European Conference on Information Literacy (ECIL)
Saint-Malo, France
18-21 September 2017

Data Literacy Survey Implementation at ULSIT

Tania Todorova, Rositsa Krasteva and Elisaveta Tsvetkova

University of Library Studies and Information Technologies

{t.todorova, r.krasteva, e.cvetkova}@unibit.bg

Sofia, Bulgaria

Structure of the presentation

1 Introduction

2 Specific Issues on Survey Implementation in ULSIT

3 Findings

3.1 Demographics

3.2 Competence regarding management of research data

4 Outcomes and Academic Policy Reflections

1 Introduction

In the frame of ECIL 2016, researchers from the United Kingdom, France and Turkey initiated new international research on Data Literacy and Research Data Management (Chowdhury et al, 2016).

Convinced that the survey, aimed at collecting data about data literacy of academics and research students in higher education institutions, is **timely, necessary, and very useful**, a team of researchers from the University of Library Studies and Information Technologies (ULSIT) joined an international scientific group to collaborate.

Data Literacy and Research Data Management Survey WebPage (ReDaM)

Information Literacy Association

InLitAs

Search... Search

ReDaM | CoLIS | About us

Category: ReDaM

ReDaM participating countries

February 26, 2017

Categories

- CoLIS
- ReDaM

Recent Posts

- CoLIS Publications
- About CoLis
- CoLIS Team
- ReDaM participating countries
- ReDaM team

Recent Comments

Archives

- July 2017
- February 2017

2. Conducting the Survey at ULSIT

The questionnaire survey, ‘Data Literacy Survey,’ contains two groups of questions: the first group – aims to collect demographic information about the respondents, and the second group – aims to establish their competence regarding management of research data.

The implementation of the ‘Data Literacy Survey’ in ULSIT will pass through three stages.

2. Conducting the Survey at ULSIT

During the ***first stage***, January 2017, a translation of the questionnaire instruments from English language into Bulgarian was carried out and a respondents' list was created, covering 150 participation invitations to representatives of the academic community: lecturers and doctoral students.

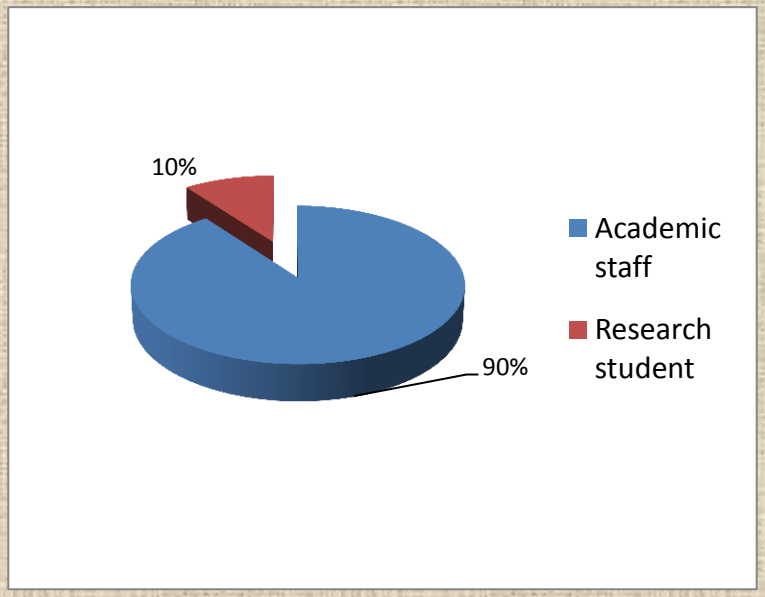
The ***second stage***, February - March 2017, includes dissemination of the participation invitation in the online based survey (Lyme Survey) and accumulation of data.

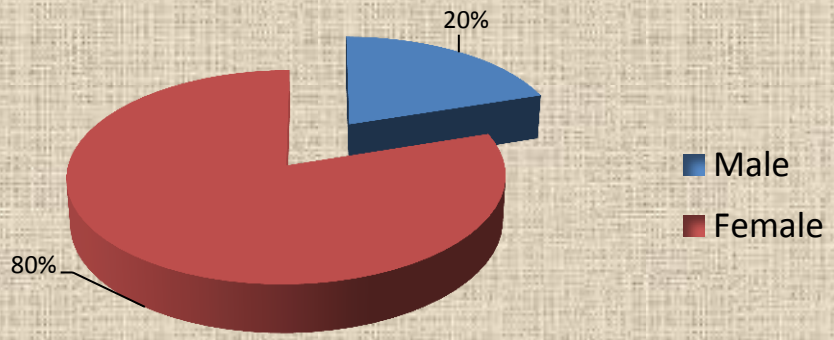
The ***third stage*** involves analysis of the results, elaboration of a document with conclusions and recommendations, as well as the preparation of a scientific article.

3. Findings

The study was conducted on the basis of systematic random sampling with stratification of 10% of all professors under basic employment contracts at UniBIT and 10% of the doctoral students currently trained.

The general group consists of 40 effectively inquired Bulgarian citizens (lecturers and doctoral students).





3. Findings

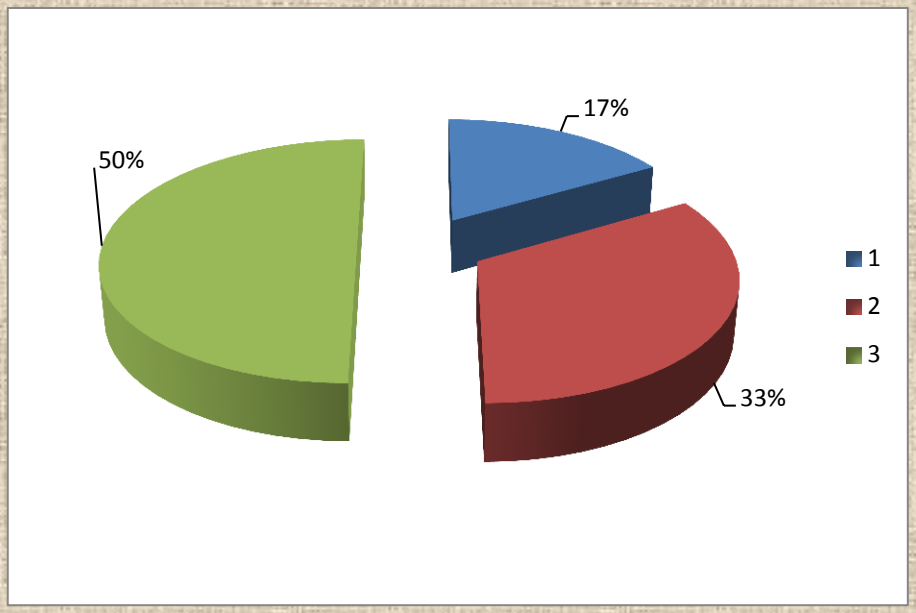


Figure 3: Distribution of the inquired subjects by scientific specialty

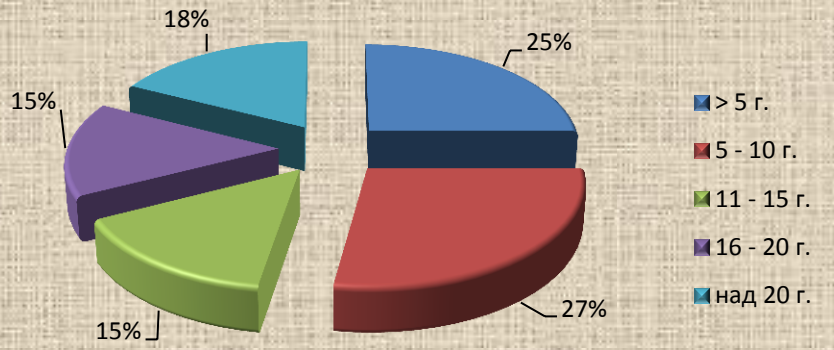


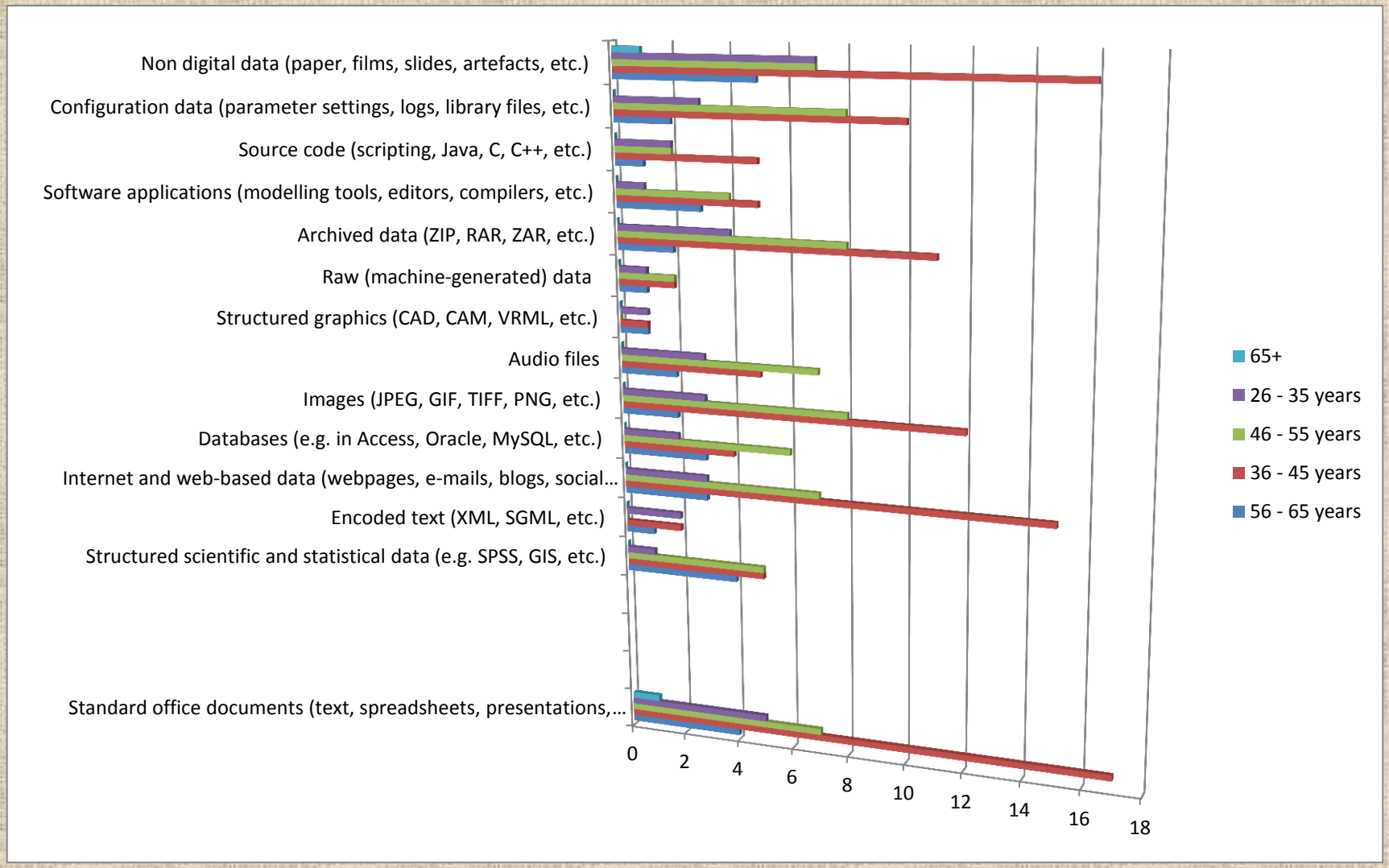
Figure 4: Distribution of the inquired subjects by internship in the research activity

Humanities

Natural Sciences: Computer and Information Sciences

Social Sciences (Media and Communications, Library and Information Sciences, Pedagogy, Sociology, Economics and Business)

Q: *Please indicate the file type of data that you normally use for your research*



Q: Which of the following better describes the volume of data you use for your research?

Factor (Research experience)	N	M	SD	F	Sig.	Min.	Max.
< 5 years	10	1,60	0,51	6,55	0,00	1	2
5-10 years	10	1,60	0,51			1	2
11-15 years	6	1,67	0,51			1	2
16-20 years	6	1,63	0,51			1	2
> 20 years	7	1,86	0,69			1	3

Q: How do you usually get the data for your research?

Factor	N	%
Create new data	24	58,5
From own research team/group at the university	20	48,8
From own research network (or personal/professional connections)	16	39,0
Always from one known source	5	12,2
Always from multiple known sources	32	78,0

Q: How do you usually use data that you get from others/outside sources?

Factor	N	%
As it is without any problems	6	14,6
With a bit of effort for some cleaning and/or modifications	7	17,3
After spending a lot of time and efforts to make it usable for the project	34	82,9

Q: What type of data do you produce from your research?

Factor	N	%
Standard office documents (text, spreadsheets, presentations, etc.)	40	100
Structured scientific and statistical data (e.g. SPSS, GIS, etc.)	8	19,5
Encoded text (XML, SGML, etc.)	1	2,4
Internet and web-based data (webpages, e-mails, blogs, social network data, etc.)	6	14,6
Images (JPEG, GIF, TIFF, PNG, etc.)	22	53,7
Audio files	7	17,1
Structured graphics (CAD, CAM, VRML, etc.)	0	0
Raw (machine-generated) data	0	0
Archived data (ZIP, RAR, ZAR, etc.)	18	43,9
Software applications (modelling tools, editors, compilers, etc.)	4	9,8
Source code (scripting, Java, C, C++, etc.)	0	0
Configuration data (parameter settings, logs, library files, etc.)	6	14,6
Non digital data (paper, films, slides, artefacts, etc.)	18	43,9

Q: Which of the following better describes the volume of data you produce from your research?

Tables 6: Volume of data generated as a result of the surveys (in units of measurement of information)

Factor (Research experience)	N	M	SD	F	Sig	Min.	Max.
< 5 years	10	1,60	0,51	1,50	0,22	1	2
5-10 years	11	1,45	0,52			1	2
11-15 years	6	1,67	0,51			1	2
16-20 years	6	1,77	0,40			1	2
> 20 years	7	1,86	0,69			1	3

Q: Where do you usually store the data you produce from your research?

Factor	N	%
Your own devices (your computer, your tablet, external drive, etc.)	35	95,1
Cloud	16	39,0
Central servers/repositories of the university	12	29,3
Outside repositories	5	12,2

Q: Do you normally assign any additional information to your research data?

Factor	N	%
Administrative information (e.g. creator, date of creation, file name, access terms/restrictions, etc.)	31	75,6
Discovery information (e.g. creator, funding body, project title, project ID, keywords, etc.)	25	61,0
Technical information (e.g. file format, file size, software/hardware needed to use the data, etc.)	13	31,7
Description of the data file (e.g. file/data structure, field tags/descriptions, application rules, etc.)	13	31,7
No, I do not assign additional information to my research data	2	4,9

Q: Do you collaborate with other researchers and share data?;
Q: Which of the following applies to your research data?

More than half of respondents are willing to share data and collaborate with other researchers, both from their own institution and from other scientific organizations, including foreign ones. This fact is undoubtedly important for the creation and development of national and international research networks.

Nearly half of the respondents are more likely to freely provide their data and developments.

Asked if they have any concerns about sharing data with others, almost half of respondents (43.9%) say that their research data is freely available to everyone interested, i.e., they have no worries.

A finding that deserves attention is, however, the concern of more than one third of the respondents (34.1%) of the lack of appropriate copyright protection policies.

The last questions in the questionnaire refer to specific documents and activities for managing research data at the university and aim to establish the level of awareness of the respondents

The analysis of the results shows that the picture of the individual respondents' opinions is quite colorful.

The respondents' responses are located across the whole range of opportunities provided.

Therefore, when planning, organizing and implementing future metadata trainings and/or Data Management Plan (DMP), to which respondents are generally well-minded, trainees will not be able to be approached as a homogeneous group of knowledge and understanding.

The last questions in the questionnaire refer to specific documents and activities for managing research data at the university and aim to establish the level of awareness of the respondents

Respondents are not well informed about the existence of the Data Management Plan (DMP) at the university, as well as about the implementation of such a plan.

Encouraging is the fact that most of them (63.4%) believe that the Data Management Plan (DMP) effectively helps scientists manage their research data.

Positive is also the opinion of the respondents on the metadata training. According to 80.5% it would be useful for the research data management.

The last questions in the questionnaire refer to specific documents and activities for managing research data at the university and aim to establish the level of awareness of the respondents

In Bulgaria, standards are widely known and used regarding citation of the sources used.

By 2011, a national standard was valid, and currently, following the harmonization as a national standard of ISO 690: 2010 "Information and documentation - Guidelines for bibliographic references and citations to information resources", the international standard is applied.

In this regard, understandable and natural is the shared opinion of 90.2% of the respondents on the use of any specific standard for citing and referring to research data.

Q: In your opinion who should pay for storage and public access to the data set that you created?

Q: Where should the data be stored for long term access?

The survey found that, according to most of the respondents, financial provision should not be the concern of the researchers themselves or of their teams. In this activity should be engaged the university and other research funding bodies and / or any national authority.

Most of the researchers (65.9%) believe that the scientific data for provision of long-term access is reasonable to be stored at the university or in an external repository, free of charge (48.8%).

Question № 24 and №25 of the questionnaire aim to establish have formal trainings been conducted on the various components of research data management activity and whether respondents would be involved in future trainings

Factor (topic of training)	N	%
Data Management Plan (DMP)	3	7,3
Metadata	2	4,9
Consistent file naming	1	2,4
Version control of data sets	2	4,9
Data citation styles	13	31,7
No, I'm not trained on any topic	22	53,7

From the data presented it becomes clear that trainings on these topics have been conducted, but not always the respondents have been involved. More than half of the respondents (53.7%) have not participated in such trainings so far.

The survey found that the respondents show a keen interest and desire to conduct similar trainings in the future.

4 Outcomes and Academic Policy Reflections

The purpose of the analysis is to collect data for making assessments and conclusions with further reflection on the institutional policy and the specific improvements of the Quality of Education Management System (QEMS) at ULSIT. We consider that important recommendations on the necessity of establishment of institutional policy for the implementation of the Data Management Plan will loom large. Recommendations may call for the establishment of a prescribed metadata set for uploading data into a university repository and of specific guideline for citing data and others. The systematized conclusions will be provided to all individuals and groups responsible for the quality standards design at the university. Understanding the current levels of awareness and gaps in knowledge of the research community at ULSIT will help us to suggest the appropriate data literacy training in the near future, based on international experiences and standards.

European Conference on Information Literacy (ECIL)
Saint-Malo, France
18-21 September 2017

Thank you for your attention!

Data Literacy Survey Implementation at ULSIT

Tania Todorova, Rositsa Krasteva and Elisaveta Tsvetkova

University of Library Studies and Information Technologies

{t.todorova, r.krasteva, e.cvetkova}@unibit.bg

Sofia, Bulgaria