Management of research data in image format: an exploratory study on current practices

Miguel Fernandes[1], Joana Rodrigues[1,2][0000−0002−1309−2122], and Carla Teixeira Lopes[1,2][0000−0002−4202−791X]

[1] Faculty of Engineering of the University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
[2] INESC TEC, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
{miguelfernandes197,joanasousarodrigues.14}@gmail.com, ctl@fe.up.pt

Abstract. Research data management is the basis for making data more Findable, Accessible, Interoperable and Reusable. In this context, little attention is given to research data in image format. This article presents the preliminary results of a study on the habits related to the management of images in research. We collected 107 answers from researchers using a questionnaire. These researchers were PhD students, fellows and university professors from Life and Health Sciences, Exact Sciences and Engineering, Natural and Environmental Sciences and Social Sciences and Humanities. This study shows that 83.2% of researcher use images as research data, however, its use is generally not accompanied by a guidance document such as a research data management plan. These results provide valuable insights into the processes and habits regarding the production and use of images in the research context.

Keywords: Research Data Management, Image Management, Image as research data.

1 Introduction

Recent technological and scientific developments gave rise to the appearance of new methods, instruments, and research tools. These changes led to an increase in the volume, complexity and importance of research data. Combined with the increase in computing and digital storage capacity, data collection, dissemination and analysis are increasingly intensive. This new feature of science has led to data-intensive science [4, 5]. This paradigm shift has caused changes and challenges in the way data are stored, preserved, accessed, and shared in the context of scientific activity [15].

The rapid development of processing capacity, image management and the ease of replication and dissemination increased the access and value of image collections [9]. In the context of research, various image capture devices have emerged. With the increased use and volume of image collections, new challenges and opportunities have arisen in image research data management. Vejvoda, Burpee and Lackie [21] give preliminary recommendations for image management in the research context through recommendations established for numerical data sets. In our search for related works, we didn't find other studies focused on image management in the research context.

This study is motivated by the lack of knowledge concerning the production processes, use, and management of images in the research context. As images constitute a valuable informational element for research, it becomes necessary to include them in research data management processes. Therefore, it is essential to know the processes and habits in the production and use of images in research to produce recommendations for appropriate management. This work does not focus only on digital images as research data, although the importance of the technological development of capture devices is mentioned here. Analog images are also included, as they are also used (albeit in smaller numbers) and are equally important in research projects.

This article presents the preliminary results of a study on the habits related to the management of images in research that will later lead to guidelines on how researchers should manage their images.

2 Literature Review

Research data management involves a set of practices that include planning, documentation, organization, storage, dissemination, and preservation of research data [11]. It aims to prolong the life of the data during and after the end of the investigation, as well as to encourage data sharing and reuse [6]. Data management forms the basis for applying the Findable, Accessible, Interoperable, and Reusable (FAIR) and open science principles, which are often required by funding agencies [11]. Several models allow the creation of abstractions, the definition of concepts, key moments, and activities of the research data lifecycle to guide the planning and implementation of research data management.

The research data lifecycle consists of a simple, understandable, and organic way of visualizing the different phases of research data management through a descriptive model [7]. These key concepts depend on the scientific area, the type of data, among other factors [10].

Research data can be defined as the factual records used as primary sources in scientific research, accepted by the scientific community, and indispensable for validating research results. These records can be textual, numeric, images, or sound records [14]. Research data is collected and produced in various formats, from digital spreadsheets to compilations of questionnaires, images, and objects [13]. Research data is considered the input of the investigation and not the output. Thus, the figures produced for articles and other publications are not the focus of this article.

An image conveys information or meaning differently from text. While text transmits information through conventional and arbitrary symbols, the image carries information through the representation and similarity of the objects as they are. The fact that the image is used in conjunction with text suggests that the image itself carries information different from the text, managing to transmit things that the text cannot [12]. It can be understood as something that depicts, it consists of a form of representation of which photography, video, drawing and painting are part, among many others [1]. The scientific community currently accepts the use of photographs, videos, and other similar resources. It is a common practice in several areas of research, such as Astronomy, Anthropology, Geography, History, Social Sciences, and Health Sciences [3, 16, 17].

3 Methodology

We used a questionnaire to study the practices and habits in the management of research data in image format. The structure of the questionnaire and the elaboration of the questions were informed by the research data lifecycle from Data Documentation Initiative [18], DataONE [2] and UK Data Archive [20]. Questions were grouped by stages of the research data lifecycle: planning, creation/compilation, quality assurance, processing/analysis, description, storage and sharing.

The questionnaire underwent several revisions by the authors where the adequacy and formulation of the questions were discussed and reflected. Before dissemination, we tested our questionnaire with an external researcher to analyze how each item question was interpreted. The objective was to assure that the questions were well understood by researchers not specialized in the subject.

The questionnaire opened on 20 February 2020 and accepted answers until 26 March 2020. The questionnaire was distributed by email at the University of Porto research community[3], namely i3S and INESC-TEC, and in research units outside the University of Porto funded in 2019 by the Foundation for Science and Technology[4]. Namely, Centre for Informatics and Systems of the University of Coimbra, Cardiovascular Centre of the University of Lisbon, Centre for Philosophical and Humanistic Studies of Universidade Católica Portuguesa and Center for Mathematics and Applications of Universidade Nova de Lisboa. An email was sent to the coordinators of each group, asking the dissemination of the questionnaire among their members.

Respondents were people with research experience, namely PhD students, fellows, and university professors. Answers were given anonymously, not allowing the identification of the participants.

We used quantitative methods to analyze closed-ended questions and content analysis for open-ended questions. There were two types of closed-ended questions. A kind of question collected answers on a 5-point Likert scale (Never, Rarely, Occasionally, Often and Always), generating ordinal variables. The other type, involved the selection, or not, of provided answer options, creating nominal variables.

For each question, we analyzed the general tendency in the overall set of answers and conducted comparisons between 4 scientific areas (Life and Health

[3] https://www.i3s.up.pt/; https://www.inesctec.pt/en

[4] https://www.cisuc.uc.pt/; http://ccul.pt/; https://cefh.braga.ucp.pt/; https://www.cma.fct.unl.pt/

Sciences - LHS; Exact Sciences and Engineering - ESE; Natural and Environmental Sciences - NES; Social Sciences and Humanities - SSH) and between 3 rates of image use in research (low - less than 25% of the projects; moderate - about 50% of the projects; high - more than 75% of the projects). For ordinal variables, we used Kruskal-Wallis to detect if there were differences among the groups. In cases where differences were found, we have used the Pairwise Test Mann-Whitney with Bonferroni correction to identify the location of the differences. For nominal variables, we used a Chi-square test for equality of proportions. When reporting our results, we use * to indicate results significant at $\alpha = .05$ and ** to indicate results significant at $\alpha = .01$.

4 Results

We collected answers from 107 researchers. The questionnaire, answer data and detailed statistical results are available at a data repository (https://doi.org/10.25747/7ma9-9132).

From the respondents, 41 (38.3%) work in the Life and Health Sciences, 30 (28%) in the Exact Sciences and Engineering, 12 (11.2%) in the Natural and Environmental Sciences, and 24 (22.4%) in the Social Sciences and Humanities field. Figure 1 relates the use of images as data with the research domain. We can observe a greater tendency to use images as research data in the Life and Health Sciences domain.

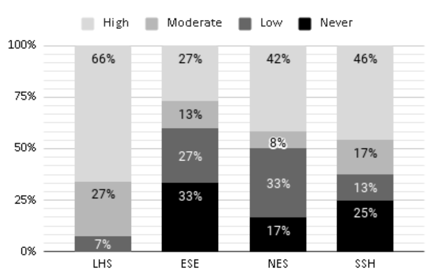

Fig. 1: Image use by domains

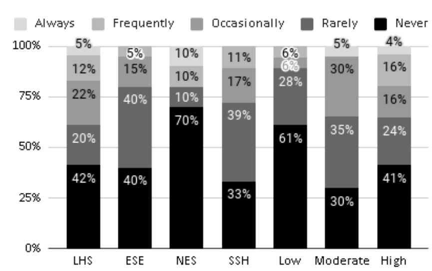

Fig. 2: Guide by area and frequency

Results are presented according to the stages of the research data lifecycle.

4.1 Planning

When asked about the existence of a document to guide the use and production of images during the research, 69.7% answered never or rarely, 16.9% occasionally and only 13.5% replied frequently or always. Even the researchers more accustomed to using images (High group) rarely make a document to guide the production and use of images, as can be seen in Figure 2.

We did not find significant differences between scientific areas or rates of image use.

Regarding the way and frequency in which researchers acquire and produce images, most of them produce images (73%) always or frequently, although about 59.5% of the researchers also consider that they occasionally or frequently use images from past projects. It should be noted that 89.9% of the respondents state that they never buy images from others and 61.8% of them say that they rarely or never acquire images from third parties, even images with no associated cost.

We found that the Life and Health Sciences domain is associated with a significantly higher production of images in the research context when compared with the Natural and Environmental Sciences and the Social Sciences and Humanities fields. This is visible in Table 1 that shows the significant differences between scientific areas.

Table 1: Significant comparisons in planning

Description	**LHS>**	**H>**	**M>**
Images produced in the research	NES* SSH*	L**	L**
Images come from past research		L**	

Likewise, those with low use of images, produce fewer images that researchers with moderate and high use of images in research. This is visible in Table 1 that shows the significant differences between rates of image use.

4.2 Creation/Compilation

Images are always or frequently produced by the computer (51.7%), the microscope (42.7%), the camera of the mobile phone (31.5%), and the traditional photo camera (21.3%). In Figures 3 and 4, we can see the percentage of use of these instruments by scientific area and frequency of use.

Regarding significant differences between scientific areas (Table 2), Social Sciences and Humanities researchers use the traditional camera significantly more than researchers from Exact Science and Engineering and Life and Health Science. The latter researchers also use this instrument significantly less than Natural and Environment Sciences ones. Similarly, Life and Health Sciences researchers are the ones who mostly use the microscope in comparison with the other scientific areas.

In Table 2, we can see that those who use images less often, use microscope images less often.

Regarding image edition/manipulation, 65.2% of the researchers say they always or frequently use image clipping, 47.2% always or frequently use simple adjustments to properties such as contrast, brightness and saturation, and 49.4% change the dimensions of the image, always or frequently. About 59.6% of researchers say that they rarely or never combine objects from different images to create a new one and 53.9% say they rarely or never use filters to improve image quality.

Table 2: Significant comparisons in creation/compilation

Description	LHS>	NES>	SSH>	H>	M>
Image capture via traditional camera		LHS*	LHS** ESE**		
Image capture via microscope	NES** SSH** ESE**			L**	L*
Perform simple editions to the image	NES*			L**	L**

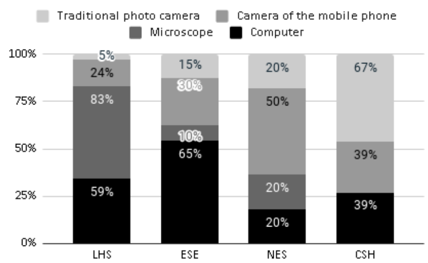

Fig. 3: Instrument of capture by scientific area

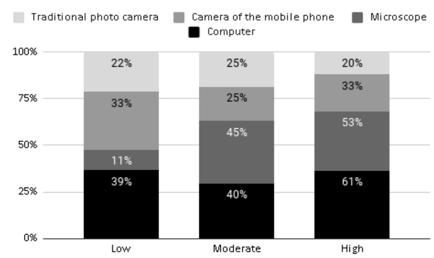

Fig. 4: Instrument of capture by frequency of use

Figure 5 relates the use of images as data with the research domain. In terms of significant differences, Natural and Environmental Sciences researchers perform significantly less simple adjustments to images comparatively to Life and Health Sciences researchers (Table 2).

Figure 6 relates the use of images as data with the frequency of use. Researchers that use images less often perform simple editions significantly less than the others (Table 2).

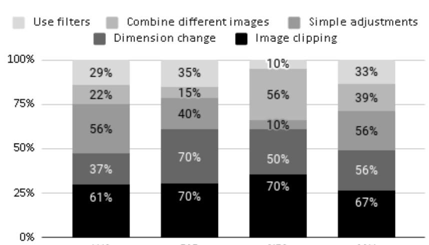

Fig. 5: Image edition by domains

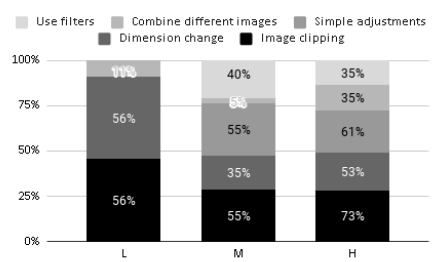

Fig. 6: Image edition by use

Of the respondents who edit the images, 57.6% indicated that they always preserve the original image, and 23.5% frequently preserves it. Only 2.4% of the researchers never keep the original image. We did not find significant differences between scientific areas or rates of image use.

In turn, documentation of editions is always or frequently done by 34.1% of respondents, and 47.1% rarely or never document. We did not find significant differences regarding research domains and rates of image use.

4.3 Quality assurance

Of the respondents, 53.9% answered that they always or frequently check the quality of their images. Only 7.9% of the researchers say they never do it, and 14.6% rarely check the quality.

Natural and Environmental Sciences researchers check the quality of images significantly less than Life and Health Sciences researchers (Table 3).

Table 3: Significant comparisons in quality assurance

Description	LHS>	H>
Check the quality of images	ESE*	L*

The same happens in researchers that use images less often. These researchers check the quality of the images significantly less than researchers accustomed to using images (High use group) (Table 3).

As for the processes mentioned by the researchers to ensure the quality of the images, these are varied. However, the following stand out: a review of images and their properties (36%), calibration of instruments (23.3%), and disposal of inappropriate ones (9%). We did not find significant differences regarding scientific areas and rates of image use.

4.4 Processing/Analysis

When asked about the most used computer programs in image processing and analysis, the researchers mostly mentioned ImageJ (33.7%), Photoshop (19.7%), Paint (10.1%), Cell Profiler (9%) and programming languages (7.9%). About a quarter (24.7%) of the researchers do not use any computer program.

It can be said that images are the object of different types of analysis. Those that stand out the most are content analysis (37.1%), mathematical calculations (18%), quantifications (11.2%), and measurements (5.6%). Of the respondents, 20.2% do not analyze images.

Finally, we noticed that researchers tend to combine manual and automatic analysis of images. There is a balance between manual mode (38.2% answered always or frequently) and automatic mode (42.7% answered always or frequently). In Table 4 we can see that automated analysis is significantly higher in the Life and Health Sciences and in the Exact Science and Engineering domains. On the other hand, Social Sciences and Humanities researchers are the ones who use manual analysis less in contrast with Life and Health Sciences.

Researchers that rarely use images use significantly less automatic and manual modes to analyze images in comparison with those who highly use images (Table 4).

Table 4: Significant comparisons in processing/analysis

Description	LHS>	ESE>	H>	M>
Manual image analysis	ESE*		L*	
Automatic image analysis	SSH** NES**	SSH** NES*	L*	
Document the conducted analysis	NES* SSH** SSH**		L**	L**

We found that 46% of the researchers said that they never or rarely document their analysis of the images. Life and Health Sciences researchers are the ones that document more frequently the steps taken in the analysis of images in comparison with the other scientific areas (Table 4). The vast majority of researchers who do not use images often, document their analysis significantly less that researchers who do a more intensive use of images (moderate and high use groups), as can be seen in Table 4.

4.5 Description

There is a slight tendency to associate annotations with individual images (44.9% always or frequently), instead of annotations in the set of images (37.1% always or frequently). We did not find significant differences between scientific areas. Researchers with low image use, do not annotate sets of images as often as those who highly use images do (Table 5).

Regarding the support where the annotations are made, the use of documents other than the image stands out. 32 (35.9%) of the respondents said that they always or frequently write it down on a paper document, about 36 researchers annotate (40.5%) on a digital document. Only 16 (18%) of the researchers write in the image, always or frequently. Exact Sciences and Engineering researchers are the least likely to take notes on a paper document other than the image in comparison mainly with Life and Health Sciences but also with Social Sciences and Humanities. These last two have the highest percentage of researchers that annotate on a paper document, 56.1%, and 27.8%, respectively (Table 5).

There is no clear preference in the annotation support regarding the frequencies of images use in research.

It should be noted that 65 (83.1%) of the respondents say that they never or rarely use any application that helps them describe images. Only six researchers (6.7%) say they do use, always or frequently, an application to help them. We did not find significant differences between scientific areas and frequencies of image use.

Regarding the metadata standards for image description, it is clear that few researchers use and know the topic. Of the three options given in the questionnaire (Dublin Core, Common European Research Information Format and EXIF), none showed a percentage of use above 1.1% for “always” frequency and

Table 5: Significant comparisons in description

Description	LHS>	NES>	SSH>	H>
Annotate sets of images				L*
Describe image on paper	ESE**		ESE*	
Use of Author as a descriptor	ESE*	ESE*		
Use of Description as a descriptor	ESE*			
Use of Capture Instrument as a descriptor	ESE** SSH**			
Use of Methodology as a descriptor	SSH**			
Use of Sample as a descriptor	ESE** SSH*			

above 6.7% for "frequently". In turn, the average percentage for the frequency "never" is 85.7%. When posed the possibility of presenting standards options other than those of the questionnaire, only five researchers answered, showing a clear trend towards the non-use of metadata standards for the description of images. We did not find significant differences between scientific area and frequencies of image use.

When asked about the vocabulary/elements of description that researchers most use and consider relevant, the answers vary. Seen as most relevant are the elements "title" (68.5%) "author" (53.9%), "date" (57.3%) and "description" (51.7%). The following descriptors are the least seen as relevant, "rights" (39.3%), "format" (40.4%) and "capture instrument" (41.6%). Among the most used descriptors are "title" (83.1%) "author" (62.9%), "date" (62.9%) and "description" (64%). With the lowest utilization percentages are also the descriptors "rights" (24.7%), "format" (39.3%) and "capture instrument" (39.3%).

Figures 7 and 8 show the descriptors used and seen as relevant by the scientific area.

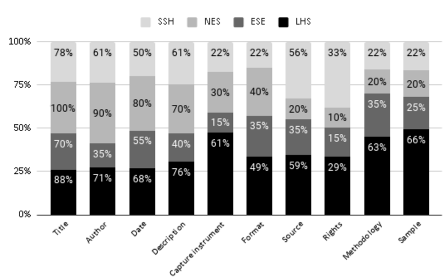

Fig. 7: Descriptors use by research domains

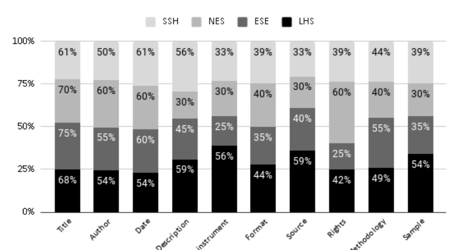

Fig. 8: Descriptors relevance by research domains

Exact and Engineering Sciences researchers are the least likely to use the *author* in comparison with Life and Health Sciences and Natural and Environmental Science as well as the *description* in comparison with Life and Health Sciences. In the opposite direction, Life and Health Sciences researchers are more likely to use the descriptor *capture instrument* and *sample* in comparison with the Exact and Engineering Sciences and the Social Sciences and Humanities domains. Life and Health Sciences are also more likely to use the descriptor *methodology* than Social Sciences and Humanities (Table 5). Regarding the relevance assigned to descriptors, we did not find significant differences between scientific areas and frequencies of use of images as research data.

4.6 Storage

Regarding storage location, the computer stands out (95.5% always or frequently), followed by external disk (73%), pen drive (41.5% always or frequently), and cloud (40.4% always or often). We did not find significant differences between scientific areas and frequencies of image use.

The most used formats to store are TIFF (74.2%), JPEG (83.1%) and PNG (61.8%). Although with very low usage percentages, RAW (13.5%), BMP (7.9%), SVG (2.2%) and PDF (3.4%) formats are also used. Figure 9 shows the storage format preference by the research domain. Regarding significant differences, Exact and Engineering Sciences researchers are the least likely to use the TIFF format to store their images when equated to Life and Health Sciences. Similarly, Life and Health Sciences researchers are less likely to use the PNG format than Exact and Engineering Sciences researchers (Table 6).

Table 6: Significant comparisons in storage

Description	LHS>	ESE>	H>
Save images as TIFF	ESE**		L*
Save images as PNG		LHS*	

Figure 10 shows the storage format preference by frequency of use. Regarding significant differences, researchers who rarely use images are less likely to use the TIFF format to store images than those who frequently use images (Table 6).

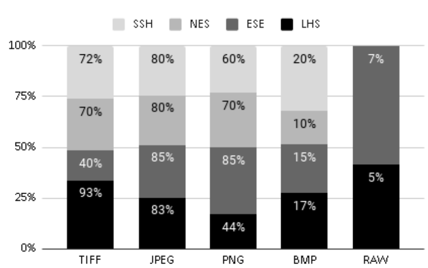

Fig. 9: Storage format by domains

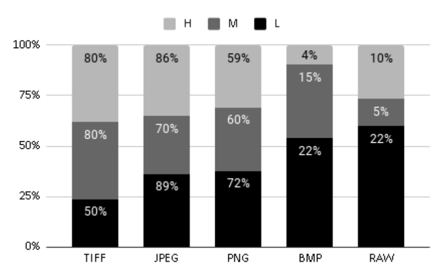

Fig. 10: Storage format by use

Regarding the volume of images stored during a research project, about 58.4% of the researchers said they were not able to quantify it. We did not find significant differences relative to the scientific areas and frequencies of image use.

When asked about the frequency with which they made backup copies, 40.4% responded monthly, 15.8% weekly, and 19.7% daily. About 18% of researchers do not regularly back up and 19.1% do not back up. When researchers make backup copies, they do it manually (65.2%), only 13.5% answered that they were done automatically. In these two questions, we did not find significant differences relatively to scientific areas and frequencies of image use.

4.7 Sharing

Image sharing occurs mainly at the end of the investigation (61.8% answered always or frequently) or during the investigation (59.5% answered always or frequently). Despite this, 41.6% of the researchers responded that they often share the images sometime after the research project is finished. In this question, we did not find significant differences between scientific areas and frequencies of image use.

To the answer about who promotes image sharing, the most frequent answer was the principal investigator (PI) (71.9%), followed by the investigator producing the image (44.9%), institution (31.5%) and funding entity (11.2%).

The PI is the one who promotes the sharing of images most significantly in the Life and Health Sciences relative to Exact Sciences and Engineering (Table 7). No significant differences were found between the frequencies of image use.

Table 7: Significant comparisons in sharing.

Description	**LHS>**	**ESE>**	**SSH>**	**L >**
Sharing promoted by PI	ESE**			
Include images location in scientific articles		LHS**	LHS*	H*

Regarding rights of use, the researchers replied more frequently that they leave the images restricted in access and use (35.4% always or frequently), followed by free to access and use (27.7% always or frequently) and free access but restricted use (20.7% always or frequently). When asked whether they shared images in archives, institutional or thematic repositories, the researchers' response was never or rarely (85%). About 4.5% replied that they shared in scientific articles. In these two questions, no statistically significant differences were found concerning the scientific areas and frequencies of use.

Most researchers (77.5% never or rarely) do not mention the location where the images are stored in scientific articles. Only 14.6% answered that they always or frequently mention the storage location in scientific articles. Life and Health Sciences researchers are less likely to mention the place where images are stored in scientific articles compared mainly to Exact Sciences and Engineering but also Social Sciences and Humanities (Table 7). Surprisingly, researchers who use images in more than 75% of their research projects are also the least likely to mention the location where the images are stored (Table 7).

Regarding who they share their images with, 67.1% of the researchers that use images said they were always or frequently shared with the research group, 40.7% with the research institute/center, 30.3% with a restricted community of researchers and 13.1% with the public in general. We did not find significant differences regarding the scientific areas and frequencies of image use.

5 Discussion

In this study, we collected information about the processes related to the use of images in research. We found that researchers often use them as research data, mainly in the Life and Health Science domain. Although they are widely produced and used in the research context, there are no guidelines that contribute to the standardization and orientation of their use. These conclusions are in agreement with a study carried out at Arab universities [8] focused on research data management in general.

In the creation process, there is a wide use of digital instruments, which may suggest that the digital revolution was an essential factor for the greater use of images as research data. The most significant use of the microscope in Life and Health Sciences and the traditional camera in Social Sciences and Humanities can be explained by the fact that this use is strictly related to the object of study and methodologies employed in each area.

The processes related to ensuring the quality of the photos are diverse and not all researchers carry out this activity. Likewise, the processing and analysis are heterogeneous, with no existence of standards. This can due to the multiple analysis options that an image may be subject of and with the existence of different methodologies in the various scientific domains.

Regarding the description, researchers do not use metadata models to assist them in this task. The same result was found in past research [8, 19]. When asked about the vocabulary/elements of description that researchers most use and consider relevant, the answers are varied. Only three descriptors were used and viewed as relevant by most researchers.

Regarding the storage location, the computer is the location chosen by the majority. This can be explained by easy access, familiarity, and recurring use of the computer in research projects. Elsayed and Saleh [8] found that most of the research stored their research data on their personal devices. It should also be noted that researchers are unaware of the volume of images produced and used during their research projects.

Although researchers said that they share the images used during the investigation, it was found that they do not do it by depositing the images in repositories that would ensure their preservation and sharing. A similar result was found by Elsayed and Saleh [8] were the least preferred way to make data electronically available was open data repositories. These results are also in line with the results obtained in a study [19] that found that researchers want to share their research data, but often find the process difficult.

6 Conclusions and future work

With this study, we identified patterns and habits in the creation, description, storage, and sharing of images. We have also compared scientific areas in terms of pratices and analyzed if the habit of using images affected habits. The phases where there was more diversity were quality assurance and processing/analysis, due to the heterogeneity of the methodologies used by the different domains.

Since no articles are addressing this subject, the results presented are useful as they provide valuable insights into the processes and habits regarding the production and use of images in the research context. Although they are preliminary results, we were able to verify that images are used as research data across all research areas. Many practices are common to all areas and some differ by research area.

Next, we will deepen our study by conducting interviews with researchers from different research domains. With the information collected through the questionnaire and interviews, guidelines for the management of research data in image format will be developed.

Acknowledgements. Joana Rodrigues is supported by a research grant from FCT - Fundação para a Ciência e Tecnologia: PD/BD/150288/2019. Special thanks to Professor Luis Teixeira for his help in the validation of the questionnaire. Thanks also to the Master in Information Science at FEUP for supporting the registration in this conference.

References

[1] Catharine Abell. "The epistemic value of photographs". In: *Philosophical Perspectives on Depiction*. Ed. by Catharine Abell and Katerina Bantinaki. Oxford University Press, 2010, pp. 82–103. ISBN: 9781351517577. DOI: 10.4324/9780203790762.

[2] Alex Ball. *Review of Data Management Lifecycle Models*. Tech. rep. Bath, UK., 2012, p. 15. URL: https://purehost.bath.ac.uk/ws/files/206543/redm1rep120110ab10.pdf.

[3] Marcus. Banks. *Using visual data in qualitative research*. Ed. by Uwe Flick. 1st. London: SAGE Publications, 2007, p. 161. ISBN: 9780857020260.

[4] Gordon Bell. "Foreword". In: *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Microsoft Research, 2009. Chap. Foreword, pp. xi–xvii. URL: https://www.immagic.com/eLibrary/ARCHIVES/EBOOKS/M091000H.pdf.

[5] Christine L. Borgman, Jillian C. Wallis, and Noel Enyedy. "Little science confronts the data deluge: Habitat ecology, embedded sensor networks, and digital libraries". In: *International Journal on Digital Libraries* 7.1-2 (2007), pp. 17–30. ISSN: 14325012. DOI: 10.1007/s00799-007-0022-9.

[6] Kristin Briney. *Data Management for Researchers: Organize, Maintain and Share your Data for Research Success*. Exeter, UK: Pelagic Publishing, 2015, p. 189. ISBN: 978-1-78427-030-8.

[7] Andrew Martin Cox and Winnie Wan Ting Tam. "A critical analysis of lifecycle models of the research process and research data management". In: *Aslib Journal of Information Management* 70.2 (2018), pp. 142–157. ISSN: 20503814. DOI: 10.1108/AJIM-11-2017-0251.

[8] Amany M. Elsayed and Emad I. Saleh. "Research data management and sharing among researchers in Arab universities: An exploratory study". In: *IFLA Journal* 44.4 (2018), pp. 281–299. ISSN: 17452651. DOI: 10.1177/0340035218785196. URL: https://doi.org/10.1177/0340035218785196.

[9] Peter Enser. "Visual image retrieval: seeking the alliance of concept-based and content-based paradigms". In: *Journal of Information Science* 26.4 (2000), pp. 199–210.

[10] Filipe Furtado, Pedro Príncipe, and José Carvalho. *Kit sobre dados de investigação*. Tech. rep. RCAAP, 2017, pp. 1–34. URL: http://repositorium.sdum.uminho.pt/handle/1822/46351.

[11] Rosie Higman, Daniel Bangert, and Sarah Jones. "Three camps, one destination: the intersections of research data management, FAIR and Open". In: *Insights the UKSG journal* 32 (2019). ISSN: 2048-7754. DOI: 10.1629/uksg.468.

[12] Sara Shatford Layne. "Some Issues in the Indexing of Images". In: *Journal of the American Society for Information Science* 45.8 (1994), pp. 583–588.

[13] OECD. *Making Open Science a Reality*. Paris, 2015. DOI: 10.1787/5jrs2f963zs1-en.

[14] OECD. *OECD Principles and Guidelines for Access to Research Data from Public Funding*. 2007.

[15] Cristina Ribeiro et al. *Os repositórios de dados científicos: estado da arte*. Tech. rep. 2010, p. 54. URL: https://repositorio-aberto.up.pt/handle/10216/23806.

[16] M. A. Sandweiss. "Image and Artifact: The Photograph as Evidence in the Digital Age". In: *Journal of American History* 94.1 (2007), pp. 193–202. ISSN: 0021-8723. DOI: 10.2307/25094789.

[17] James Roberto Silva. "Fotografia e ciência: A utopia da imagem objetiva e seus usos nas ciências e na medicina". In: *Boletim do Museu Paraense Emilio Goeldi:Ciencias Humanas* 9.2 (2014), pp. 343–360. ISSN: 21782547. DOI: 10.1590/1981-81222014000200006.

[18] Structural Reform Group. "DDI Version 3.0 Conceptual Model". 2004. URL: https://ddialliance.org/sites/default/files/Concept-Model-WD.pdf.

[19] Carol Tenopir et al. "Data sharing by scientists: Practices and perceptions". In: *PLoS ONE* 6.6 (2011), pp. 1–21. ISSN: 19326203. DOI: 10.1371/journal.pone.0021101.

[20] UK Data Archive. *Research data lifecycle*. 2019. URL: https://www.ukdataservice.ac.uk/manage-data/lifecycle.aspx (visited on 12/19/2019).

[21] Berenica Vejvoda, K. Jane Burpee, and Paula Lackie. "Image Management as a Data Service". In: *IASSIST Quarterly* 40.27-34 (2016).