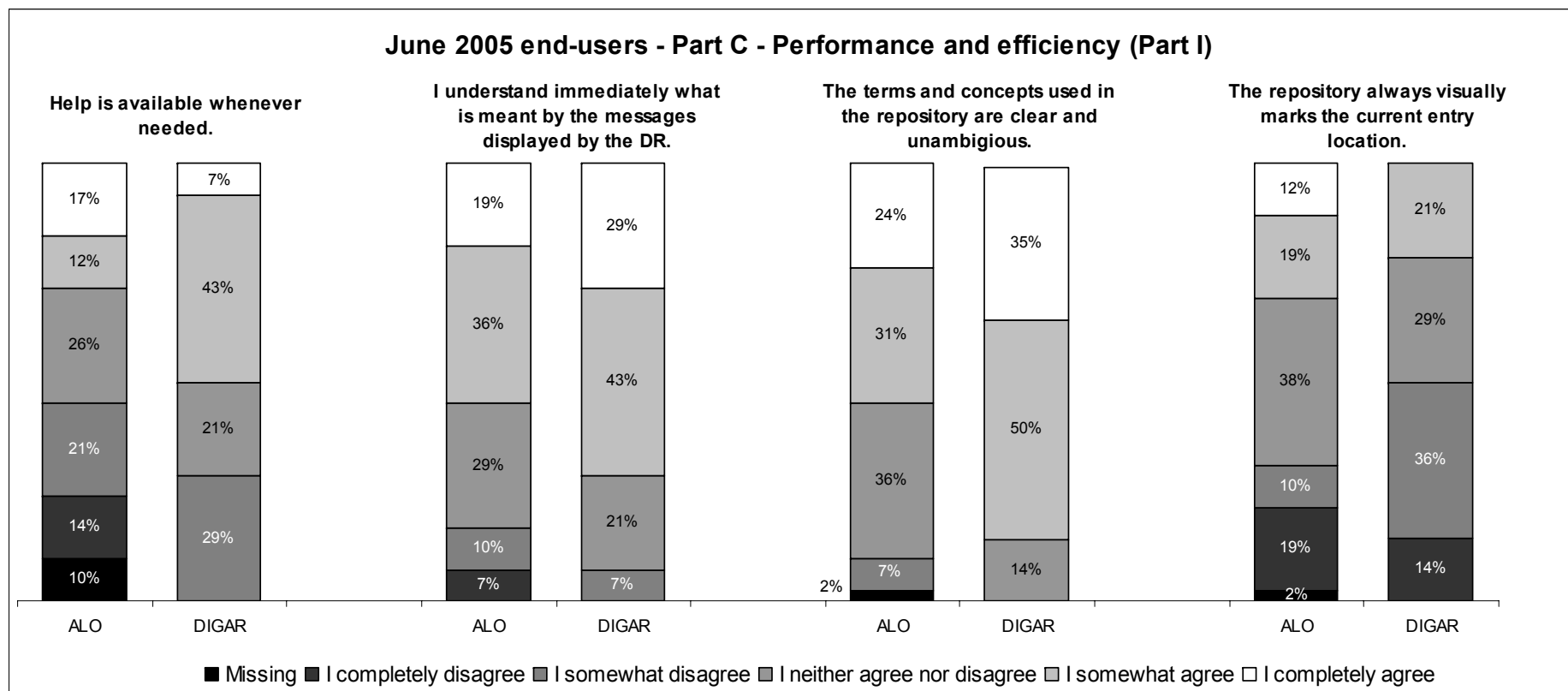


Figure 25: Comparison of results from part: Performance and efficiency (part I) among ALO, DIGAR and EDOC end-users in the 1st evaluation period (June 2005)

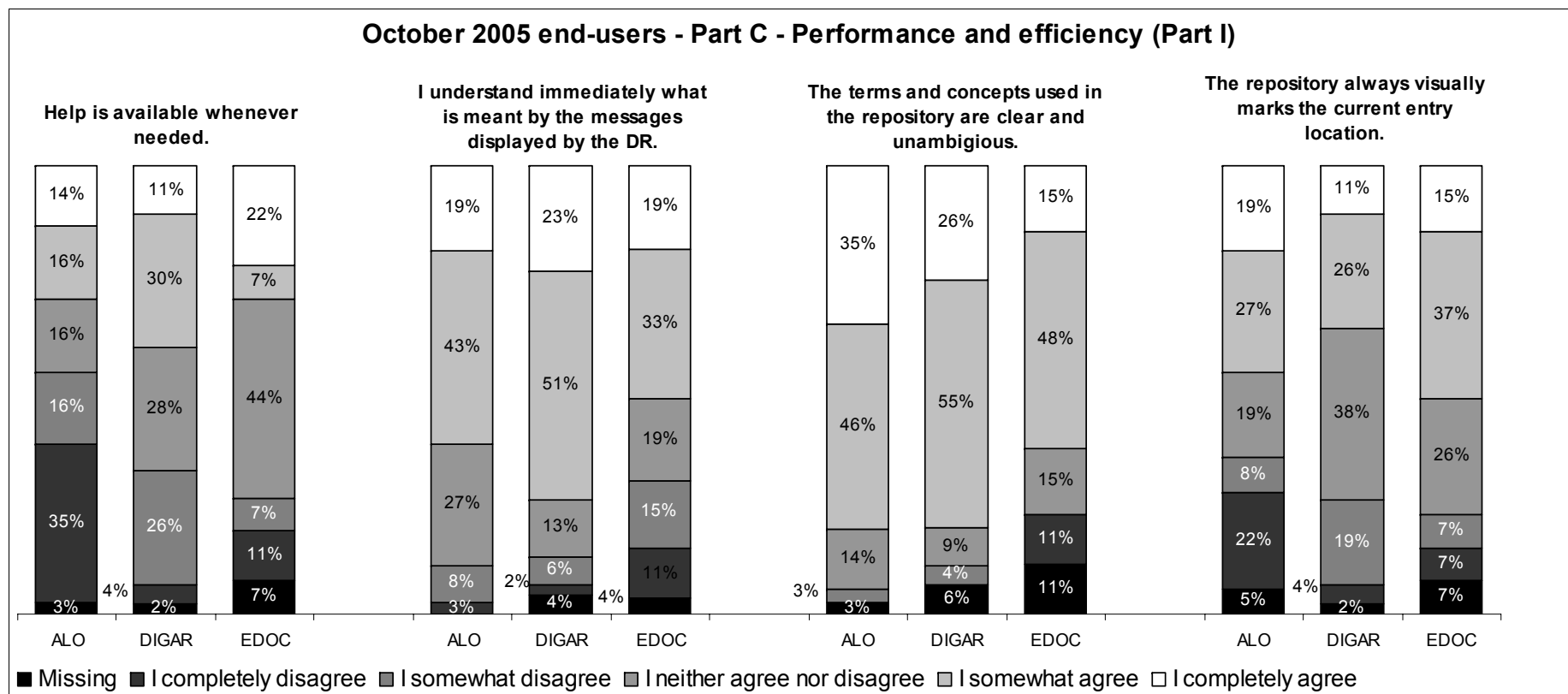


Figure 26: Comparison of results from part: Performance and efficiency (part I) among ALO, DIGAR and EDOC end-users in the 2nd evaluation period (October 2005)

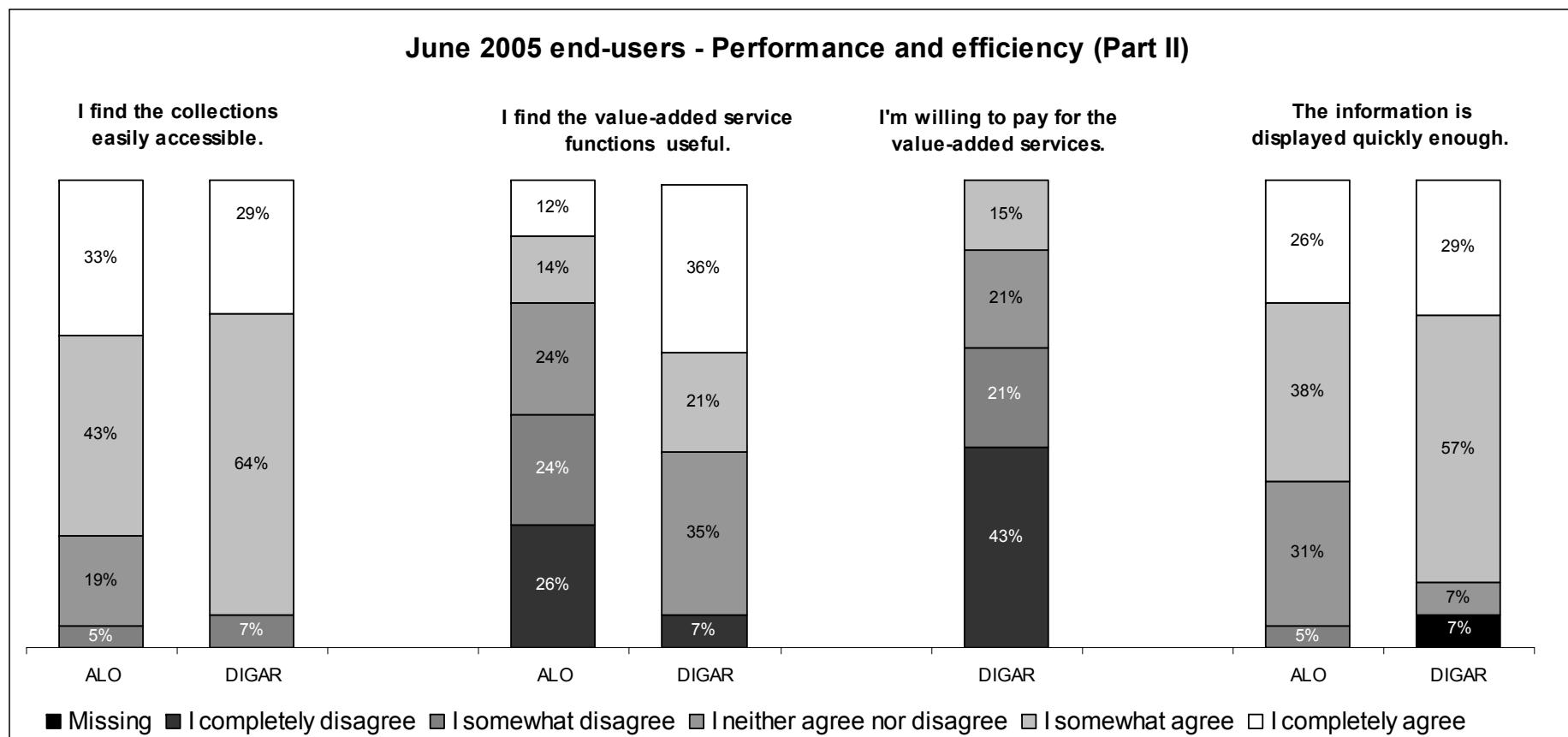


Figure 27: Comparison of results from part: Performance and efficiency (part II) among ALO, DIGAR and EDOC end-users in the 1st evaluation period (June 2005)

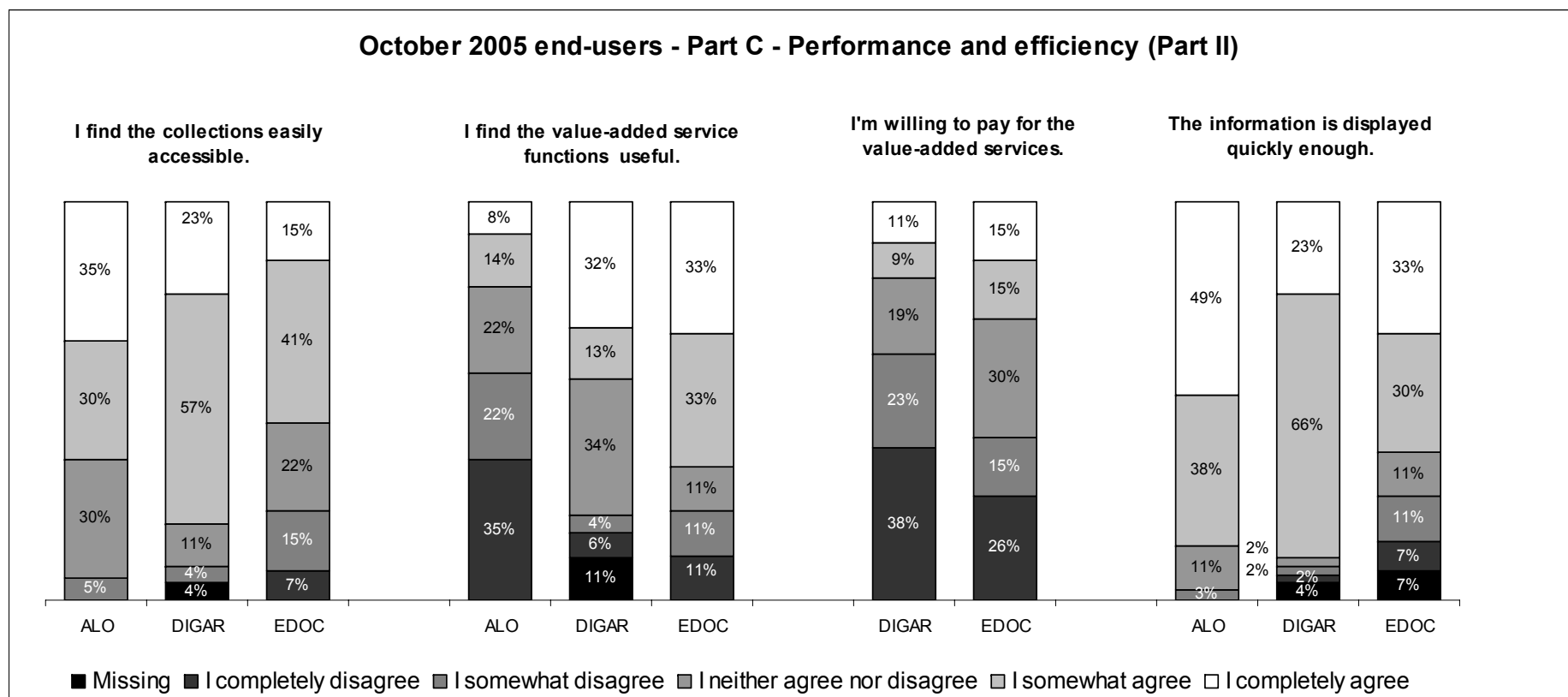


Figure 28: Comparison of results from part: Performance and efficiency (part II) among ALO, DIGAR and EDOC end-users in the 2nd evaluation period (October 2005)

Raising the awareness and demonstrating the advantages offered by value-added services would for definitely pay the effort and would be worth trying. It could at least essentially improve the repository, if not also benefit the organization. Organizations could offer several kinds of value-added services, depending on the users' needs and the nature of the repository (national/university/institutional).

Many of respondents agree that information in the repository is displayed quickly enough.

Figure 29 shows that 5-10 seconds on average (in both) repositories were needed to get the requested items In the June evaluation. Some end-users needed less, others more. Of course, documents with pictures or scanned books take more time than simple digital text.

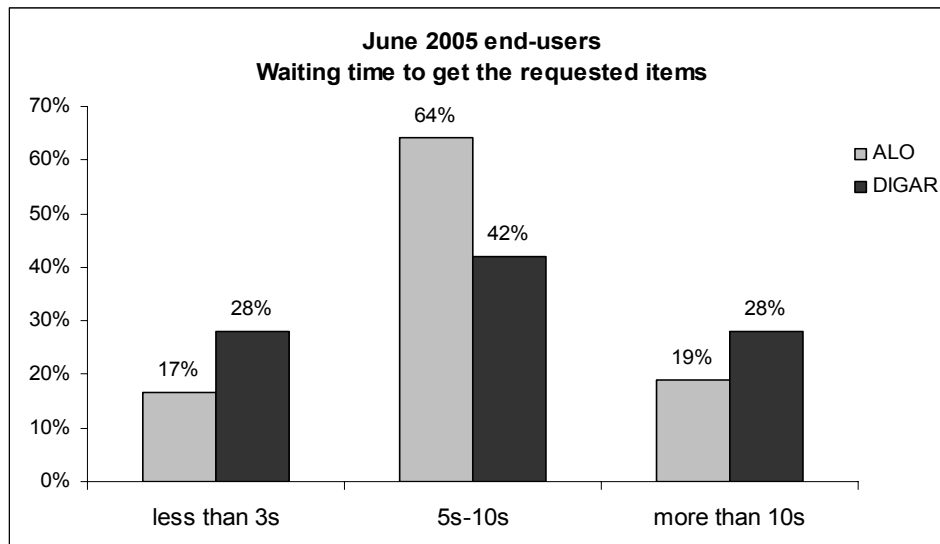


Figure 29: The amount of time needed to get the requested item (1st evaluation period, June 2005)

In October the results were similar, most of the users needed 5 to 10 seconds. Comparison showed that in October DIGAR and ALO end-users waited less time than in June. No one complained about the waiting time being too long or about there being problems with the opening of the documents (see **Figure 30**).

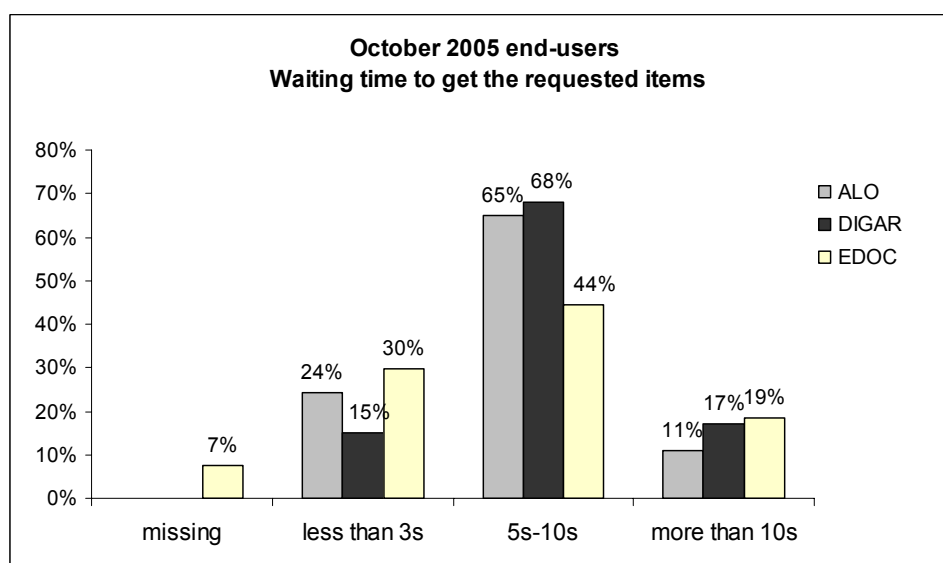


Figure 30: The amount of time needed to get the requested item (2nd evaluation period, October 2005)

We also wanted to know the amount of time needed to find the proper information. In June most of ALO end-users needed around 5 minutes, while others needed more – from 5 to 15 minutes, but no one claimed to need more than 15 minutes. In DIGAR 35% of end-users needed less than 3 minutes, 29% needed 5 minutes and 36% end-users remained undecided. However, one user claimed that might need 60 minutes for a complete search. For details see **Figures 31 and 32**.

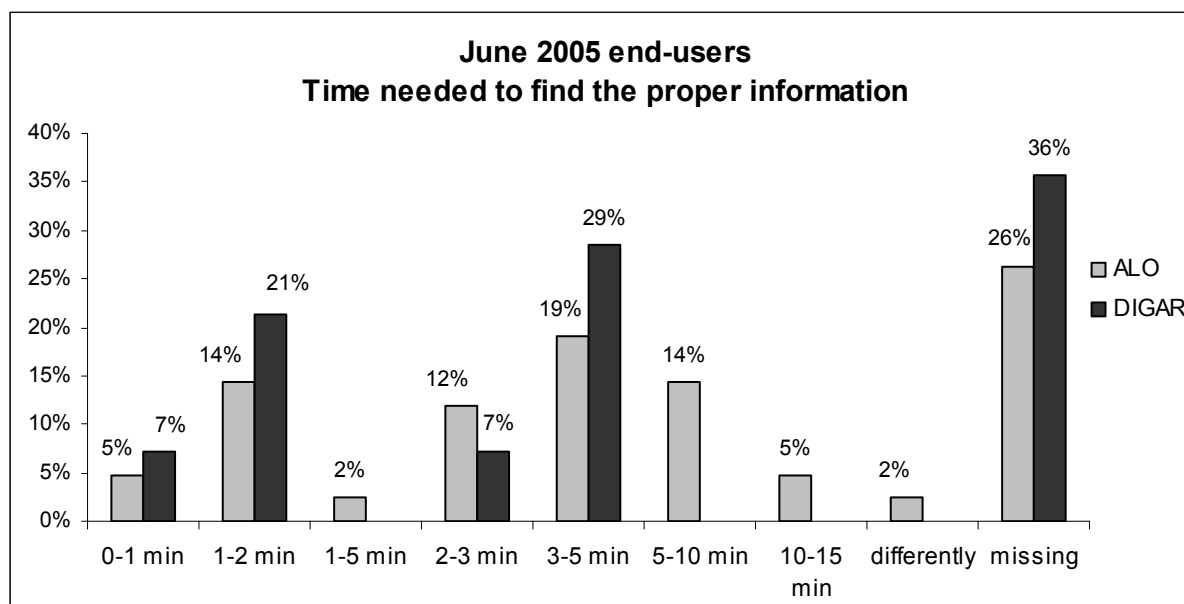


Figure 31: The amount of time needed to find the proper information (1st evaluation period, June 2005)

Unfortunately, we didn't receive any data from EDOC about this question and so a detailed analysis was not possible. October results were similar to June results; however, we could see that 2 users (a scientist and a librarian) from ALO needed as much as 45 minutes to find the proper information, but they didn't wrote any comments.

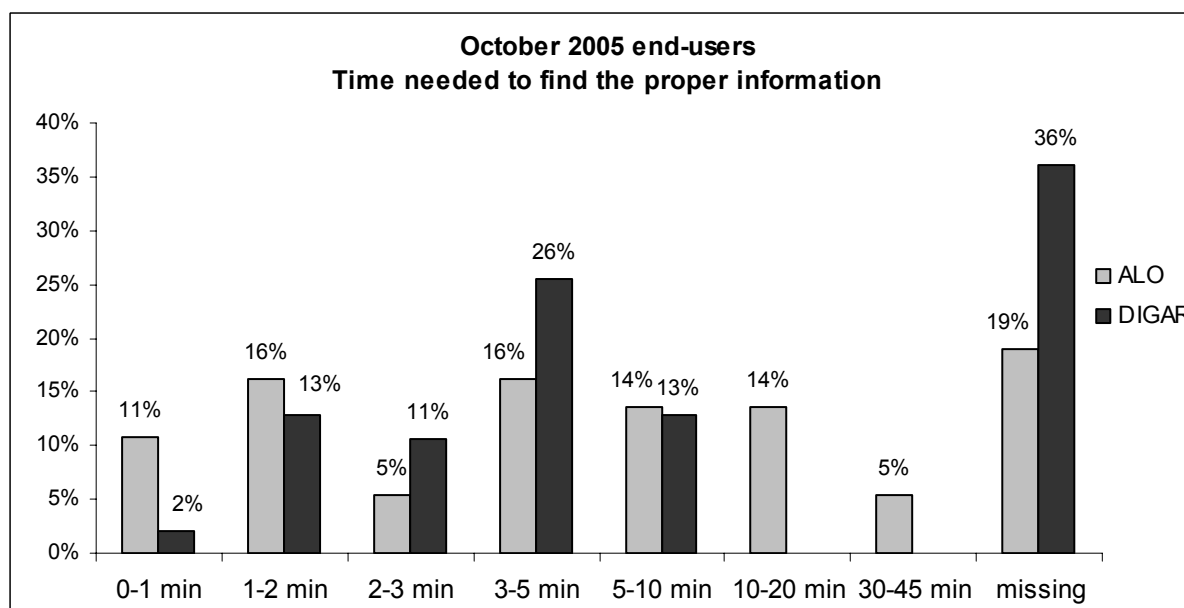


Figure 32: The amount of time needed to find the proper information (2nd evaluation period, October 2005)

3.4.5 PERSONAL AND SUBJECTIVE PERCEPTION

This section deals with the end-users' subjective perception of the interface, services, system,...etc. Generally we were interested in their personal opinion about the digital repository as a whole. The opinion reflects the level of motivation and expectations a user has from a given system. The response rate for these auto-reflexive questions is much lower among EDOC end-users than among the other two repositories' users.

First of all, we would like to stress that many respondents had very positive comments about repositories and were very satisfied to be able to use it. In this section we will try to present an overall picture of their opinions, wishes, eventual disagreements and comments.

It is clear that (in both periods) all end-users found the structure easy to use. However, in October there appeared a suggestion for DIGAR's improvement; an end-user wished to have a list of the whole content arranged by subject (of all archived documents) on the interface. A similar situation occurred in EDOC, where one end-user suggested sorting qualification papers via subject and the other suggested that it would be nice to have a list of the recent documents acquired in the last month. Perhaps the structure of the content at ALO could be designed more simply and clearly (**Figures 33 and 34**).

Although some of the end-users didn't complain of having problems with the security system and logging, many of them didn't express their feelings about it (technical data show us that there is no limitation access in any repository). As before, end-users complained about the numerous details and steps needed in order to use the repository properly (especially in EDOC).

Many end-users from all repositories were satisfied and familiar with the designations such as icons, signs, etc.... They were content with the messages and repository texts and liked the overall website design. Most of them were so satisfied with the repository that they intend to come back when in need of further information. The end-users are at all times aware of their precise location of the repository website

It is interesting how ALO, DIGAR and EDOC end-users differ in getting used to a certain repository. It looks like DIGAR end-users spent less time learning about the repository use than ALO and EDOC users. Either the former is more simple to use or it is that DIGAR offers more training courses.

Particularly DIGAR end-users agreed that the website design was nice and pleasant and didn't have problems in communicating with the repository. In ALO there was one blind end-user, who was disappointed because there were only a few OCR scanned text in the digital repository. More comments about both repositories will be described at the end of this section (**Figures 35, 36, 37 and 38**).

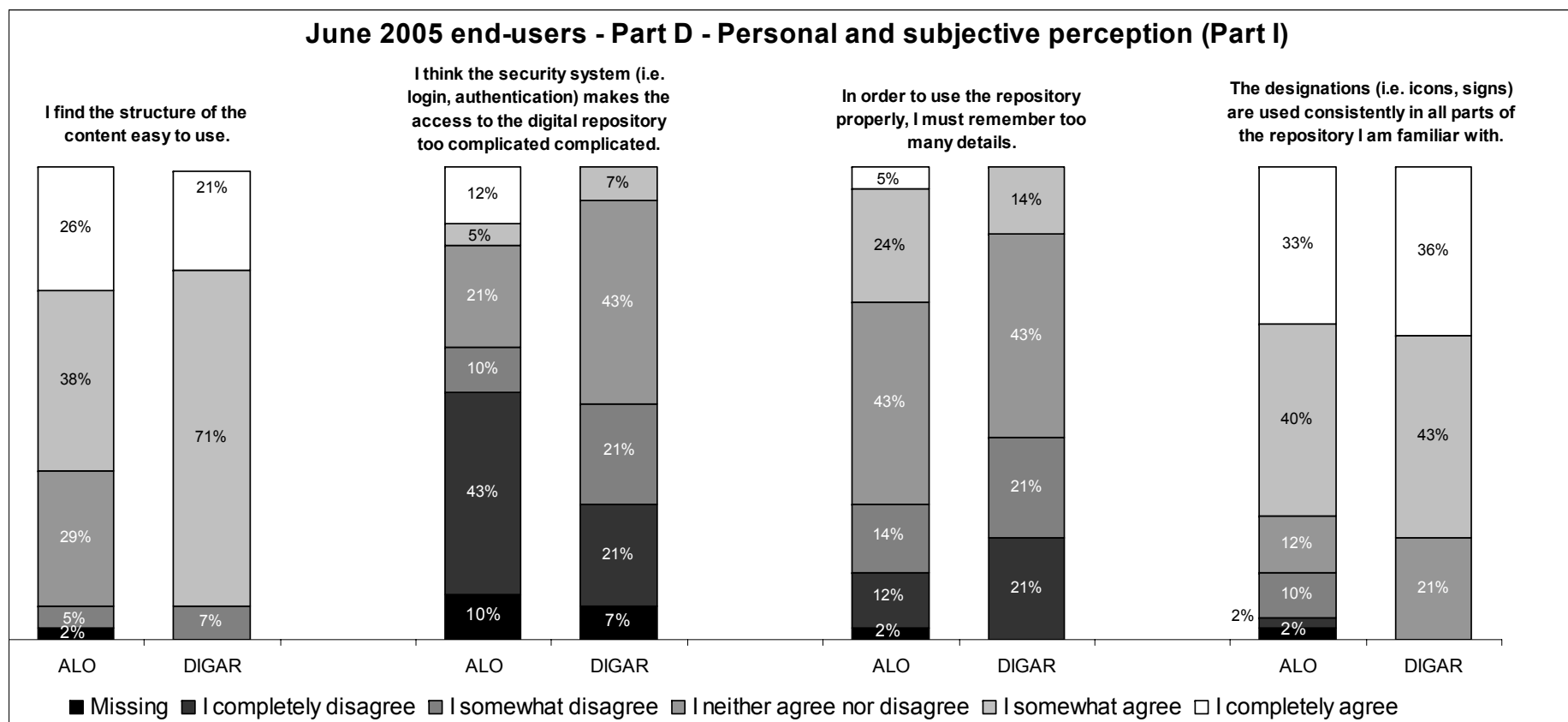


Figure 33: Comparison of results from part: Personal and subjective perception (part I) among ALO and DIGAR end-users in the 1st evaluation period (June 2005)

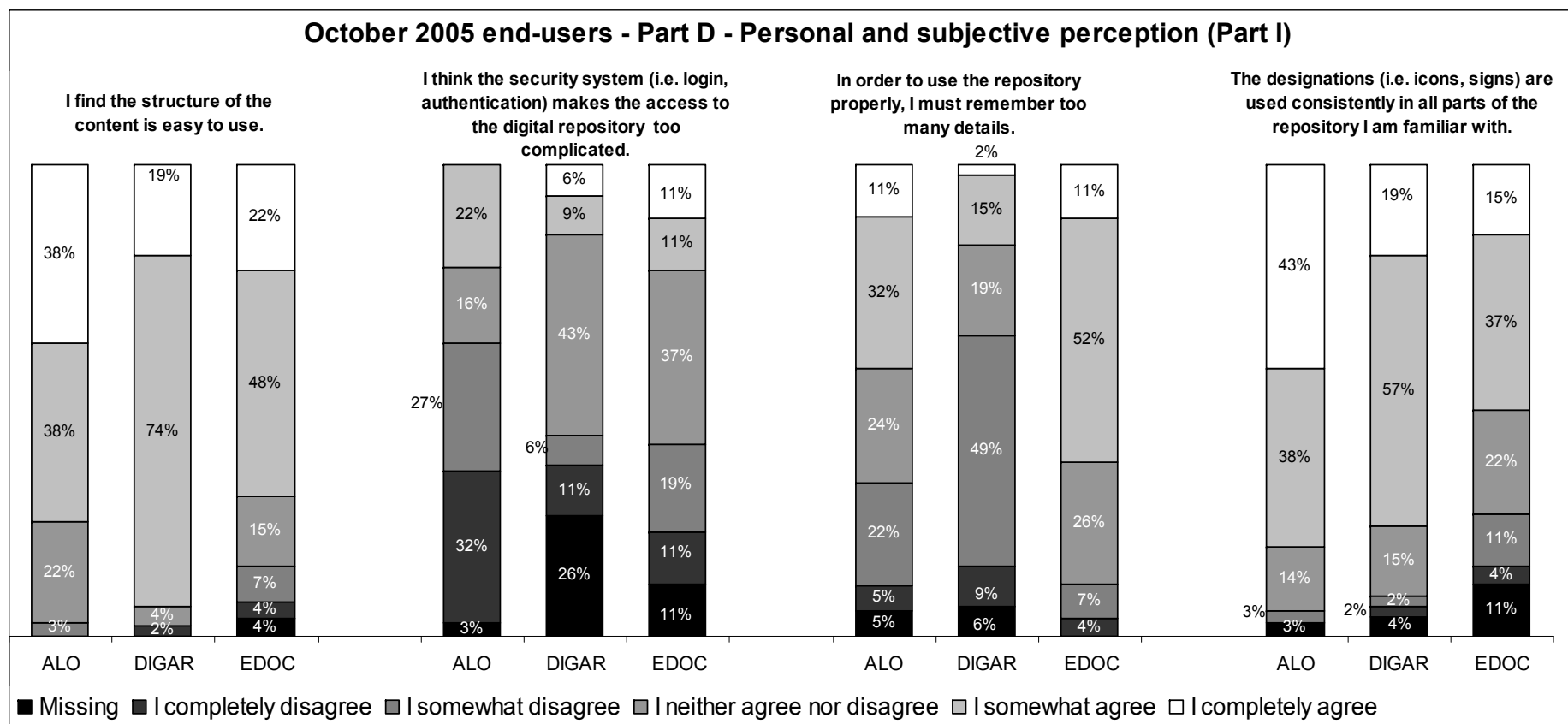


Figure 34: Comparison of results from part: Personal and subjective perception (part I) among ALO and DIGAR end-users in the 2nd evaluation period (October 2005)

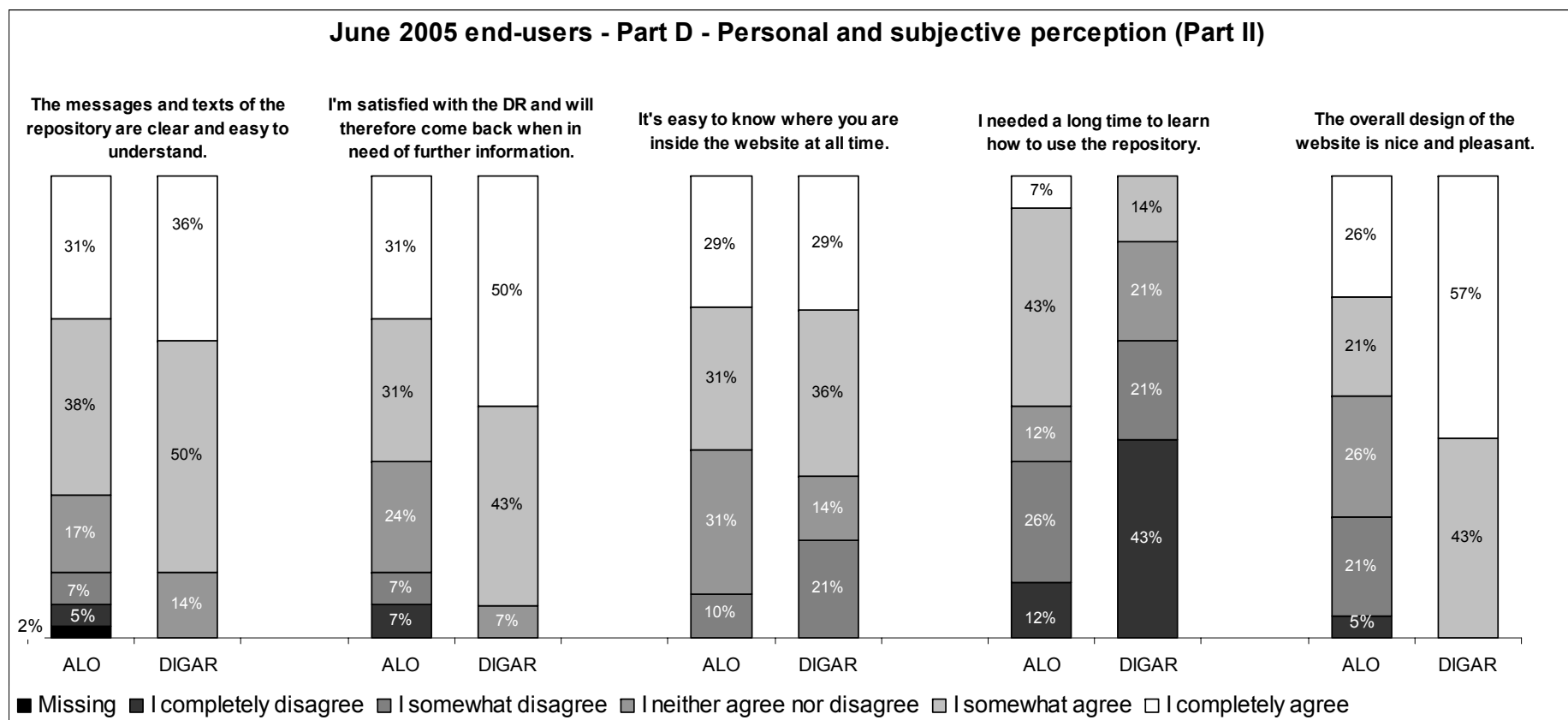


Figure 35: Comparison of results from part: Personal and subjective perception (part II) among ALO and DIGAR end-users in the 1st evaluation period (June 2005)

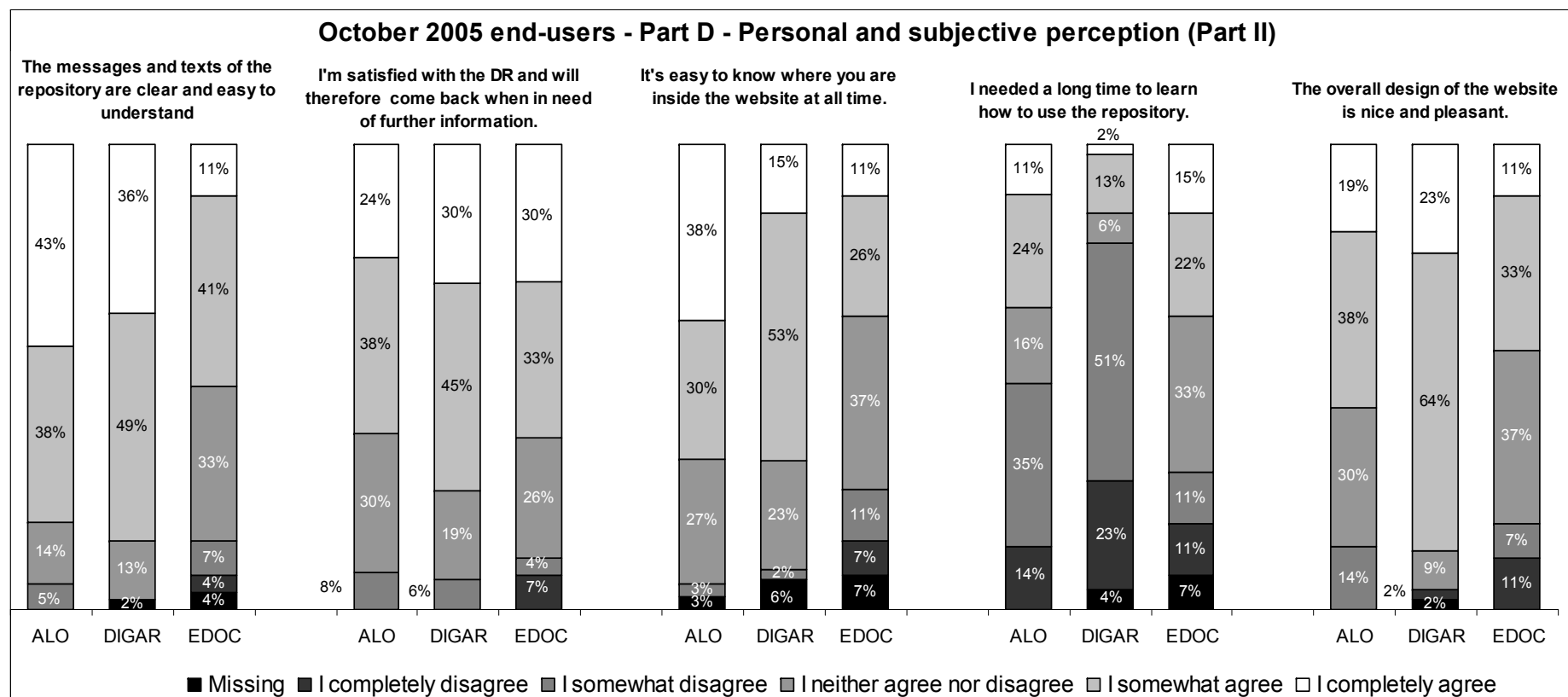


Figure 36: Comparison of results from part: Personal and subjective perception (part II) among ALO and DIGAR end-users in the 2nd evaluation period (October 2005)

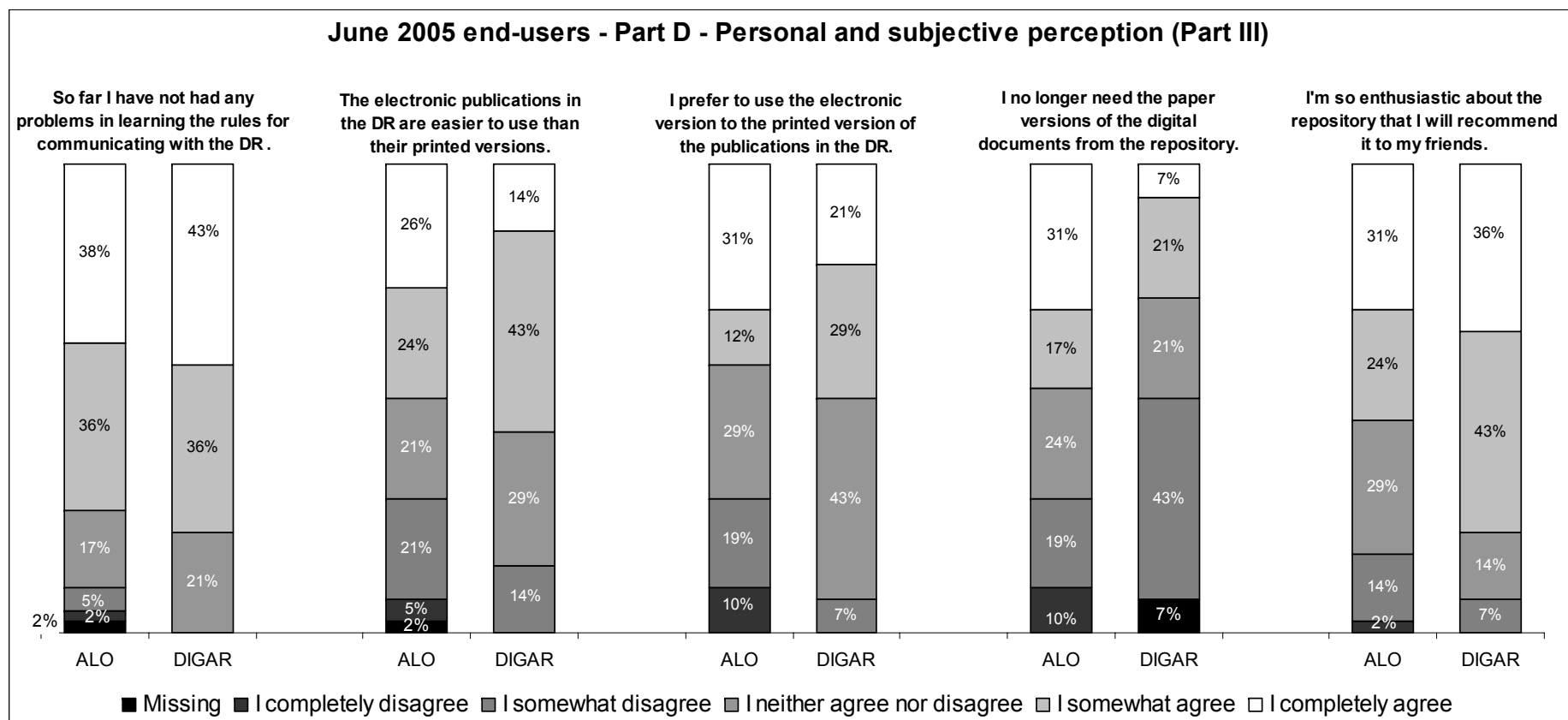


Figure 37: Comparison of results in part Personal and subjective perception (part III) among ALO and DIGAR end-users in the 1st evaluation period (June 2005)

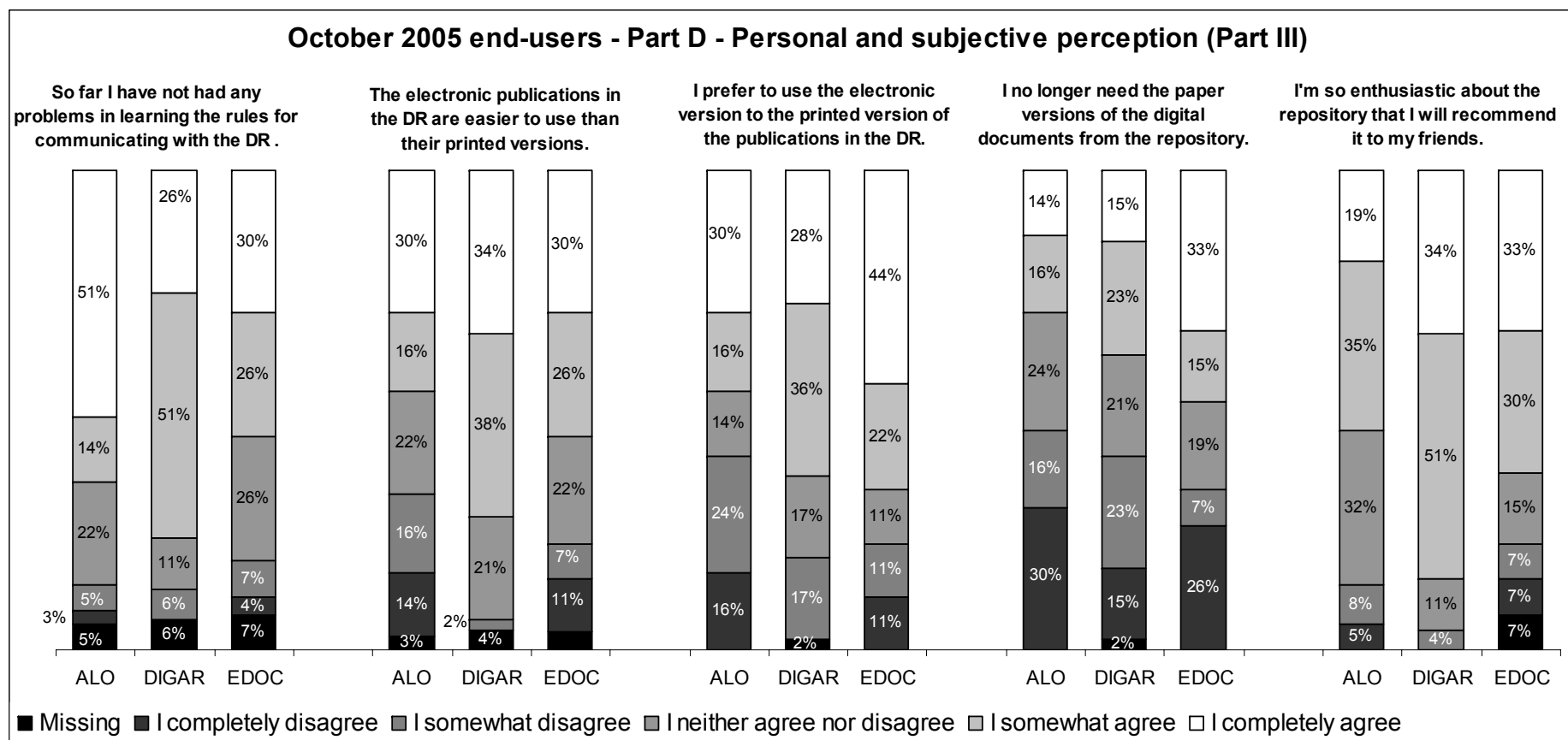


Figure 38: Comparison of results from part: Personal and subjective perception (part III) among ALO and DIGAR end-users in the 2nd evaluation period (October 2005)

We also wanted to know what end-users thought about e-documents and whether they preferred e-documents rather than paper ones. By comparing both evaluations, it became evident and encouraging as well, that in the last survey, which took part in October, more end-users agreed that the e-documents were easier to use than the paper ones. However, the users with a background in human sciences or architecture were not so enthusiastic about the e-documents. Correlation with the age of the respondents shows that older users (born in the 40s and 50s) and surprisingly also younger (born in the 80s), don't think that e-version is easier to use than the printed one. We got very similar results when we asked if they preferred e-version to printed material. Why? Because it is hard to read long texts on the screen? An important factor might also be that e-documents with full text and printed material do not rule out one another, but are rather two types of output for different use cases.

Regarding the statement that paper version will no longer be needed the end-user perception was quite similar. The major exception could be found among the DIGAR end-users born in the 70s, which is maybe due to the fact that end-users from the National library were more heterogeneous and didn't use material only for studies and research. See details in **Figures 39, 40 and 41**.

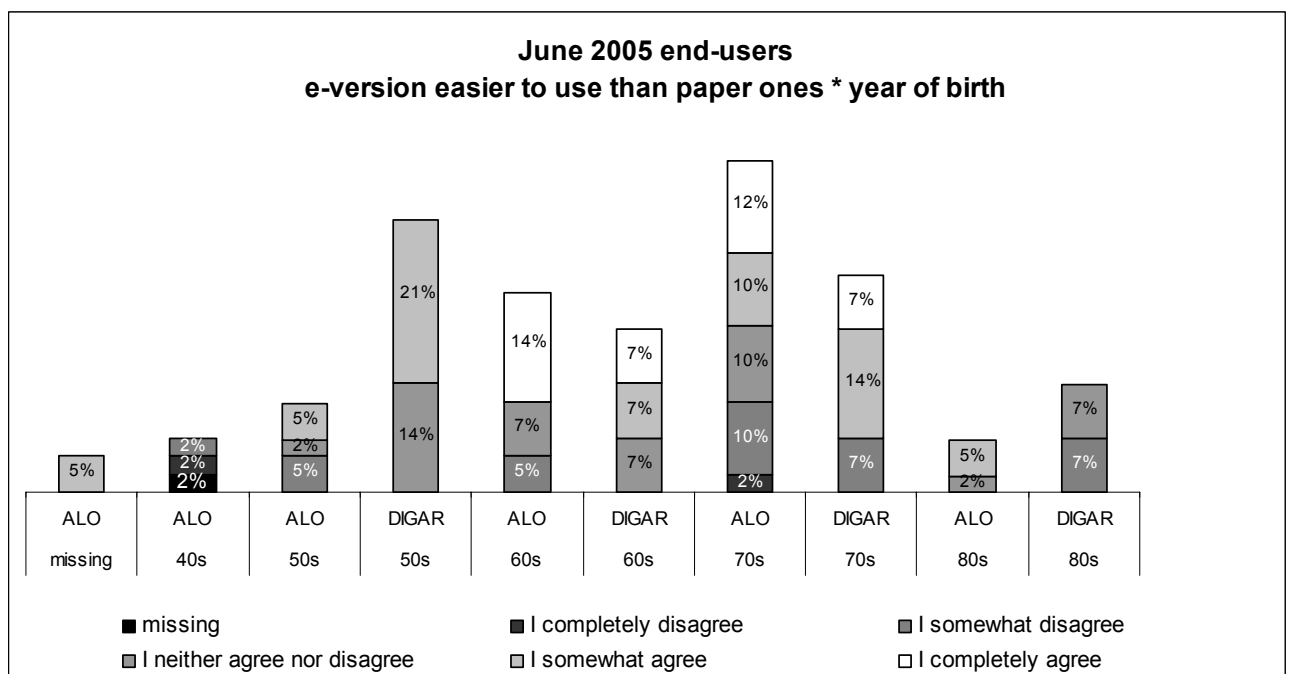


Figure 39: Correlation (*) between use of e-documents and age from the 1st evaluation period (June 2005)

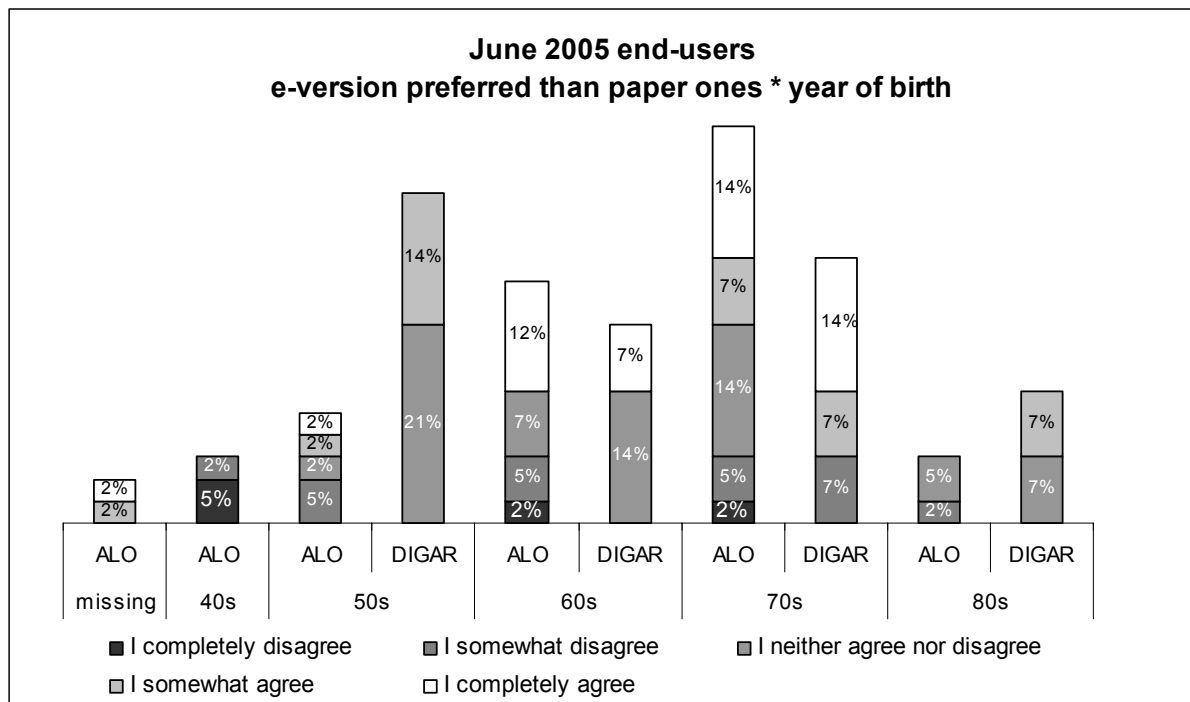


Figure 40: Correlation (*) between preferences of e-documents and age from the 1st evaluation period (June 2005)

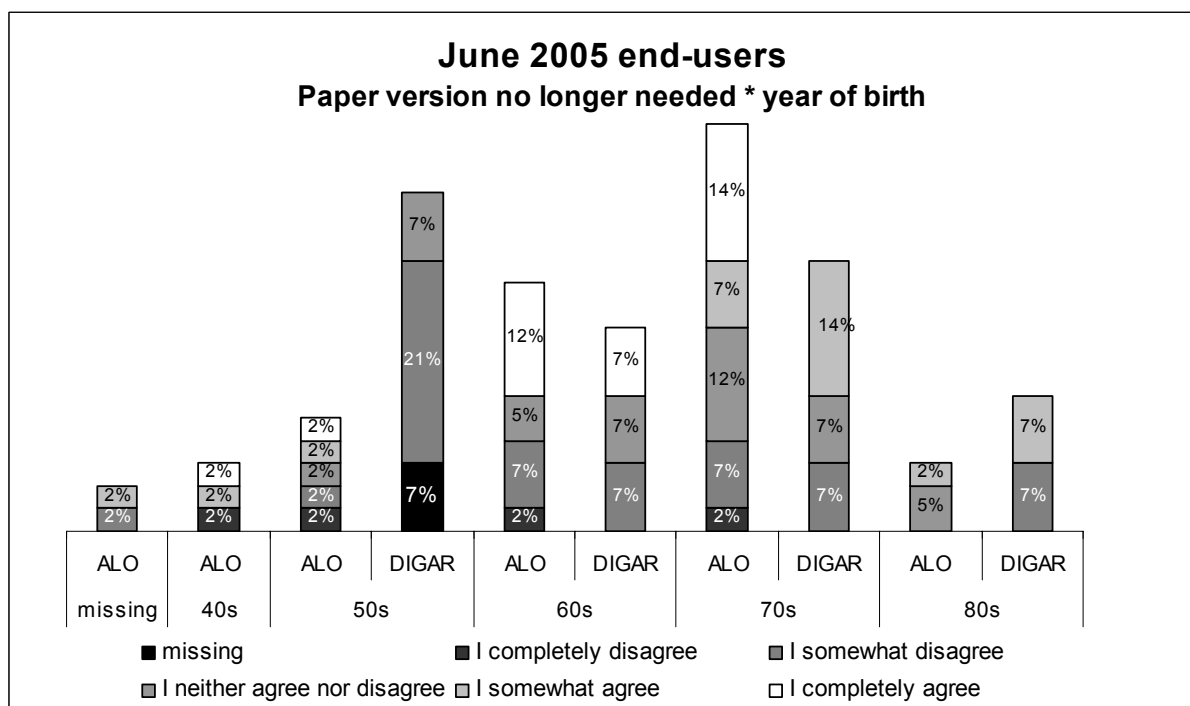


Figure 41: Correlation (*) between future use of e-documents and age from the 1st evaluation period (June 2005)

The last statement was meant more for fun but it shows that there are quite a lot of end-users – more in DIGAR than in ALO - who are still so enthusiastic about the repository that they would recommend it to their friends.

At the end of this section end-users could express their comments. A lot of them were happy to have a digital repository (at least the idea of it) and to use it at their work or in studies, but

they wished to find more objects from different topics. Some comments for improvement were already mentioned in previous chapters.

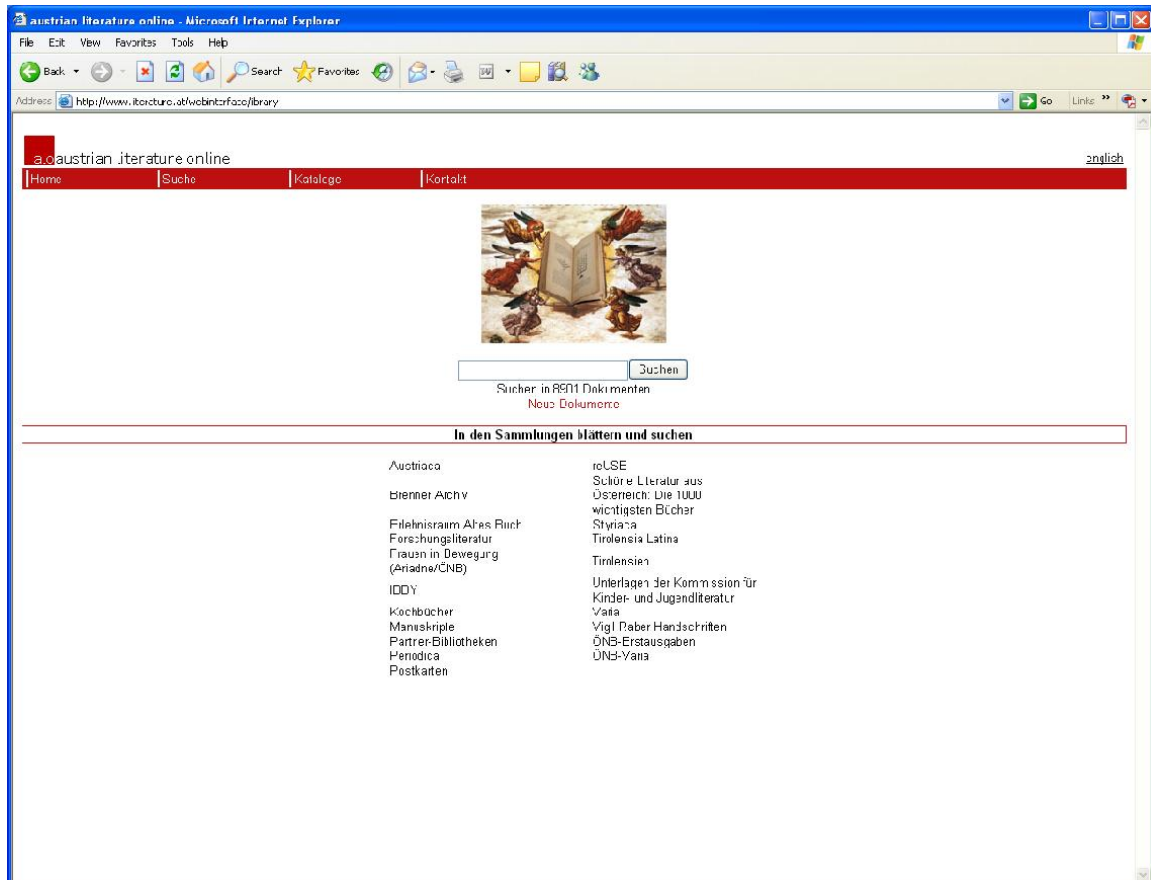


Figure 42: Interface of the ALO repository.

A few respondents from different universities in Austria claimed that the repository was nice and practical but that it was not promoted enough; especially in the University of Graz there were complaints about ALO repository not being known well among users (the ALO repository interface is shown in **Figure 42**). It is interesting that in ALO end-users, who wished the repository included the Bible, books and old literature, were not very satisfied with it. They missed more full text documents and complained that the texts and pictures were not very useful. The repository should also be improved for blind users. As mentioned in the section "*Personal and subjective perception*" the current repository's content is not very friendly to blind users. More OCR scanned text should be available. One or two end-user mentioned that, there were some layout problems with the screen and that the text was unmanageable. Is the layout suited for different size of screens? Some users from both evaluation periods complained that the searching criteria was inappropriate and that the keywords were not indicated within the hits. On the other hand, end-users generally appreciated that a lot of work and energy had been already put in the building of the repository and they are encouraging developers to continue with the project, especially with gathering full-text documents from different topics. The need for additional professional publications was also mentioned.

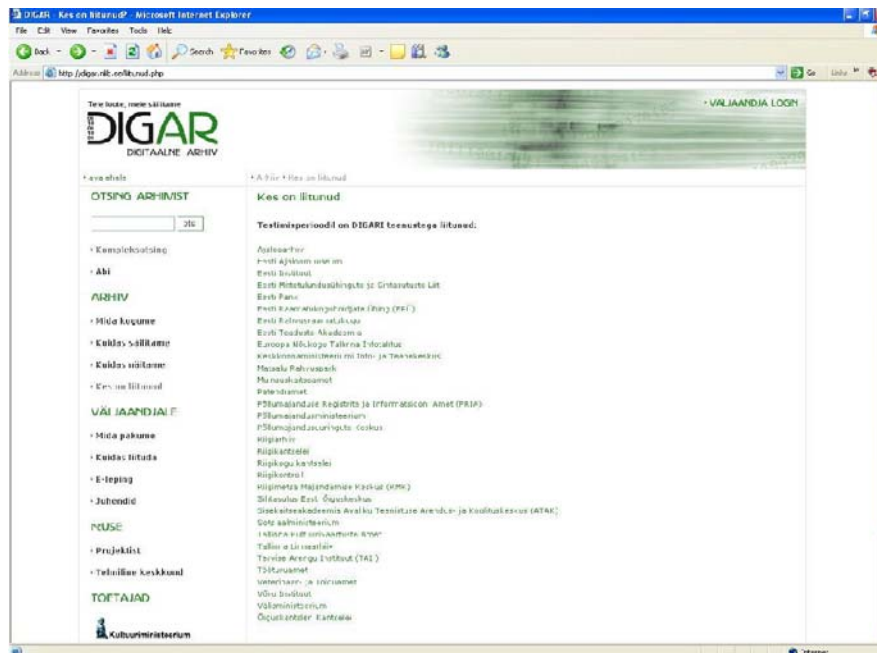


Figure 43: The interface of the DIGAR repository

DIGAR end-users are satisfied as well (the DIGAR repository interface is shown in **Figures 43 and 44**). They also wish to have more documents and one of them suggested that the publications on the repository home page should be linked to the publishers. It would be useful to have a list of subjects contained in the archive and the search and display of the results should be improved in order to be more accurate. The functionality of search by subject should be developed. A help function may be included in order to guide the users in their search by author, subject index, etc. Results could be sorted by date or by field. The text font on the archive home page should be larger. One respondent wished to have an e-mail notification on new acquisitions, but only on the chosen subjects. End-users from the National library also expressed the need to have more scientific literature (specially articles), which was published in Estonia.

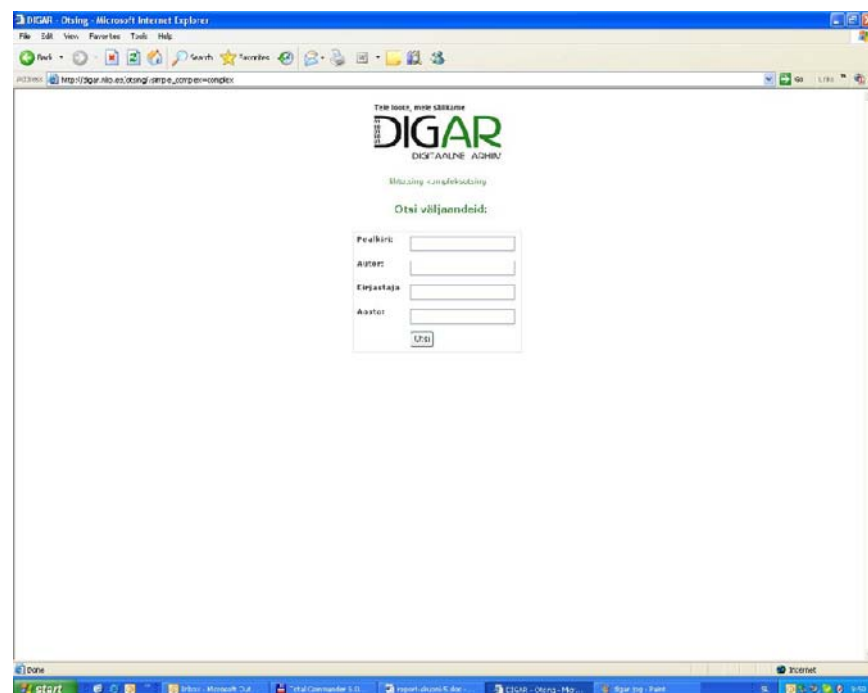


Figure 44: The searching engine at DIGAR digital repository

Some respondents from EDOC complained that it was very difficult to find the website (**Figure 45**); one wrote that it was very difficult to find it via the HU- and Charite-pages. Although there are quite a lot of digital documents in the repository, in order to make it even more usable, users wish to have access to more recent documents from different scientific fields (specifically reference to economy was made). It would be useful to ensure search functionality separately per subject as well as a list of the new documents of the last month. Some users would appreciate if the qualification/habilitation papers could be sorted by subject.



Figure 45: Interface of EDOC digital repository.

3.4.6 ERROR TOLERANCE AND SUITABILITY FOR INDIVIDUALIZATION

In the last set of questions we asked users about the repository error tolerance and their wish for individualization. Many respondents considered this topic to be of minor importance and the number of unanswered question in this section was very high.

We can notice differences between ALO, DIGAR and EDOC. In the 1st evaluation there were some ALO end-users who claimed that even small mistakes could cause serious consequences. However, it looks like ALO has already improved the repository and therefore no major claims appeared in the 2nd evaluation. There are, on the other hand some end-users from EDOC (30%), who have this kind of troubles. See **Figures 46 and 47**.

A lot of end-users from all evaluated repositories would like to be able to adjust the amount of on-screen information to their needs. They expressed this in their comments.

Regarding mistakes produced during their search process, a majority of end-users do not have any difficulty with restoring everything to its previous state or undoing the last operation. A few of them reported system errors while working with the repository, but they didn't write any comments.

According to the responses it is clear that some ALO and DIGAR end-users' entries have already caused errors. More importantly, there are some end-users (more in ALO than in DIGAR) who would like to be informed by e-mail about new entries in their digital repository while many end-users feels it is important that the same e-document can be found also in their library catalogue. Details can be found in **Figures 48 and 49**.

At present the linkage between the e-document and the e-catalogue of the library is offered only by DIGAR. If a record already exist in OPAC, they just add the link to DIGAR, and if not, they make a pre-record (with a link to DIGAR) using the information they received from the pre-print file. If the library receives a book after the link to DIGAR was established, the Cataloguing Department completes the record according to the edition in paper-form.

ALO is planning to link reUSE documents to the bibliographic records of the ALO library catalogue. EDOC plans to do the same, but has only new digitized books, which can be accessed via the library catalogue.

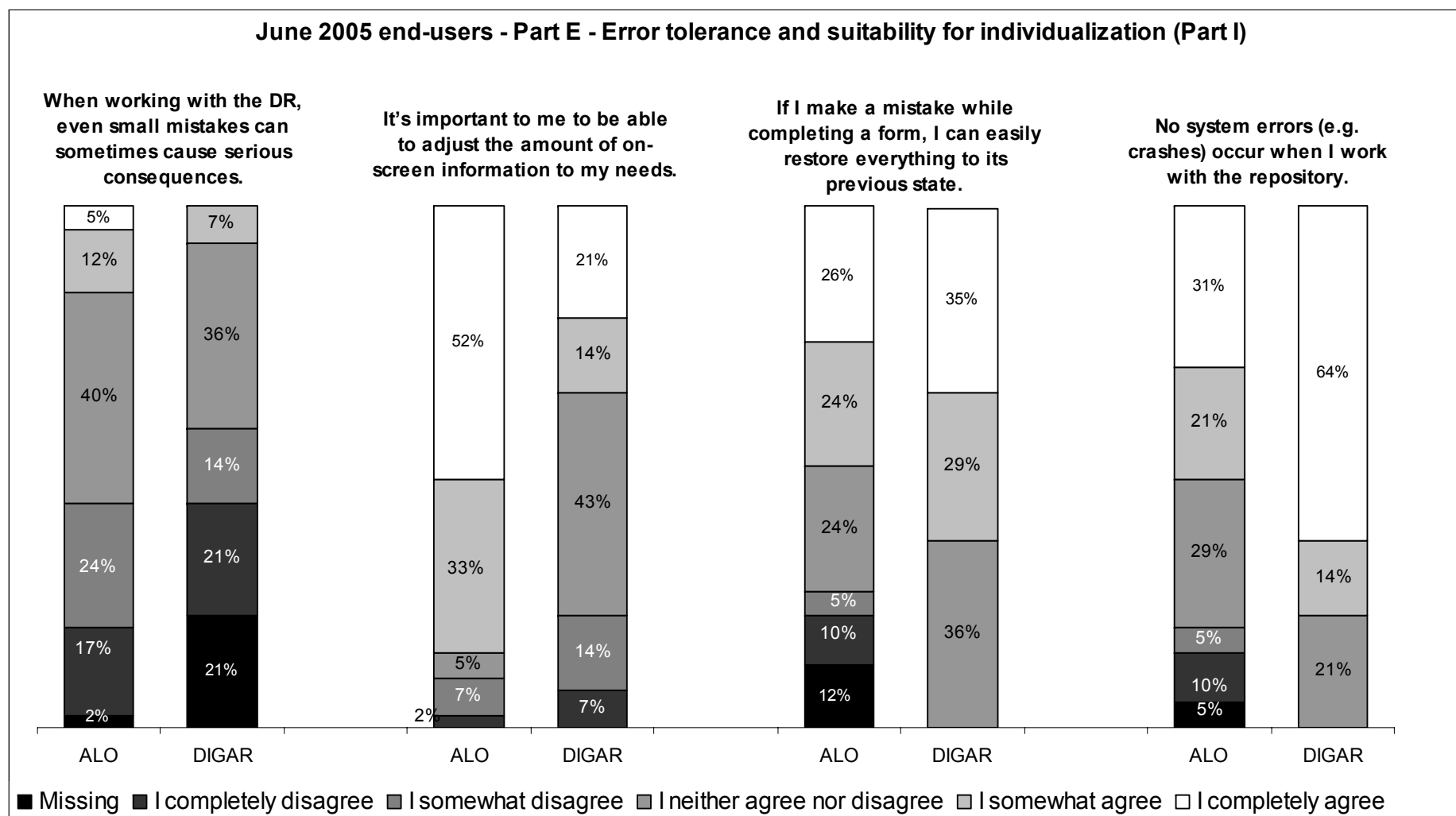


Figure 46: Comparison of results from part: Error tolerance and suitability for individualization (part I) among ALO and DIGAR end-users in the 1st evaluation period (June2005)

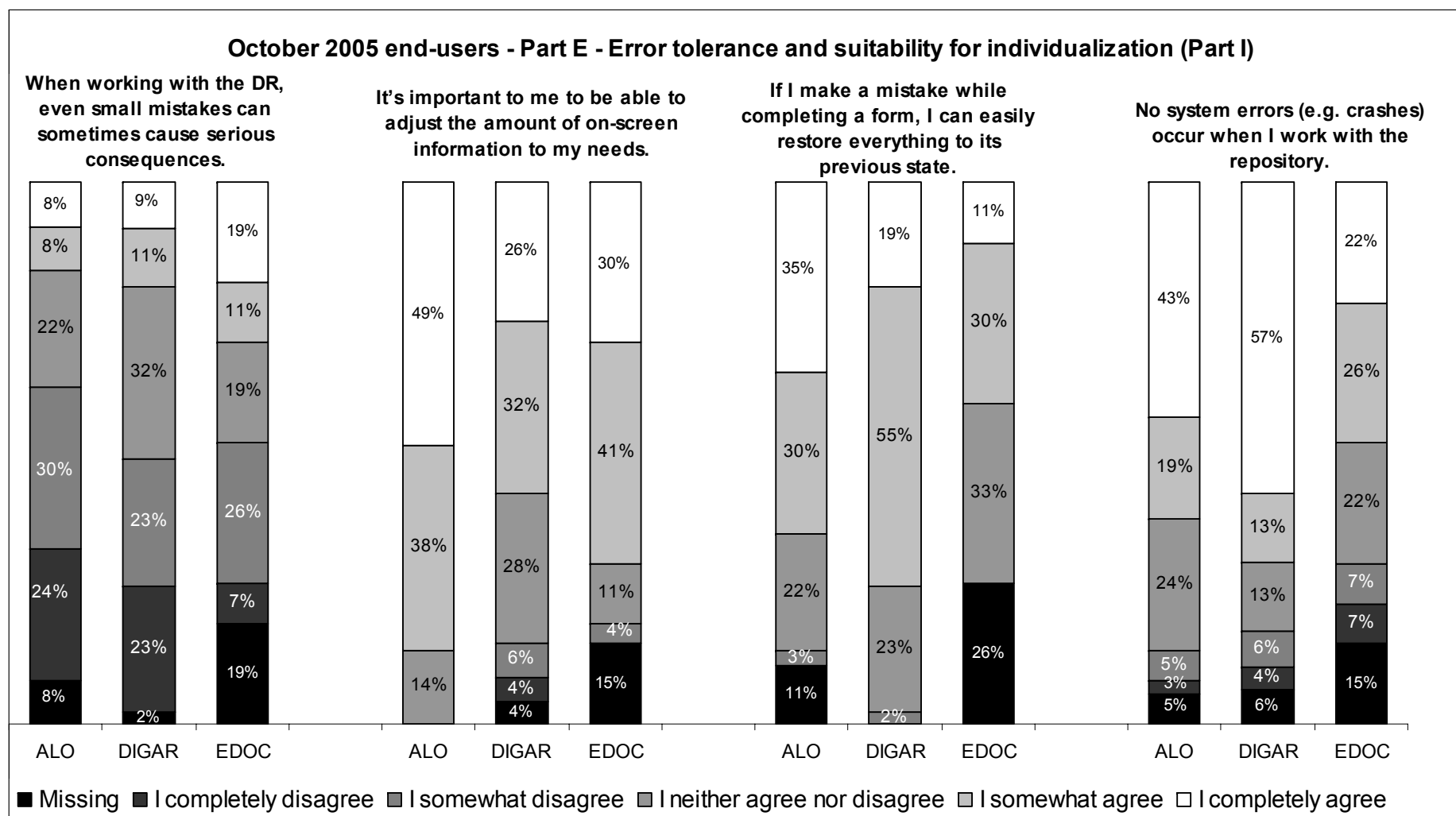


Figure 47: Comparison of results from part: Error tolerance and suitability for individualization (part I) among ALO and DIGAR end-users in the 2nd evaluation period (October 2005)

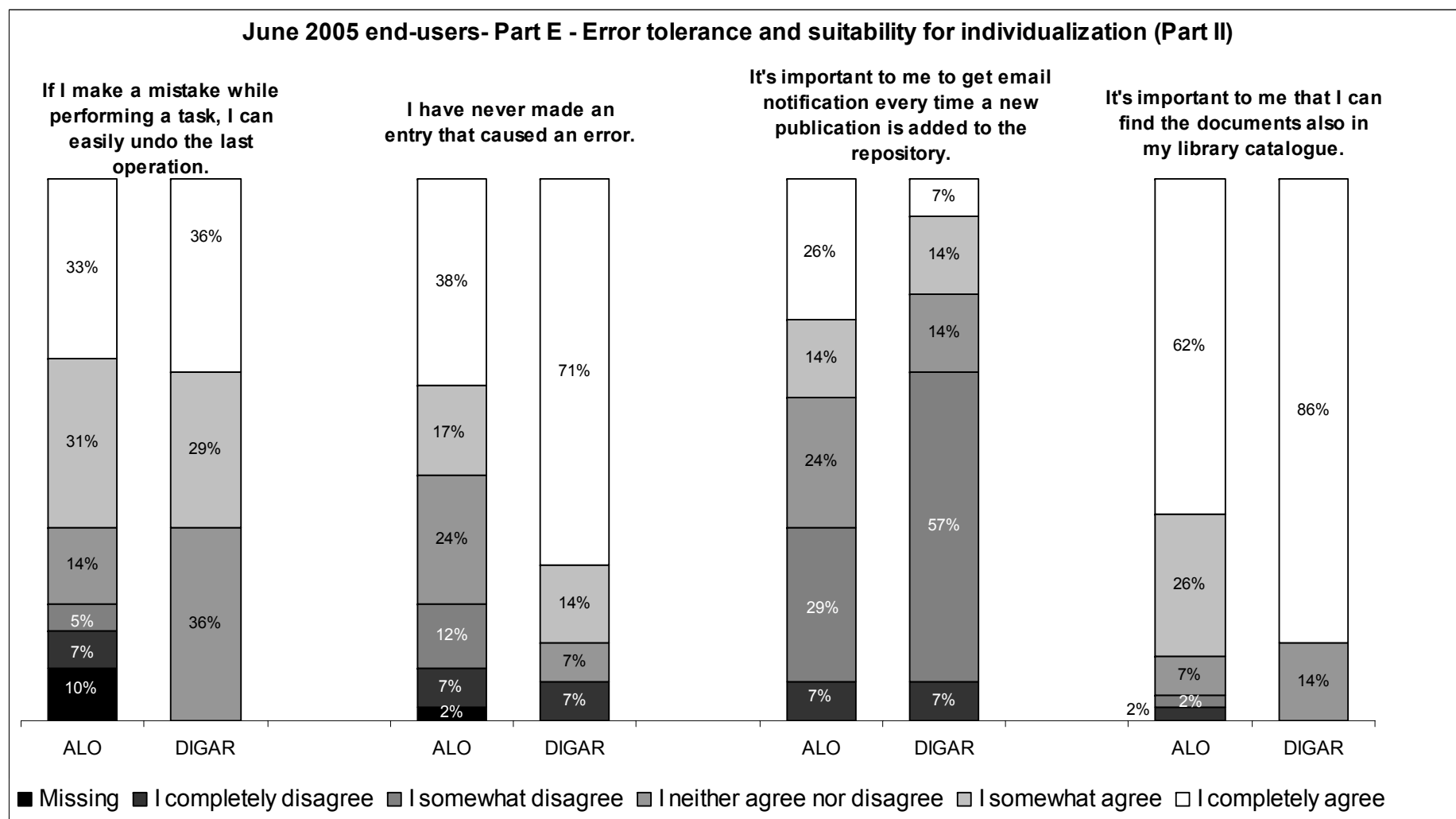


Figure 48: Comparison of results from part: Error tolerance and suitability for individualization (part II) among ALO and DIGAR end-users in the 1st evaluation period (June 2005)

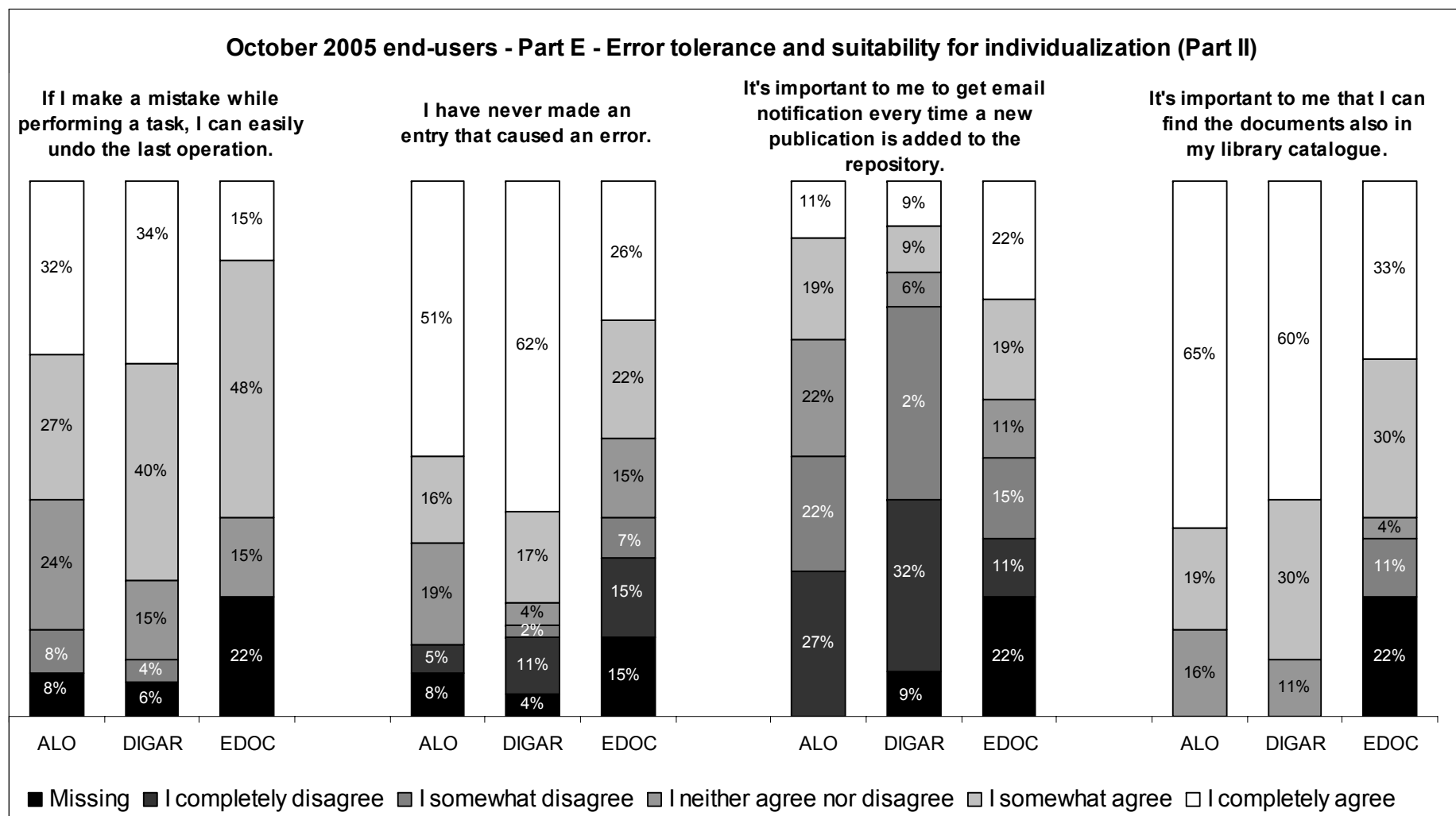


Figure 49: Comparison of results from part: Error tolerance and suitability for individualization (part II) among ALO and DIGAR end-users in the 2nd evaluation period (October 2005)

3.4.7 ADVANTAGES OF THE TRUSTED DIGITAL REPOSITORY

At the end of the questionnaire we asked end-users to rank the list of given advantages of the reUSE digital repository. Similarly as in the section on search criteria and format preferences, many respondents didn't follow the instructions. To get at least some of the most preferred advantages, we used the same methodologies for these analyzes as we did in other responses (e.g. 1 represented most wanted and 5 least wanted).

In June it turned out that the most important advantage in ALO was availability, followed by better access and full-text. On the contrary, the most important advantage for DIGAR end-users was to have full-text then availability and better access. In both cases the e-form of the document as such is not important. (See **Table 21**, and **Figures 50 and 51**)

Table 21: Ranged advantages among ALO and DIGAR end-users in the 1st evaluation period (June 2005)

	ALO	DIGAR
better access	2	3
availability	1	2
full-text search	3	1
e-form	4	4
other	5	5

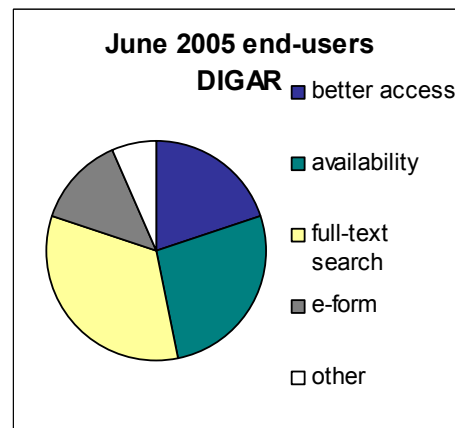
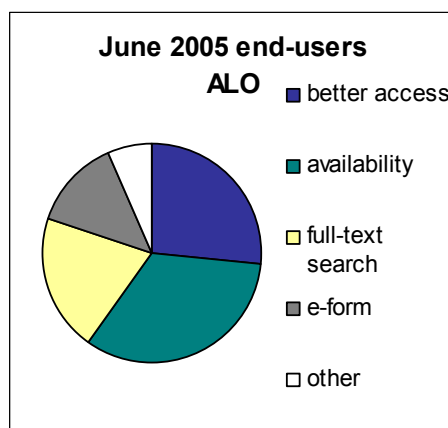


Figure 50 and Figure 51: While the most preferred advantage for ALO end-users is document availability, answers show that among DIGAR end-users the most important advantage is full-text.

In October, when also the opinion of EDOC respondents was taken into account, a new pattern could be found. While ALO and DIGAR end-users still think that the most preferred advantage is availability. EDOC respondents surprisingly considered e-form of the document equally important as its availability. (See **Table 22**, and **Figures 52, 53, and 54**)

Table 22: Ranged advantages among ALO, DIGAR and EDOC end users in the 2nd evaluation period (October 2005)

	ALO	DIGAR	EDOC
better access	2	2	4
availability	1	1	1
full-text search	3	3	3
e-form	4	4	1
other	5	5	5

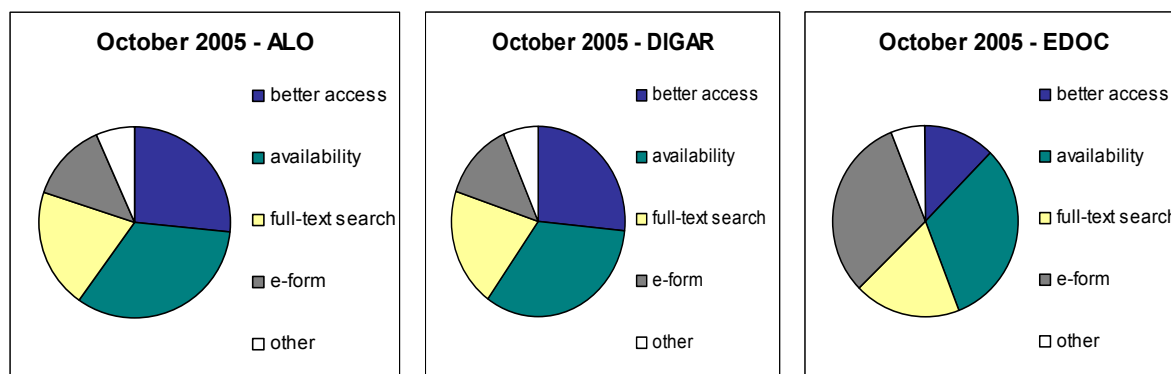


Figure 52, Figure 53 and Figure 54: The preferred advantages for each repository obtained in the 2nd evaluation phase (October 2005)

3.5. NON-USER EVALUATION

As already mentioned in the report, we also considered non-users in our project. Although their feedback is not as important as the end-users' feedback, it was interesting to analyse their answers. Somehow we hoped our survey could include some of the non-users who already had experience with well known full-text commercial providers (such as Elsevier, SCI, etc). Their opinion would be very helpful. Unfortunately, we didn't succeed in this.

One of the most important outcomes is probably the fact that many of them (especially in ALO and DIGAR) confused the e-catalogue offered by the library with the digital repository and many of them probably evaluated the e-catalogue. The results of different questions are presented in **Figures 58 and 59**. A majority of non-users have intermediate experience in using computers and usually they use them when a new research needs to be done or before going to the library. Some differences between the 1st and the 2nd evaluation can be noticed. Concerning the use, it is interesting that in the 1st evaluation period almost no one chose the option of preferring to use the digital repository instead of going to the library. In the 2nd evaluation the percentage of those who didn't like the evaluated digital repositories decreased. Once again we can see the difference between different types of libraries, which implies not only the age of non-users but also as to the reason of using such kind of digital repository. The age span is similar as with end-users. Cross tabulation research, however showed that within non-users, men are older (born already in the 50s and 60s) than women. Men are also more likely to use the digital repository.

Age is a significant factor in the non-users' perception. It is interesting to point out that younger non-users, born in the 70s and 80s (the Google generation) claimed that they will not use the digital repository in the future (**Figures 55, 56, and 57**). Further cross tabulation research also showed that many non-users, who had never before used a digital library, were born in the 70s and 80s. All mentioned facts indicate that some improvements should be done in order to extend the end-user group and attract also the younger generation.

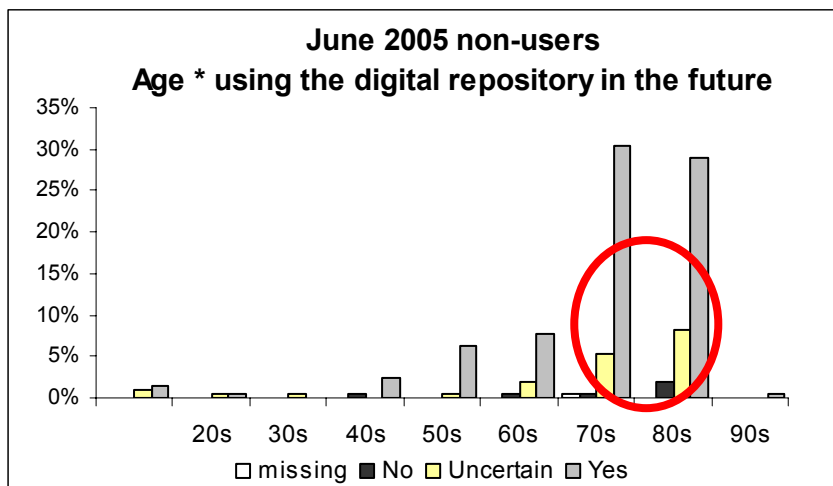


Figure 55: Correlation (*) between age and future use of the digital repository in the June evaluation

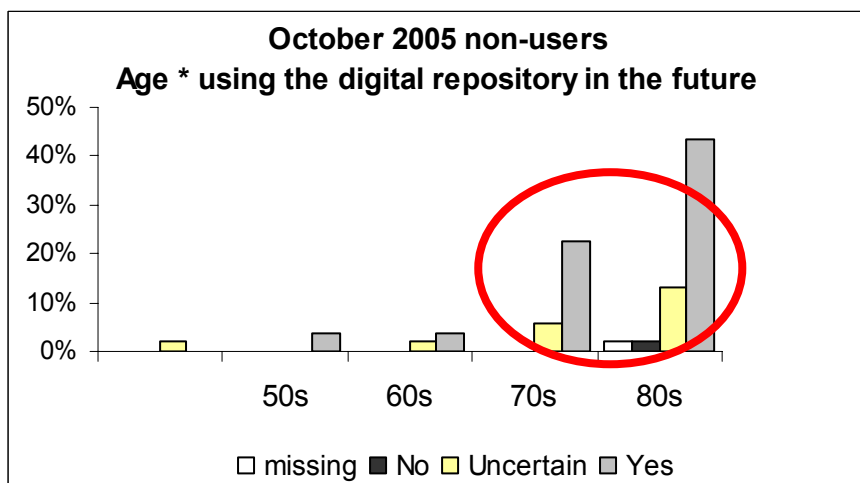


Figure 56: Correlation (*) between age and future use of the digital repository in the October evaluation

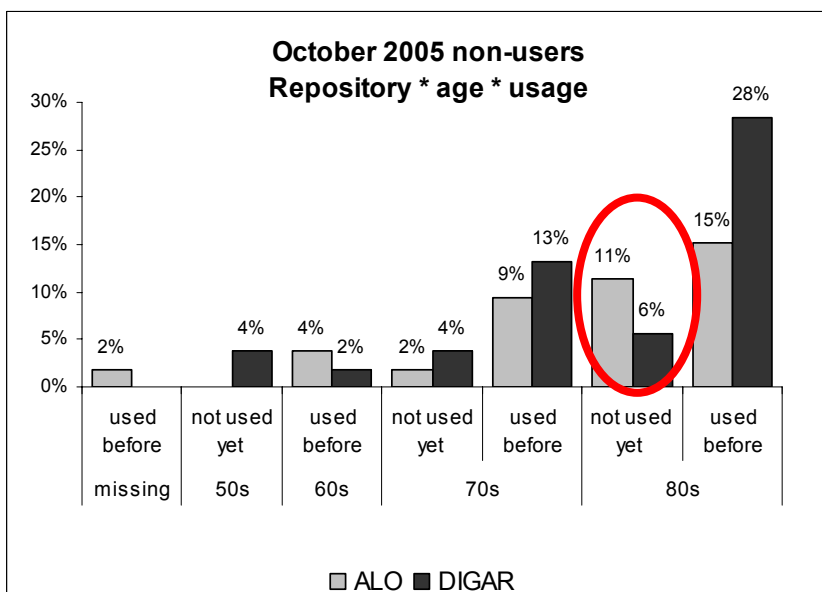


Figure 57: Correlation (*) between repositories, age and use by repositories in the October evaluation

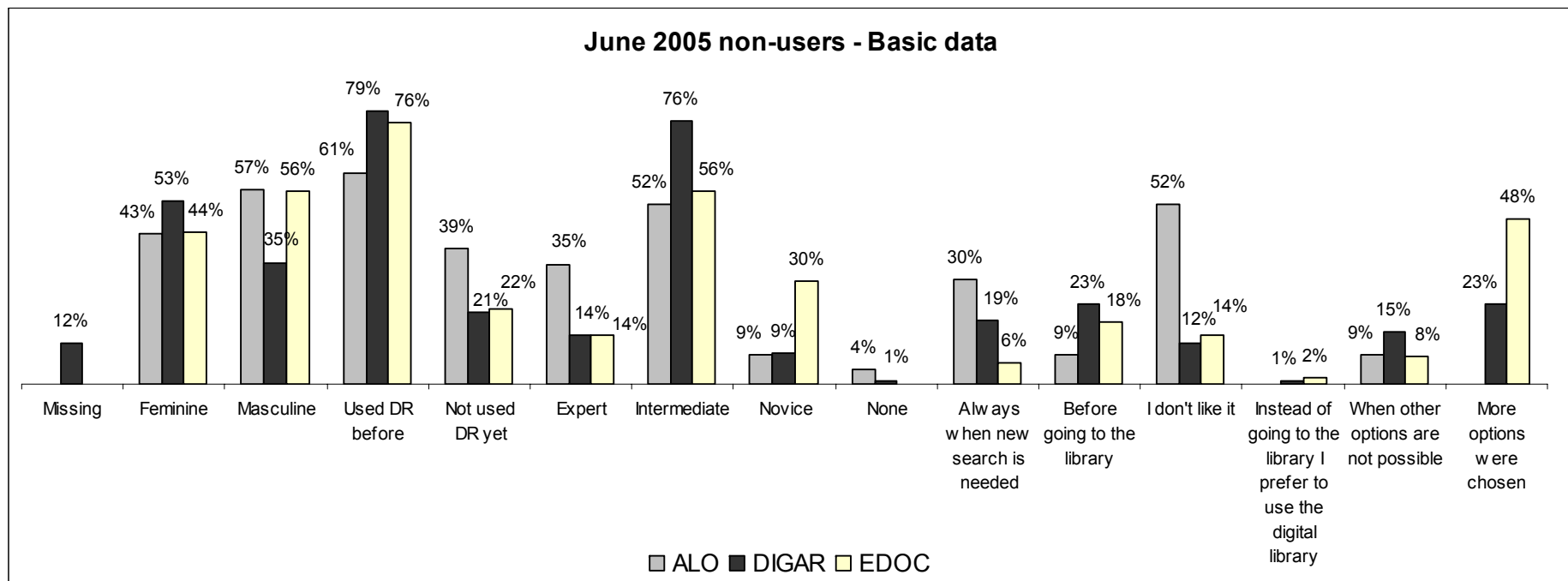


Figure 58: Detailed basic data for the non-user survey in June 2005

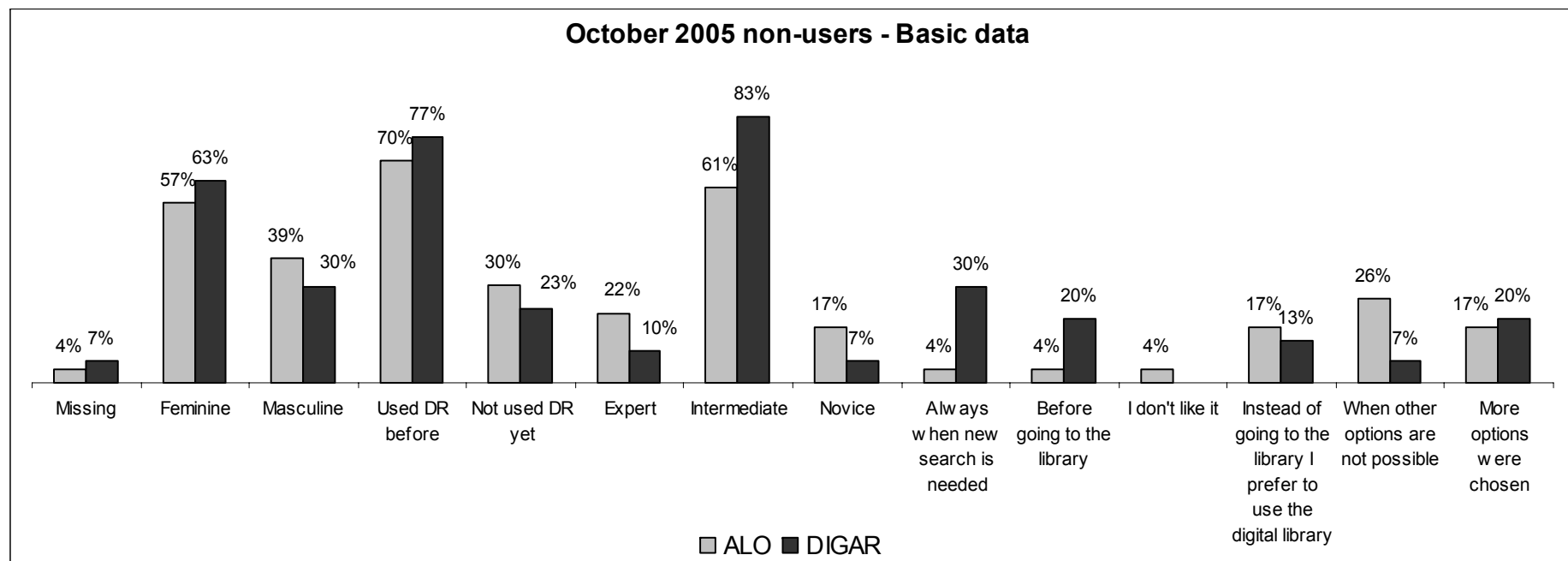


Figure 59: Detailed basic data for the non-user survey in October 2005

At the end we asked non-users if they would now, that they had learned about the digital repository, come and try to use it. While the results gathered in the 2nd period are quite encouraging for DIGAR, in ALO the number of uncertain non-users increased a little bit (Figures 60 and 61).

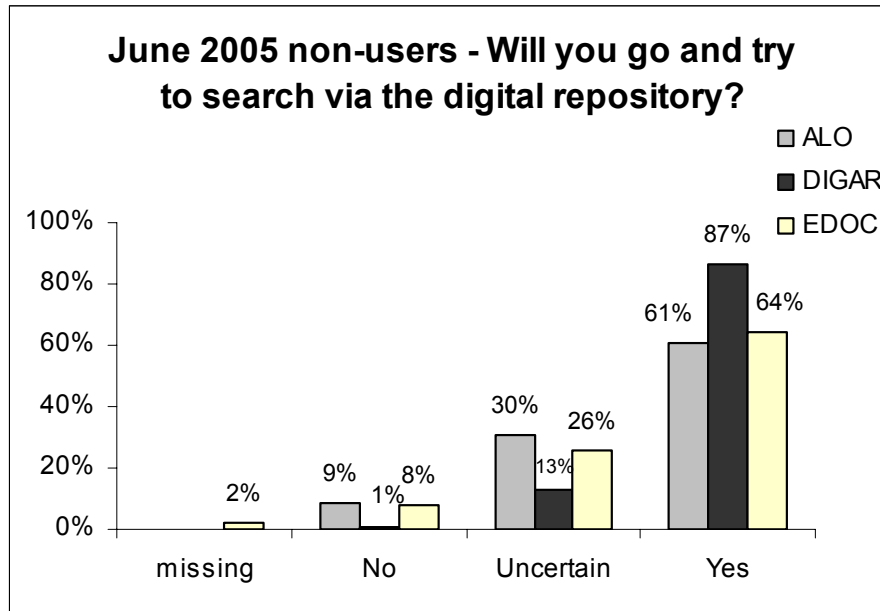


Figure 60: *The future use of the digital repositories among non-users from the 1st evaluation period (June 2005)*

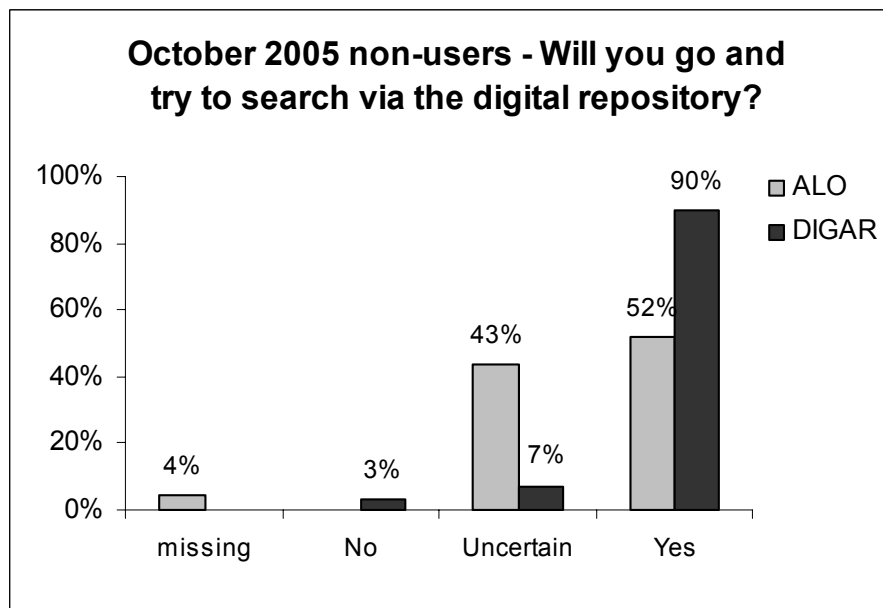


Figure 61: *The future use of the digital repositories among non-users from the 2nd evaluation period (October 2005)*

3.6 REVIEW OF ALL THE FINDINGS IN THE EVALUATION OF THE DIGITAL REPOSITORIES FROM THE USERS' PERSPECTIVE

In the research described above we examined the usability of three different digital repositories from three different countries. Due to great differences it was impossible to compare them with one another. Instead, we rather searched for interesting qualities and specifications of a certain repository and outlined them as advantages, which might be useful for different developers and implementers.

As already highlighted, end-users are an important element, which influences the vitality and success of the digital repository. Many different opinions were expressed by the respondents' in their comments, which should serve as a guide for implementers in improving their systems. Only satisfied and happy users will return to the library.

Our findings from both evaluation periods show that it is very important what purpose the repository serves and who accesses it. For example, while in digital repositories run by university libraries more recent scientific and professional publications should be available, there should be more heterogeneous publications for different kinds of purposes presented in the repositories run by national libraries. The fact that users predetermine the collection should be highly considered; implementers should understand and acknowledge that who and why predetermines what and how! The key challenge then is to know the target group and their primary needs.

Based on the survey responses, we would like to suggest some improvements, which might increase the usability of the repositories.

ALO

The abbreviation ALO stands for Austrian Literature On-line and is run by Universities of Innsbruck, Graz and Linz. Therefore, their most important users are students, professors, scientists, librarians... and thus more professional contents should be provided for them. As already shown in the report, they also have blind users and in order to facilitate their access to the documents (such as more OCR texts...) some improvements need to be considerate. One or two users mentioned problems with layout and displaying documents on different screens – so this problem has to be solved. It would be useful if the keywords were be indicated (by highlights or contrasting colour) with the displayed hits.

DIGAR

DIGAR is the repository of the National library of Estonia with much broader population in terms of age, profession, and use habits (indicated in **Figures 12 and 13** – Basic data) and should provide various kinds of material and publications. Since it has only been running for a few months, it is understandable that some elements of the interface still need to be developed. Users expressed their wish for a better searching engine, which would enable them to search by subject categories as well. Another requirement expressed by end-users' comments might be to provide automatic-file sorting by date and by field. There was also a suggestion to improve the help system, which would provide information about navigation, search criteria and subject indexes. The help system should be in Estonian, English and Russian. Regarding the interface text, the font is rather small and the display of the documents should also be developed. Since the users are of all ages, the layout of the interface should be clear and simple enough to be usable for different kinds of users.

EDOC

EDOC is run by Humboldt University of Berlin and is the oldest digital repository evaluated by the reUSE project. Similarly as at ALO, end-users come from the university environment with very specific needs for scientific material and scholarly publications. Therefore, it is not surprising that its end-users are probably among the most critical at all, which was also indicated by the obtained results. Many of them requested to have more scientific and professional documents in the digital repository.

Additionally to all three evaluated repositories, there is a strong need to increase the awareness among present and potential end-users and especially to show them the advantages of added-value services offered by the specific digital repository. Many end-users also stressed the need to have a list of repository content by subject and a list of new documents from the last/previous month.

4 COST-BENEFIT ASPECTS

By Alenka Kavčič-Čolić

4.1 BACKGROUND

Another issue that pertains to organizational aspects' evaluation is the cost-benefit dimension. We know that the benefits of the services of such digital repositories, as well as most of the library services, do not have short-term, but long-term manifestations. They will show up in a few years after the reUSE project's conclusion. That is why we were not able to describe the benefits in this evaluation but rather the anticipated. Another problem is that most of the financial information, like salaries or cost of infrastructure, is not publicly available and depends on the financial models of the different types of organizations involved. Consequently, we had to focus on a qualitative evaluation of cost-benefit aspects.

The cost-benefit dimension affects content providers and demonstrators. On the one hand, content providers should benefit from cooperating with demonstrators, and on the other hand, demonstrators should have more benefits than expenses in the maintenance of the digital repository.

Therefore, we focused on two groups: content providers and the librarians in charge of the repositories.

4.2 EVALUATION METHODOLOGY

For the survey on content providers we prepared a questionnaire with 15 questions (See **Annex 3.3**), which each demonstrator distributed to content providers in November 2005. We decided on that time, because we expected that by then the content providers would already be able to assess the usefulness and benefits of their cooperation with reUSE repositories.

The number of respondents was more representative than thorough. The number of organizations that responded to our questionnaire varies from 23%-46% of demonstrator content providers (**Figure 62**).

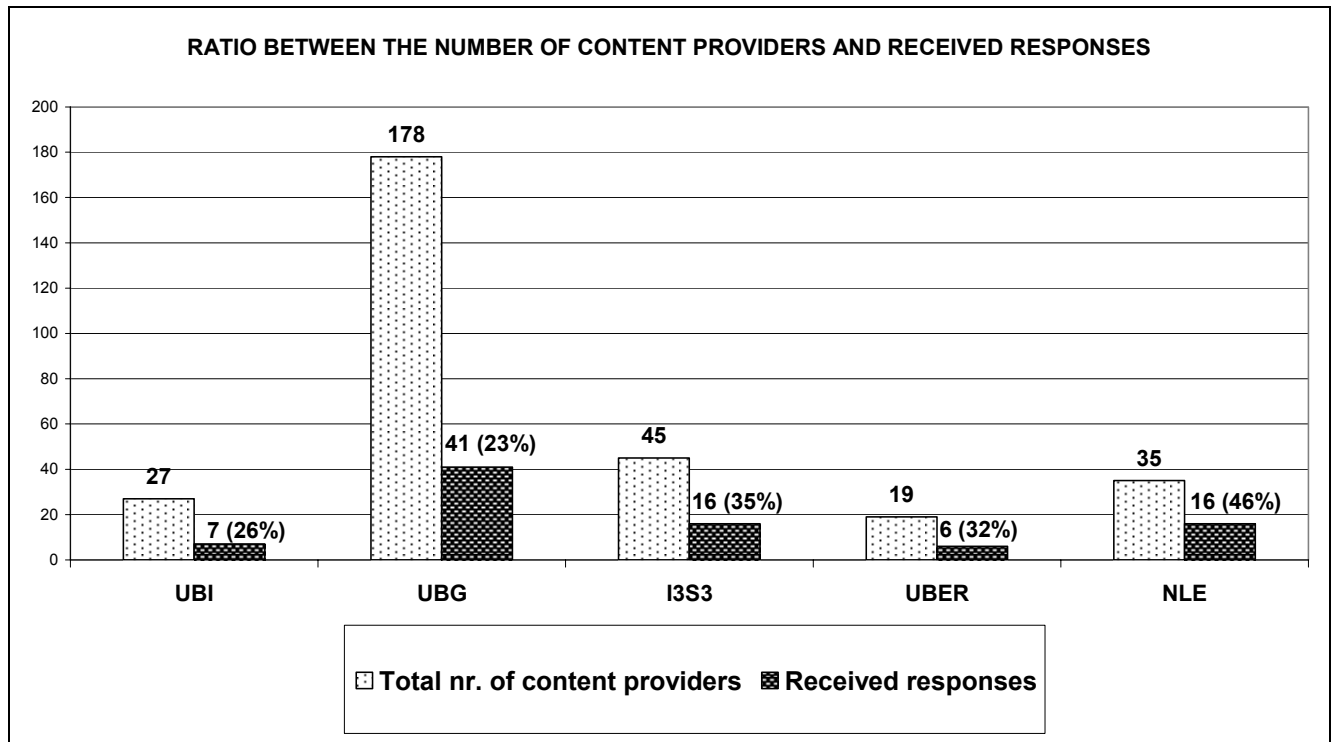


Figure 62: Ratio between the total number of content providers and received responses

We received 86 responses from different types of organizations as shown in **Figure 63**. NLE and UBG have more content providers from the municipal and government institutions, while at UBI, UBG, I3S3 and UBER they mostly come from the academic environment.

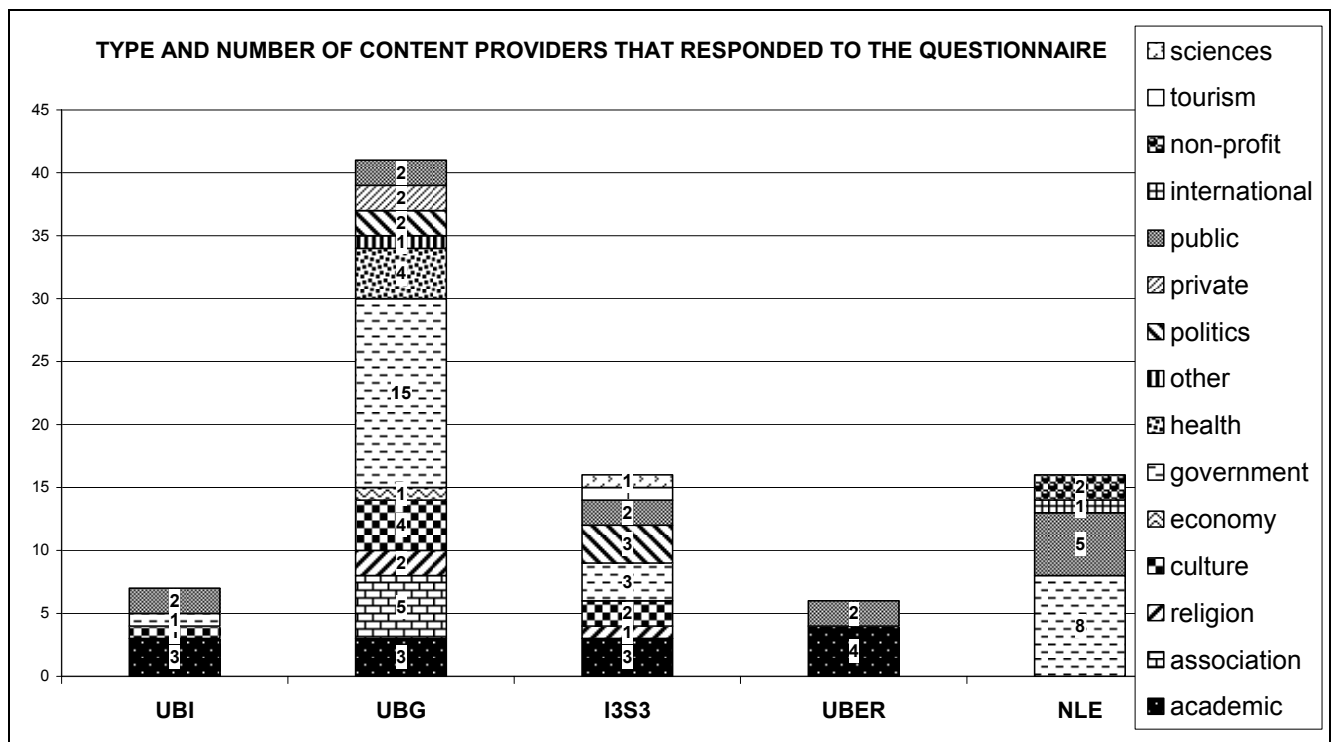


Figure 63: Type and number of content providers that responded to the questionnaire

Concerning the librarians in charge of the operation of the repository, we distributed one questionnaire (See **Annex 3.4**) to each demonstrator. We received four answers concerning the three repositories. We tried to get as much information as possible on different organizational and cost-benefit aspects. Thus the questionnaire included some of the questions already used in the White Paper survey, since we wanted to compare the responses at the beginning of the project to the responses that followed the development of the digital repositories. We obtained very interesting information which showed us that concerning organizational matters a big progress had in fact been achieved. We added some questions related to the ingest process of the repositories, which helped us gather information on cost-benefit aspects.

4.3 CONTENT PROVIDERS' SURVEY

Content providers were first asked about their motivation for contributing to the digital repository. We listed nine possible reasons for their cooperation and in front of everyone it was possible to choose the number on a scale from 1 to 5, where 1 meant complete agreement and 5 complete disagreement with the statement. We also added the possibility of additional comments.

4.3.1 Content providers' interest in increasing the access to their publications

There were 4 questions related to the access to their publications. The received answers are shown in **Table 23**:

Table 23: Access and dissemination aspects that produced interest among the content providers in cooperation with reUSE repositories

<i>The main motivation for your collaboration in the digital repository has been:</i>	reUSE demonstrator	<i>I completely agree</i>	<i>I somewhat agree</i>	<i>I neither agree nor disagree</i>	<i>I somewhat disagree</i>	<i>I completely disagree</i>
1.1. Easier and faster access to our publications	UBI	57%	43%			
	UBG	49%	22%	17%	5%	7%
	I3S3	81%	13%	6%		
	UBER	66%	17%	17%		
	NLE	75%	25%			
1.2. The possibility for disabled people to access our publications	UBI	29%	43%	14%	14%	
	UBG	29%	24%	22%	15%	10%
	I3S3	19%	25%	44%	6%	6%
	UBER	17%	50%	33%		
	NLE	56%	19%	25%		
1.4. Easier and broader distribution of our printed documents	UBI	29%	29%	29%	14%	
	UBG	70%	23%	5%	3%	
	I3S3	88%	6%	6%		
	UBER	100%				
	NLE	63%	31%	6%		
1.9. To increase public use of our publications	UBI	57%	29%	14%		
	UBG	41%	39%	15%	5%	
	I3S3	94%			6%	
	UBER	83%	17%			
	NLE	94%	6%			

The responses regarding access and dissemination of their publications were very highly ranked. **Figure 64** shows the respondents' assessment of this statements. If we analyze the assessment of each of the four statements we can get interesting results.

As shown in **Figure 64**, the motivation of most respondents for contributing their contents to reUSE repositories was to provide access to, broader distribution and an increased use of their publications among the public. They were more sceptical of the possibility to provide access for disabled people.

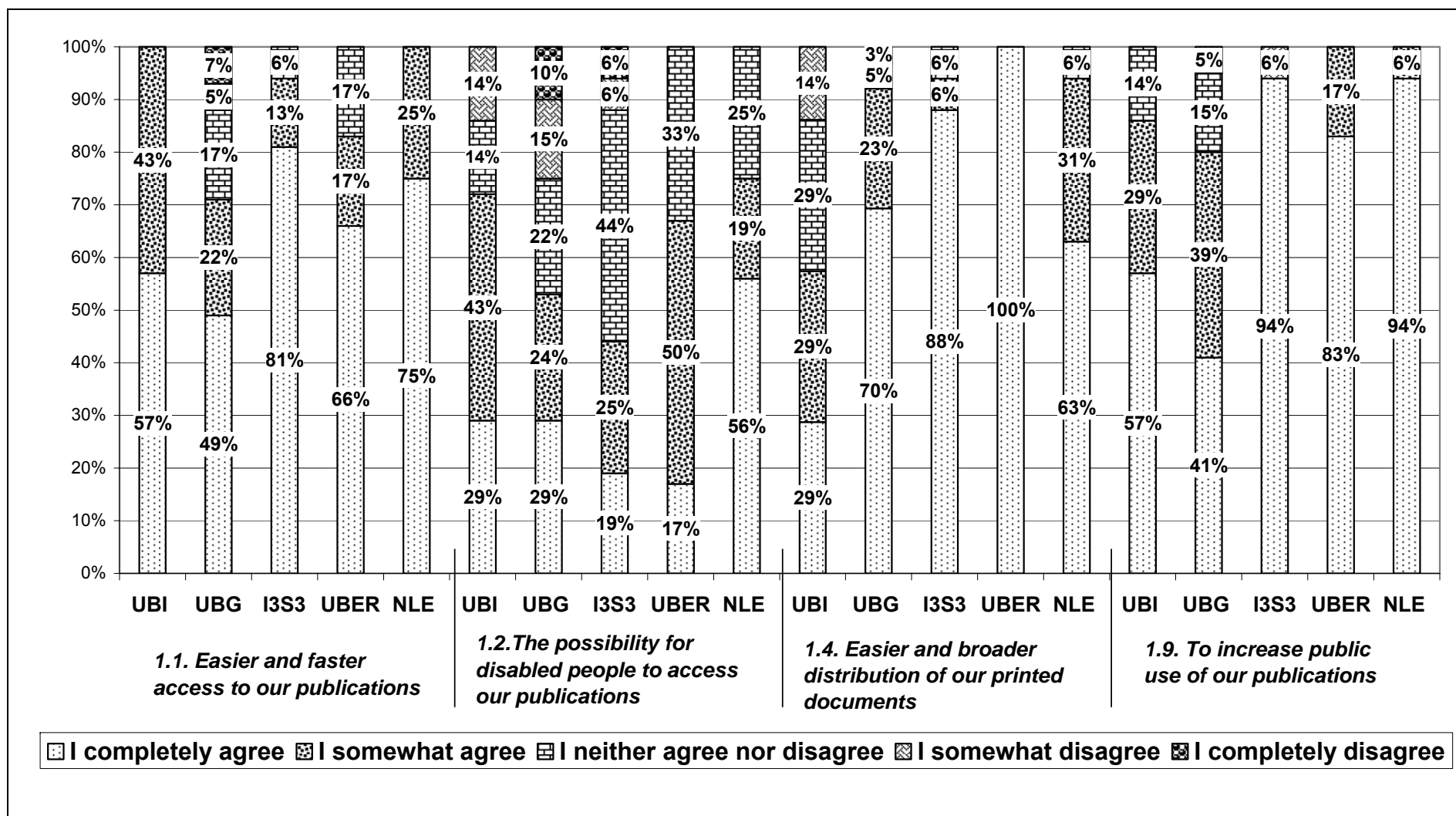


Figure 64: Main motivation for collaborating in the reUSE digital repository related to the access and dissemination of publications

If we analyse the assessment of each of the four statements we can get interesting results, which are described below.

As shown in **Figure 65**, most of the respondents decided to contribute their contents to reUSE repositories because they wanted to enable an easier and faster access to their publications. This answer was ranked higher with 81% at I3S3, 75% at NLE and 57% at UBI. At UBG 12% of the content provider respondents seem to disagree with this being an important reason for collaboration.

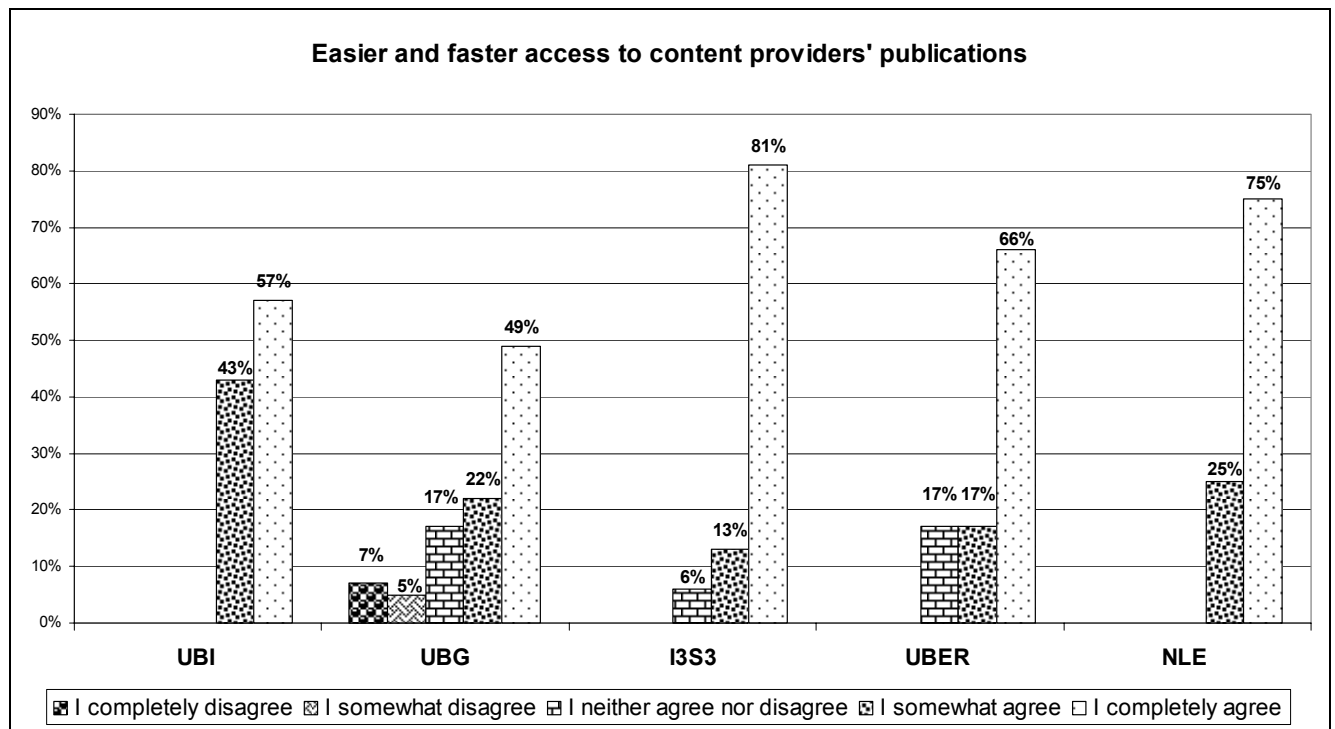


Figure 65: Easier and faster access to their publications as one of the motivation factors of the content providers for contributing to the reUSE repository

Content provider respondents also had to assess the statement that disabled people were also enabled to access their publications. Most importance to this issue was given by the NLE content providers (**Figure 66**). This is also a very important aspect at UBER, since 50% of the respondents somewhat agreed with this statement. At UBG and UBI 29% and at I3S3 12% of the respondents completely or somewhat disagreed with this statement. UBG and I3S3 should pay more attention to this issue.

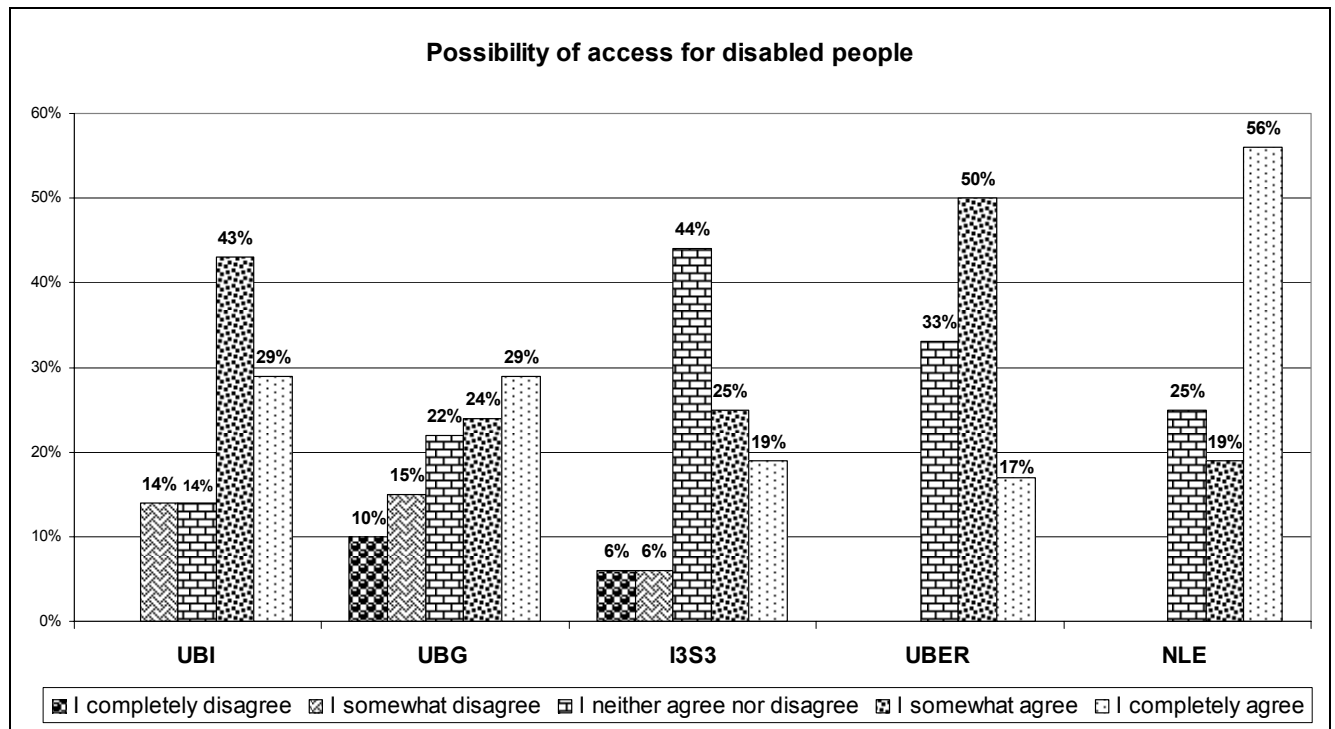


Figure 66: Possibility for disabled people to access to content providers' publications as one of the motivations for contribution to the reUSE repository

The third statement claimed that content providers were motivated to contribute to reUSE repositories in order to achieve an easier and broader distribution of their publications. All UBER's content providers completely agreed with this statement (**Figure 67**). It means that they hope for a community that has complete access to their repository. At I3S3 88% of all respondents completely agreed with this statement, while at UBG 93% of respondents expected to use the repository as a channel for distribution of their publications. The results of NLE respondents is very similar, since 63% completely agreed and 31% somewhat agreed, making a total of 94% of content providers. At UBI most of the content providers that responded to the questionnaire did not see this to be such an important reason for their cooperation. Maybe a bigger sample would have shown different results.

To sum up, by allowing content providers a broader access to their publications, they want to increase public use of their publications. **Figure 68** shows that most of the respondents completely agreed with this statement: at NLE and I3S3 94%, at UBER 83%, at UBI 57%, and at UBG 41%. Together with the second score it makes a 100% at NLE and UBER, 86% at UBI and 80% at UBG. At UBG there are 5% and at I3S3 6% of content provider respondents, who did not completely agree with this statement.

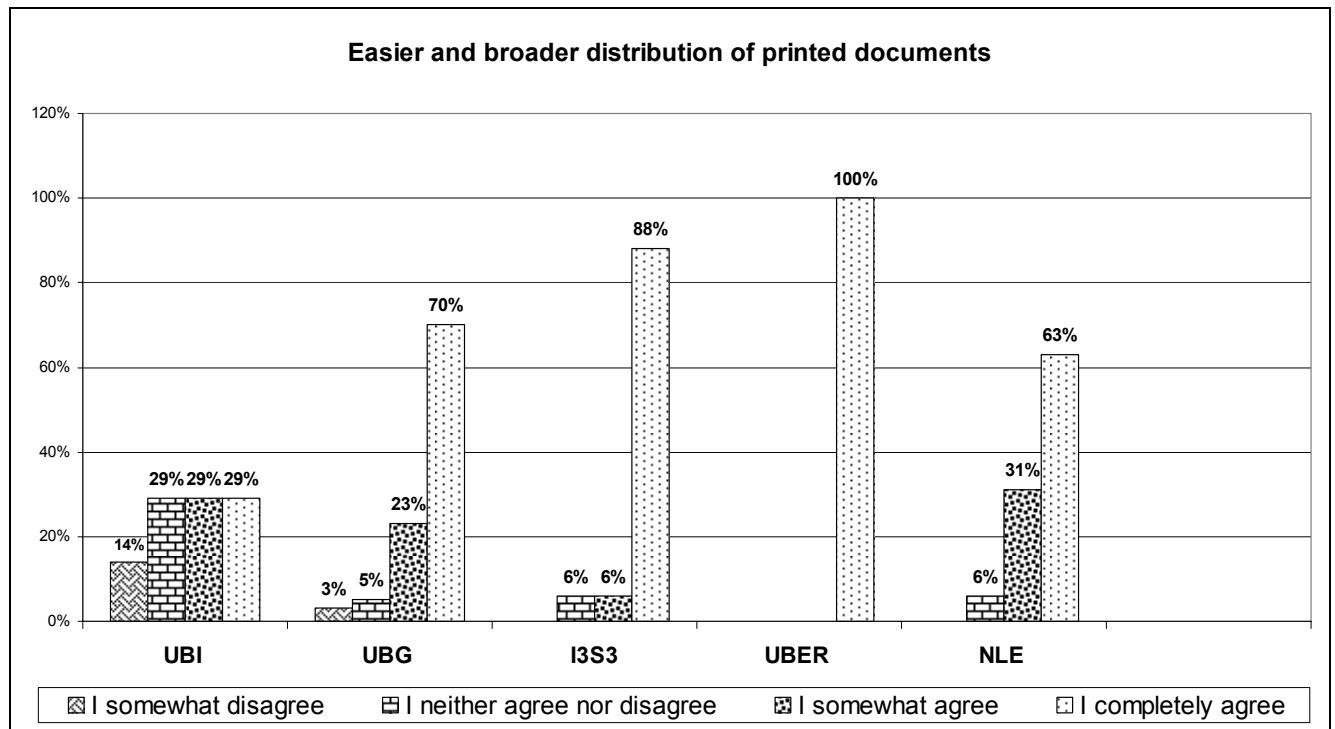


Figure 67: Easier and broader distribution of content providers' printed documents as one of the motivation factors for their contribution to the reUSE repository

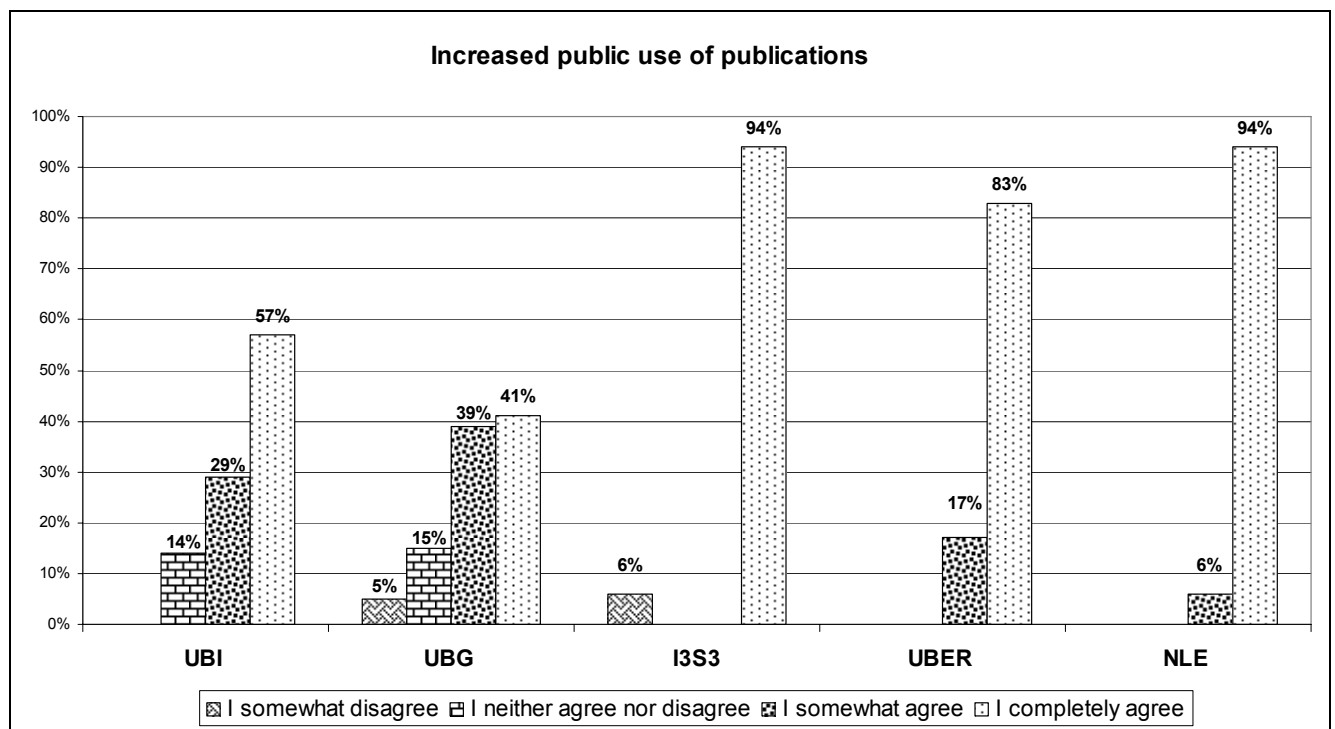


Figure 68: Increased public use of content providers' publications as one of the motivation factors for their contribution to the reUSE repository

4.3.2 Long-term preservation

One of the most relevant issues in the last few years is long-term preservation of digital publications. As shown in previous chapters, one of the goals of reUSE was to build trusted digital repositories in order to keep and preserve the access to the digital contents on a long-term basis. This topic was also used by demonstrators during the marketing of their repositories and their attracting of content providers. The survey on long-term preservation of digital documents in European libraries done by ARGE showed that specially national libraries and deposit institutions paid a lot of attention to this issue. Others are very aware of the importance of the deposit institutions activities for preserving the digital documents for the future.

The reUSE content providers do not lag behind the deposit institutions in their awareness regarding long-term preservation of digital documents. **Figure 69** shows a 100% of UBI respondents contributed to the repository in order to provide for the preservation of their digital documents. At NLE and I3S3 81%, at UBG 78% while at UBER only 50% opted for this factor. The reason that at UBER long-term preservation is not the main motivation factor for cooperation for 50% of the content provider respondents lies in the fact that they are more focused on providing public access to their contents.

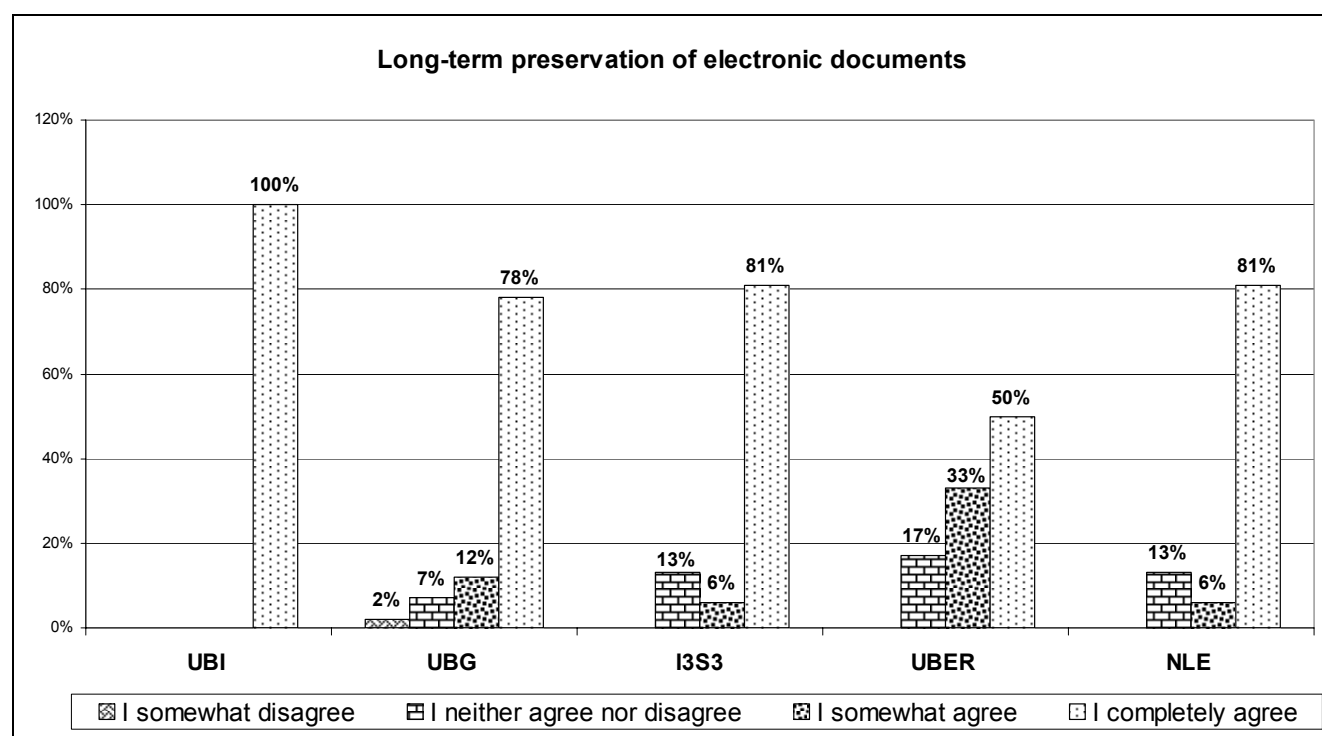


Figure 69: *Long-term preservation of content providers' electronic publications as one of the motivation factors for their contribution to the reUSE repository.*

4.3.3 Strategic reasons for cooperation with reUSE demonstrators

With increased dissemination of their publications the work and activities of the content providers are more visible to the public and their role in the community becomes more important. They might even attract new sponsors. We supposed that these aspects could also raise desire for cooperation with digital repositories. We dedicated four statements to

this issue to be assessed by content providers. The responses are shown in **Table 24** and in **Figure 70**.

Table 24: Strategic aspects quoted as reason for cooperation with reUSE repositories

<i>The main motivation factor for your collaboration with the digital repository has been:</i>	<i>reUSE demonstrator</i>	<i>I completely agree</i>	<i>I somewhat agree</i>	<i>I neither agree nor disagree</i>	<i>I somewhat disagree</i>	<i>I completely disagree</i>
1.5. To increase the interest in our organization	UBI	43%	43%	14%		
	UBG	51%	29%	15%	5%	
	I3S3	69%	19%	13%		
	UBER	33%	33%	17%	17%	
	NLE	25%	63%	6%		6%
1.6. To increase the dissemination of our activities and work	UBI	14%	14%	71%		
	UBG	54%	29%	10%	7%	
	I3S3	88%	6%		6%	
	UBER	50%	25%	25%		
	NLE	63%	31%	6%		
1.7. To increase the recognition of our organization among our stake-holders	UBI		29%	43%	14%	14%
	UBG	44%	29%	15%	10%	2%
	I3S3	13%	38%	25%	25%	
	UBER		33%	17%	17%	33%
	NLE	31%	44%	19%		6%
1.8. To attract new sponsors	UBI	14%		29%	29%	29%
	UBG	10%	20%	32%	20%	20%
	I3S3	25%	25%	19%	19%	13%
	UBER		17%	17%	17%	50%
	NLE		25%	13%	38%	25%

The strategic issues are not as important to the content providers as the access to the publications or long-term preservation of digital documents, although they are still a very important reason for cooperating with the digital repository. Attracting new sponsors is the least expected reason.

The next page provide a detailed review of content providers' assessment.

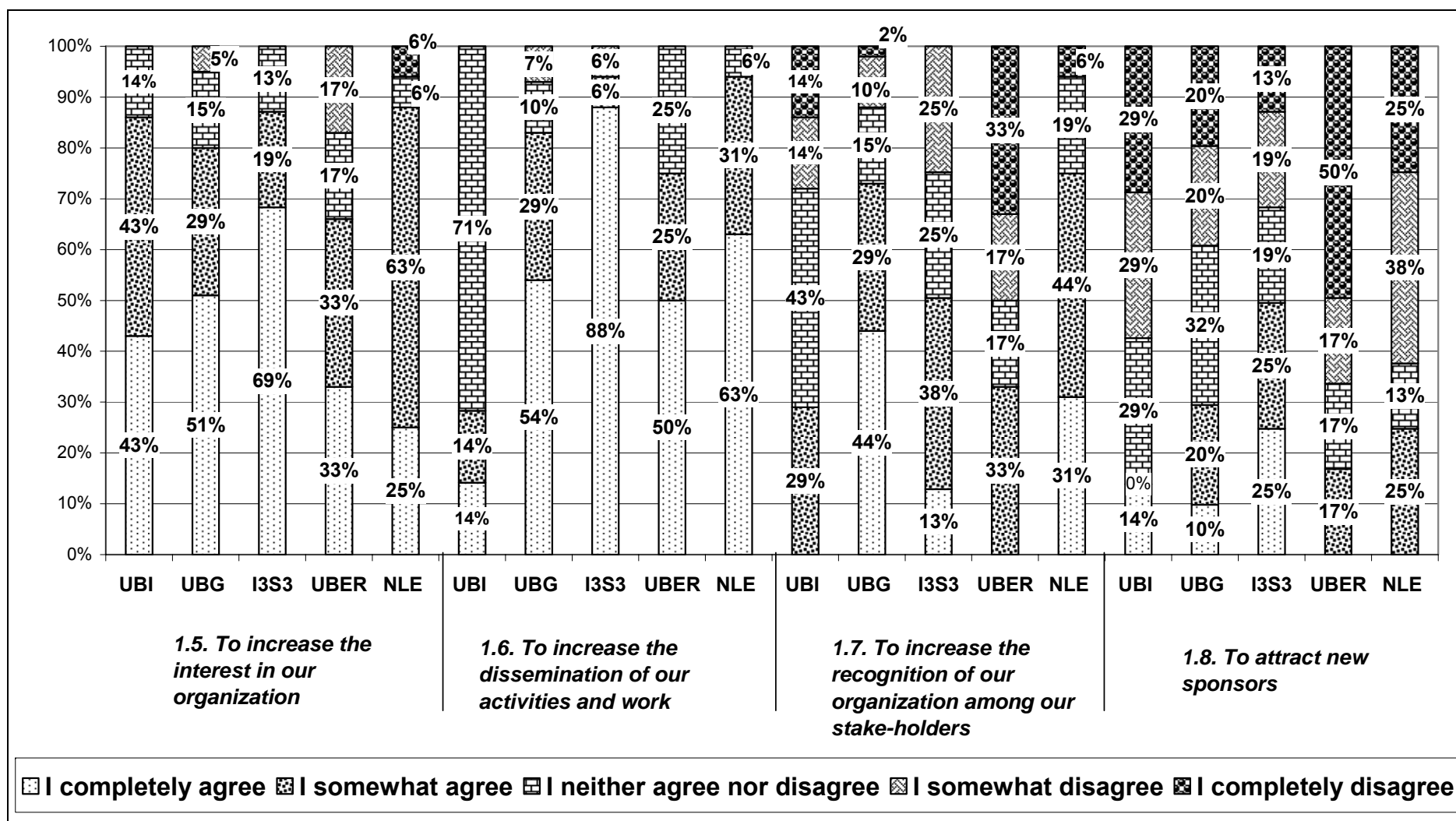


Figure 70: Main motivation factor for collaborating with the reUSE digital repository related strategic issues

A relatively high percentage of respondents see their cooperation with the reUSE repository as a way of increasing public interest in their institution (**Figure 71**). At I3S3 69%, at UBG 51% and at UBI 47% of all respondents completely agreed with this statement. The percentage of content providers, who only partly agreed is much higher at NLE (63%), while at UBER and UBI it is the same as the percentage of those, who completely agreed with this statement (43% and 33% respectively). Even it was not the main motivation factor for cooperation, it is evident that repositories were well enough promoted and that content providers understand their role in the community.

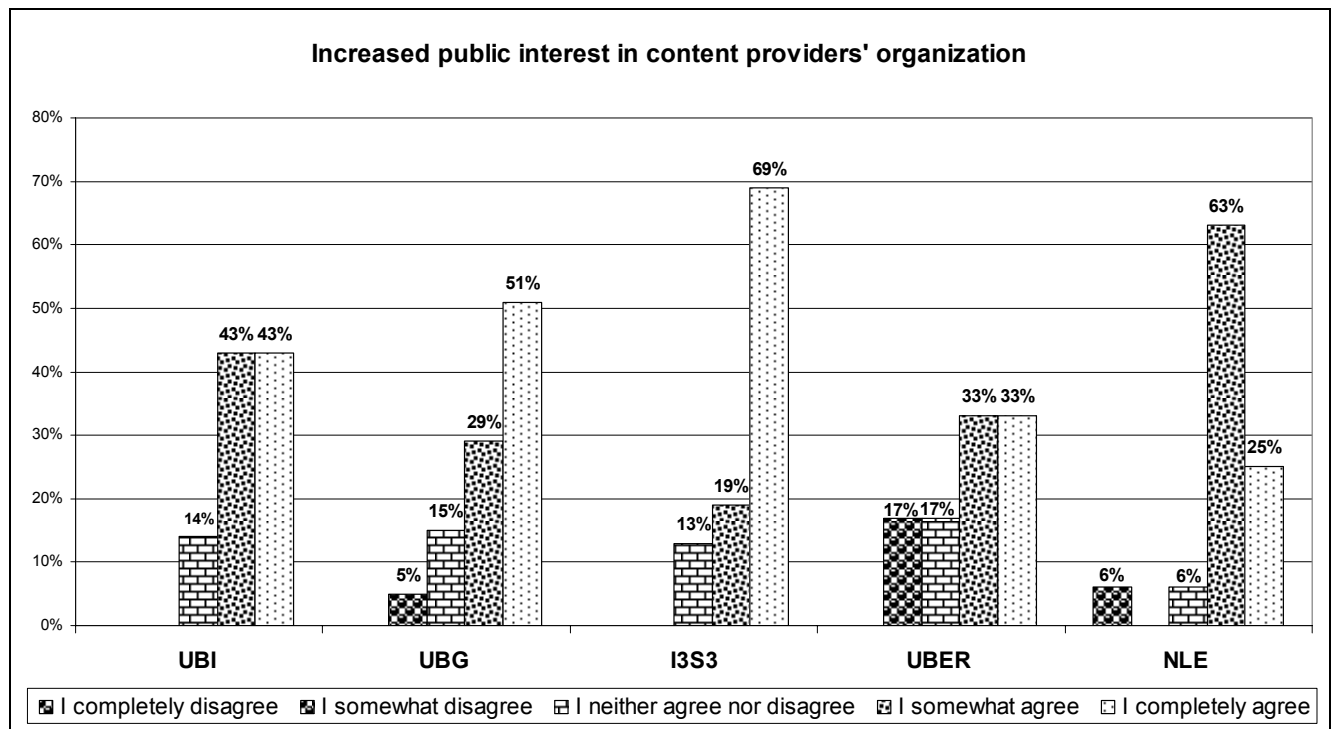


Figure 71: Increased public interest in content providers' organization as one of the motivation factors for their contribution to the reUSE repository

The respondents' assessment of the statement that digital repositories could be an instrument for the dissemination of their work and activities was very similar (**Figure 72**). At I3S3 88%, at NLE 63%, at UBG 54%, and at UBER 50% of the respondents completely agreed with this statement. More sceptical in this regard were the content providers of UBI (71%). At UBG 7% and at I3S3 6% of the respondents even somewhat disagreed with this statement.

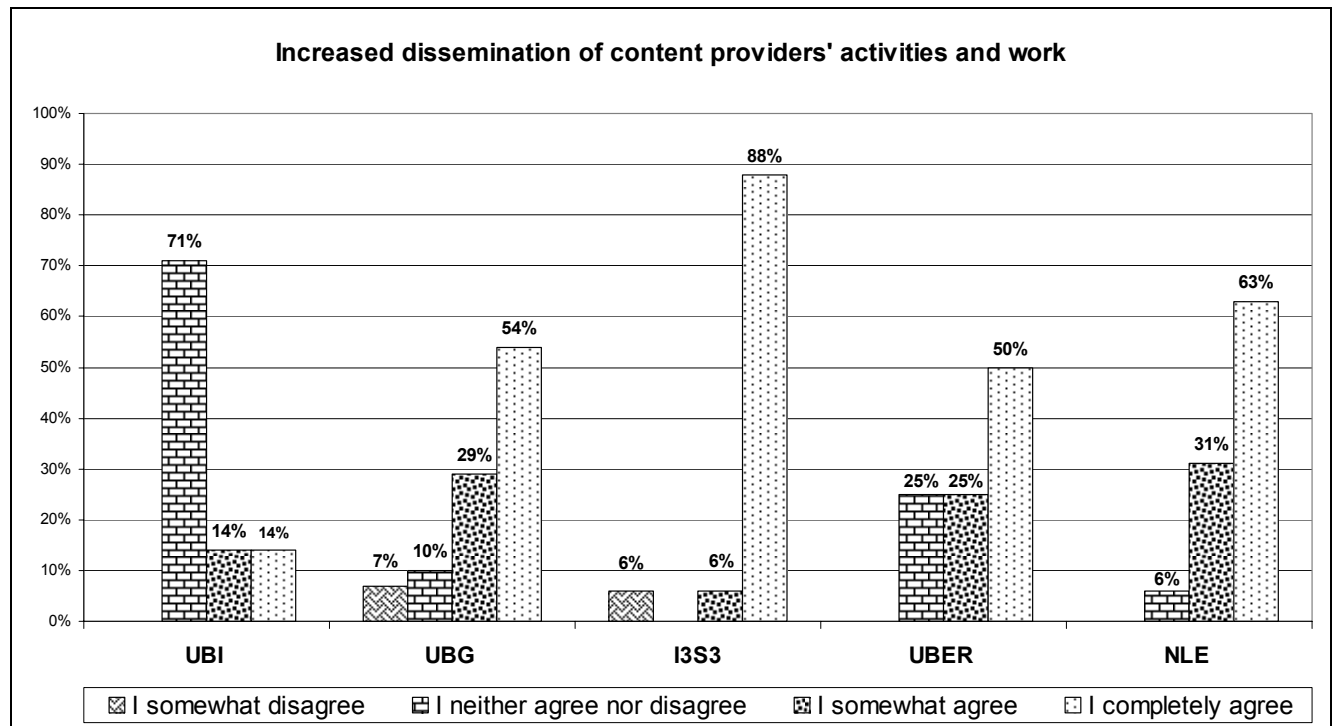


Figure 72: Increased dissemination of content providers' activities and work as one of the motivation factors for their contribution to the reUSE repository

A lower percentage of respondents think that their cooperation with digital repositories could increase the recognition of their organization (**Figure 73**). At UBG 44% of the respondents, at NLE 31% and at I3S3 13% of the respondents completely agreed with this statement. If we include the content providers that somewhat agreed, the percentage becomes 73% for UBG, 75% for NLE, and 51% for I3S3. It is interesting that at UBER 50% of the respondents completely or somewhat disagreed with this statement, while only a low percentage those, who completely or somewhat disagreed could be found at UBI (28%), UBG (12%) and NLE (6%). There is a higher percentage of respondents, who do not have an opinion on this issue (UBI 43%, I3S3 25%, NLE 19%, UBER 17%, and UBG 15%).

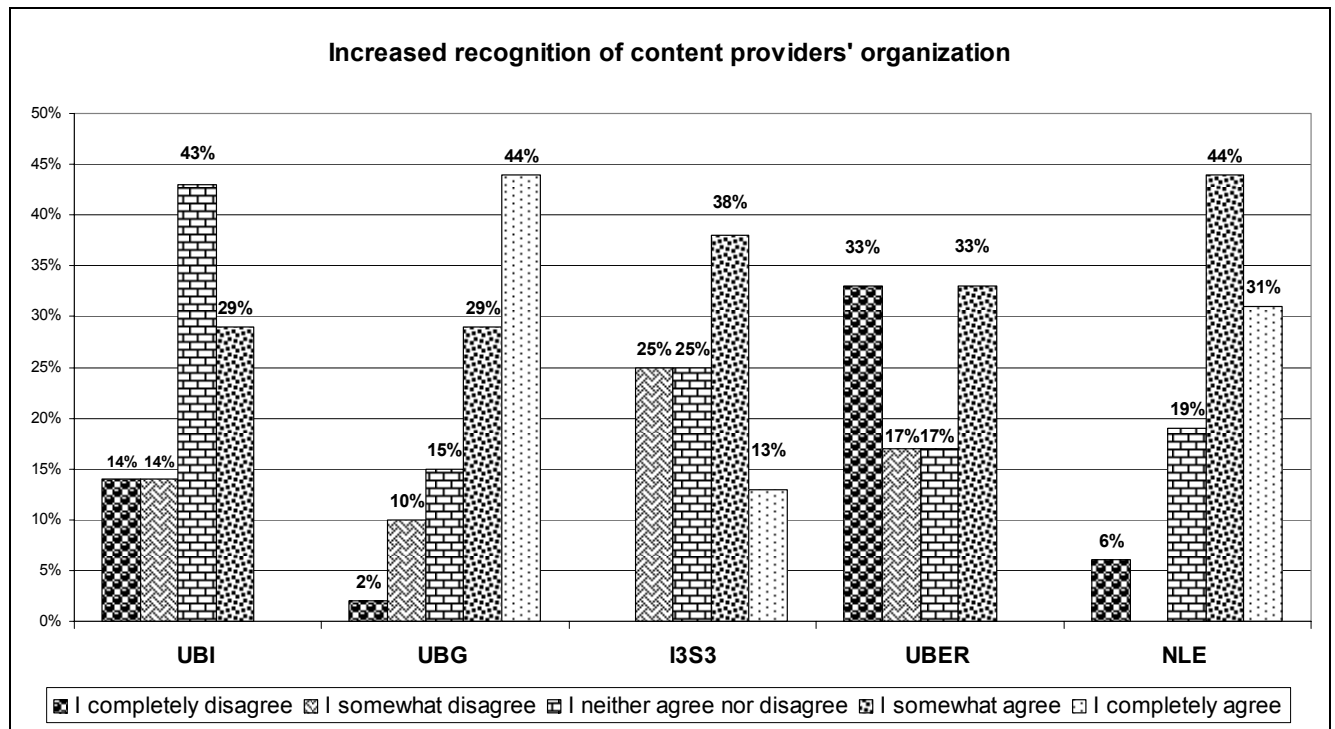


Figure 73: Increased recognition of content providers' organization among their stake-holders as one of the motivation factor for their contribution to the reUSE repository

All organizations wish to attract new stake-holders and sponsors. However, this was not the main reason for cooperation with the reUSE digital repositories (**Figure 74**). A relatively high percentage of respondents completely disagreed with this statement (UBG 20%, UBI 29%, NLE 25% and UBER 50%). At UBG only 10% and at UBI 25% completely agreed with the statement that this cooperation could bring them new sponsors. At I3S3 the content providers were more enthusiastic about attracting new sponsors, since together 50% completely or somewhat agreed with this statement.

Some of the respondents from UBG and NLE gave additional reasons for cooperation. Their answers to the question: ***"Is there any other reason for your collaboration with the digital repository?"*** are listed below:

UBG:

- Logical (future oriented) parallel-action to legal-deposit-copy-system.
- Safe and long-term archiving. Our journal distribution on a larger scale. It is impossible to ignore a professional offer to archive publications in a digital way. We are interested to see to what extent this possibility will be used by our members, partners and sponsors. My question: do you have any figures on the number of entries to different pages and statistical material. I am always surprised when I see the number of entries to our homepage!
- To offer other persons (incl. students) information and stimuli, which they can use for projects, theses, etc.
- Help to make information accessible; socio-political and location (advantage for this economic region, where something is happening).
- To collaborate with other cultural institutions.
- To get a better view of our own publications and the publications of others.
- It is the best way to archive publications long-term and to save them for future readers.

- The reUSE idea is very attractive, with a lot of new topics.
- Great thing and small costs.
- The Styrian server for official village homepages has been taken over by Telekom Austria. This caused problems as far as the web-presentation of our journal is concerned. For that reason, the reUSE offer came at the right time.
- To reduce postage expenses. We have members in Germany, Switzerland, Hungary, Bulgaria, Iran, Brasil,... and do not have to send them a printed copy.

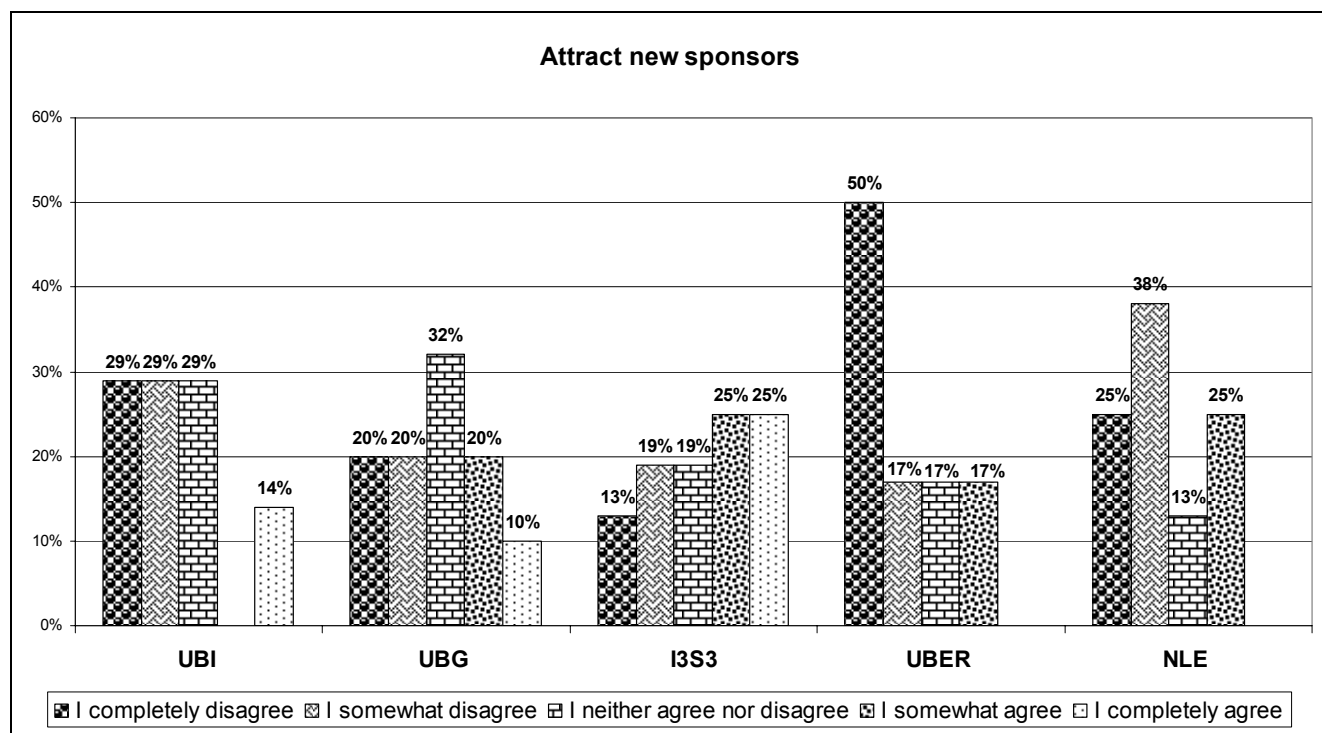


Figure 74: Attraction of new sponsors as one of the motivation factors of the content providers for contributing to the reUSE repository

In the responses above we can see that content providers are most interested in long-term preservation and in access to digital collections. The last three topics could be understood as a cost-benefit issue, since the content providers see the repository as an instrument for saving money.

I3S3

- Storage place - they use the repository as a storage place for their digital contents.
- A service for former teachers and pupils.

NLE

- As the archive is in its initial state now, it's very hard to say how many publishers and users will join and use all these value-added services. We hope that they do, but if they don't the digital archive is valuable anyway.
- Too little experience so far.
- To help arrange our publications and our files; to have an overview of all our publications.
- The archive is still too new to see the real benefits.
- We see the main reason for the archive in easier access to our publications.

At NLE it is too soon for the content providers to see the possible interest in the digital repository. We believe this will become clear at least within a year of repository's operating.

The second part of the questionnaire asked the content providers about their opinion on the benefits they expected from their contribution to the digital repository. The questions were formulated in a way that we could find indirect benefits or costs, which could be valuable for the demonstrators in their promotion enterprise. The method of answering to the questions was the same as in the first part of the questionnaire. The responses are shown below.

29% of UBG, 13% of I3S3, 17% of UBER and 19% of NLE respondents expect the digital repository to benefit their stake-holders, which are much more important than its costs (**Figure 75**). There is a higher percentage of respondents that only somewhat agreed with this statement. Although the opinions are more or less equally divided, at UBER the percentage of those, who completely disagree with this statement is higher than the rest (33%). At UBG and I3S3 opinions are also divided, but there are more positive answers in favour of the repository services.

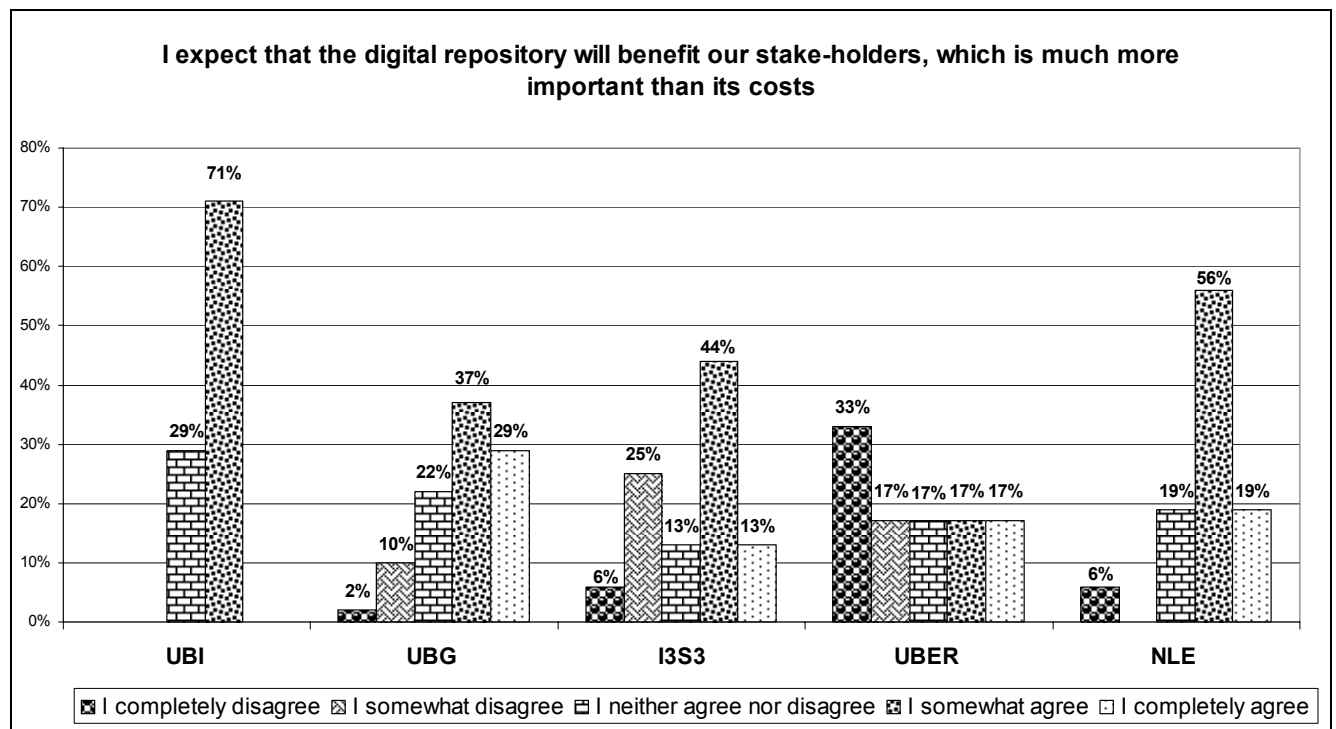


Figure 75: Content providers respondents' expectation concerning the benefits of the digital repository for their stake-holders

Do the content providers expect any profit from this collaboration? We can say that monetary advantages could come indirectly via other channels. Their documents publication in a digital repository can do a lot in this regard. The answers show that most of the respondents do not expect any profit. I3S3 mostly disagrees with this statement (81%). 50% of the content provider respondents from each demonstrator somewhat disagree or completely disagree with this statement, and less than 25% support this statement (**Figure 76**).

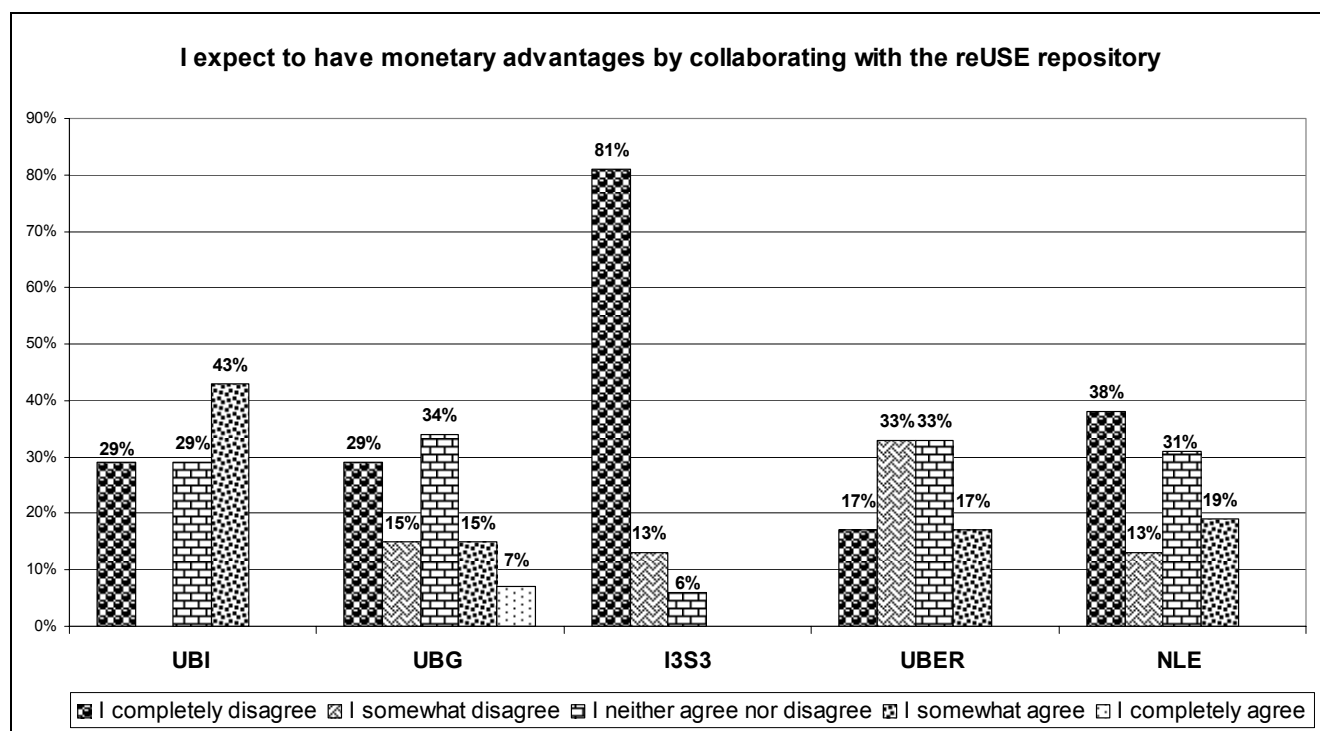


Figure 76: Expectations concerning monetary advantages by collaborating in the reUSE project

As the digital repository grows, it is expected that costs as well as investments will increase. An interesting business model where some of the content providers would be prepared to pay for the services could be established. This means that they would also contribute financially to the digital collection development. Thus, the content providers were asked if they would be prepared to cover the repository expenses. Considering the high percentage of respondents that completely agreed or somewhat agreed with this statement, this really shows that they are interested in the services provided by the repositories. At UBG and UBER 10% and 17% (respectively) of the respondents care about the expenses (**Figure 77**).

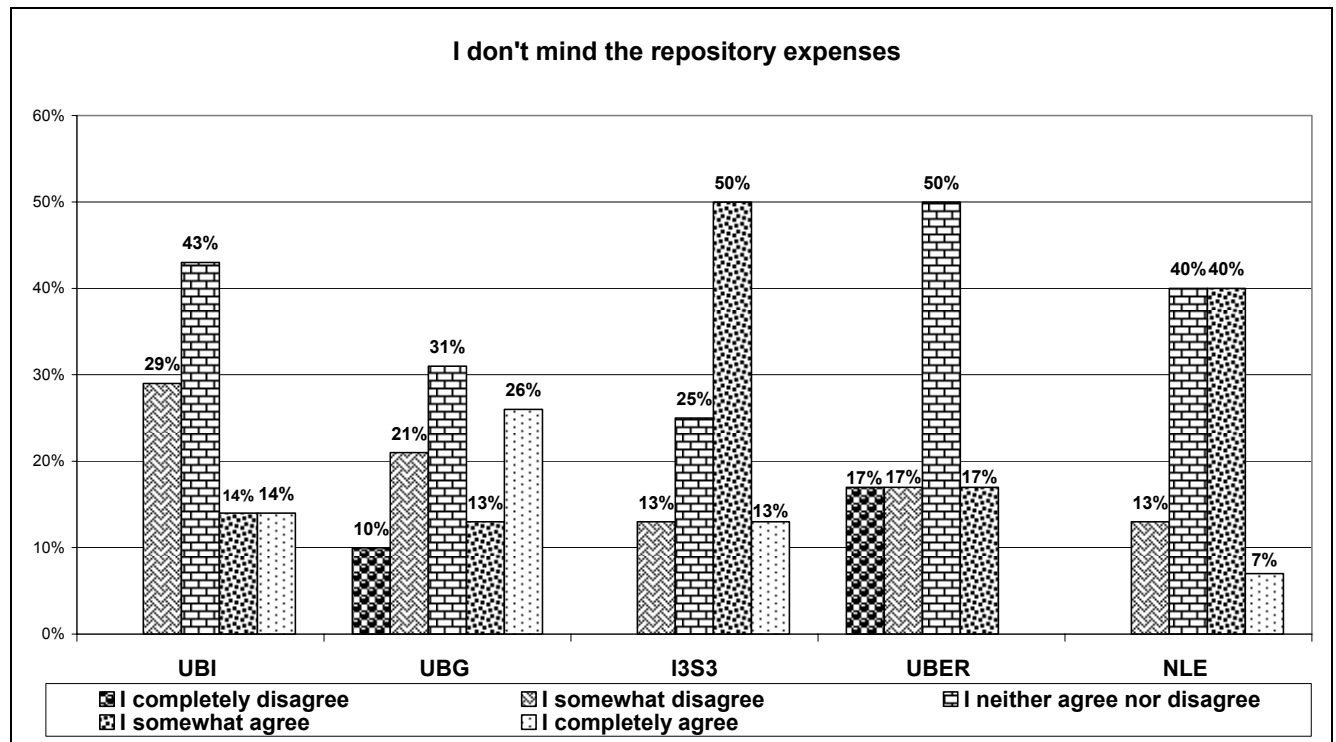


Figure 77: Opinion of the content providers respondents on the expenses for the repository.

Nevertheless content providers are expected to have to invest additional staff or efforts in the cooperation with reUSE demonstrators. This aspect also seems to show in the organizational costs. At UBI 29%, at UBG 34% at I3S3 31% and at NLE 20% of all respondents completely disagreed with this statement (**Figure 78**). At UBG 20%, at I3S3 38% at UBER 33% and at NLE 13% of respondents somewhat disagreed. At UBER 34% of respondents, at UBI 43%, at UBG 22% and at NLE 47% expected to save money by cooperating in this project.

These responses can be understood as a consequence of the fact that content providers expect to have expenses from collaborating in the reUSE project, but that the benefits they receive are more important to them than the costs.

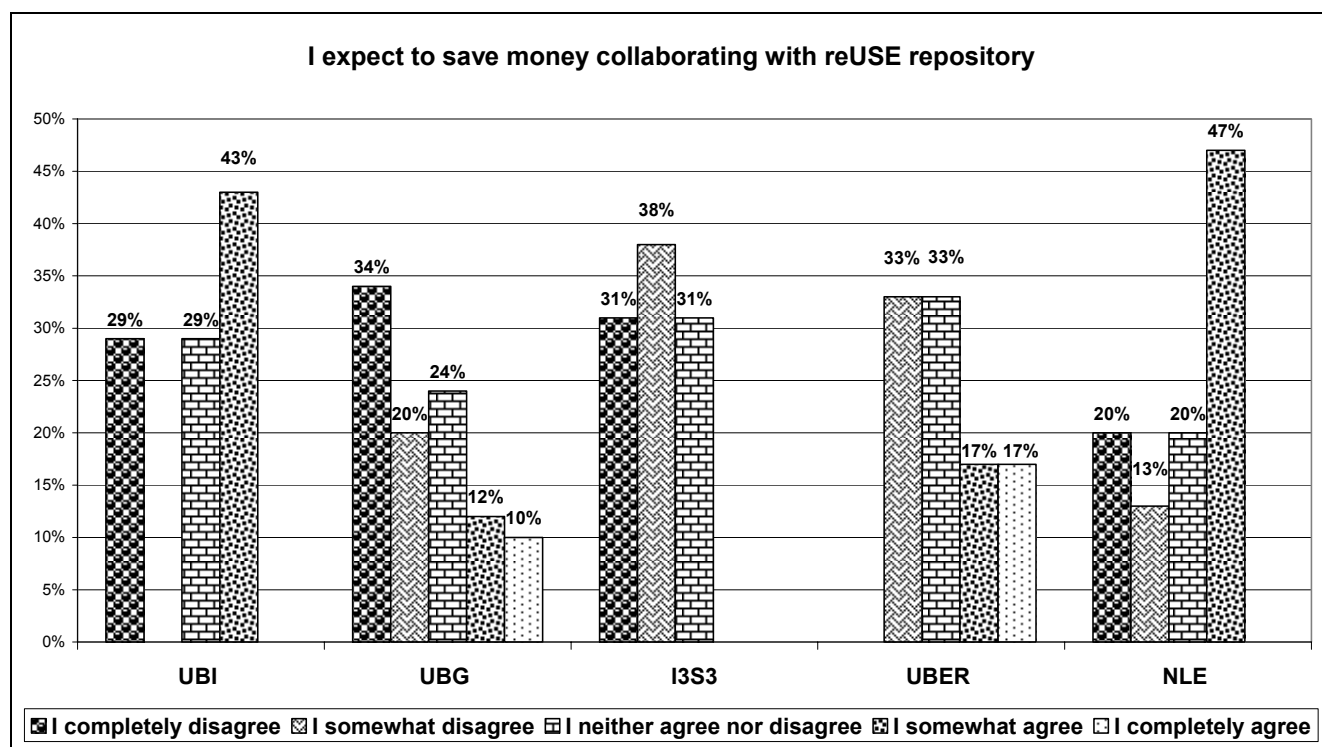


Figure 78: Content provider respondents' expectations concerning saving money with the reUSE project

What about the new workflows that are expected from the content providers? Is it the same as with the printed publication? At UBG 44%, at I3S3 38%, at UBI 29%, and at NLE 13% of the respondents are convinced that they do not need additional time to produce electronic versions of their printed materials. The reason lies in the fact that these organizations help content providers in preparing digital contents for their ingest in the repository. At UBER about 34% of the respondents think that this could be an important issue (**Figure 79**), due to the very strict rules for content delivery that the content providers from the academic sector have to follow. UBER rejects digital publications that are not prepared according to their guidelines or their templates.

Are the negotiation processes between content providers and demonstrators too complicated? It doesn't seem so. 50% of respondents and more declared that they completely disagreed with this statement. (**Figure 80**)

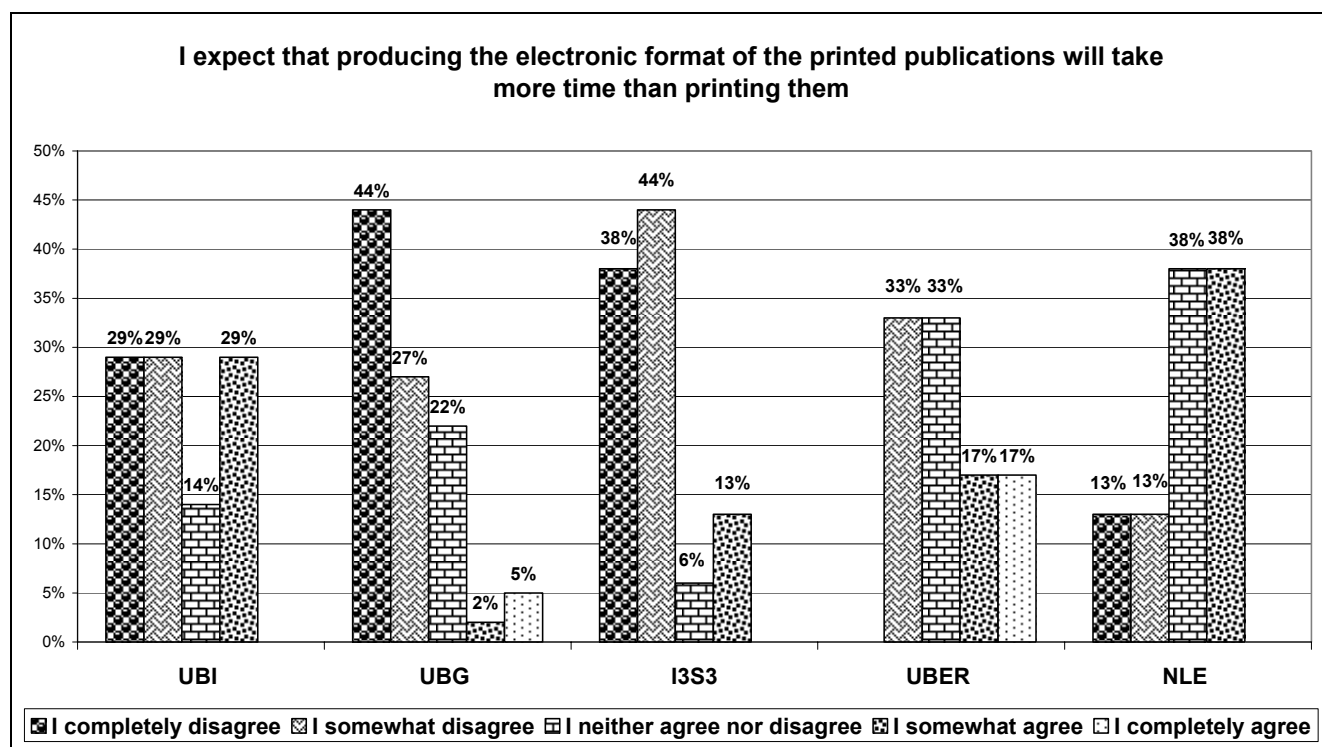


Figure 79: Content provider respondents' expectations regarding investing more time in producing the electronic format of printed publications

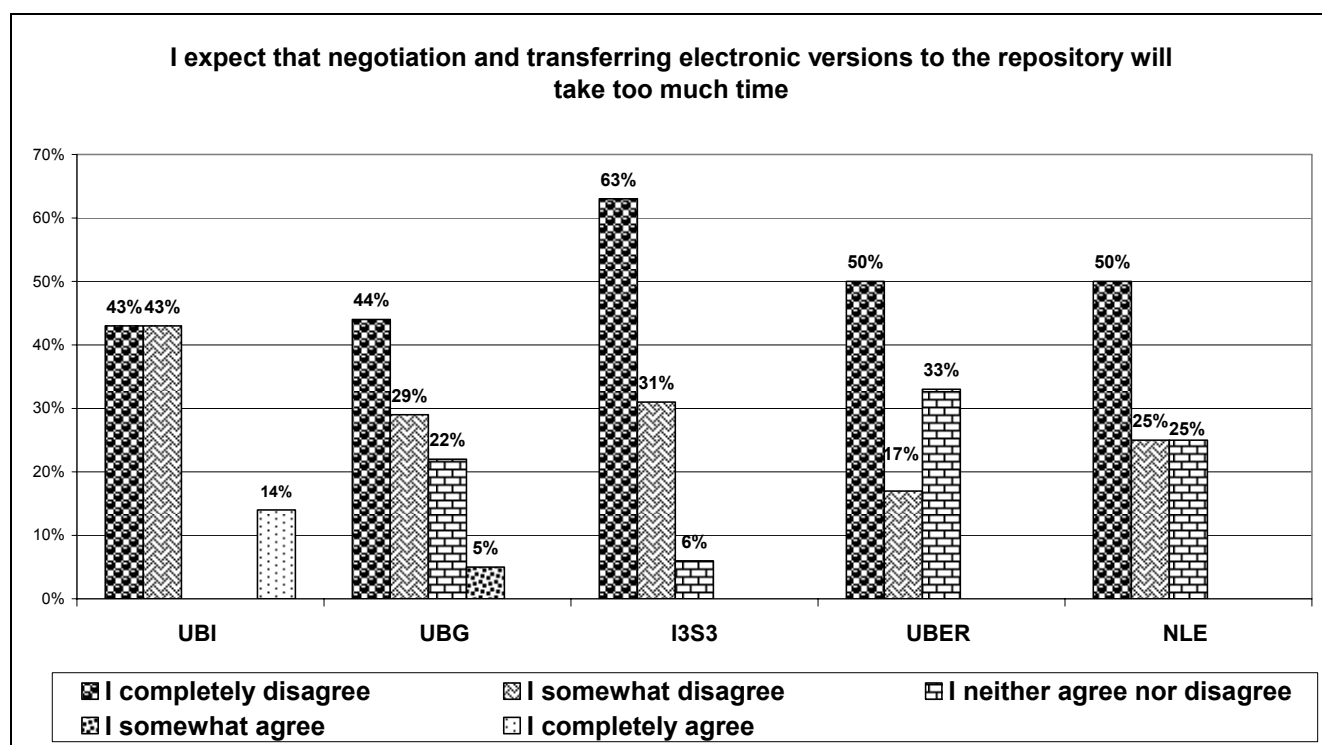


Figure 80: Content providers respondents' expectations concerning negotiations and transfer of electronic versions

4.4 LIBRARIANS IN CHARGE OF DIGITAL REPOSITORIES

We evaluated the time and staff needed for the delivery and processing of the electronic documents in the repository as well as the time needed to access it.

NLE and UBI consider that the digital repository requires additional work and time from their staff. For NLE this means additional activities concerning negotiating and acquiring electronic master files of printed materials from content providers, processing the documents in the repository, etc. However, they claimed to need only 8 hours for processing a publication in the digital repository and one whole week for processing a printed book.

UBI deems personal contacting and negotiations with content providers the most time consuming.

I3S3 feels that loading documents in the repository (metadata creation) and getting documents from publishers and institutions requires additional work and time from their staff. They need approximately 4 minutes for processing an electronic publication, which is much less than in other libraries.

When asked what kind of benefit they expected from the new digital repository, they replied following:

NLE: Better access to collections, increased cooperation with content providers.

UBI: Satisfied organisations, contribution to setting-up digital libraries.

I3S3: Increased services for library readers and a better reputation of the library.

4.5 REVIEW OF THE FINDINGS OF THE COST BENEFIT ISSUES EVALUATION

4.5.1 FINDINGS OF THE SURVEY ON CONTENT PROVIDERS

- The majority of content provider respondents contributed their contents to reUSE repository mainly to establish the access and dissemination of their printed documents. Access related services should therefore be promoted and developed.
- The access to publications for disabled people is a very important motivation factor for approx. 50% of the respondents. This could be a very important value-added service provided by the partners.
- Most of the respondents are aware that long-term preservation is a very important task. From 80% to 100% of all respondents consider this their main reason for contributing to the reUSE repository.
- The content provider respondent do not see reUSE repositories as a way of achieving strategic goals but rather of providing access to their contents. However, the possibility of increasing interest in their organizations as well as increasing dissemination of their activities and work was very highly assessed. They are less convinced that reUSE repositories would be able to increase the recognition of their organization among their stake-holders and furthermore that they would attract new sponsors.
- Most of the respondents do not expect any profit. They think that the benefits for their stake-holders would be more important than the costs. They neither expect any monetary advantages or money saving. This means that to some of the content

providers this collaboration represents benefits more important than its costs. Broader access and dissemination as well as long-term preservation are very good reasons for collaboration. Demonstrators could think about attracting sponsors and commercial partners.

- Delivery system and negotiations between the demonstrators and the content providers are functioning very well and do not require much additional time and effort from the content providers. This is also a very good basis for attracting new content providers.

4.5.2 FINDINGS OF THE SURVEY ON THE LIBRARIANS

- According to the data received it seems that the processing of electronic documents requires less time than the processing of printed materials. We think that this is due to the fact that printed materials usually follow a processing line through specialized departments (acquisition, cataloguing, documentation... etc.) while electronic documents are ingested faster and the metadata can even be automatically extracted. Where the metadata schemas are not as complex as the printed materials the processing could last even 4 minutes.
- Negotiations and document delivery require additional work and time.

5. SWOT ANALYSIS

By Alenka Kavčič-Čolić, Mateja Šmid, Darko Majcenović

ORGANIZATIONAL ASPECTS AND COST-BENEFIT ASPECTS	
ALO	
STRENGTHS	<ul style="list-style-type: none"> • The repository is jointly built and supported by several university libraries. • The libraries have a regional deposit function and as such hold a very important role in their local environment. • They were able to promote the basic concepts of reUSE and attract many content providers. • There is a high political awareness in Austria which supports long-term preservation and is positive in regard to the activities in this field. • The function of a "trusted digital repository" could be a competitive advantage over other digital deposit institutions. • The repository fulfils most mandatory responsibilities of OAIS archives. • Top management is very supportive of the reUSE concept. Some are prepared to support the repository financially and with additional staff. They also have all the necessary infrastructure, which will permit them to develop the system. • They use metadata basic standard schemas, which enable the exchange of metadata with other systems. • Content providers, especially those coming from the academic field have found ALO very attractive. • The delivery system and negotiations between ALO repositories and content providers are functioning very well and do not require much additional time and effort from content providers. This is also a very good basis for attracting new content providers
WEAKNESSES	<ul style="list-style-type: none"> • ALO is not yet a "trusted digital repository". They fulfil most of the mandatory responsibilities, but the overall policy is still access oriented, while long-term preservation is partly neglected. There is no clear digital preservation planning policy in the repository institution (UBI), which is more important, since all content providers expressed this to be the main motivation factor for their contribution to the repository.
OPPORTUNITIES	<ul style="list-style-type: none"> • Collaboration with other university libraries could be an opportunity to establish a more secure back-up system by producing mirror sites for digital contents. • After the end of the reUSE project ALO contents will be transferred into a new commercial software. This could be a good opportunity to improve the public services. • Additional service for access for disabled people could be developed, which could increase the number of stake-holders in the repository. • For content providers the access to and dissemination of their publications seem to be more important than other services. This could be an important aspect for marketing the repository. • Most of the content providers think that the benefits for their stake-holders would be more important than the costs. Knowing this, the demonstrator partners of ALO could give some thought to attracting sponsors and commercial partners. They can also consider the possibility of new business models for providing services to content providers that could cover the main expenses of the repositories.

	<ul style="list-style-type: none"> • Simplification of negotiations with content providers can save ALO staff time and efforts.
THREATS	<ul style="list-style-type: none"> • Some of the top managers are prepared to support the repositories financially and with additional staff. However, some of them declared that additional external funds should be localized in order to finance the repositories after the completion of the reUSE project. These funds could not be enough for normal operation and maintenance of the repository. • A new software could increase the operational and maintenance costs as well as diminish the quality of services.
EDOC	
STRENGTHS	<ul style="list-style-type: none"> • EDOC is building a thematic repository, which is maintained by the University of Berlin. • It has been certified by DINI and has the competitive advantage as digital deposit institution. • EDOC has a very well defined preservation planning policy. • The repository fulfils most mandatory responsibilities of OAIS archives. • Top management is very supportive of the reUSE concept. • They use metadata basic standard schemas, which enable the exchange of metadata with other systems. • UBER has been very successful in cooperating with content providers. They mainly come from the academic sector and are very interested in contributing to EDOC. • They have very precise standards for document delivery, and thus organize content providers' training.
WEAKNESSES	<ul style="list-style-type: none"> • Very clear and strict delivery standards could require more time and efforts from content providers.
OPPORTUNITIES	<ul style="list-style-type: none"> • UBER has a very clear designated community; it is easier to adjust its services and fulfil their needs. • Additional service for access for disabled people could be developed, which could increase the number of stake-holder in the repository. • For content providers the access to and dissemination of their publications seem to be more important than other services. This could be an important aspect for marketing the repository. • Most of the content providers think that the benefits for their stakeholders would be more important than the costs. Knowing this, UBER could give some thoughts to attracting sponsors and commercial partners. They can also consider the possibility of new business models for providing services to content providers that could cover the main expenses of the repositories. • Simplification of the negotiations with content providers can save the EDOC staff much time and effort.
THREATS	<ul style="list-style-type: none"> • In spite of the strong support from the top management, additional funds are expected to be localized from external sources like EU or DFG. They could not be enough for the maintaining and operating of the repository.
DIGAR	
STRENGTHS	<ul style="list-style-type: none"> • As the national library it is the main deposit institution for all media publications, this fact is supported by the Legal Deposit Law. • The repository was built from scratch and represented a very good opportunity to design the organizational policy and strategy from the beginning. NLE has followed all the goals set by the White Paper and the valid standards set by the digital preservation expert community. • The organizational structure is flat and top management is more embedded in the digital repository activities, which has a big impact on the strategy and budget planning processes.

	<ul style="list-style-type: none"> NLE has a very well defined preservation planning policy. The repository fulfils most of the OAIS archives mandatory responsibilities. NLE top management is very supportive of the reUSE concept. They are prepared to support it with additional staff, information technology and software. A special budget is foreseen for the funding of the repository after completion of the reUSE project. They use metadata basic standard schemas, which enable the exchange of metadata with other systems. The delivery system and negotiations between DIGAR and content providers are functioning very well and do not require much additional time and effort from the content providers. This is also a very good basis for attracting new content providers
WEAKNESSES	<ul style="list-style-type: none"> As an open source system a lot of efforts and more staff members, training and information technology has to be invested in DIGAR. They should establish a very strong IT department to implement, develop and support the system.
OPPORTUNITIES	<ul style="list-style-type: none"> Content providers expressed a strong interest for disabled people access to the digital contents. Should it be developed, it could increase the number of stake-holders in the repository. For content providers the access to and dissemination of their publications seem to be more important than other services. This could be an important aspect for marketing the repository. Most of content providers think that the benefits for their stake-holders would be more important than the costs. Knowing this NLE could think about attracting sponsors and commercial partners. They can also have in mind the possibility of new business models for providing services to content providers that could cover the main expenses of the repositories. Simplification of the negotiations with content providers can save DIGAR staff time and effort.
THREATS	<ul style="list-style-type: none"> External threats could come from change in the national policy regarding long-term preservation of digital documents. It could result in lower budget support for these activities. Internal threats could come from change of NLE management, which could set different priorities and therefore change the financing schema of the repository.
TECHNICAL ASPECTS	
ALO	
STRENGTHS	<ul style="list-style-type: none"> Good disaster readiness planning. Good limitations of access to servers and databases. 3 partners (UBI, UBG and i3s3) cooperate in the project, which means higher contribution to the development of the system Access to the repository via OPAC. ALO is designed as an open-source package available for free. It does not allow submitters to work directly with the repository system which guarantees more security and reliably ingested content. It provides print-on-demand and CD delivery as added-value services. It offers multiple language support and indexation by search engines.
WEAKNESSES	<ul style="list-style-type: none"> Backup locations in one closely related area in Innsbruck. Web interface design is outdated. Lack of online help for the repository use and functions. Limited accessibility for disabled people. Users cannot customize their profile and the user interface to suit their

	<ul style="list-style-type: none"> needs. No UNICODE extended character support. Too few fields to search with.
OPPORTUNITIES	<ul style="list-style-type: none"> It could provide backups on at least two highly separate locations because of participating partners. It could achieve better user satisfaction by providing customization of user interface, search results and saving personal preferences for each user. Could have better informed users by sending email news of new additions to the repository. Could implement a list of pending submissions and approved submissions so that the contributors can easily see the progress of submission process. Could provide more ample advanced search with more fields and display options.
THREATS	<ul style="list-style-type: none"> What happens when contents providers begin providing multimedia content? What happens if the contents providers wish to store their webpage in the repository and have it periodically web-harvested? There currently is no such option available. Possible future problems with file format conversion and migration because of acceptance and storage of most file formats contributed. There is currently no additional support for disabled users. Value-added functions should be extended to utilize repository contents and also to create profit to help fund the repository.
EDOC	
STRENGTHS	<ul style="list-style-type: none"> Good disaster readiness planning. Good limitations of access to servers and databases. A long-term preservation strategy based on predefined contents templates allows file types and content organization and control. Separate institutions for contents selection and technical solutions allowing division of work for better efficiency. Value-added print-on-demand service and CD delivery and content can be indexed by external search engines Web interface to the repository is compliant with Web Content Accessibility Guidelines 2.0, level 1. Basic level of accessibility. Advanced user management. Anti-virus, data format and digital signatures checks included into ingest workflow. Ability to generate a variety of file formats from XML based archival object. Full-text search capabilities (including highlighting search terms).
WEAKNESSES	<ul style="list-style-type: none"> Template based contents provision can mean more work for contents providers. Users cannot customize their profile or the user interface to suit their needs. Too few fields to search in.
OPPORTUNITIES	<ul style="list-style-type: none"> Could achieve better user satisfaction by provision for customization of user interface, search results and saving personal preferences for each user. Could have better informed users by sending email news of new additions to the repository. Could implement a list of pending submissions and approved submissions so that the contributors can easily see the progress of submission process. Could provide more ample advanced search with more fields and display options.

THREATS	<ul style="list-style-type: none"> • Separate institutions managing the repository can mean delays in updates and developing solutions. • What happens when contents providers begin providing multimedia content? How will contents templates be adapted? • What happens if the content provider wishes to store their webpage in the repository and have it periodically web-harvested? There are currently no such options available. • There is currently no additional support for disabled users. • Value-added functions should be extended to utilize repository content and also create profit to help fund the repository.
DIGAR	
STRENGTHS	<ul style="list-style-type: none"> • Good disaster readiness planning. • Good limitations of access to servers and databases. • Use of open source software solution for the repository means building on a pre-existing base. • Advanced user management and different views on objects for different user types • Considers the user suggestions in updates for the next development phase. • Allows submitters to view their history of submissions which indicates the status of submissions.
WEAKNESSES	<ul style="list-style-type: none"> • Presentation of search results lacks sorting. • Users do not use advanced search method. • Accessibility for disabled people. • Users cannot customize their profile and the user interface to suit their needs. • Few dissemination formats. • To few fields to search in.
OPPORTUNITIES	<ul style="list-style-type: none"> • Improve user satisfaction with better search results presentation and more promotion of advanced search method. • Could achieve better user satisfaction by provision for customization of user interface, search results and saving personal preferences for each user. • Could have better informed users by sending email news of new additions to the repository. • Could implement a list of pending submissions and approved submissions so that the contributors could easily see the progress of the submission process. • Could provide more ample advanced search with more fields and display options.
THREATS	<ul style="list-style-type: none"> • Users do not get the exact data they requested due to use of to simple search methods. • How will the repository behave in the future with increased workload and number of users? • What happens when contents providers begin providing multimedia contents? • What happens if the contents providers wish to store their webpage in the repository and have it periodically web-harvested? There currently is no such option available. • There is currently no additional support for disabled users. • Value-added functions should be extended to utilize repository content and also to create profit to help fund the repository.
REPOSITORIES FROM THE USERS' PERSPECTIVE	

ALO	
STRENGTHS	<ul style="list-style-type: none"> • The collections are easily accessible. • Information in the repository is displayed quickly enough. • Many respondents were very positive about the repository. They said that the web site was nice, pleasant and practical. • The structure of the contents is easy to use. • Designs, such as icons, signs, etc. are clear and understandable (end-users like the design of the website). • Half of the end-users are so satisfied with the repository that they intend to come back when is need of further information. • In case of mistakes, there are no difficulties with restoring everything to its previous state or undoing the last operation. • Most end-users are quite satisfied with the contents. • The ALO digital repository also offers some full-texts adapted for blind users.
WEAKNESSES	<ul style="list-style-type: none"> • Too many details and too many steps/clicks are needed in order to perform a task. • There are not enough OCR texts for blind end-users in the ALO digital repository (should be improved for blind users - more OCR scanned texts should be available). • The ALO repository was not familiar enough among users-especially among young people in UBG, therefore more efficient publicity and dissemination should be done. • More full-text documents from different topics must be added (especially handbooks, manuals, reference books, professional and scientific literature). • Sometimes layout problems occur (depending on the screen). Is the layout suited for different size of screens? • Digital repository could be improved in order to be more simple, efficient (regarding key-words search) and user-friendly. • There is no pop-up-information on mouse over. • Some end-users already made such an entry that caused errors.
OPPORTUNITIES	<ul style="list-style-type: none"> • The repository could visually mark the current entry location (and hits as well) by highlights or contrasting colour. • Rising awareness of the advantages offered by value-added services (i.e. contents accessible to blind people) would definitely pay off. It could essentially improve the repository, and even benefit the organization. Since the digital repository already has some blind users, developing this kind of services should be taken as an opportunity. • To provide an option for adjusting the amount of information displayed on screen it would be beneficial to end-users. • Users would appreciate e-mail notification about new entries in the digital repository. • The fact that the same e-documents can be found also in the library catalogue should be considered in order to extend the end-user group. • In order to attract more end-users, results could be sorted by date or by field. Subject and full-text search would be beneficial as well. • Since the main target group is academic staff, researchers and students, the digital repository should focus on the scientific and professional contents. • A help system or guide for efficient searching, navigating and printing might be helpful to some end-users. • Both versions of documents, on electronic and paper media, do not necessary rule out each other and thus leave many opportunities for specific formats and various contents. The implementers should recognize which format suits what contents; it is of great importance to find a niche for e-documents.

	<ul style="list-style-type: none"> We recommend a complete list of subjects contained in the archive as well as a list of new documents for the previous month.
THREATS	<ul style="list-style-type: none"> The end-users with a background in human sciences or architecture are not so enthusiastic about the e-documents. Different formats for different contents are needed and need to be developed. Because of the lack of content there is a risk of losing many end-users. More full-text documents from different topics must be added (especially handbooks, manuals, reference books, professional and scientific literature). The interface is not simple enough for young end-users (the so-called Google generation). University libraries could start losing their key role and the influence they used to have on their end-users. Due to the lack of OCR scanned texts, there is a threat to lose blind end-users.
EDOC	
STRENGTHS	<ul style="list-style-type: none"> The collections are easily accessible. Information in the repository is displayed quickly enough. The EDOC digital repository already offers the following value-added services: <ul style="list-style-type: none"> different types of hosting, print-on-demand service (ProPrint), technology, software and know-how transfer. Many respondents had a very good opinion about the repository. They said the web site was nice, pleasant and practical. the structure of the contents is easy to use. The designs such as icons, signs, etc are clear and understandable (end-users like the design of the website). Approx. half of the end-users are so satisfied with the repository that they intend to come back when is need of further information. In case of mistakes, there are no difficulties with restoring everything to its previous state or undoing the last operation. EDOC provides efficient help system, which contains basic information and assistance concerning navigating the EDOC repository, viewing and printing PDF documents, and searching.
WEAKNESSES	<ul style="list-style-type: none"> There is no list of the whole content by subject field on the interface. Too many details and too many steps/clicks are needed in order to use the repository properly. It is hard to find the EDOC web site; more publicity should be done. More full-text documents from different scientific fields must be added. Separate search functionality per subject would be needed. Some of the EDOC end-users claimed that even small mistakes could cause system errors. The digital repository could be improved in order to be more simple, efficient (regarding key words search) and user-friendly.
OPPORTUNITIES	<ul style="list-style-type: none"> The repository could visually mark the current entry location (and hits as well) by highlights or contrasting colour. Sorting via subject for the qualification/habilitation papers would be needed in order to attract more end-users. We recommend a complete list of subjects contained in the archive as well as and a list of new documents for the previous month. It would be beneficial to end-users to provide an option for adjusting the amount of information displayed on-the screen. Users would appreciate e-mail notification about new entries in the

	<p>digital repository.</p> <ul style="list-style-type: none"> • The fact that the same e-documents can also be found in the library catalogue should be considered in order to extend the end-user group. • Users would appreciate ensuring separate search functionality per subject. • Since the main target group is academic staff, researchers and students, the digital repository should focus on the scientific and professional contents. • Both versions of documents, on electronic and paper media, do not necessary rule out each other and thus leave many opportunities for specific formats and various contents. The implementers should recognize which format suits what contents; it is of great importance to find a niche for e-documents.
THREATS	<ul style="list-style-type: none"> • The interface is not simple enough for young end-users (the so-called Google generation). • Due to the lack of contents there is a risk of losing many end-users. More full-text documents must be added (especially professional and scientific literature). • University libraries could start losing their key role and the influence they used to have on their end-users.
DIGAR	
STRENGTHS	<ul style="list-style-type: none"> • Many respondents had a very good opinion on the repository. They said the web site was nice, pleasant and practical. • The collections are easily accessible. • The structure of the contents is easy to use. • Information in the repository is displayed quickly enough. • The designs such as icons, signs, etc. are clear and understandable (end-users like the design of the website). • Most of the end-users are so satisfied with the repository that they intend to come back when is need of further information. • In case of mistakes, there are no difficulties with restoring everything to its previous state or undoing the last operation. • The same e-documents can also be found in the library catalogue. • It's easy to learn how to use the digital repository (trainings are offered by the library staff). • It is simple to use the digital repository. • Most end-users are quite satisfied with the contents.
WEAKNESSES	<ul style="list-style-type: none"> • There is no list of the whole contents by subject field on the interface. • Many details and too many steps/clicks are needed to open the documents. • the digital repository doesn't contain enough full-text documents from different topics. • Search and display of the results should be improved in order to be more accurate. • Searching by subject is missing. • Subject index is missing. • The text font on the archive home page is too small. • DIGAR could be improved in order to be more efficient (regarding key-word search) and user-friendly. • Search by author was not successful. • Some end-users already made such an entry that caused errors.
OPPORTUNITIES	<ul style="list-style-type: none"> • Rising awareness of the advantages offered by value-added services (i.e. contents accessible for disabled people) would definitely pay off. It could essentially improve the repository, and even benefit the organization. • Publications on the repository home page should be linked to the

	<p>publishers.</p> <ul style="list-style-type: none"> • Users would appreciate a complete list of subjects contained in the archive as well as a list of new documents for the previous month. • A help function could be included in order to guide end-users in their search by authors, subject, etc. It should be available also in Russian. • In order to attract more end-users, results could be sorted by date or by field. Subject (keywords) and full-text search would also be beneficial. • Repository could visually mark the current entry location (and hits as well) by highlights or contrasting colour. • Due to the diversity of end-users it would be wise to provide an option to adjust the amount of information displayed on screen for the end-users' needs. • Users would appreciate e-mail notification about new entries in the digital repository. • National libraries still play a key role and have an important influence on their end-users and non-users. This enables them to provide training on the use of the digital repository and keep users satisfied with their services. • Many different kinds of documents can be added (for more details see the final report– section: End-users perspective: what end-users wish to find in the repository). • Both versions of documents, on electronic and paper media, do not necessary rule out each other and thus leave many opportunities for specific formats and various contents. The implementers should recognize which format suits what contents; it is of great importance to find a niche for e-documents.
THREATS	<ul style="list-style-type: none"> • The end-users with a background in human sciences or architecture are not so enthusiastic about e-documents. Different formats for different contents are needed and need to be developed. • Due to the lack of contents there is a risk of losing many end-users. More full-text documents from different topics must be added (especially handbooks, manuals, reference books, professional and scientific literature). • If the contents in the digital repository of the national library is not heterogeneous enough, some end-users could be lost.

6 REFERENCES

REPOSITORIES

- CCSDS (Consultative Committee for Space Data Systems) (2002): *Reference Model of an Open Archival Information System (OAIS) : Recommendation for space data system standards* : CCSDS 650.0-B-1.Blue Book, January 2002. - Washington : CCSDS Secretariat. URL: http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html.
- NEDLIB Project homepage: <http://www.kb.nl/coop/nedlib/>
- Steenbakkens, J. (2000): *The Nedlib Guidelines : Setting up a Deposit System for Electronic Publications*. - Den Haag : Koninklijke Bibliotheek. - (NEDLIB Report series ; 5).
URL: <http://www.kb.nl/coop/nedlib/results/NEDLIBguidelines.pdf>.
- Werf, T. van der (2000): *The Deposit System for Electronic Publications: A process Model*. - Den Haag : Koninklijke Bibliotheek. - (NEDLIB Report Series ; 6). URL: <http://www.kb.nl/coop/nedlib/results/DSEPprocessmodel.pdf>.
- *Trusted Digital Repositories: Attributes and responsibilities : an RLG/OCLC Report*. (2002) – Mountain View, CA : RLG. – URL : <http://www.rlg.org/longterm/repositories.pdf>.
- FEDORA (Flexible Extensible Digital Object and Repository Architecture). Mellon Foundation Project, University of Virginia and Cornell University, <http://www.fedora.info/>.
- White Paper on Digital Repositories (reUSE! reUSE Digital Master Files of Printed Material!), http://www2.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf.

METADATA

- CCSDS (Consultative Committee for Space Data Systems) (2002): *Reference Model of an Open Archival Information System (OAIS) : Recommendation for space data system standards* : CCSDS 650.0-B-1.Blue Book, January 2002. - Washington : CCSDS Secretariat. URL: http://nssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html.
- Day, M. (1998): CEDARS: Digital preservation and metadata. *Sixth DELOS Workshop: Preservation of Digital Information, Hotel dos Templários, Tomar, Portugal, 17-19 June 1998*. URL: <http://www.ukoln.ac.uk/metadata/presentations/delos6/cedars.html>.
- Dublin Core: <http://www.dublincore.org/documents/dces/>
- IFLA: <http://www.ifla.org/II/metadata.htm>
- Lupovici, Catherine and Masanès, Julien (2000): *Metadata for Long Term Preservation*. - Bibliothèque Nationale de France. - Den Haag : Koninklijke Bibliotheek, 2000. - (NEDLIB Report series ; 2).
URL: <http://www.kb.nl/coop/nedlib/results/NEDLIBmetadata.pdf>

- METS (Metadata Encoding and Transmission Standard: <http://www.loc.gov/standards/mets/>)
- MODS (Metadata Object Description Schema): <http://www.loc.gov/standards/mods/>
- OCLC/RLG Working Group on Preservation Metadata (2002): *Preservation Metadata and the OAIS Information Model. A metadata framework to support the preservation of digital objects.* http://www.oclc.org/research/projects/pmwg/pm_framework.pdf
- PANDORA: <http://www.nla.gov.au/metadata.html>
- RLG & OAIS <http://www.rlg.org/longterm/oais.html>
- UNIMARC: <http://www.ifla.org/VI/3/p1996-1/sec-uni.htm>
- Open Archives Initiative Protocol for Metadata Harvesting, <http://www.openarchives.org/>

LONG-TERM PRESERVATION

- National Library of Australia (2003): *Guidelines for the preservation of digital heritage.* - Paris : Unesco. URL: <http://unesdoc.unesco.org/images/0013/001300/130071e.pdf>

STANDARDS

Digital repository standards:

- ISO 14721:2003. See: CCSDS (Consultative Committee for Space Data Systems) (2002): *Reference Model of an Open Archival Information System (OAIS) : Recommendation for space data system standards* : CCSDS 650.0-B-1. Blue Book, January 2002. - Washington : CCSDS Secretariat. URL: http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html (last viewed 5.4.2004).

User centered design standards:

- (<http://www.usability.serco.com/trump/resources/standards.htm#9126-2>)
- ISO/IEC 9126 for internal SW quality
- ISO 9241-11 for usability
- ISO 13407 I for human centred design processes for interactive systems

DINI Certificate

Dobratz, Susanne: DINI Institutional Repository Certification and Beyond. <http://conference.ub.uni-bielefeld.de/2006/programme/abstracts/dobratz.htm>)

OTHER

- Clavel-Merrin, Genevieve: List of NEDLIB Terms. - Den Haag : Koninklijke Bibliotheek, 2000. - (NEDLIB Report series ; 7). Available at URL: <http://nedlib.kb.nl/results/NEDLIBterms.pdf> (last accessed 2006-3-10)

- Krimbacher, Monika, Neuhauser, Michael, Vogl, Martina: Survey on the long-term preservation of digital documents in European libraries 2005. - Innsbruck : ARGE Kulturconsulting - Medienkonzepte -Marktforschung ; Innsbruck : reUSE Project, 2005.
- A Workshop on Cost Models for preserving digital assets organized by DCC/DCP took part at the British Library on 26 July 2005. The report and some of the papers are available at URL: <http://www.dpconline.org/graphics/events/050726workshop.html>.
- Universal Preservation Format Glossary, <http://info.wgbh.org/upf/glossary.html>
- VARNER, Philip E. The Economics of Open Source Software.
http://www.cs.virginia.edu/~pev5b/writing/econ_oss/advantages.html (last accessed 2005-11-24)
- Watkins, Thayer: An introduction to cost benefit analysis.
<http://www2.sjsu.edu/faculty/watkins/cba.htm>]

ANNEX 1

USED ABBREVIATIONS

AIP – Archival Information Package

ALO – Austrian Literature Online

CRC – Cyclical Redundancy Check

DBMS – Data Base Management System

DC – Dublin Core

DDB – National Library of Germany (<http://www.ddb.de/>)

DELOS - DELOS Network of Excellence on Digital Libraries (<http://www.delos.info/>)

DGF – German Research Foundation

DIGAR – Digital Archive of The National Library of Estonia

DINI – Deutsche Initiative für Netzwerkinformation

DIP – Dissemination Information Package

DSEP – Deposit System for Electronic Publications

EDOC - edoc servers of the Humboldt University of Berlin (<http://edoc.hu-berlin.de>)

FEDORA - Flexible Extensible Digital Object and Repository Architecture

FTE – Full Time Equivalent

GIF - Graphics Interchange Format

GPL – The GNU General Public License

I3S3 – Interuniversitäres Institut für InformationsSysteme zur Unterstützung Sehgeschädigter Studierender / Austrian Institute for Information Systems and Support Services for Blind and Visually Handicapped University Students (<http://www.integriert-studieren.jku.at/>)

ISBN – International Standard Book Number

ISO – International Organization for Standardization (<http://www.iso.org/>)

LJU – Faculty of Civil and Geodetic Engineering of the University of Ljubljana (<http://www1.fgg.uni-lj.si/>)

LMER – Langzeitarchivierungsmetadaten für elektronische Ressourcen [Long-Term Preservation Metadata for Electronic Resources], <http://www.ddb.de/standards/lmer/>.

LTP – Long-Term Preservation

NEDLIB – Networked European Deposit Library (<http://www.kb.nl/coop/nedlib/>)

NESTOR - Network of Expertise in Long-Term Storage of Digital Resources

NGO – Non-Government Organization

NLA – The National Library of Australia,

NLE – The National Library of Estonia (<http://www.nlib.ee/html/inglise/indexi.html>)

NLNZ – The National Library of New Zealand

NUK – The National and University Library of Slovenia (<http://www.nuk.uni-lj.si/>)

MAB2 – Maschinelles Austauschformat für Bibliotheken

METS - Metadata Encoding & Transmission Standard

OAIS – Open Archival Information System Reference Model (ISO 14721/2003)

OCLC – Online Computer Library Center (<http://www.oclc.org/>)

ONB – The Austrian National Library (http://www.onb.ac.at/index_eng.htm)

OAI-PMH – Open Archives Initiative-Protocol for Metadata Harvesting

PDI – Preservation Description Information

RAID - A redundant array of independent disks

RLG – Research Libraries Group, Inc. (<http://www.rlg.org>)

SciX – Open, self organising repository for scientific information exchange.

SIP – Submission Information Package

UBG – The University Library of Graz (<http://ub.uni-graz.at/>)

UBI – The University Library of Innsbruck (<http://www.uibk.ac.at/index-en.html>)

UBER – The University Library of Berlin (<http://www.ub.hu-berlin.de/>)

UML – Unified Modelling Language

UNICODE – Universal Code

VNC – Virtual Network Computing

W3C - World Wide Web Consortium

WWW – World Wide Web

XML – Extensible Mark-up Language

ANNEX 2

GLOSSARY

Access control

The process of determining whether an identity is permitted to perform some action, such as accessing a resource. It requires prior identification and authentication of the identity. A single user may have several identities with differing permissions.

[NEDLIB List of Terms, p.3, <http://www.kb.nl/coop/nedlib/results/NEDLIBterms.pdf>]

Archival Information Package

An Information Package, consisting of the Content Information and the associated Preservation Description Information (PDI), which is preserved within an OAIS.

[Reference model for an open archival information system (OAIS), p. 1-7

<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>]

Archive

An organization that intends to preserve information for access and use by a Designated Community. [Reference model for an open archival information system (OAIS), p. 1-7

<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>]

Authentication

A process to establish the authenticity of digital material stored in a Deposit System for Electronic Publications for long-term preservation and access purposes. The authentication process is part of preservation handling processes that alter characteristics of the original publication. Preservation handling processes may alter characteristics of the original document. The authentication process incorporated in the preservation handling process enables such changes to be monitored and authenticity degradation to be measured or established.

[NEDLIB List of Terms, <http://www.kb.nl/coop/nedlib/results/NEDLIBterms.pdf>, p.4]

Content providers

Individual authors, publishers and/or institutions that provide their contents to be included in the digital repository.

Context information

The information that documents the relationships of the Content Information to its environment. This includes why the Content Information was created and how it relates to other Content Information objects.

[Reference Model for an Open Archival Information System (OAIS), p. 1-8,

<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>]

Cost-Benefit Analysis

Cost-Benefit Analysis (CBA) estimates and totals up the equivalent money value of the benefits and costs to the community of projects to establish whether they are worthwhile. These projects may be dams and highways or can be training programs and health care systems.

[Thayer Watkins: An introduction to cost benefit analysis.

<http://www2.sjsu.edu/faculty/watkins/cba.htm>]

Data

A reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing. Examples of data include a sequence of bits, a table of numbers, the characters on a page, the recording of sounds made by a person speaking, or a moon rock specimen.

[Reference model for an open archival information system (OAIS), p.1-9,
<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>]

Descriptive information

The set of information, consisting primarily of Package Descriptions, which is provided to Data Management to support the finding, ordering, and retrieving of OAIS information holdings by Consumers.

[Reference model for an open archival information system (OAIS), p. 1-9,
<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>]

Designated community

An identified group of potential consumers who should be able to understand a particular set of information. The designated community may be composed of multiple user communities.

[Reference model for an open archival information system (OAIS), p. 1-10,
<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>]

Digital archive

An organization that intends to preserve digital information for access and use by a designated community.

Digital object

An object composed of a set of bit sequences.

[Reference model for an open archival information system (OAIS), p. 1-10,
<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>]

Digital repository

A network-accessible storage system in which digital objects may be stored for possible subsequent access or retrieval.

[adjusted from NEDLIB List of Terms, p. 14, <http://nedlib.kb.nl/results/NEDLIBterms.pdf>]

DINI Certificate

Certificate assigned by the Deutsche Initiative für Netzwerkinformation. The *DINI certificate for document and publication repositories* aims to network document and publication repositories by pushing the use of standards and promoting interoperability and cooperation between German higher education institutions that run digital repositories. The certificate shows potential users and authors of digital documents that a certain quality level in operating the repository is guaranteed and that this distinguishes it from common web servers of institutions. The certificate consists seven criteria. In addition, *DINI* sees its certificate as an instrument to support the Open Access concept.

[Susanne Dobratz: DINI Institutional Repository Certification and Beyond.
<http://conference.ub.uni-bielefeld.de/2006/programme/abstracts/dobratz.htm>]

Dissemination

The transfer from the stored form of a digital object in a repository to a client.

[www.cs.cornell.edu/wya/DigLib/MS1999/glossary.html]

Dissemination Information Package (DIP)

An information package that contains parts or all of one or more Archival Information Packages and that is distributed to the Consumer as requested.

[Reference model for an open archival information system (OAIS), p. 1-10,
<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>]

Dublin Core

The Dublin Core is a metadata element set. It includes all DCMI terms (that is, refinements, encoding schemes, and controlled vocabulary terms) intended to facilitate discovery of resources. The Dublin Core has been in development since 1995 through a series of focused invitational workshops that gather experts from the library world, the networking and digital library research communities, and a variety of content specialties.

[<http://dublincore.org>]

Electronic publication

A document disseminated in machine-readable form. It includes off-line (physical format) publications distributed on CD-ROM or any other portable medium, and on-line publications stored using digital technology. Some electronic publications are 'born digital' (i.e. are created in digital form) and some are created originally in another form e.g. print on paper, photograph, LP and have subsequently been digitised.

[NEDLIB List of Terms, p.7, <http://nedlib.kb.nl/results/NEDLIBterms.pdf>]

Fixity information

The information which documents the authentication mechanisms and provides authentication keys to ensure that the Content Information object has not been altered in an undocumented manner. An example is a Cyclical Redundancy Check (CRC) code for a file.

[Reference Model for an Open Archival Information System (OAIS), p.1-10,
<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>]

Format

The arrangement of data for computer input or output, such as the number and size of data fields in a logical record or the spacing and letter size used in a document.

[Universal Preservation Format Glossary, <http://info.wgbh.org/upf/glossary.html>]

Identifier

A character or sequence of characters designed to uniquely designate a document. Examples of identifiers are URN (Uniform Resource Name), DOI (Digital Object Identifier), ISBN (International Standard Book Number), SICI (Serial Item and Contribution Identifier).

[NEDLIB List of Terms, p.8, <http://nedlib.kb.nl/results/NEDLIBterms.pdf>]

Information Package

The Content Information and associated Preservation Description Information which is needed to aid in the preservation of the Content Information. The Information Package has associated Packaging Information used to delimit and identify the Content Information and Preservation Description Information.

[Reference model for an open archival information system (OAIS), 1-11,
<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>]

Ingest

The OAIS entity that contains the services and functions that accepts Submission Information Packages from Producers, prepares Archival Information Packages for storage, and ensures that Archival Information Packages and their supporting Descriptive Information become established within the OAIS.

[Reference model for an open archival information system (OAIS), p. 1-11,
<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>]

Long-term preservation (LTP)

The act of maintaining information, in a correct and Independently Understandable form, over the Long Term.

[Reference model for an open archival information system (OAIS), p. 1-11,
<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>]

Multi-attribute evaluation

An evaluation that is consisted of multiple examined attributes of the research subject.

Metadata

1. Structured information, perhaps contained in an attached header, that describes other resources. Catalogue records for library materials are a common example of metadata. While the resources are interesting to the end user, the metadata is interesting to the people or programs that have to manage the information. Metadata assists in the process of retrieving information by enabling users to initially discover the existence of the information, to locate it and then to determine if it is the information that the user wants. Usually the metadata describes the contents, physical description, location, type and form of the information, and information necessary for management including migration history, expiry dates, security, authentication, and file formats.
[Universal Preservation Format Glossary: <http://info.wgbh.org/upf/glossary.html>]
2. Metadata are structured, encoded data that describe characteristics of information-bearing entities to aid in the identification, discovery, assessment, and management of the described entities.
[<http://www.ala.org/alcts/organization/ccs/ccda/tf-meta3.html>]

Open Archival Information System (OAIS)

OAIS is an archive, consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community.
[Reference model for an open archival information system (OAIS), 1-11,
<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>]

Provenance information

The information that documents the history of the Content Information. This information tells the origin or source of the Content Information, any changes that may have taken place since it was originated, and who has had custody of it since it was originated. Examples of Provenance Information are the principal investigator who recorded the data, and the information concerning its storage, handling, and migration.

[Reference Model for an Open Archival Information System (OAIS), p.1-12,
<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>]

Reference information

The information that identifies, and if necessary describes, one or more mechanisms used to provide assigned identifiers for the Content Information. It also provides identifiers that allow outside systems to refer, unambiguously, to a particular Content Information. An example of Reference Information is an ISBN.

[Reference Model for an Open Archival Information System (OAIS), p.1-12,
<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>]

Representation information

The information that maps a Data Object into more meaningful concepts. An example is the ASCII definition that describes how a sequence of bits (i.e., a Data Object) is mapped into a symbol.

[Reference Model for an Open Archival Information System (OAIS), p. 1-13,

<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>

Submission Information Package (SIP)

An Information Package that is delivered by the Producer to the OAIS for use in the construction of one or more AIPs.

[Reference Model for an Open Archival Information System (OAIS), p. 1-13,
<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>]

SWOT analysis

A SWOT Analysis is a strategic planning tool used to evaluate the Strengths, Weaknesses, Opportunities, and Threats involved in a project or in a business venture. Strengths and weaknesses are internal to an organization. Opportunities and threats originate from outside the organization. A SWOT analysis, usually performed early in the project development process, helps organizations evaluate the environmental factors and internal situation facing a project.

[<http://en.wikipedia.org/wiki/SWOT>]

Trusted digital repository

Trusted digital repository is the one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future.

[Trusted Digital Repositories: Attributes and responsibilities : an RLG/OCLC Report, p.5,
<http://www.rlg.org/longterm/repositories.pdf>]

Verification

The act of checking the physical and logical consistency of a document (are all bits, files correct; are all files present). Checking of the physical medium aims at testing the integrity of the support and the data stored on it. Content checking is concerned with the logical integrity of a document and the authentication of its source.

[NEDLIB List of Terms, p.15, <http://nedlib.kb.nl/results/NEDLIBterms.pdf>]

ANNEX 3

DISTRIBUTED QUESTIONNAIRES

3.1 Questionnaires enquiring users and non-users about repository services

3.1.1 Questionnaire for end-users

(See separate file: annex3_1_1_endusers.pdf)

3.1.2 Questionnaire for non-users

(See separate file: annex3_1_2_nonusers.pdf)

3.2 Questionnaire for the developers enquiring on repository's technical aspects (See separate file: annex 3_2_techasp.pdf)

3.3 Questionnaire for content providers

(See separate file: annex3_3_contprov.pdf)

3.4 Questionnaire for the librarians-operators of the repository enquiring about organizational aspects

(See separate file: annex3_4_libroper.pdf)

3.5 Questionnaire for top managers of main repository's institution enquiring about organizational aspects.

(See separate file: annex3_5_topman.pdf)