Grant Agreement Number: 777483 / **Acronym:** ICEDIG
Call: H2020-INFRADEV-2017-1 / **Type of Action:** RIA
Start Date: 01 Jan 2018 / **Duration:** 27 months

REFERENCES:

Deliverable D2.2

Work package **WP 2**/ Lead: **Willemse**
Deadline M13

ICEDIG.EU

Innovation and consolidation for large scale digitisation of natural heritage

Prioritising scientific and societal needs for data of private collections

Deliverable D2.2

Emily van Egmond (Naturalis)
Luc Willemse (Naturalis)
Veljo Runnel (UTARTU)
Hannu Saarenmaa (LUOMUS)
Anne Koivunen (LUOMUS)
Kari Lahti (LUOMUS)
Laurence Livermore (NHM)

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

ICEDIG.EU

Summary

Specimens held in private natural history collections form an essential though often neglected part of the specimens held worldwide in natural history collections. The suitability of physical specimens for research purposes or the importance of data they hold clearly does not depend on the size of the collection they are kept in. Therefore, when engaging in regional, national or international projects, programs or activities aimed at increasing the accessibility of biodiversity data, it is paramount to include private collections as much and often as possible. Compared to larger collections, private collections by their very nature are small, relatively numerous, more or less anonymous, and diverse in all aspects of collection management. This poses various challenges to actually be able to include private collections in Europe's digitisation efforts of natural history collections.

The current task (Task 2.2 within ICEDIG), using a survey, first of all set out to gain more insight in private collections in Europe like their size and composition, the use being made of the collection, the degree of digitisation and the motivation of collection owners regarding managing and digitising collections and their willingness to share data and information. Although results of the survey were in some ways difficult to interpret and most likely skewed, they provide sufficient information to allow for valuable, general statements. The mere fact alone that the combined size of the collections of the respondents have been estimated to lie between 9 and 33 million specimens, this being only part of the total volume present in private collections in Europe, underlines the importance of private collections. Digitising and sharing collection data are activities that overall are considered important, are already being undertaken or for which there is a keen interest among private collection owners. Also, the survey showed that for those who have not yet started digitising their collection various kinds of support (e.g. tools, platform, training) would be of great help. The fact that lack of time is often mentioned as reason not to have started with digitisation, indicates that deployment of volunteers may offer an important means to assist private collection owners.

A second aspect tackled in this task relates to a protocol to keep an inventory of private collections up to date. Several options are described but overall, a website seems the most obvious option. Within DiSSCo efforts should be directed towards comparing several online platforms (ECSA, GBIF, CETAF, GRBio) and examine whether these are suitable to store basic collection information from privately owned collections.

A third goal of this task centred on a communication strategy to raise awareness and offering guidance for digitising privately owned collections. Six essential elements are recognised and described in more detail: stakeholders, collection owners, message, means, goals, resources and evaluation. One of the recommendations provided states that as the

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

target audience, their collections and the degree of digitisation is highly mixed, a flexible approach using different means of communication is needed.

Overall, it appears that privately owned natural history collections hold a large potential to add to the growing amount of digitised specimen data. Private collections will need to receive more attention in the future to identify how these can be incorporated within initiatives such as DiSSCo.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

Table of contents

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

1. Introduction

When people think about natural history collections, they mostly refer to large collections kept in natural history and university museums that are publicly funded and openly accessible. Such institutions hold large numbers of specimens in their collections, harbouring the unique biodiversity of the majority of past and present life, and geological history on Earth that we know existed (Kemp 2015). Globally, it is estimated that there are between 1.2 and 2.1 billion specimens present in natural history collections (Ariño 2010). European natural history collections have been estimated to contain a large part of these specimens, up to 1.5 billion specimens (see DiSSCo.eu).

Historically, private collections provided the basis of many natural history institutions (Stearn 1971). Up to the present day, private collections are being donated to natural history institutions and contribute an important inflow of specimens. Privately owned natural history collections may harbour highly valuable specimens for a specific taxon and/or (local) geographic range. Furthermore, on a European scale, these private collections combined are expected to form a significant part of the natural history collections as a whole. For example, it was found that private herbaria in Italy harbour over 150.000 specimens and together rank 16th in size of the 68 institutional public herbaria (Roma-Marzio et al. 2017). The owners of these private collections also play various important roles in the national and international biodiversity collection and research landscape, such as being a (world) specialist for certain plant or animal groups, member of associations or (co-) author of scientific papers. Therefore, it is important to engage private collection owners and to include private collections in Europe's digitization efforts of natural history collections.

In principle, private collections eventually become part of a larger national history collection within an institution. These private collections often add to the already existing digitisation backlog within natural history institutions when they have not been previously digitised. As a consequence, private collections commonly become part of a larger collection without digitisation. When these additions are not swiftly digitised, it may take a long time before the information about this collection becomes available online for the wider (scientific) community, hence hampering the spread and sharing of data. More importantly, specific knowledge about the collection and its specimens may be lost when the private collection owner passes away. To tackle this issue, private collections may be digitised before they become part of a larger natural history collection. This will require an aligned, well-informed strategy on a European, national and institutional level.

To this end, this deliverable D2.2 'Prioritizing scientific and societal needs for data of private collections' as part of Task 2.2 will address how private collections can be included in

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

Europe's digitisation efforts. First, it is necessary to identify the current volume, scope and level of digitisation of private collections within Europe to have a better view of the background and inform our next steps. Secondly, a protocol to keep this inventory up to date in the future is proposed. Finally, we provide a communication strategy to stimulate awareness and participation among private collection owners about digitisation.

Thus, this deliverable consists of three main parts:

I. Inventory of private collections in Europe
II. Protocol to keep the inventory of private collections up to date
III. Communication strategy to engage private collection owners

In a final chapter conclusions and recommendations are presented followed by seven appendices which present additional information about the survey and the communication strategy.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

2. Part I: Inventory of private collections in Europe

2.1 Objectives

The first objective of the current task under Work package 2 of ICEDIG was to get information about the volume, scope and degree of digitisation of private collections within Europe. Secondly, to be informed about incentives and required help needed by private collection owners to perform digitisation. To our knowledge this is the first attempt to take stock of private collections across a range of (taxonomic) groups on a European scale.

2.2 Methods

Survey design

A survey was conducted to collect individual responses from owners of private natural history collections using a questionnaire consisting of 32 questions (Appendix 1). Due to the applied logic, none of the respondents answered all questions since certain questions were skipped based on previously given answers. Some questions were therefore represented twice in the questionnaire, but answered by different subsets of people and always answered only once by a respondent (Figure 1). Only the final question was optional, where respondents could give their name and email address to be among the first to receive a summary of the results of this questionnaire. This question was included to entice respondents to fill in the questionnaire and have them benefit from their contribution as well. If they wished to remain anonymous, respondents could skip this question.

The first question 'Do you have a private collection of natural specimens?' was intended as a test to make sure all respondents without a private collection were quickly separated from those who did. Of the 1382 responses collected in total, 1145 respondents indicated to have a private collection, while 90 respondents indicated to collect only for institutional needs and 147 respondents did not collect specimens at all. Respondents that did not have a private collection were excluded from the analysis. Up to question 9, general questions regarding the private collection itself were posed. At question 9 - 'How do you manage your collection?' - private collection owners were asked how they managed their collection (either digitally, by hand or both). This resulted in a distinction between private collection owners based on management practices and a slight difference in subsequent questions. For respondents that did manage their private collection, another distinction was made at

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

question 12: 'Do you share your collection information?'. For respondents that did not manage their collection, we wanted to single out those respondents that were both considering to manage their collection (Q24) and to share their collection information (Q25).

The questionnaire was prepared using the online survey tool SurveyMonkey (surveymonkey.com). This allowed for an easy design and distribution of the questionnaire, including the advantage of adding multiple languages in which the questionnaire could be completed, while all data was added in one combined file.

Target group

In order to reach as many private collection owners as possible, we chose to ask amateur associations to help distribute our questionnaire. The implicit assumption made here is that many private collection owners are a member of at least one amateur association focused on either nature, science or natural history. The survey included 7 associations at international level, 178 associations active at a national level and 317 regional associations (Appendix 2). These contacts were found through Google searches for associations in the domain of natural history and contact information was subsequently obtained from the association's public website. In addition, we had a list of some 70 individual private collection owners that we obtained from people within the ICEDIG consortium that (personally) knew private collection owners. An effort was made to include associations from different groups (e.g. botany, insects and geology) and different European countries. It is important to note that there are no responses (except for two) from any private collection owners in the United Kingdom. The amateur associations approached in the UK were reluctant to circulate surveys to their members as in their view benefits for private collection owners were unclear. This results in an underestimation of the volume of private collections, especially since the UK has a rich history in natural history collections and is expected to have many privately owned collections.

Survey distribution

To make it as easy as possible for people to understand the purpose of this survey and complete the survey itself, we decided to translate the English emails and questions to eight other languages: French, German, Spanish, Italian, Dutch, Estonian, Finnish and Swedish. These languages were all implemented in the online survey to be chosen from a drop-down menu. Countries in the contact list that did not have any of the above languages as their native language, received the emails and questionnaire in English.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

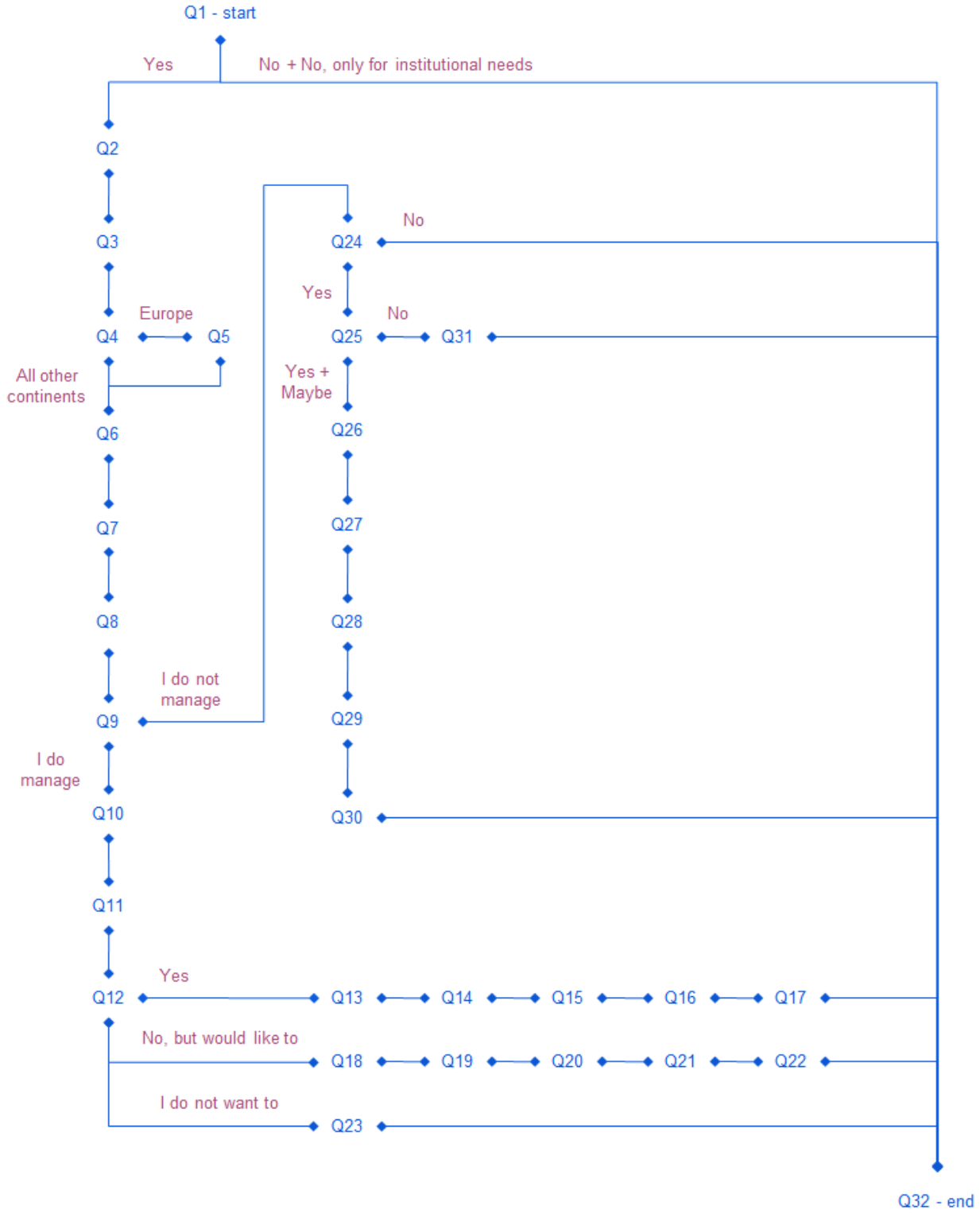

Figure 1. Flow diagram of questions in the distributed questionnaire, representing the applied logic. Question numbers refer to the questions as presented in Appendix 1.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

Our approach was to first send an announcement mail to the presidents and/or secretaries of each association in our contact list (Appendix 3). Our aim with this mail was to inform and prepare amateur associations for the upcoming survey and gouge their willingness to participate. If an association did not wish to participate, they had the opportunity to request exclusion from our contact list. None of the associations indicated beforehand that they did not wish to participate, and instead, we received several positive reactions to the upcoming survey. After the announcement, the questionnaire was distributed by sending an email containing an open link to the questionnaire in SurveyMonkey (Appendix 4). Language-specific links to the questionnaire were used for each of the different languages. All emails were sent from the Naturalis account. The questionnaire was distributed on the 31st of July 2018. A first reminder was sent on the 23rd of August and a second reminder was sent the 10th of September 2018. The questionnaire was closed on the 17th of September 2018, after which data were processed and analysed.

Data cleaning and analysis

Not every response that we received was complete, as some respondents only answered question 1 with 'yes', without giving any further details about their private collection. We therefore eliminated these responses from further analysis. This resulted in a total of 1027 responses to be used in the analysis, most of which were fully complete with only a small share being partly complete.

For each bar graph, we summed the total number of responses for each of the answer options. In the figure captions, N always refers to the number of respondents, not responses. To calculate upper and lower boundaries of the total number of specimens in the private collections sampled here, the number of private collections was multiplied by both the upper and the lower boundary of each size class. All values across size classes for the upper and lower boundaries separately were summed, to calculate the upper and lower estimate of the total number of specimens. Percentages for question answers where multiple answers were possible were calculated by using the total number of respondents (and not responses). To test the relation between collection size and the percentage of the collection that has been digitised at specimen level, we performed a chi-square test. All data and statistical analysis were performed using the program R (R Development Core Team, 2018).

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

2.3 Results

General

In total, we received responses from 25 different European countries. Based on the current results most private collections are kept in the Netherlands (284 responses or 28%), followed by Germany (143 responses or 14%) and France (137 responses or 13%) (Figure 2, Q2). It is important to note that this is only a subset of the total private collections present within Europe and needs to be kept in mind when reading the rest of the results.

Owners of private collections are often a member of an association focussed on either nature, science or natural history, with 89% of respondents (or 841 respondents) being a member of one or more associations (Q6). The majority of private collections is being used for scientific research (636 responses or 67%), where collections are mainly either solely used by the private collection owner himself (300 responses or 32%) or by both the collection owner himself and others (303 responses or 32%) (Figure 3, Q8).

Volume

The size of private collections is most often in the range of 501-5000 specimens per collection and thus of intermediate size (295 collections or 31%) (Figure 4, Q7). There were 60 private collections (or 6%) containing over 100.000 specimens, hence representing a very large collection. Considering the European private collections included in the results of this survey, the total number of specimens in these private collections is estimated to be between 8.8 and 32.6 million specimens (Table 1).

Scope

Most private collections included insects (444 responses or 43% of respondents), followed by mollusks (263 responses or 26% of respondents) and vascular plants (182 responses or 18% of respondents) (Figure 5, Q3). From individual responses it became clear that although most private collections focus on one taxonomic group (710 respondents or 69%), there are also private collections that cover more than one taxonomic group and thereby a much wider range (317 respondents or 31%).

Almost all European-based private collections contain at least some specimens collected within Europe (998 responses or 97% of respondents), while not surprisingly there were few private collections that contained specimens from Antarctica (41 responses or 4% of respondents) (Figure 6, Q4). Most private collections (624 responses or 61% of respondents)

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

contained specimens from one geographic range, while 403 private collections (or 39% of respondents) covered more than one geographic range. Of the private collections that contained specimens that were collected within Europe, most specimens were collected in the Netherlands (194 responses or 20% of respondents), followed by France (173 responses or 18% of respondents) and Germany (116 responses or 12% of respondents) (Figure 7, Q5).

Digitisation

The majority of private collection owners manage data pertaining to their collection (814 responses or 86%). This is mostly done by hand and digitally at the same time (321 responses or 34%), closely followed by doing this only digitally (294 responses or 31%) (Figure 8, Q9).

Of the private collection owners that manage the data connected to their collection, data was most often recorded at the level of individual specimens (468 responses or 60% of respondents) and taxonomic group (423 responses or 55% of respondents) (Figure 9, Q10). Recording information per storage unit was less common among private collection owners (156 responses or 20% of respondents). Most private collection owners recorded information at one level only (502 responses or 65% of the respondents). In general, private collection owners either recorded specimen information for (almost) the complete collection (392 responses or 50% of respondents) or a very small part of the collection (175 responses or 22% of respondents) (Figure 10, Q11).

Data sharing

Almost all of the private collection owners are interested in sharing collection data (responses or 92%). Over half of them are at present sharing data about their collection (442 responses or 57%), while 272 private collection owners (or 35%) would like to do so in the future (Q12). A small number of private collection owners is not interested in sharing collection data at all (64 responses or 8%).

Of the private collection owners that are at present sharing collection information, supporting scientific research was by far the most important reason to do so (341 responses or 84% of respondents) (Figure 10, Q13). Examples of other reasons to share collection data, which were given under 'Other', include: 'For educational purposes', 'To check whether identifications are correct' and 'To support (local) nature conservation'. Also private collection owners that do not share collection data yet, but would be interested to share in the future, found supporting scientific research the most important reason to do so (Figure 11, Q18 and 26). The order of reasons for sharing collection digitisation was similar between Figure 10 and 11.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

Regarding the question on how private collection owners currently share collection information, different methods of dissemination are used to a relatively similar degree (Figure 12, Q14). However, publishing information in print was the most common method to share collection information (219 responses or 54% of respondents). There were two main other ways to share collection information, as mentioned in the survey under 'Other': 'Via email/social media, upon request' and 'Via a larger, shared database (not necessary accessible online)'. Interestingly, when asking private collection owners that do not share collection data at the moment, how they would like to share collection data in the future, a clear preference was shown for sharing data on a website (193 responses or 61% of respondents), followed by publishing in print (119 responses or 38% of respondents) (Figure 13, Q18 and 27). Under 'Other', a couple of collection owners indicated that they did not know yet how they would like to share their collection data.

Finally, we asked private collection owners that are (thinking about) sharing collection data whether they would be interested in a new, European website where they can register their own private collection. Almost all respondents were (potentially) interested in a central, online place to share collection information (656 responses or 91%) (Figure 14, Q15, 20 and 28).

Of the private collection owners that are not sharing collection data, time limitation was the most important reason not to do this (32 responses or 28% of respondents), followed by the view that data sharing is not necessary (25 responses or 22% of respondents) (Figure 15, Q23 and 31). Examples of other reasons not to share collection data as indicated in the survey under 'Other' include: 'To prevent over-exploitation of sampling locations', 'Because the collection is constantly being rearranged' and 'I regard this as personal/confidential information as I still work on the collection myself'.

Necessary digitisation support

We asked what kind of support private collection owners would need to be able to digitise their collection at specimen level, either at home or outside the home at a nearby institute. At home, most private collection owners indicated they need tools, such as a dataset template or a web based digitisation platform, (397 responses or 55% of respondents) followed by information (257 responses or 36% of respondents) and physical equipment (195 responses or 27% of respondents) (Figure 16, Q16, 21 and 29). Other kinds of support that would be useful according to respondents under 'Other' include: 'Help with determination of specimens' and 'Help with handling computers in general (computer illiterate)'. Also 'more time' is often mentioned here, but this is ofcourse not something a natural history institute/amateur association can offer directly, only in the form of funding, equipment and/or (volunteer) staff. Of the respondents indicating they do not need any

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

support to digitise their collection at home, based on the input under 'Other' we know that some private collection owners have already performed digitisation, while others do not know what they would need to perform digitisation.

Regarding the question on the kind of support needed to digitise the collection outside the home at a nearby institute, the majority of private collection owners did not wish to perform specimen level digitisation outside of their home (324 responses or 45% of respondents) (Figure 17, Q17, 22 and 30). Those who did consider to do so, mainly needed to use the digitisation facilities on site (228 responses or 85% of remaining respondents), closely followed by the use of human resources and digitisation expertise on site (198 and 177 responses, or 50% and 45% of the remaining respondents, respectively). Under 'Other', a couple of respondents indicated they were too old to digitise their collection by themselves outside the home or the nearest institute was too far away. Some other respondents indicated they had already digitised and shared the majority of the collection, so they do not need any support.

Relation between collection size and percentage digitised

When testing the relation between the collection size and the percentage of the collection digitised at specimen level, we found that these are significantly dependent on each other (Chi-square test, df=20, χ^2=41.9 , p=0.002). There is a strong positive association between medium-sized collections (501-5000 specimens) and a high digitisation percentage (81-100%) (Figure 18). Also, there is a positive association between very large collections (>100,000 specimens) and a low digitisation percentage (1-20% and 21-40%), while these collections are negatively associated with a high digitisation percentage (81-100%).

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

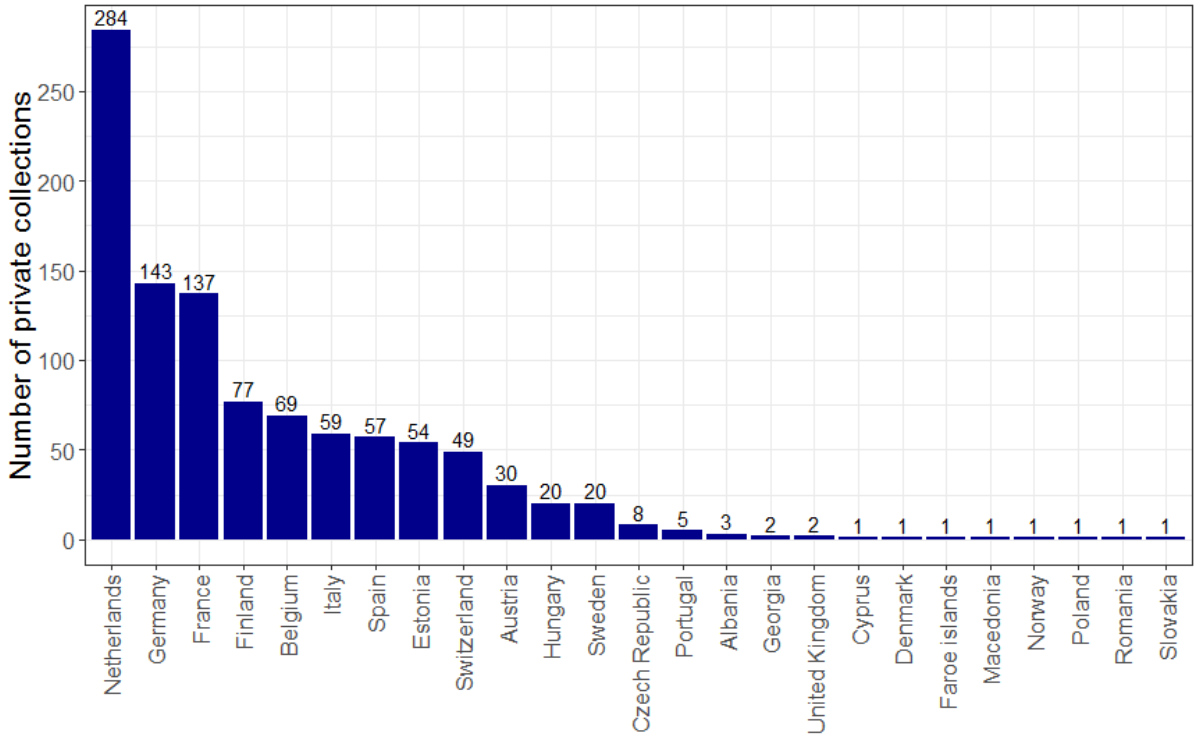

Figure 2. Number of private collections identified within each European country within this survey (Q2, n=1027).

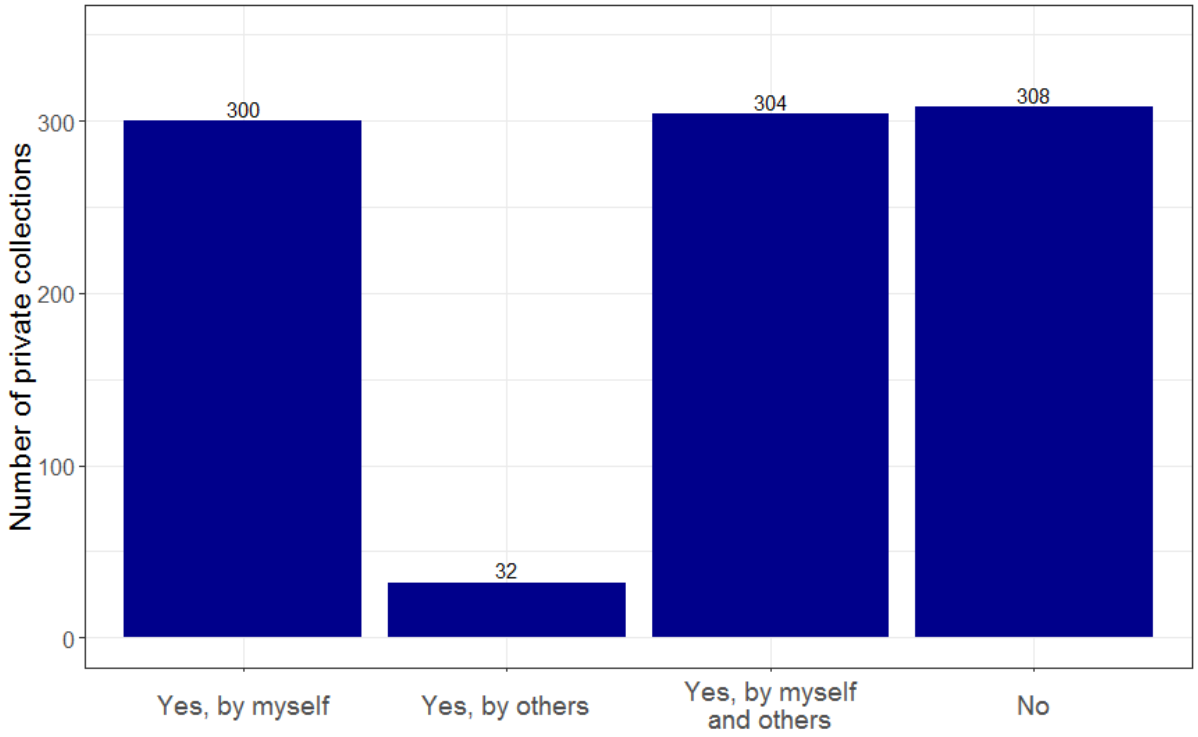

Figure 3. Number of private collections that are or are not being used for scientific research (Q8, n=945).

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

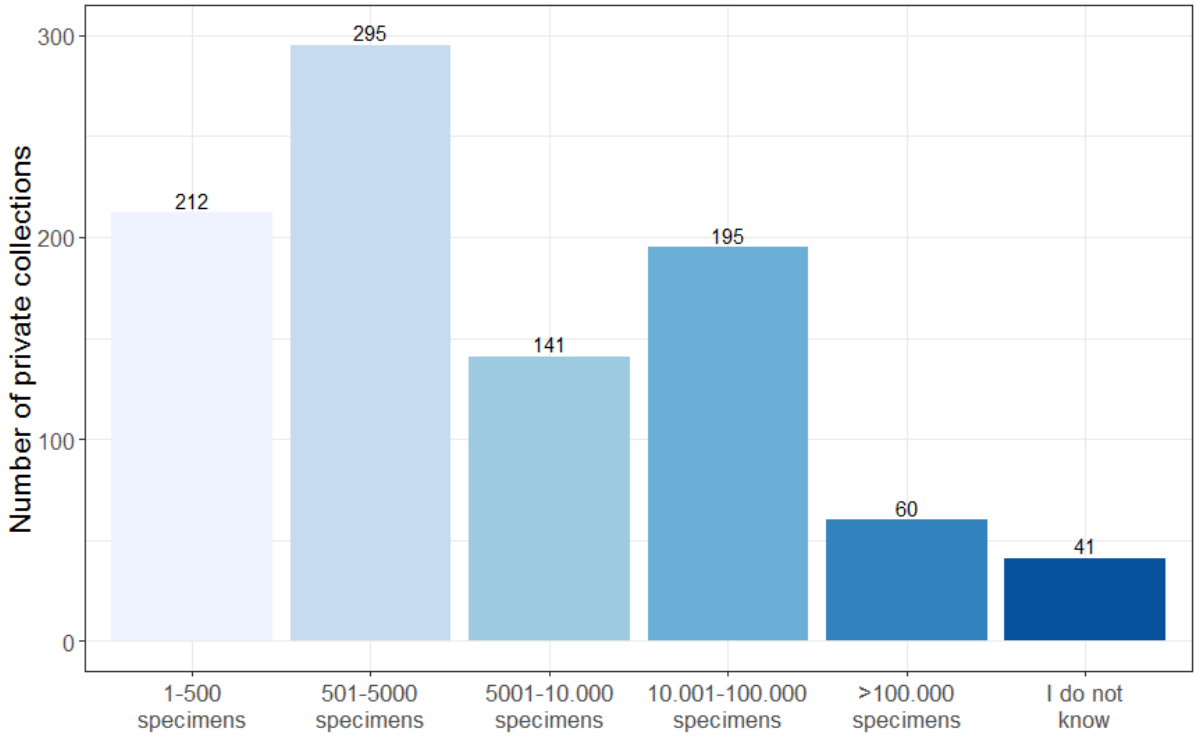

Figure 4. Number of private collections per volume category representing the number of specimens in the collection (Q7, n=944).

Table 1. Number of collections in each collection size category and upper and lower bound of the total number of specimens held in all private collections identified within this survey combined (Q7, n=944).

	1 - 500 specimens	501 - 5000 specimens	5001 - 10.000 specimens	10,001 - 100.000 specimens	more than 100.000 specimens	I do not know	total
Number of private collections	212	295	141	195	60	41	944
Lower bound	212	147.795	705.141	1.950.195	6.000.000	41	8.803.384
Upper bound	106.000	1.475.000	1.410.000	19.500.000	6.000.000	4.100.000	32.591.000

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

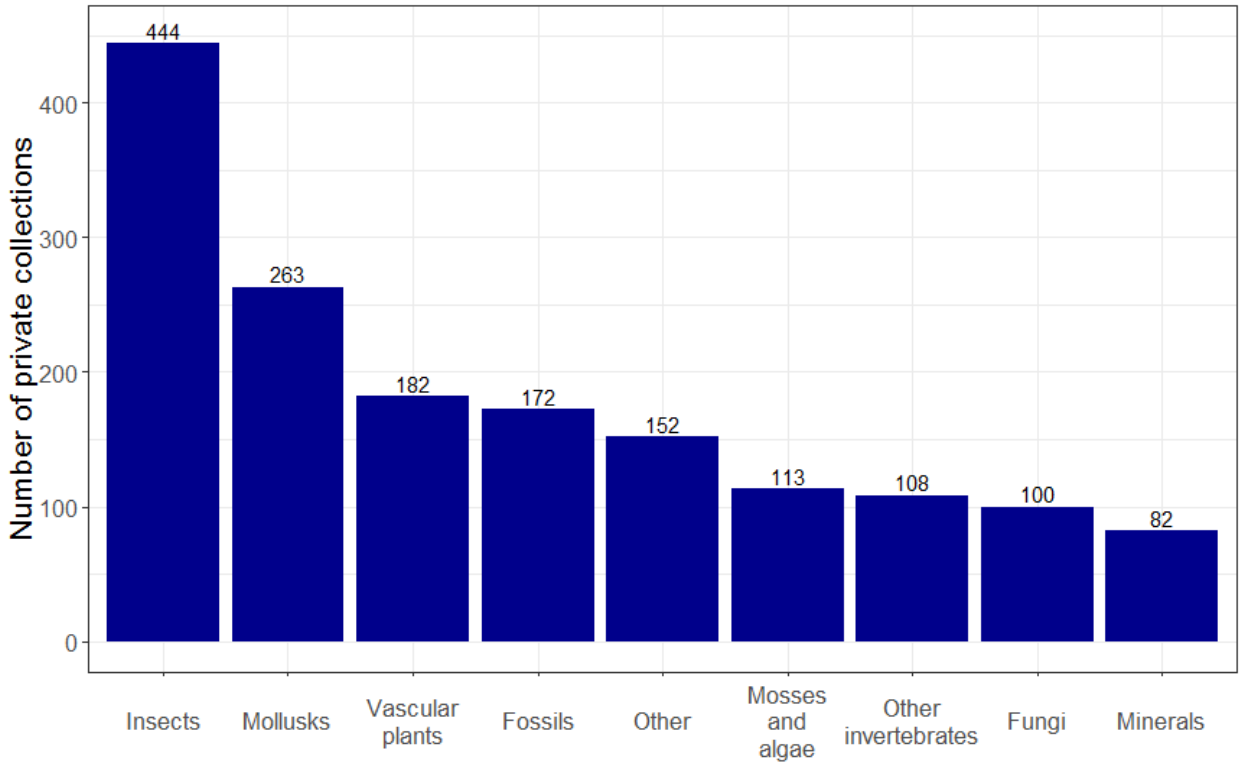

Figure 5. Number of private collections containing at least some specimens from a (taxonomic) group (Q3, n=1027, multiple responses were possible.

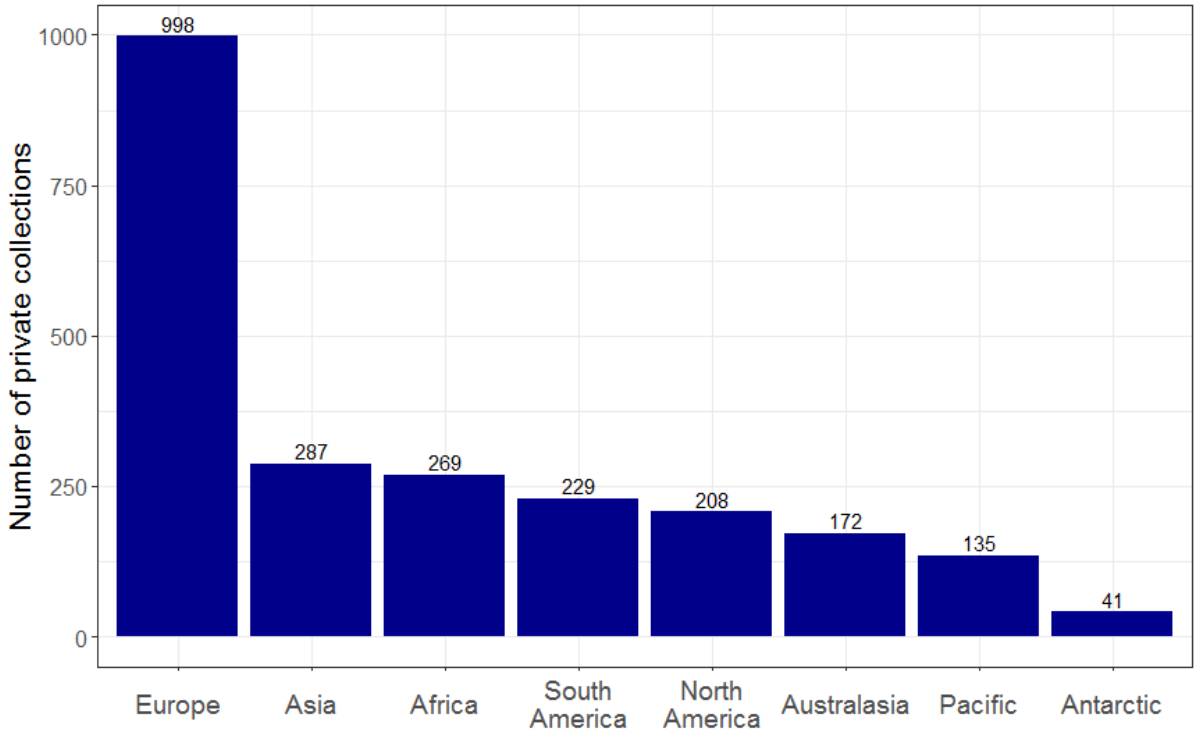

Figure 6. Number of private collections containing at least some specimens from a geographic range (Q4, n=1027, multiple responses were possible).

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

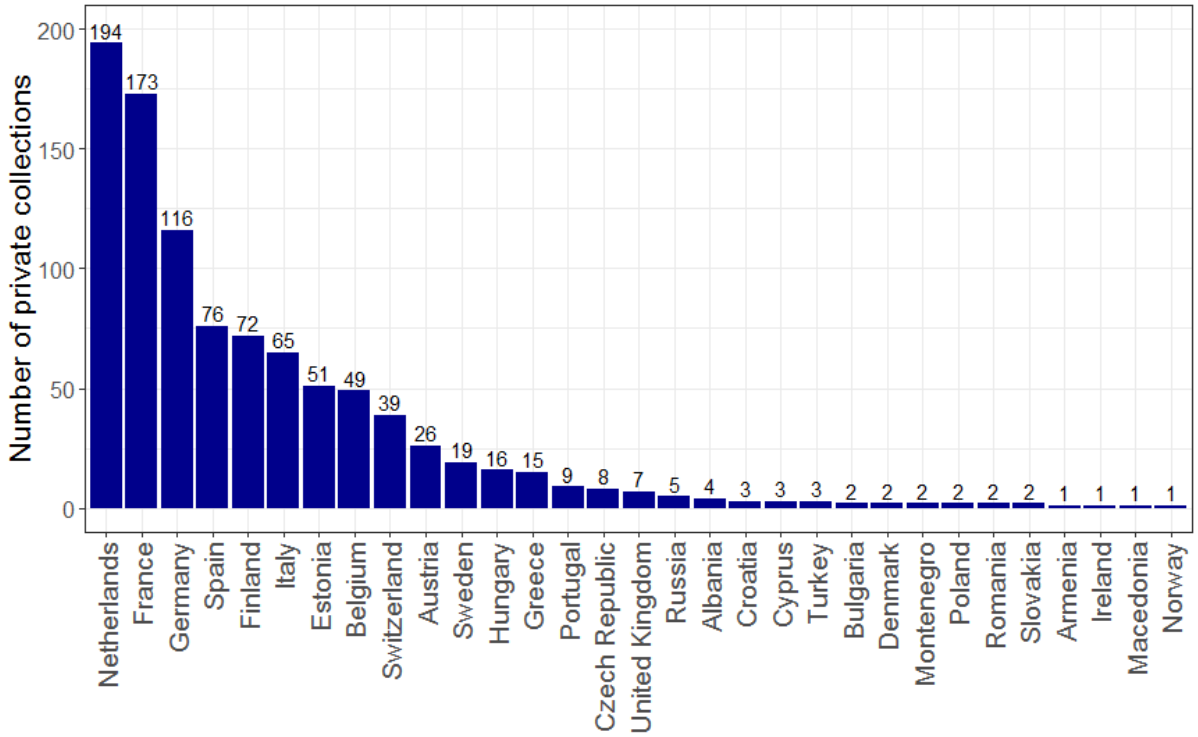

Figure 7. Number of private collections per European country from which the main part of the collection originates (Q5, n=969).

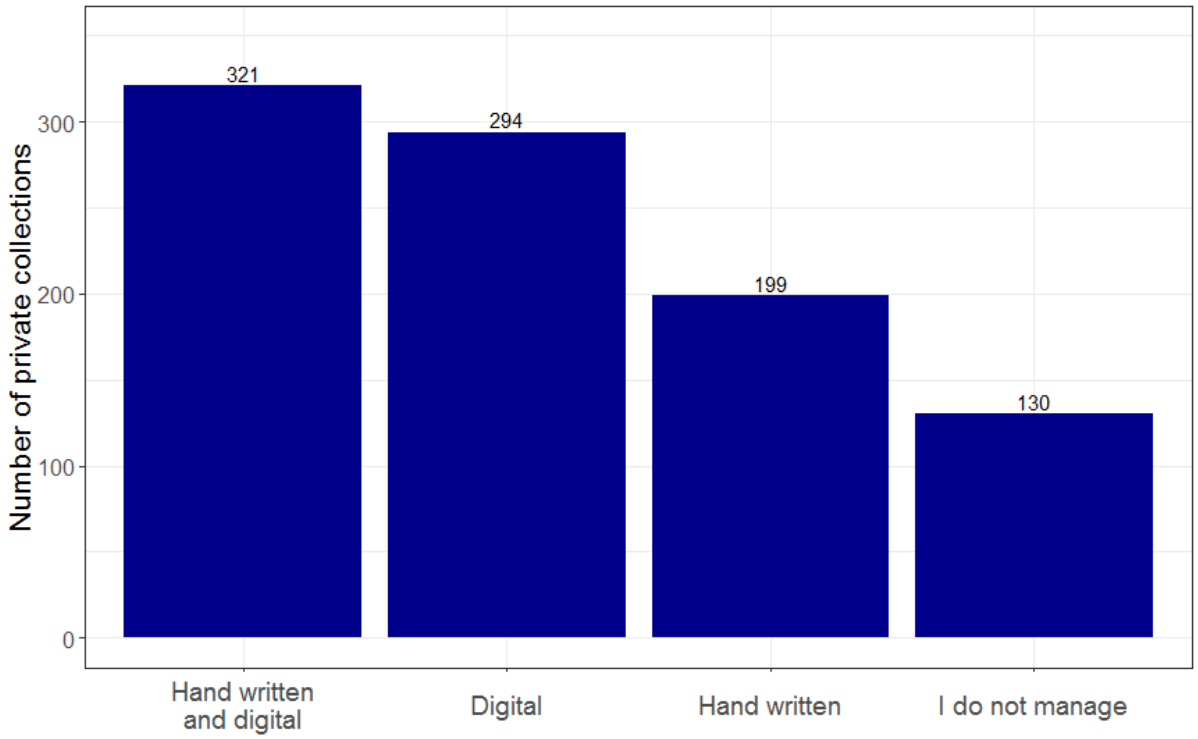

Figure 8. Number of private collections for which the data is being managed in a certain way by the private collection owner (Q9, n=944).

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

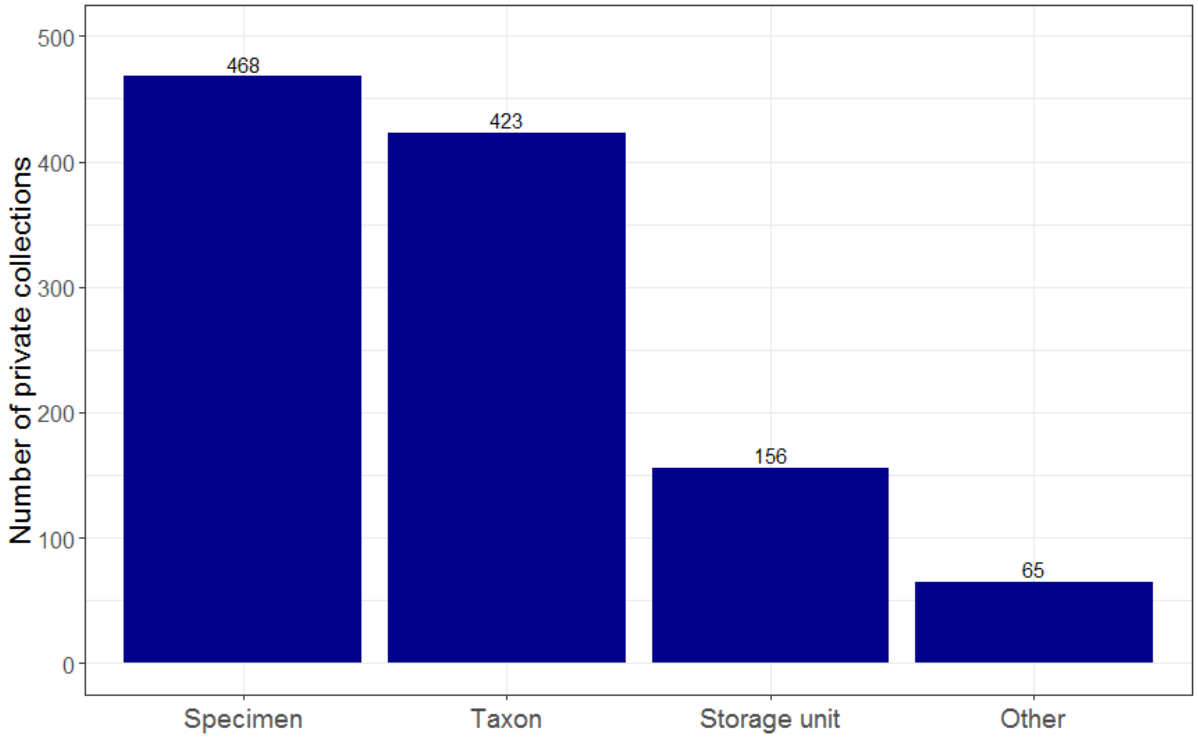

Figure 9. Number of private collections for which data has been recorded at a certain level by the private collection owner (Q10, n=775, multiple responses were possible).

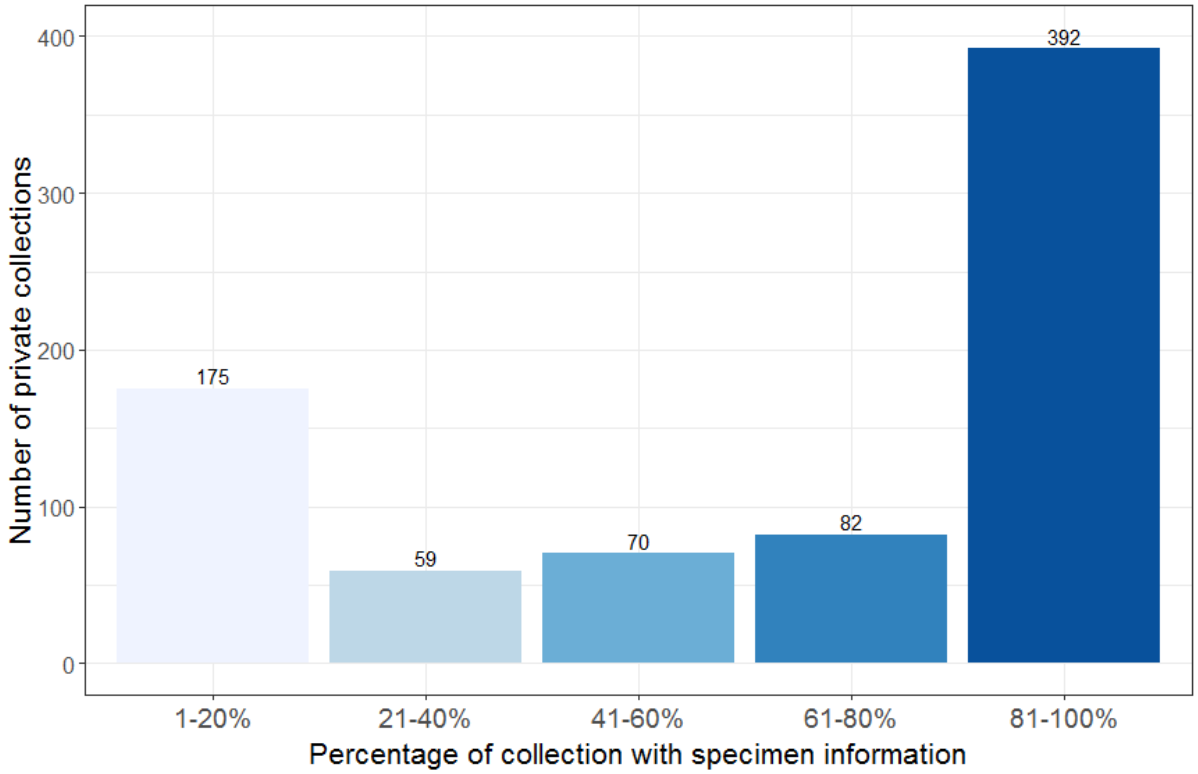

Figure 9. Number of private collections for which specimen information has been recorded up to a certain percentage of the entire private collection (Q11, n=778).

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

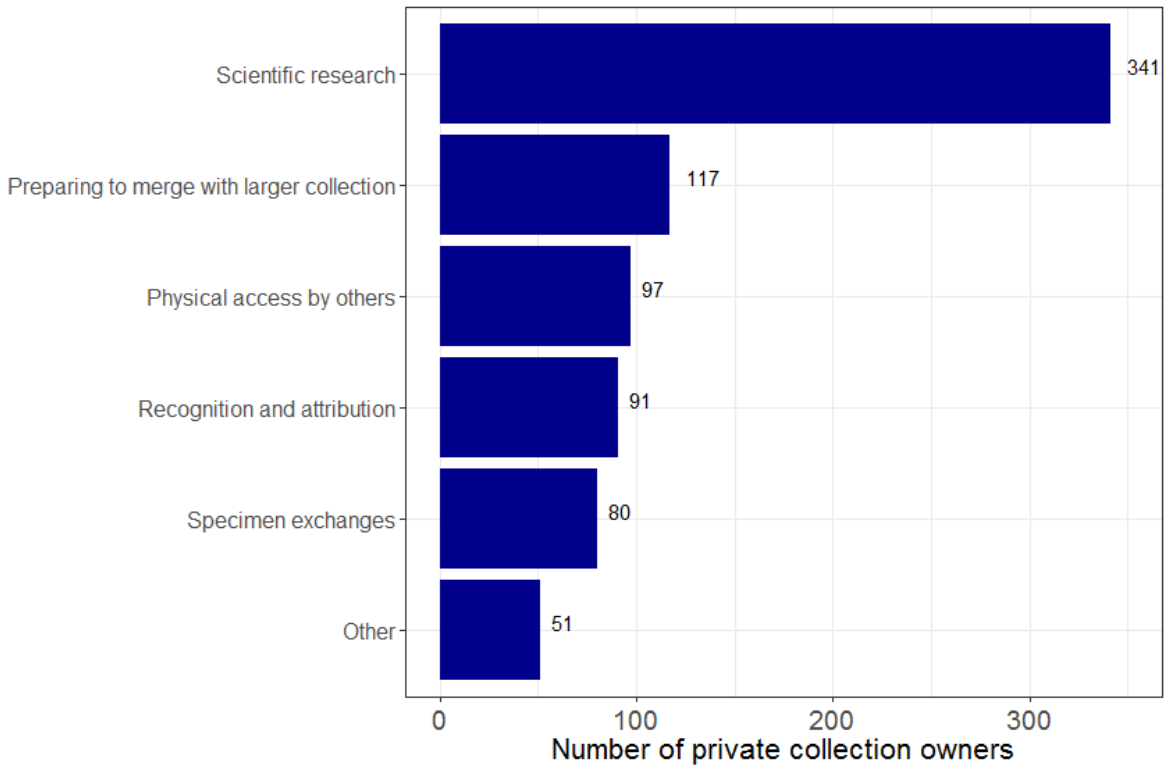

Figure 10. Number of private collection owners that indicated a certain reason to share collection data (Q13, n=404, multiple responses were possible).

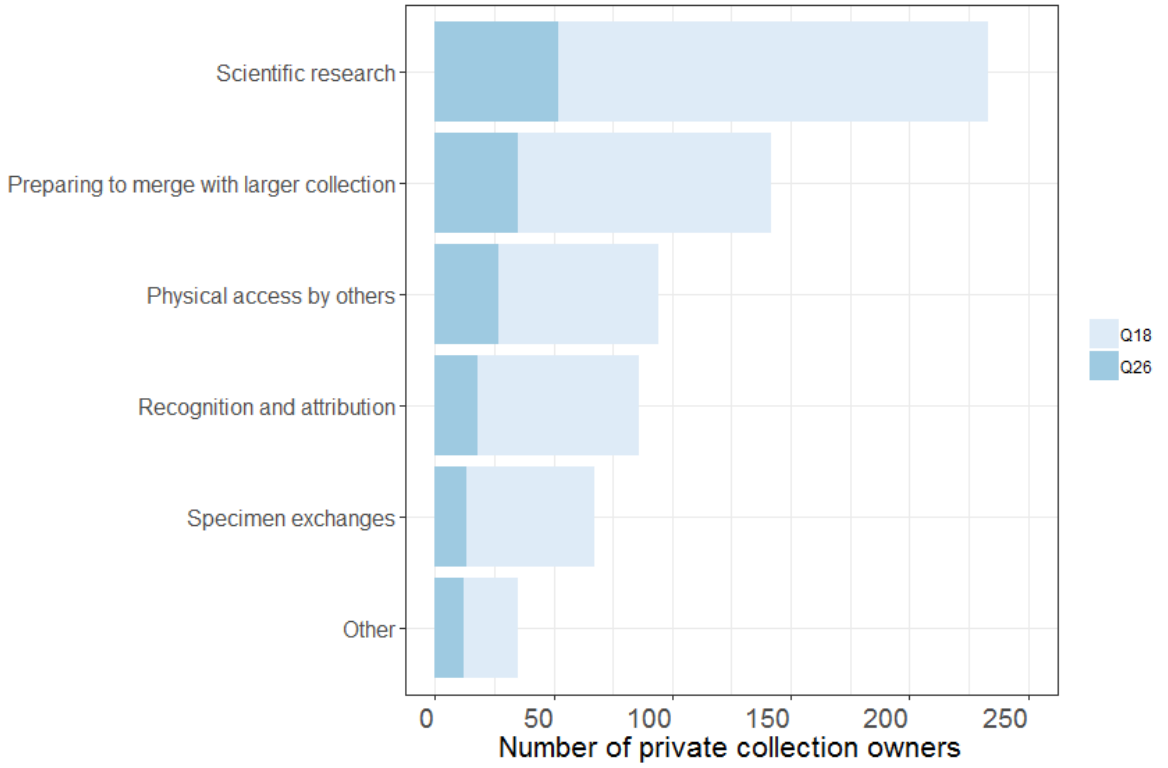

Figure 11. Number of private collection owners that indicated a certain reason to be sharing collection data in the future (Q18, n=243; Q26, n=73; total n=316, multiple responses were possible).

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

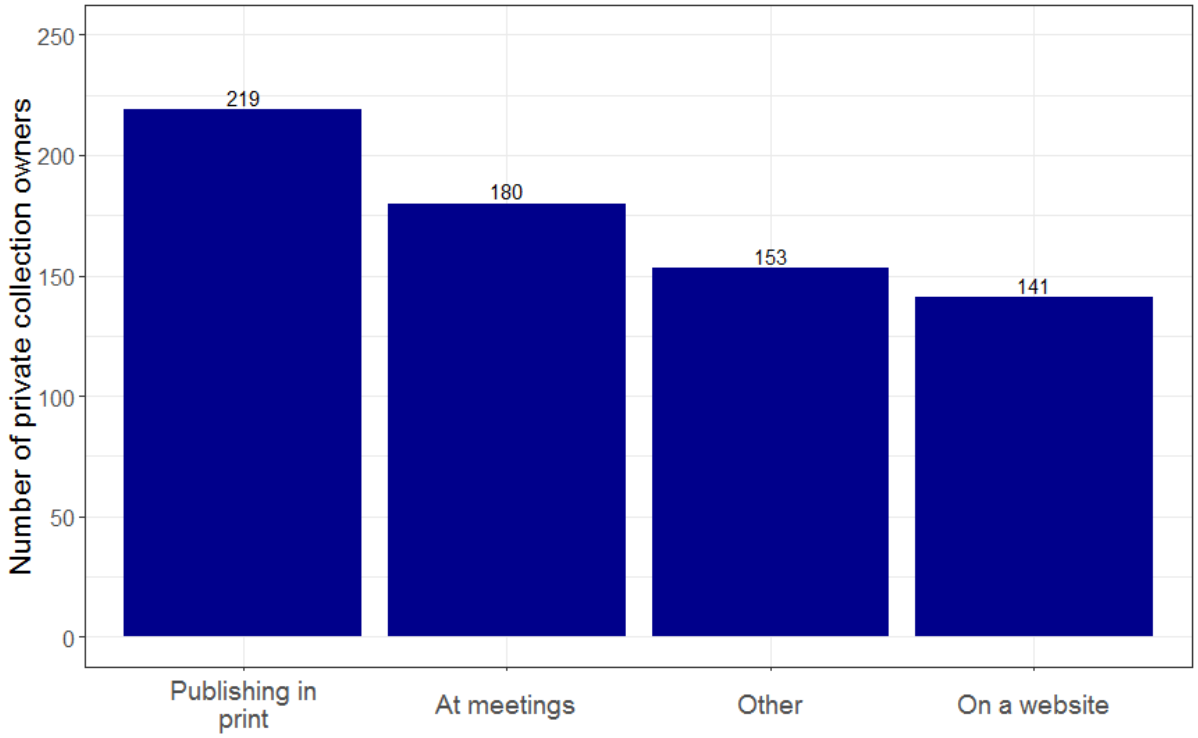

Figure 12. Number of private collection owners that indicated a certain method to share collection data (Q14, n=404, multiple responses were possible).

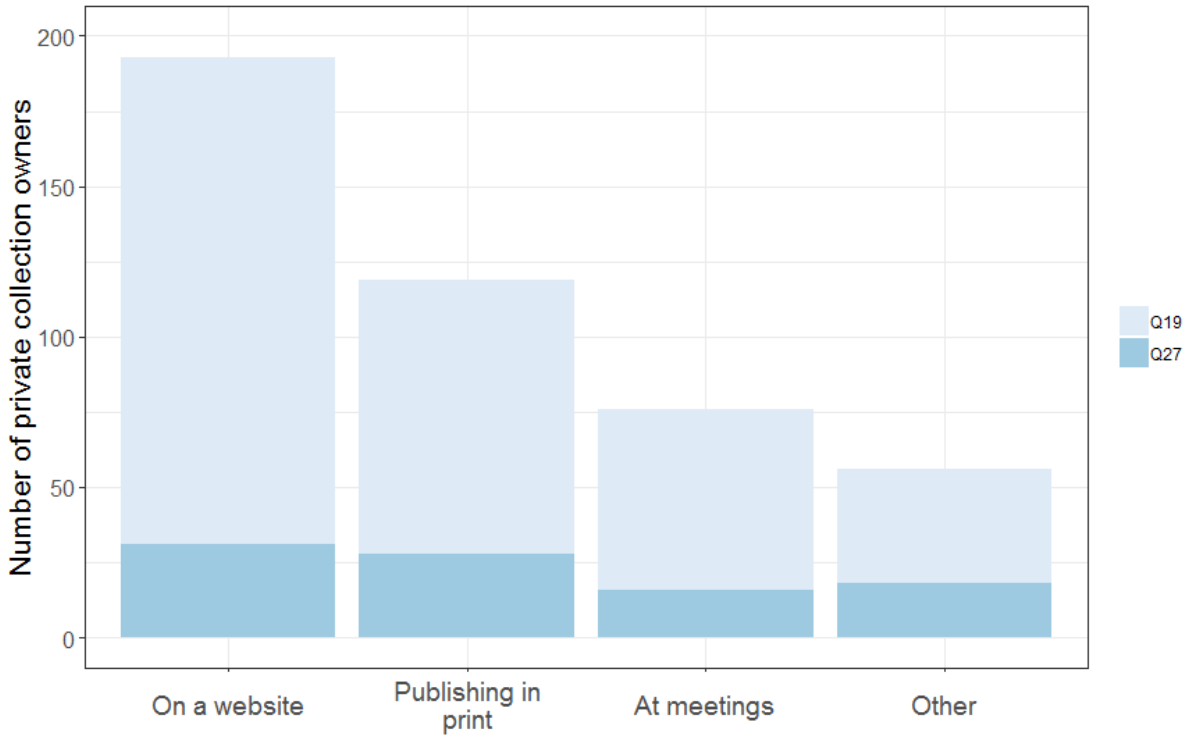

Figure 13. Number of private collection owners that indicated a certain method to be sharing collection data in the future (Q19, n=243; Q27, n=73; total n=316, multiple responses were possible).

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

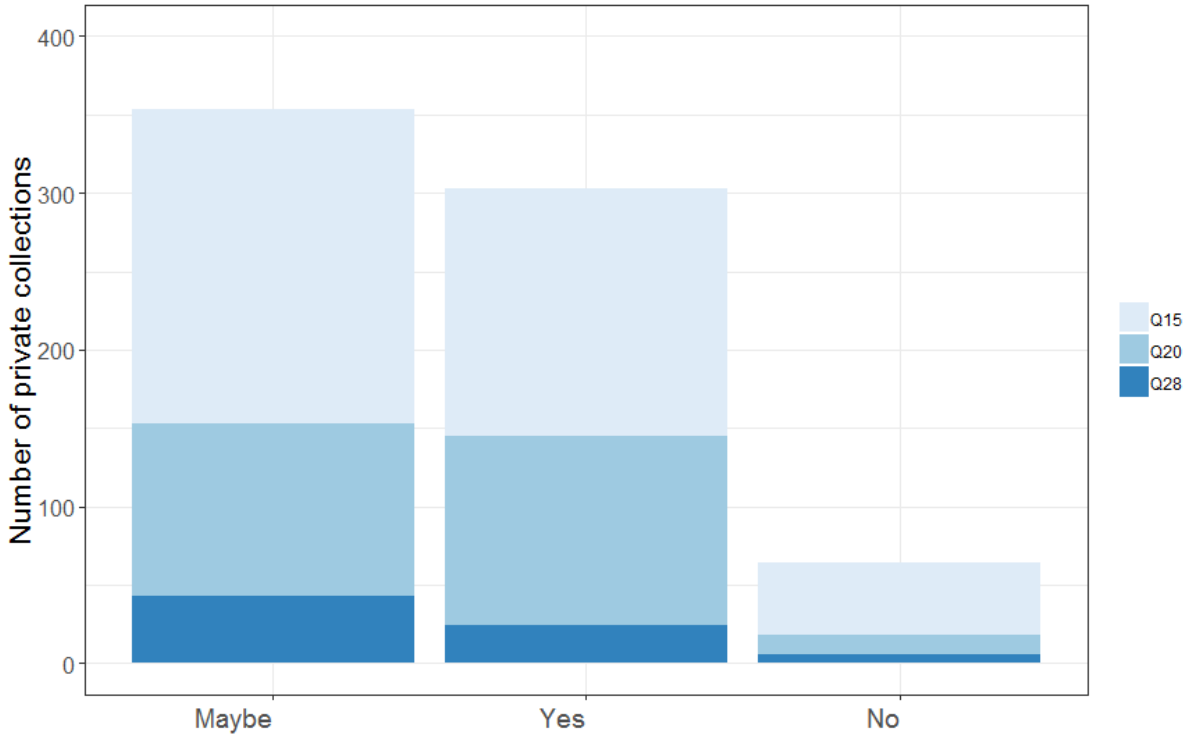

Figure 14. Number of private collection owners that indicated whether they were interested in a new website for registration of private collections, for three questions together (Q15, n=404; Q20, n=243; Q28, n=73; total n=720).

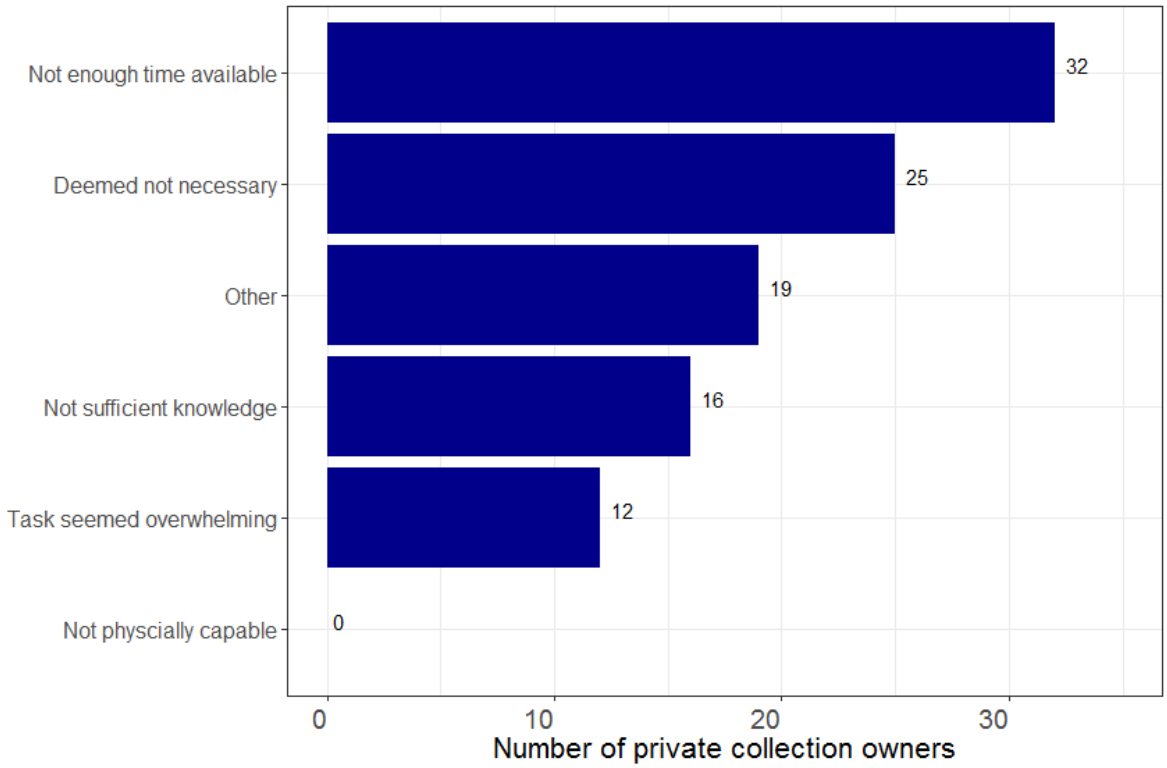

Figure 15. Number of private collection owners that indicated the reasons why they are not interested in sharing collection data, for two questions combined (Q23, n=62; Q31, n=3; total n=65, multiple answers were possible).

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

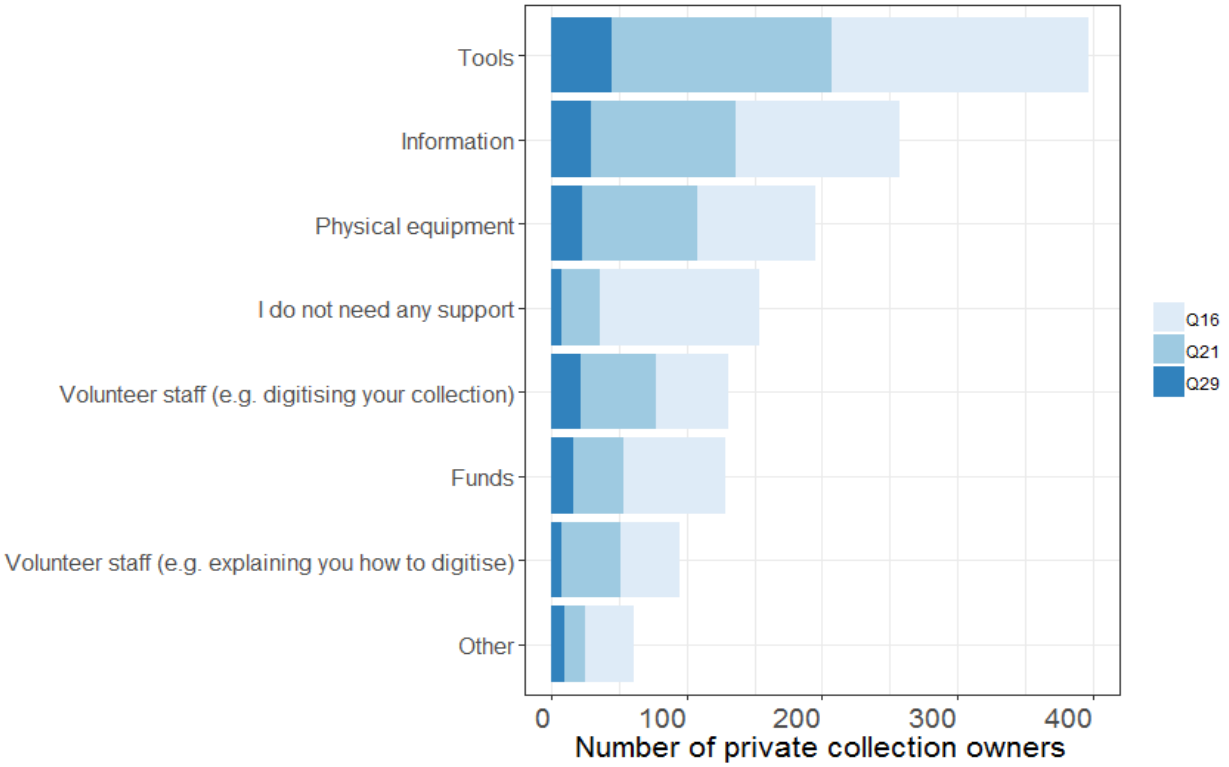

Figure 16. Number of private collection owners that indicated the kind of support they would need to digitise their collection at home (Q16, n=404; Q21, n=243; Q29, n=73; total n=720, multiple answers were possible).

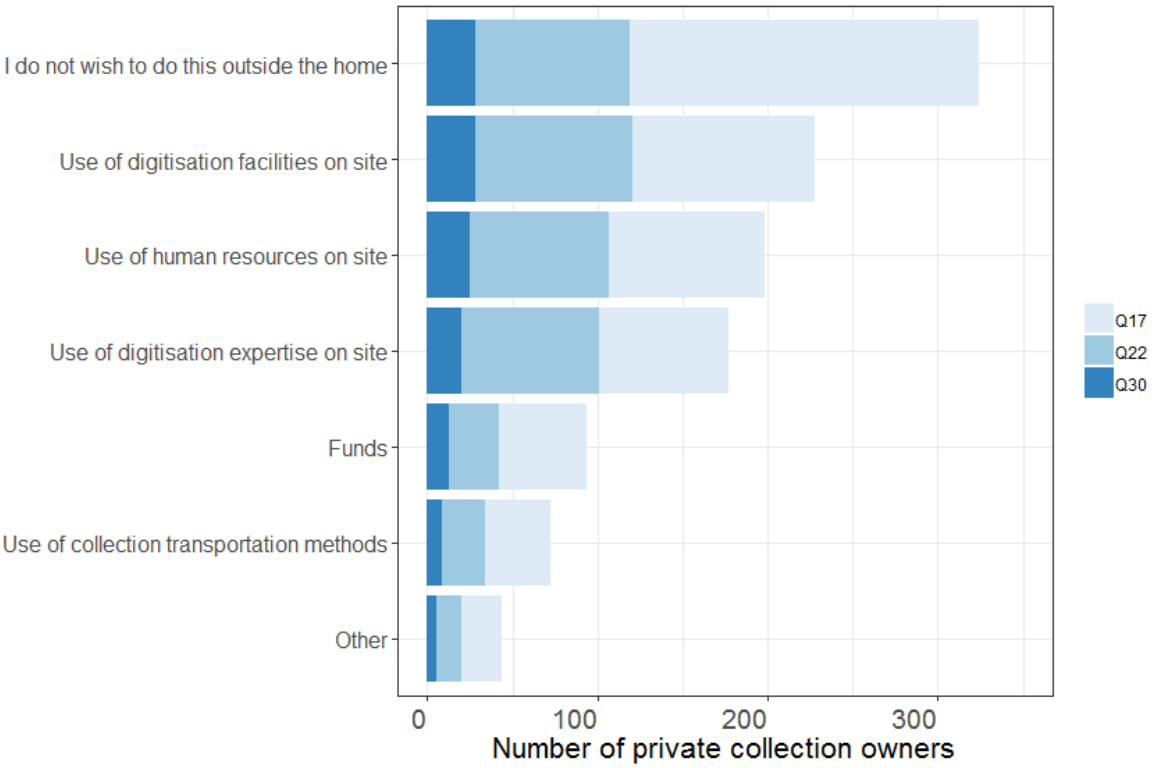

Figure 17. Number of private collection owners that indicated the kind of support they would need to digitise their collection outside their home at a nearby institute (Q17, n=404; Q22, n=243; Q30, n=73; total n=720, multiple answers were possible).

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

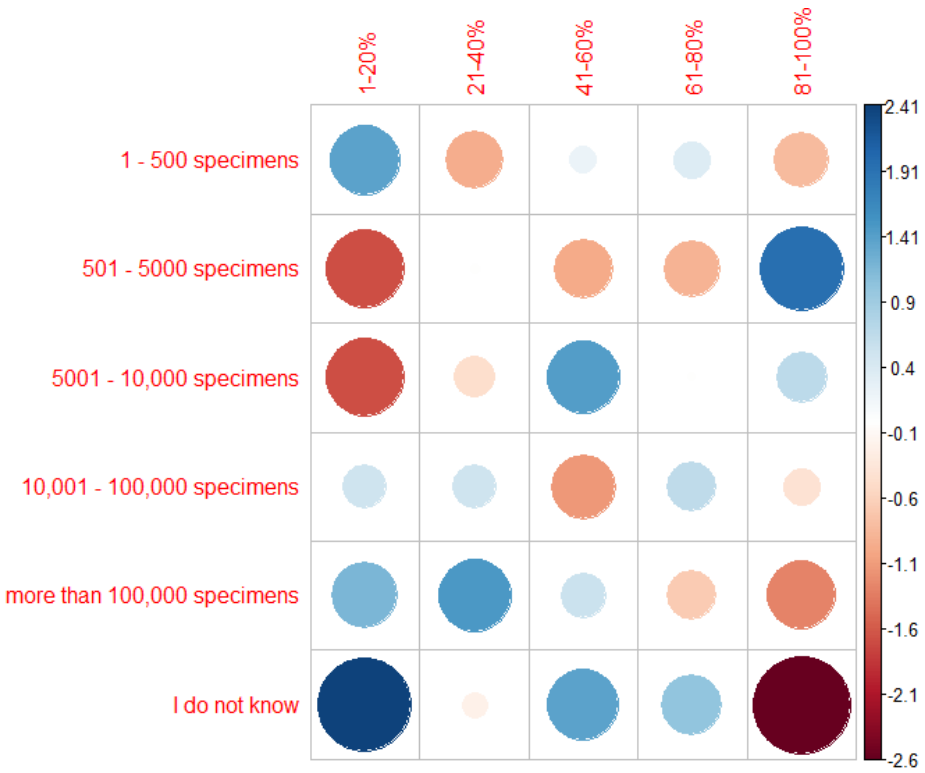

Figure 18. Relation between collection size and the percentage of the collection digitised at specimen level as indicated by private collection owners, represented by the residuals from the Chi-Square test. Large circles: strong association, small circles: weak association; Red circles: negative association, blue circles: positive association. Colours are represented as a scale.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

2.4 Discussion

This survey was intended to gather information about the volume, scope and digitisation of private natural history collections within Europe. Below we will indicate and discuss our main findings.

Volume

We found that the total volume of the sampled private collections is estimated to be between 8.8 and 32.6 million specimens. This is, however, expected to only be the tip of the iceberg, as for example the UK was not included and based on our personal experience there should be many more private collections within Germany and France. The total estimated number of specimens within Europe, as identified within this survey, thus resembles that of a larger national history collection. For comparison, Naturalis Biodiversity Centre belongs to the top 5 of natural history museums in size and has about 40 million specimens in its collection.

Scope

Regarding the scope of the identified private collections, we found that the most popular taxonomic groups that have been collected are insects, mollusks and vascular plants. These groups are dissimilar in the difficulty to digitise these collections, especially when creating specimen pictures is included. Creating pictures is easier for flat herbarium sheets than for multiple mollusks in a pot filled with ethanol, and may be easier for mounted and stretched butterflies than for very small pinned beetles. Since most private collections are focused on one taxonomic group, this will mostly require the set-up of one main digitisation strategy, instead of multiple, adding to the ease of digitisation. It will be important to acknowledge these differences when starting digitisation of private collections and adapt a suitable digitisation strategy.

Not surprisingly, most private collections contained specimens collected within Europe and only few in the Antarctic region. When comparing the ranking of the countries where the collection is kept with the ranking of in which countries most specimens are collected, these are roughly similar. In other words, most private collections appear to centre on collecting specimens from the country in which the private collection is based (e.g. a private collection in the Netherlands likely holds specimens collected in the Netherlands). There may be exceptions, however, where a collection is held in one country, and most specimens within this collection has been collected in another European country (e.g. on holidays). This would confirm the notion that private collections may hold unique, local biodiversity data that may

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

be interesting for related scientific research and nature conservation within Europe (Roma-Marzio et al. 2017).

Digitisation

The majority of private collection owners either uses a combination of managing collection data by hand and digitally, or does this completely digitally. It may therefore be that quite some collection data is already (partly) in a digital format. This is supported by our finding that private collection owners either had not or had just started digitisation of their collection, or were almost or completely finished with digitisation. It will be important to also explore digitisation strategies that focus on completing digitisation where needed and sharing this data online, which requires another strategy than starting digitisation of a private collection from the beginning. Thus, we need digitisation strategies that focus on both the beginners in digitisation as well as those more advanced in digitisation that may need guidance to share their collection data.

In general we found that the larger the collection (i.e. larger number of specimens), the lower the percentage of the collection is that has been digitised at specimen level. When a private collection contains many specimens, it is expected that it will take a lot of time and effort to complete digitisation at specimen level for the entire collection, hence resulting in a low digitisation percentage. Our findings seem in accordance to this expectation. An exception is the finding that the smallest collections (1-500 specimens) are positively associated with a low digitisation percentage (1-20%). It may be that private collection owners of very small collections want to digitise their collection but lack e.g. tools and information on how to do so, or do not see the need to digitise their collection (in the near future).

Data sharing

Interestingly, almost all private collection owners are interested in sharing collection data with others, specifically to support scientific research and to prepare a collection to be included in a larger natural history museum. When communicating with private collection owners, it will be important to emphasise these two aspects, especially the former, and how these are achieved when a private collection is digitised and shared. This may be a challenge, as there is currently not often a clear link between the data and scientific research that used this data. Linking data directly to publications to acknowledge private collection owners for their work may be one option to consider.

Private collection owners that currently share collection data mostly do so by publishing in print, while private collection owners that want to share collection data in the future have a

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

preference to do this online at a website. Our survey also showed that a new website to share collection-level data of private collections would be valued and considered useful according to survey respondents. Ideally, the new GRBio website, currently to be developed within GBIF, would not only be suitable to register natural history collections from institutions, but be expanded to be able to register private collections as well. It will be needed, based on the outcomes of Task 5.3, to identify what parameters need to be recorded for private collections and how these may differ compared to institutional collections.

The few private collection owners that are not interested in sharing data most often did not want to do this because they viewed it to be time consuming or unnecessary. The matter of not enough time available could for example be tackled by using volunteers to do the actual digitisation work, reducing the time effort of the private collection owner considerably. For those private collection owners that find digitisation unnecessary, information may help to shift this perspective (see Appendix 7).

Study limitations

We found that most of the identified private collections are kept in West-European countries. These happen to be the countries for which the survey has been translated to the national language (with the exception of Hungary), suggesting that language could pose a barrier to private collection owners when having to complete a questionnaire. On the other hand, this result may suggest that private collection owners of other European countries are either less interested/motivated to take part in a European-wide survey or there really are less private collections in these countries. It was found that most private collections are presently kept in the Netherlands which is likely biased, perhaps due to the degree of engaging amateur associations and translating the questionnaire to different languages. It is important to note that this is a small subset of all private collections that are kept in Europe (from some countries no response at all, we do not know how many private collections there are in total, we do not know which associations have forwarded our survey and hence do not know to how many individuals the survey was sent out etc.). Nevertheless, we believe this study provides an interesting first insight in the volume and scope of private collections across Europe. In the future, it will be necessary to enhance the participation of other European countries in surveys like these in order to have a more balanced view of private collections within Europe.

Digitisation resources for the future

To be able to digitise specimens at home, private collection owners mostly need tools and information. Tools, such as data set templates or web based digitisation, can be (further)

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

developed to support private collection owners in the digitisation process. Information regarding digitisation can be given in different ways, such as via workshops and guidelines. Amateur associations and natural history museums may work together in providing information, while larger initiatives may help to develop useful tools for private collection owners. As most private collection owners do not wish to digitise their collection specimens outside the home, it will be better to focus on improving and supporting digitisation of private collections at home.

2.5 Conclusion

Privately owned natural history collections show great potential to add valuable data and knowledge to the larger pool of European natural history collections. Future efforts to help private collection owners to digitise their collections at home should focus on providing and/or developing appropriate (online) tools and providing information on how to tackle digitisation (where to start?). In particular, there is a large interest in a new European-wide website where private collection owners can register and subsequently share collection-level data.

Besides motivation, practical and human resources there are other issues that crop up in the context of digitising private collections not dealt with in the survey and not elaborated on in this report. These issues pertain to matters like (future) transfer of ownership of the collection to an institute and linked to that the use of institutional registration codes, ownership of data, recognition of a collection and collection owner, lack or absence of proven legality of certain specimens, online publication of sensitive data, privacy and maintenance of datasets. Dealing with such issues often is complex and relates to institutional policies or (inter-)national rules, regulations or agreements and is beyond the scope of this task.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

3. Part II: Protocol to keep the inventory of private collections up to date

Besides a small group of collection owners that were contacted directly, within the current survey contacted private collection owners through secretariats of their associations. At this point in time, when there is very little information about European private collections available, we think this was the best strategy to reach private collection owners. The current approach, however, also carries some disadvantages: It makes it more difficult to get a response of all persons that own a private collection and to keep data in the future up to date. More particularly, in the current survey, issues related to obtaining a complete response and keeping information up-to-date were:

- Not all relevant associations may have been included; due to, amongst others, the language barriers when trying to find correct contact details of relevant associations online.
- Although not a single association secretariat explicitly refused to participate in the survey, there is no way of telling if each secretariat indeed forwarded the questionnaire to its members; other than sending a mail to ask secretariats for a confirmation that the questionnaire was distributed (which we decided against as this may turn out to be counterproductive due to the total number of emails sent) there was no mechanism in place to verify whether or not the survey had been forwarded and if so, to how many people.
- One country (UK) did not participate in the survey at all, except for two responses.
- Not all approached members may have responded to the survey; results from a previous similar survey carried out in the Netherlands in 2017 among entomologists indicated that based on actual knowledge of collection owners and the list of respondents at least half of the private collection owners had not responded. The proportion non-respondents/respondents on the total population of private collection owners across Europe is thus unknown; it may vary considerably between collections with different species groups or between countries.
- A share of the private collection owners is not a member of an association The proportion non-member/members on the total of collection owners is unknown as well and may again differ considerably between collections centering on different species groups or between countries.
- To keep an inventory up to date, the responses to a questionnaire have to be accompanied by information that uniquely links data to a specific collection; anonymous responses provide meaningful input for an overall analyses but are less adequate to assess changes over time.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

3.1 Options

Despite the above issues, a survey using a questionnaire was the only feasible option within the time-frame of the current Task and the means available. In order to overcome the above issues and strive for a complete and up-to-date overview of private collections several strategies can be envisaged to be employed in the future:

- A first option would be to more or less stick to the current approach and resend the same questionnaire every couple of years. This will require again and again a commitment of the amateur associations to participate. It also requires someone responsible to execute the questionnaire and to analyse the data obtained. If every time the same questionnaire is used, combining and analysing data will require limited time. The question remains what meaningful results can be obtained by repeatedly resending a questionnaire if respondents are anonymous and/or there is no way to know what proportion has responded. All in all, this strategy seems less suitable.
- Another option would require amateur associations to collect information about the private collections of their members following a protocol and then add this information on a website. Added value for amateur associations would be that they can see how their association compares to other associations and perhaps share their expertise regarding digitisation (e.g. via a forum). This will be further investigated in Task 5.3, which builds upon the work of this Task.
- The most promising strategy, it seems for now, would be to use a (European-centered) website where private collection owners can register information about their own collection themselves. In its simplest form it would be a website with rather high-level collection information that anyone can query. Results could be presented in a simple dashboard, showing how many collections and what kind of collections exist (similar to results shown in Figures 2-7). If user-friendly and offering collection owners full access right to their own data and providing interesting and useful information, such a website might become popular among private collection owners and a much larger community of stakeholders. If such a website becomes popular and starts receiving many hits, this may form an incentive for those collection owners whose data are not yet included to participate, resulting in a positive feedback loop.

A website to register private collections may be expanded to an online European platform where private collection owners can go to find information on how to digitise their collections. Amateur associations could also use this website to find information on what and how to communicate about the value of digitisation of private collections for scientific and social purposes. Developing and sustaining such a website requires a strong network,

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

ideally connected in some way to larger institutes or European initiatives like DiSSCo (http://dissco.eu), the European Citizen Science Association (ECSA; https://ecsa.citizen-science.net), CETAF (https://cetaf.org) or GBIF (http://gbif.org). Regarding the latter two, the development of a website for private collections could for instance be coupled to the development of the new GRBio website, likely as a part of GBIF, or the passports showing institutional collection information as used by CETAF (see for instance Naturalis).

On the former GRBio site, not only collection level information of natural history institutes was represented, but also private collection owners were able to register their own collection. The main challenge with this method would be how to keep collection owners motivated to keep their collection information up-to-date. An advantage about this method would be that all collections, both public and private, can be found on a single website, creating a more complete overview of natural history collections within Europe.

3.2 Conclusion

Developing an efficient protocol that brings together data and information of a large proportion of privately owned natural history collections across Europe is one of the bigger challenges for the immediate future. Taking into account among others the efficiency of such a protocol, the amount of human resources required to maintain it and the possibility to keep data and information up to date, a website is the most obvious option to implement such a protocol. Whether or not a completely new website is required or whether existing websites can be adjusted is still a matter which has yet to be investigated.

3.3 Recommendations

- An inventory of websites hosting information about collections is required to assess whether or not one of these can also be used to store information about private collections or whether a new website needs to be developed.
- Developing a protocol to store data and information from private collections would benefit enormously from collaboration between international organisations (CETAF, GBIF), musea, institutes, associations and private collection owners themselves.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

4. Part III: Communication strategy to engage private collection owners

In order to increase the degree of digitisation in private collections, collection owners need to be aware of the fact that digitising collection specimens and sharing data online is important for the global research community and contributes considerably to widening and increasing the use and general interest of collections. Being aware of this is a critical first step, but ideally collection owners should be entirely convinced of the necessity of digitisation. It will furthermore be equally important to clearly indicate how private collection owners can benefit from digitising and sharing their collection themselves, as private collection digitisation cannot solely rely on intrinsic, idealistic motivation of individuals (e.g. performing digitisation to help advance science). Collection owners who are convinced of the mutual benefits of digitisation are more likely to look for ways to actually begin with the digitisation process. The digitisation process itself could be started by the collection owners themselves. They may seek help from organisations (musea, institutes) or initiatives (DiSSCo) or they may become involved in national or international projects. A communication strategy to engage them should zoom in on both aspects: raise awareness and offer guidance for the process.

This communication strategy is intended to serve as a guideline that can be used to encourage and motivate private collection owners to participate in Europe's overall digitisation efforts. In particular, this strategy aims to indicate how communication with private collection owners can be used to:

- Increase awareness of the need for digitisation of natural history collections among private collection owners, including private collections and sharing the digitised information.
- Encourage private collection owners to take action and join current and future digitisation activities at institutional, national and international levels.
- Make private collection owners aware about existing knowledge and experience related to digitising collections and the fact that advice and assistance related to the digitisation process can be provided in many forms, including solutions for a lack of time.

The communication strategy described here, which is aimed at raising awareness and offering guidance for digitising privately owned collections, should take the following seven elements into consideration:

1. Stakeholders: the network surrounding private collection owners.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

2. Private collection owners: composition.
3. Message: which questions, needs, wishes and demands need to be addressed/tackled; what are the bottlenecks and incentives.
4. Means: what are the best channels for sending the message to the audience.
5. Goals: what do we aim for.
6. Resources: what is required in terms of time, manpower and information.
7. Evaluation: how do we measure the effectiveness of this communication strategy

4.1 Stakeholders

Private collection owners operate in a landscape that includes various other stakeholders who are involved in, or value digitisation of, natural history collections. A communication strategy should carefully consider this landscape and the network that connects the various players. Private collection owners may be a member of an amateur association, may be more or less actively involved in a regional or national museum and may even be active in national or international initiatives, projects and programs. Due to the variety of stakeholders which may interact with private collection owners, it may be wise to opt for a tiered communication strategy in order to strengthen the overall message aimed at the private collection owner. Depending on factors like the nature of the collection, the (natural) role of an association or museum as well as the international context, a communication strategy could be more efficient if designed to be tiered and not only directed at the individual private collection owner. This means that there needs to be both, but different, communication towards individual private collection owners and amateur association and museums.

4.2 Private collection owners

Despite the fact they all have a collection, private collection owners are not a homogeneous group by any standards. There are four key areas related to private collection owners that a communication strategy should take into account:

- Motivation: why does someone have a collection? What aspects are important? Is there a social aspect to it (meeting people), is it used for research leading to scientific publications or is the collection owner a ‘stamp collector’ (Lee and Trace, 2009; Saridakis and Angelidou, 2017).
- Person: personal aspects like age or computer literacy may play a role. This was indicated several times in our survey to explain why some private collection owners had not started digitising their collection.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

- Collection: the nature of the collection (plant/animal, dry/spirit, samples/specimens, big/small, well curated or not) affects the owners attitude towards (aspects of) digitisation and sharing data. For instance, a small group of collection owners that in the survey indicated an unwillingness to share information explained this with the sheer size of the task and with the time required to do it.
- Network: a private collection owner may operate completely on his/her own, may be a member of an association or may have a global network and be involved in all kinds of (inter-)national projects and initiatives. This is expected to shape his/her ideas regarding digitisation.

A communication strategy thus has to take into account the fact that collection owners are a very diverse group. In order to be successful, a communication strategy should therefore be flexible and allow for more than just one message and communication channel. A single message and a single communication channel are never sufficient to reach the whole community and bring the message across.

4.3 Message

Although messages have to be tailored when addressing specific audience groups, there are a number of general statements to be made that apply to all private collection owners. Each message should emphasise the importance of private collections for research, highlighting the benefits of digitising and sharing collection data. It should include statements like:

- Although it may not always be evident, efforts required to build a private collection and carrying out research related to this collection are highly appreciated by the society in general and the research community in particular.
- Private collection owners most certainly are and should consider themselves part and parcel of the community of staff and researchers associated with natural history collections.
- Private collections are important in addressing taxonomic and geographic gaps in the overall collection (Deborah Paul, personal communication);
- Private collections may in fact constitute a significant part of the entire natural history collection (both institutes and private collections combined) within a country (Roma-Marzio et al. 2017);
- Private collections may hold unique and crucial temporal and spatial data.
- Information, knowledge and lessons learned about digitising (private) collections are available and can be shared.
- Private collections constitute the museum collections of the future; at one point in time private collections always end up in a museum or a similar organisation that receives external funding.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

Clarification of the term 'digitisation'

Digitisation is a container concept that, depending on the context or situation, may have different meanings. For communication purposes it is crucial that both the sender of a message and the receiver of that message attach exactly the same meaning/definition to the term digitisation. If there is a mismatch between sender and receiver regarding the term digitisation this may result in misunderstanding which may lead to frustration, resistance and unwillingness at both sides. Appendix 6 gives a brief summary of six aspects of digitisation that may differ between private collections and between private collection owners and have to be taken into account during communication.

Importance of digitisation

The survey indicated that only few of the respondents (14%) had not managed their collection information neither hand written nor digitally. About one-fifth (21%) only had hand written information about their collection, whereas the majority (65%) has collection information digitally available to some degree. While only few collection owners do not see a need to digitise collections (64 responses or 8%), about half of the private collection owners indicated they have digitised 80-100% of their collection at specimen level. The survey did not allow to make a more detailed assessment of the exact nature of the term 'digitisation'. It is therefore not clear which data individual collection owners included in their digitisation effort, at what level their collection was digitised and whether or not for instance images were included. Overall the survey showed that anno 2018 digitising collections is either already being tackled by most collection owners or, if this is not the case, collection owners do realise its importance and the need to start looking into it. This means that a communication message could be less focused on why it is important to perform digitisation, but more on tackling aspects of time limitation and lack of the required knowledge and/or skills to digitise their collections. Especially time is a valuable and limited resource, and it may well be that private collection owners rather spend time working on their collections instead of digitising and sharing their private collection. Working on the collection is naturally their main interest.

The few private collection owners that do not see the need to digitise collections may have different reasons. Perhaps they do not understand how their collection can benefit a wider community or what benefits there are for themselves. Perhaps they do not feel a member of a community that could benefit from digitisation, they are "stamp collectors" that have no affiliation with research or their collection is still in disarray making it impossible to start digitisation. As this is a small and diverse group of private collection owners, it will be difficult to change their ideas about digitisation. Possibly it helps to show them positive cases where private collection owners successfully digitised their collection.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

Importance of sharing data

The prime goal of digitising collections is to allow data and information held in a collection to be shared online with the world. When asked about media or methods collection owners use to share collection information, traditional means like meetings and publications are roughly considered equally important as the use of websites. Although most of the private collection owners are interested in sharing collection information on a new website, there is clearly some doubt. The survey indicated that there is still a need to communicate why digitised data of private collections need to be shared, and in particular how this could be done in practical terms. Although not included in the survey, when emphasising the importance of publishing digitised collection information online, three aspects need to be taken into account:

- Sensitive data: specimen data may contain details that are privacy related (for instance the whereabouts of collectors at specific dates). An example on how to deal with this subject is the Privacy Policy used by GBIF:

  > *GBIF has legitimate interests in collecting and maintaining the information needed to provide biodiversity-related evidence that supports scientific research and policy. Beyond that, we collect the minimum amount of personal information needed to fulfil the purpose of your interactions with us. We don't sell this information to third parties, and we process it only as described in this Privacy Notice. As an EU-based body, we comply with the General Data Protection Regulation (GDPR). But regardless of where you come from, where you are or where you live, we apply the same standard of privacy protection to all our users.*

  Another category of sensitivity relates to very accurate locality data of rare and endangered species. A report dealing with this kind of data was published by GBIF in 2008 by Chapman and Grafton. This document reviews current approaches for obscuring or generalising such data, with a set of recommendations to guide data holders to develop their own policies. A third category of sensitivity relates to specimens collected in countries or areas requiring general permits at the time of collection or worse, specimens collected in protected areas or national parks where collecting is forbidden. It may not always be clear when handling specimens from private collections whether these were collected while adhering to national and/or local regulations, e.g. the collector being in the possession of relevant permits (see for instance permit requirements for Australia).
- Ownership: although the ownership of physical specimens in a private collection may be rather straightforward, private collection owners may raise a question about

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

ownership of digital data. This question has been subject of a white paper in the context of the European data economy (Van Asbroeck et al. 2018):

> *Against a background where the EU strives towards a data-driven environment in which both citizens and companies can reap the benefits of novel data technologies, but also against a background where the current legal framework does not sufficiently tackle all the issues related to data and where actors involved in the data value chain have no certainty as to the ownership of the data they have gathered, created, analysed, enriched or otherwise processed, we conclude that a more solid and legally secure solution is needed.*
> *In such context, we suggest the creation of a non-exclusive, flexible and extensible ownership right in data(sets), with a data traceability obligation as a safeguard. In addition we discuss the specificities of said right and obligation, their interaction with the other existing rights in data, their incidence on civil law, and their possible reflection in contractual arrangements.*

- Open access: across the globe and across the research community there are different views on the self-evidence of sharing data and information. Differences in view may be principal (digital information originating from biodiversity collection should or should not be available for everyone without restrictions) or pragmatic (certain data or information should not be published online as it is still being used in current research).

Rather than emphasising the importance of digitisation itself, which seems superfluous (see above), a communication strategy aimed at engaging private collection owners should emphasise on the importance of sharing information via websites and how this can be accomplished. Appendix 7 summarises some specific arguments/answers that could be used/given when discussing the pros and cons of digitising private collections and sharing the data.

Incentives for digitising and sharing data

To raise awareness, increase motivation and to entice private collection owners to actually start digitising their collection and share data, incentives can be quite useful. Incentives first and foremost are (in-)direct advantages linked to digitisation which are beneficial to the private collection owner. Appendix 7 mentions four such advantages.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

As many private collection owners are interested to support scientific research by sharing data from their private collections, it may be a good strategy to stress how private collections contribute to (scientific) research. This could for example be done by showing publications that make use of private collection data. Another possibility can be to demonstrate how private collection data are applied to deal with social, agricultural issues or issues related to nature conservation. An example of the latter is the possibility to use private collections and the information these contain to assist in preparing Red List Assessments (RLA) for the IUCN. RLA's form the basis for conservation actions directed toward endangered species. Taking part in such initiatives may increase their engagement and determination to digitise their collection:

> *"For the Butterfly Conservation Assessment the Lepidopterist Society of Southern Africa was a major partner in the project and their members were involved in all aspects of the project. They were involved as authors in the project – so received full credit. The data were all integrated – from museums and amateurs for the analyses of conservation status."*
> (Michelle Hamer, personal communication)

Unfortunately, the (scientific) goals of DiSSCo are currently (2018) too general and too far in the future to be directly appealing to private collection owners. Questions they may have could be: What research could the data support and who is going to use the data of my private collection? It may be useful to stress that by giving unique identifiers to the collection as a whole and/or individual specimens, these can always be cited in (scientific) publications.

To summarize, based on the survey results, a communication strategy aimed at digitizing private collections should contain messages and incentives directed towards the process (the how) and the sharing of data rather than the motivation (the why) behind digitisation. Incentives can be about tools (e.g. a dataset template, web based digitisation) or information (e.g. in the form of a guideline, workshop or help desk) required to digitise or options to reduce time required by collection owners to digitise (for instance the deployment of volunteers). In Task 5.3 of the ICEDIG project a number of pilots will be carried out to test various approaches to cater for the numerous requirements or conditions that would enable collection owners to start digitising their collection.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

4.4 Means

As many private collection owners are a member of regional, national or international associations, these associations form a very important channel through which collection owners can be reached. Communication strategies should therefore not only be addressed directly to private owners but also to their associations and possible other stakeholders (see also 3.1). Notwithstanding the importance of associations as a channel for communication, it is important to keep in mind that only part of the members of associations own a private collection and that not all collection owners are a member of an association. Reliable estimates for both groups are not available. Although messages to associations and private collection owners may be similar, a range of methods are required to reach the largest possible part of the private collection owners (NatSCA report 2018 (Advocacy Toolkit), Deborah Paul, personal communication):

- Presenting at conferences, symposia and meetings of amateur associations
- Organising dedicated workshops and meetings
- Communicating through international programs like DiSSCo
- Posting messages on websites of amateur associations
- Distributing information through newsletters
- Using social media (Facebook, Twitter, Instagram etc.)
- Spreading the word on fairs used by some associations (fossils/stones/minerals)
- Writing articles for popular magazines or research journals

Amateur associations of course can advocate the importance of digitisation of private collections once they are convinced of the benefits and thus play an important role in delivering the message.

By far the biggest obstacle to determine the most efficient means to deliver messages about digitising private collections is the fact that there simply is no complete, up-to-date list of private collection owners. A survey like the one carried out in this task depends completely on the willingness of associations and private collection owners to participate. In the hypothetical scenario that all associations are known and do cooperate and also all their members likewise do cooperate and return a questionnaire, there are still the private collection owners that are not a member of any association. As a result, communication strategies have to be based on an incomplete overview of the target audience at best as far as overall numbers, addresses and basic information about collections goes. Obtaining complete and up-to-date lists, if possible at all, may be very time consuming. So far we have not been able to come across an approach for achieving this goal. In conclusion, it seems likely that a diversity of communication means will be needed to reach the diverse target group.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

4.5 Goals

There are three obvious goals linked to a communication strategy centering on digitising private collections, aimed at engaging private collection owners:

- To increase awareness of the importance of digitising collections
- To increase the overall digitisation effort taking place in private collections and
- To increase the volume of data shared online originating from private collections

4.6 Resources

Any communication strategy aimed at getting private collection owners actively involved in digitisation requires time and effort. Associations or musea do not have the means nor the urge to develop elaborate and dedicated communication strategies to get private collections digitised and their data shared on internet. Musea or associations deal with the digitisation of private collections on an ad hoc basis rather than based on a long term vision: if a request or question pops up, it is dealt with there and then. As far as we know, there have not been any projects or programs in Europe or outside Europe for that matter specifically directed at digitising private collections nor are there any currently taking place.

Rather than a local, regional or national issue, digitising private collection is an activity that is best tackled in an international context. Initiatives like DiSSCo that aim to create an European research infrastructure are likely candidates for such an enterprise. They may have the necessary resources not only to more firmly incorporate amateur researchers in the professional research community but also in making data and information from privately owned collections accessible.

4.7 Evaluation

Evaluating the success and efficiency of a communication strategy is not easy. This is caused by:

- Not having a complete and up-to-date list of collections that are privately owned. This makes it difficult to assess the response rate or the percentage of the target audience that is being reached.
- The absence of detailed information on private collection owners. Therefore it is difficult to measure the effect of any communication campaign: how many collection owners changed their attitude towards digitisation, were persuaded to start digitisation or sharing their data online, etc.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

The question is if and when evaluation of a communication strategy is meaningful and worth the time and energy spend on it. Considering the ever-advancing insight in among others the degree of digitisation, differences between regions, countries, cultures and dissimilarities between collections centered on different groups, an evaluation and identification of lessons learned may change over time as well.

4.8 Conclusion

Based on the survey results, some of the first conclusions regarding a communication strategy aimed at engaging private collection owners are:

- Only very few owners of private collections are not yet convinced of the importance of digitising their collection.
- Details regarding sharing digital information and data is a subject that requires more attention in future communication directed towards private collection owners.
- Providing more information about the practicalities of digitisation appears to be another topic that should receive a lot of attention in any communication campaign directed toward private collection owners.
- The biggest issue for any communication strategy is the lack of detailed information of the target audience which makes it difficult to optimally align communication to the targeted audience nor does it make an evaluation of the success rate of a communication strategy any easier.

4.9 Recommendations

Considering the above statements regarding the communication strategy the following recommendations can be made:

- The survey results indicate that a communication strategy should tackle understanding the implementation of digitisation and sharing data online rather than raising awareness among private collection owners.
- When planning a communication strategy, it is important to bear in mind that besides private collection owners there are other stakeholders (associations, musea, international organisations) that could play an important role.
- Any communication strategy should take into account that there is no such thing as an average private collection owner. Personal aspects, motivation or the taxonomic group of their collection are for example factors which make private collection owners as diverse as the taxa they have in their collection.
- It is of utmost importance in any communication strategy to be clear about the definition/meaning of the term ‘digitisation’.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

- As there is currently no complete, up-to-date list of private collections, this may interfere with choosing the channels for communicating or with evaluating a communication strategy.
- As there are many uncertainties regarding the target audience, their collection and degree of digitisation it is probably best to opt for a flexible means of communication including as many options as possible.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

5. Conclusions and next steps

The aim of this deliverable 'Prioritizing scientific and societal needs for data of private collections' (D2.2) was to address how private collections can be included in Europe's digitisation efforts. Here we identify the main conclusions and our recommendations regarding the current state and digitisation of private collections in the future.

First, we identified the current volume, scope and level of digitisation of private collections within Europe to have a better view of the background and inform our next steps. According to the performed survey, there is a fast body of specimens to be found within private collection owners and this is expected to be the tip of the iceberg only. This means that there are potentially many specimens that will become part of a larger natural history collection in the coming decades. To not add to the already existing backlog in digitisation of specimens at natural history institutions and to not lose valuable information that can only be given by the private collection owner himself, it will be crucial to tackle digitisation beforehand, together with the private collection owner. Private collection owners indicated in the survey that they are interested in digitisation and sharing collection data, but prefer to digitise at home. Tools and information are most often wished for by private collection owners, so these need to be developed in the future to enhance digitisation of private collections. For this purpose, there needs to be clear and effective communication between (inter)national initiatives and natural history institutions, and private collection owners.

Secondly, we discussed and proposed a protocol to keep this inventory of private collections up to date in the future. Due to the nature of the survey, it will not be possible to sample the exact same subgroup from the entire population of private collections. Instead, it is recommended to develop a single platform where private collection owners can register their collection and perhaps also find information regarding the process of digitisation (e.g. by including a digitisation protocol). This survey in particular provides strong support to develop a website where private collection owners can register collection-level information. Future activities centered on engaging private collection owners in Europe's digitisation efforts should be focussed on developing a website where private collections can be made visible and discoverable to the wider community. Ideally, this should be incorporated with existing initiatives to ensure a coherent and complete overview of natural history collections within Europe, such as integration within the new GRBio website as is expected to be developed within GBIF. This will need to be further explored together with (inter)national initiatives and private collection owners.

Finally, we explored the elements of a communication strategy to stimulate awareness and participation among private collection owners about digitisation. In order for private

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

collection owners to digitise his or her collection, essentially three requirements need to be met:

- The private collection owner should be convinced of the necessity/usefulness of digitisation
- The private collection owner should have the knowledge and skills required to perform digitisation
- The private collection owner is willing and able to spend the time needed to perform digitisation

The current survey indicated that by far most of the private collection owners are convinced of the usefulness of digitisation but that those that have not performed digitisation thus far either lack knowledge/skills or time. Providing guidelines and tools may be enough to tackle the lack of knowledge and/or skills, while solving the issue of time needed to perform digitisation requires more effort. This needs to be included in a communication strategy aimed at increasing engagement by private collection owners in collection digitisation. In this deliverable, we gave an outline of what needs to be included in a communication strategy targeting amateur associations and private collection owners. A detailed communication strategy can be devised and deployed when there is a better view of what private collection owners need and would appreciate during actual contact. Parts of the proposed communication strategy can be tested in Task 5.3, where communication with amateur associations and private collection owners is occuring. For example, giving a digitisation workshop at an amateur association can be used to test the communication about digitisation work flows in person.

Following this task in the ICEDIG project, another task (Task 5.3) will run a number of pilots to specifically assess in what way private collection owners can be assisted with digitisation. The result of this task will be the description of a number of protocols/scenarios that can be used for the digitisation of private collections. Depending on the taxonomic group, the private collection owner, the willingness and potential of regional or national institutes to assist etc. one or the other protocol could be used. If such protocols prove effective and can be implemented without dedicated personnel for any taxonomic group in any European country, all that is needed is a well organised communication campaign to spread the message. This will hopefully result in even more private collection owners picking up the news, becoming intrigued and joining the ongoing digitisation efforts in Europe.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

Acknowledgements

We would like to thank the following ICEDIG members that have translated the questionnaire and accompanying mails, and who are not official participants in this task: Karsten Gödderz (German, CETAF) and Marc Pignal (French, MNHN). In addition, we are very grateful for the guidance and translations of the questionnaire and accompanying mails provided by Luca Bartolozzi (Italian, CETAF representative for Italy) and Jesús Muñoz and Charo Noya (Spanish, GBIF nodes and CETAF representatives for Spain). Finally, we would like to thank all associations that distributed our survey and all private collection owners that responded to our survey, taking the time to help us gain a good overview of private collections within Europe.

References

Ariño, A.H. 2010. Approaches to estimating the universe of natural history collections data. *Biodiversity Informatics*, 7:81-92. https://doi.org/10.17161/bi.v7i2.3991

Chapman, A. D. and O. Grafton. 2008. Guide to Best Practices for Generalising Primary Species-Occurrence Data, version 1.0. Copenhagen: Global Biodiversity Information Facility, https://www.gbif.org/document/80512/guide-to-best-practices-for-generalising-sensitive-species-occurrence-data

Kemp, C. 2015. Museums: The endangered dead. *Nature*, 518:292-294. https://doi.org/10.1038/518292a

Lee, C.P. and Trace, C.B. 2009. The role of information in a community of hobbyist collectors. *Journal of the American Society for Information Science and Technology*. 60:621-637. https://doi.org/10.1002/asi.20996

NatSCA report. 2018. Advocacy Toolkit for Natural History Museums. https://naturemanchester.files.wordpress.com/2014/08/advocacy-toolkit-1-general-notes-on-advocacy.pdf

Roma-Marzio, F., Peruzzi, L. and Bedini, G. 2017. Personal private herbaria: a valuable but neglected source of floristic data. The case of Italian collections today. *Italian Botanist*, 3:7-15. https://doi.org/10.3897/italianbotanist.3.12097

R Development Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org

Saridakis, C. and Angelidou, S. 2017. A case-based generalizable theory of consumer collecting. *European Journal of Marketing*. 52:946-972. https://doi.org/10.1108/EJM-10-2016-0570

Stearn, W.T. 1971. Sources of information about botanic gardens and herbaria. *Biological Journal of Linnaean Society*, 3:225-233. https://doi.org/10.1111/j.1095-8312.1971.tb00184.x

Van Asbroeck, B., Debussche, J. and César, J. 2018. Building the European Data Economy - Data Ownership White Paper. Bird and Bird, written as part of the TOREADOR EU project. https://www.twobirds.com/en/news/articles/2017/global/data-ownership-in-the-context-of-the-european-data-economy

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

Appendices

Appendix 1. Survey questions

Questions
Q1: Do you have a private collection of natural specimens?
Q2: In which country do you keep the collection?
Q3: Which (taxonomic) groups do you have in your collection? Multiple answers possible
Q4: What is the geographic range of the collection? Multiple answers possible
Q5: From which European country has the main part of the collection been sampled?
Q6: Are you a member of an association focused on either nature, science or natural history?
Q7: How many specimens do you have in your collection in total?
Q8: Is the collection being used for scientific research?
Q9: How do you manage your collection information?
Q10: What level of information do you record? Multiple answers possible
Q11: For which percentage of the entire collection have you recorded specimen information?
Q12: Do you share your collection information?
Q13: Why do you share your collection information? Multiple answers possible
Q14: How do you share your collection information? Multiple answers possible
Q15: Would you be interested in sharing your collection information on a new website where you can register your private collection yourself (European wide)?
Q16: If you are interested in sharing collection information at specimen level, what kind of support would you need to digitise your collection specimens at home? Multiple answers possible
Q17: If you are interested in sharing collection information at specimen level, what kind of support would you need to digitise your collection specimens outside the home at a nearby institute? Multiple answers possible
Q18: Why would you like to share your collection information? Multiple answers possible
Q19: How would you like to share your collection information? Multiple answers possible
Q20: Would you be interested in sharing your collection information on a new website where you can register your private collection yourself (European wide)?

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

Q21: If you are interested in sharing collection information at specimen level, what kind of support would you need to digitise your collection specimens at home? Multiple answers possible
Q22: If you are interested in sharing collection information at specimen level, what kind of support would you need to digitise your collection specimens outside the home at a nearby institute? Multiple answers possible
Q23: Why do you prefer not to share your collection information? Multiple answers possible
Q24: Have you considered managing your collection information?
Q25: Would you like to share your collection information?
Q26: Why would you like to share your collection information? Multiple answers possible
Q27: How would you like to share your collection information? Multiple answers possible
Q28: Would you be interested in sharing your collection information on a new website where you can register your private collection yourself (European wide)?
Q29: If you are interested in sharing collection information at specimen level, what kind of support would you need to digitise your collection specimens at home? Multiple answers possible
Q30: If you are interested in sharing collection information at specimen level, what kind of support would you need to digitise your collection specimens outside the home at a nearby institute? Multiple answers possible
Q31: Why do you prefer not to share your collection information? Multiple answers possible
Q32: Please give your name and e-mail address if you want to be among the first to receive a summary of the results of this survey.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

Appendix 2. List of associations

See the separate Excel document for the list of contacted European, national and regional associations.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

Appendix 3. Announcement mail to associations

Dear association,

This mail is send to national associations focused on either nature, science or natural history across Europe and is addressed to you as a secretary or contact for one of these associations. With this mail, we would like to announce to you that we will soon distribute a questionnaire on the volume, scope and level of digitisation of private collections within Europe. In addition, we would like to know what kind of resources private collection owners need for digitisation and dissemination of their private collections. As many association members have a private collection, we hope you are willing to help us reach these private collection owners by forwarding and drawing attention to this questionnaire. The survey is scheduled to be send out by the 31st of July, contains 20 questions and takes less than 15 minutes to complete. This survey is part of the ICEDIG (Innovation and consolidation for large scale digitisation of natural heritage) project, a design study for the DiSSCo (Distributed system of scientific collections) research infrastructure (icedig.eu and dissco.eu).

When it comes to privately owned natural history collections, these are often 'hidden' and may harbor highly valuable specimens for a specific taxon or (local) geographic range. On a European scale, these private collections combined are expected to form a significant part of the natural history collections as a whole. Therefore, it is important to include private collections in Europe's digitisation efforts of natural history collections. The outcomes of the questionnaire will be used to write guidelines for the digitisation of private collections, to perform a couple of pilot digitisation pilots of private collections and to present information about private collections to policy and decision makers. If we can show that there are many interesting private collections in Europe, we can make a strong case to focus on these natural history collections as well. It is therefore of great importance to reach as many private collection owners as possible via associations.

If you are interested to participate in spreading the word about this questionnaire on private collections, you do not have to take any action right now. You will receive the questionnaire with an accompanying mail to send out to your members by the end of July. These will be offered as much as possible in the national language to compel as many people as possible to participate in the questionnaire. Otherwise, you can reply to this message and indicate you wish not to participate.

Thank you in advance.

Kind regards, ...

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

Appendix 4. Mail accompanying the survey

Dear association,

As announced earlier, please find attached an email with the link to the questionnaire on the volume, scope and level of digitisation of private collections within Europe. We would like to ask you to distribute this among the association members. This questionnaire is part of the ICEDIG project (Innovation and consolidation for large scale digitisation of natural heritage), a design study for the DiSSCo (Distributed system of scientific collections) research infrastructure (icedig.eu and dissco.eu).

Thank you in advance.

Kind regards, ...

Dear private collection owner,

When it comes to privately owned natural history collections, these are often 'hidden' and may harbor highly valuable specimens for a specific taxon or (local) geographic range. On a European scale, these private collections combined are expected to form a significant part of the natural history collections as a whole. Therefore, it is important to include private collections in Europe's digitisation efforts of natural history collections, which is currently being explored within the ICEDIG project, a design study for the DiSSCo research infrastructure (icedig.eu and dissco.eu).

With this questionnaire, we hope to get an idea of the volume, scope and level of digitisation of private collections within Europe. In addition, we would like to know what kind of resources private collection owners need for digitisation and dissemination of their private collections. The outcomes of the questionnaire will be used to write guidelines for the digitisation of private collections, to perform a couple of pilot digitisation pilots of private collections and to present information about private collections to policy and decision makers. If we can show that there are many interesting private collections in Europe, we can make a strong case to focus on these natural history collections as well. It is therefore of great importance that this questionnaire is completed by as many private collection owners as possible.

The questionnaire consists of 20 questions and requires less than 10 minutes to complete. We ask you to fill out the questionnaire before the 17th of September.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

Please use the following link to complete the questionnaire: https://www.surveymonkey.com/r/77MQCLR

Thank you in advance.

Kind regards, ...

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

Appendix 5. Raw data of survey responses

See separate Excel document for the raw, anonymised responses to the survey questions.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

Appendix 6. The meaning of digitisation

For communication zooming in on digitisation it is of utmost importance that there is no misunderstanding between the sender and receiver of a message regarding the term "digitisation". Depending on the situation "digitisation" can have several meanings and describe various situations. Below six categories are listed all describing aspects of digitisation that may differ between collections.

1. the **overall detail level** of digitisation of a collection:
    a. lowest level: only general collection information is available (MIDS-0)
    b. medium level: storage or species information is available (MIDS-1)
    c. highest level: specimen data are available (MIDS-2)
        i) The digitisation levels for private collections have been related to the proposed digitisation levels within MS35, ICEDIG Task 6.1.

2. the **percentage of specimen digitisation** of a private collection:
    a. not yet started (0% digitised)
    b. partly digitised (1-99% digitised)
        i. with the use of unique identifiers (registration codes)
        ii. without the use of unique identifiers (registration codes)
    c. completely digitised (100% digitised)
        i. with the use of unique identifiers (registration codes)
        ii. without the use of unique identifiers (registration codes)

3. the **execution of digitisation** takes place at:
    a. the collection address by the owner
    b. the collection address by a volunteer
    c. an institute by a volunteer

4. **activities included** in the digitisation process:
    a. registration of data only
    b. registration of data including images
    c. registration, validation and enhancement (e.g. georeferencing) of data

5. **online (registration and) publication** via:
    a. a webportal that also publishes observation data (e.g. LAJI.FI)
    b. a webportal designed for private collections (e.g. PlutoF)
    c. a webportal of a regional or national institute (e.g. bioportal.naturalis.nl)

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

6. **keeping online data up to date** is done:
    a. directly online in the webportal (e.g. PlutoF)
    b. indirectly via an update from an underlying dataset

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

Appendix 7. Reasons to digitise a private collection

As a private collection owner, you might wonder: 'what is in it for me if my collection is digitised? Why should one start digitising his or her private collection in general?

- **Sharing data, information, knowledge.** One could more or less reverse the question - from: why should I digitise my collection - to: why do you collect in the first place? What is it you want to achieve with your collection? Is it something you want to keep for yourself or is it something you like to share with the rest of the world? If you would like to share (parts) of your collection or knowledge gained from it, the best start would be to share basic information about your collection. This requires some form of digitisation.
- **Collection management.** Every private collection owner knows his collection inside out. Any object/specimen or storage unit has been handled many times and its place in the collection is known. Still, having digitised a collection comes in handy for a collection owner in case a survey of the collection is required or when being asked detailed questions by colleagues. Likewise, if storage units and species have been digitised, the collection becomes more accessible for visitors as they can find their way without the help of the collection owner.
- **Identification**. In case a collection is digitised, it becomes easier by sharing data and/or images to get answers to questions about the identity of specimens. If identifications for parts of the collection are tentative or only to a higher taxonomic level, it can be advantageous to have data and images online and attract attention from or actively approach (other) specialists to assist in the identification process.
- **Data paper.** Once a collection has been digitised, assuming specimens carry validated identifications, the dataset can be published as a data paper. A data paper is a peer reviewed document describing a dataset, published in a peer reviewed journal. It takes effort to prepare, curate and describe data. Data papers provide recognition for this effort by means of a scholarly article.
- **Networking.** Is your main aim with your collection to get a complete overview of all taxa from a certain area, or do you use the collection in a scientific context publishing the results in scientific papers? In both cases, it helps to have a network of people sharing similar interests or collections. Besides publishing, you can draw attention to your collection and your interests and build a network by sharing the specimen data on internet. This of course requires digitisation.
- **Incorporation in museum collection**. Is "private ownership" the best way to describe ownership of objects in your collection or are objects in your collection actually "public property"? In other words, do you see yourself as a temporary guardian/collection manager of your collection which ultimately will become part of a publicly funded collection institute? Public collections very much welcome private

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483

collections, but as a rule lack the resources to directly digitise newly acquired collections. If, in your opinion, it is imperative that your collection is being used for research purposes once it is incorporated in a larger museum collection, it helps enormously if the collection is digitised before being transferred to a museum. Musea all over the world as a rule have a shortage of space but certainly also a shortage of manpower. Digitising collections is labour intensive which if already carried out can save a lot of labour time from the museum to which the collection is donated. **Musea** are more inclined to **accept collections** that have been digitised prior to being donated than collections that have not yet been digitised.

- Your collection may contain **information that is useful for non-taxonomic research**, for instance for ecological studies or for making IUCN Red List Assessments. Regarding the latter, for example, one must realize that to impose conservation measures for species or habitats, the conservation status of species are required. To assess this, distribution data are often very important. Digitising collections and making this data available can thus provide information required for Red List Assessments, supporting biological conservation.

Funded by the Horizon 2020 Framework of the European Union
H2020-INFRADEV-2016-2017
Grant Agreement No 777483