SYNTHESYS

Synthesis of systematic resources

Project:	Synthesis of systematic resources
Project acronym:	SYNTHESYS3
Grant Agreement number:	312253
Workpackage:	4: Moving from physical to digital collections
Deliverable number:	4.5: Digitisation on demand
Deliverable title:	Digitisation on demand: a report on feasibility of a digitisation on demand service for natural history collections
Deliverable author(s):	Vincent S. Smith (NHM) Elspeth Haston (RBGE) Robyn Drinkwater (RBGE) Laurence Livermore (NHM) Jennifer Pullar (NHM) Steffen Kiel (NRM) Johannes Lundberg (NRM) Stefan Daume (NRM) Irene Bisang (NRM) Margaret Gold (NHM)
Last edited date:	22nd August 2017

Digitisation on demand: a report on feasibility of a digitisation on demand service for natural history collections

Authors:
Vincent S. Smith (NHM)
Elspeth Haston (RBGE)
Robyn Drinkwater (RBGE)
Laurence Livermore (NHM)
Jennifer Pullar (NHM)
Steffen Kiel (NRM)
Johannes Lundberg (NRM)
Stefan Daume (NRM)
Irene Bisang (NRM)
Margaret Gold (NHM)

SYNTHESYS
Synthesis of systematic resources

SEVENTH FRAMEWORK PROGRAMME

Table of contents

1. An introduction to the scope and structure of this report

Collections-based research requires access to material held across many museums, necessitating costly visits to multiple institutions across Europe. Accurate inventories of these specimens, coupled with information on how to access this material, has the potential to greatly increase the efficiency of this research, especially if these data could be brought together in a single portal that serves as a gateway to the major collections. This "union catalogue" could include integrated information on our collective museum holdings, data on their research-readiness, up-to-date contact information on how to gain physical access and a summary of the major analytical facilities present at each institution. Such a Portal might evolve to become a common interface to provide digital surrogates of specimens, allowing users to download datasets, images and 3D models from across all our institutions in a single step. Critically, it could also support requests for digitisation on demand, such that users can request digitisation of selected objects and taxa, and obtain accurate quotes for this work, by incorporating cost models that account for differences in institutional digitisation costs for each category of object (e.g. microscope slides, herbarium sheets, fossils, pinned insects).

The purpose of this report is to explore the themes, issues and barriers associated with the development of an integrated Digitisation on Demand (DoD) service across European Natural History Collections. We include a series of surveys and feasibility studies used to understand the rates of maturity reached by different European museums and herbaria with their digitisation activities and identify the barriers to adopting a DoD approach across SYNTHESYS participants. We also explore the technical DoD infrastructure needed to support a DoD service; conduct market validation to understand how willing institutions would be to adopter DOD services; look at crowdfunding as a business model to support the cost of digitisation on demand; and explore the integration of crowdsourcing data as part of a DoD service. This research is aimed at helping to cost DoD activities, including provision for a pay-as-you-go service to help prioritise digitisation activities.

This work builds on other SYNTHESYS3 outputs including best practice guides to support collections digitisation, and is separated in to three major sections:

1.1. Data usage: new users and new uses of collections

The first chapter (section 2) of this report explores the feasibility of establishing a digitisation on demand service, including results of studies looking at the demand for digitisation, the current status of digitisation capacity within our institutions and experience building and running a service within SYNTHESYS partner institutions. Natural history institutions have evolved from internally focused organisations centred around taxonomy and systematics, to more outward facing activities that combine our knowledge of natural ecosystems with other scientific sectors and society. By strengthening and developing these connections we have the potential to have greater societal impact and increase our funding opportunities. Digitised collections are central to meeting these needs. Nevertheless, working with "adjacent" user communities in addition to developing new user communities requires museums to change how they operate, adopting a "customer first" model that necessitate changes in internal priorities and staffing. Examples of how this may impact collections and their digital activities include: 1) digitisation on demand services, enabling funders to have priority access to digitised specimens; 2) digital services supporting businesses, such as information related to environmental impact assessments or automated image recognition using our collections of reference images; 3) development of a large reference database for important species, such as

invasives, crop pests, or those of medical and veterinary significance; 4) support for the long term storage of environmental and voucher samples; 5) greater investment in applied systematics, such as work on crop pests, forensics, bird strike and border protection.

This chapter examines the readiness of European institutions to support DoD as well as some of the issues involved in enhancing the capacity of institutions to meet the evolving needs of current and new user communities.

1.2. Building a DoD network

The second major section (chapter 3) explores the issues associated with building a network for DoD across Europe. In particular, we examine the barriers to adopting a DoD approach across participants, look at the market for DoD data, and the technical infrastructure necessary to support Digitisation on Demand activities. The section closes by developing a high level plan for deployment of a European networked Digitisation on Demand (DoD) service.

1.3. Collections prioritisation and business models

Section three (chapter 4) explores possible business models to support DoD and provided an indepth look at the potential of crowdfunding to prioritise collection digitisation and financially support / supplement DoD costs. In addition to an introduction to crowdfunding, the chapter covers crowdfunding models and technical platforms, explores examples of crowdfunded Museum projects, looks at the scientific literatature associated with crowdfunding business models, and surveys current SYNTHESYS partners on their crowdfunding activities. The report concludes with a series of recommendations for incorportaing crowdfunding as part of Digitisation on Demand services.

1.4. Case Study

The last section (chapter 5) of this report is a case study on 'Digitisation on Demand' (SYNTHESYS WP4, Task 4.1) at the Swedish Museum of Natural History (NRM). This aimed to explore workflows for rapid digitisation and crowdsourced label and text transcriptions that can contribute to the increasing demand for digital collections. The case study targets were an invertebrate fossil collection representative in complexity for many such collections, and a historical, handwritten herbarium, requiring a specialized community for transcription. For the fossil collection a workflow was developed that included sorting, numbering, photographing, and digital image management that remained at a cost of below 1 Euro per specimen. During the first four months of the project, c. 10,000 specimens were sorted and c. 7,000 photographed, resulting in >12,000 digital images. Of those, 1,810 images were included in a crowdsourcing project for label transcription, which was completed after 50 days, involving 125 volunteer transcribers. A Citizen Science Day was held at NRM aimed at promoting our project, citizen science, and Natural History collections in general. The project and its results were presented at two meetings: a Swedish national collection management staff meeting in Stockholm and the SYNTHESYS final meeting in London. The herbarium subproject proved more complex than anticipated, and a new concept focusing on marking and transcribing multilingual species names was developed. For this subproject, the Citizen Science Day enabled us to establish contacts with the specialized community needed to help with transcription. The project raised awareness and appreciation of the possibilities of rapid digitisation, crowdsourcing, and citizen science at NRM. In the longer term, it facilitates the means for NRM to respond to 'Digitisation on Demand' requests, and to structure and enhance our digitisation efforts.

2. Feasibility study on a "Digitise on Demand" (DoD) service

Elspeth Haston & Robyn Drinkwater (Royal Botanic Garden Edinburgh)

2.1. Introduction

The need for increased access to collection objects for an ever expanding range of scientific and cultural research is becoming more and more reliant on digital access. Whilst institutes work to digitise their collections based on the priorities, constraints and opportunities of the institute, there is another demand for digitisation coming from the priorities of external users of the collection. Where this is taking place, it is currently fragmented and inconsistent, and therefore can be difficult and time-consuming for researchers to contact each institute separately to discover relevant holdings and request material to be made accessible through digitisation.

This report presents a feasibility study on the development of a Digitisation on Demand (DoD) service for European Natural History Collections. The report includes work carried out within the Network Activity 3 (NA3) on the impact of digitisation on loans and visits (Enghoff, 2017) and the digitisation capacity of institutes (Phillips *et al*., 2014), reviewing their results in terms of DoD. Potential barriers to adopting a DoD approach have been identified. An infrastructure built on a network approach is recommended and a schematic outline of this infrastructure is included. A staged deployment plan is necessary given the ambitious scale of the service and this is outlined in the report.

2.2. Demand and loans access to specimens pre-digitisation

Specimens are traditionally loaned between Natural History Collections to allow researchers to access material from around the globe for their work. Each institute will have a loans policy that will include some restrictions which will limit the loans that can be sent out. (i.e. no type specimens on loan, only ⅓ of all specimens for one taxon, etc). There is an inherent risk of damage or loss in sending specimens between institutes as well as difficulties with clearing customs across certain borders.

A survey and report carried out as part of the SYNTHESYS3 NA3 Work Package (Enghoff, 2017) was reviewed to provide information on the existing demand for specimens, the impact of digitisation on loans, and in the context of market validation for a DoD service.

Large numbers of loans are transacted by Natural History Collections, with hundreds of loans, comprising thousands of specimens, being processed by many institutes each year. In recent years, however, loan numbers have been seen to be decreasing. This may be due to fewer taxonomists and lower levels of funding but it may also be due to the digitisation of collections. The survey looked at loan data contributed by various institutes. The data were mostly unpublished and derived from the loans officers and curators and so therefore may not be directly comparable, but showed that the number of loans being requested were decreasing. However, in one institute, where data were available, the size of the loans has remained stable.

Over the six main collection types (dry zoological, wet zoological, fossil, herbaria, mineralogical/petrological, frozen) included in this study the availability of images online decreased

the number of loans (Figs. 1a & 1b). For one institute, the Conservatoire et Jardin botaniques, Genève, data were received in the survey showing the number of requests for digital images. These showed a large overall increase over time from 2000-2015, with a sharp increase from 2011 (Fig. 2). Whilst a decrease in loans may be seen to indicate reduced demand, later sections of this report will look at the change in demand and relevant metrics. Whilst digitised specimens can fulfill some users needs, there will still be a requirement for access to the physical specimens for some taxa and research questions.

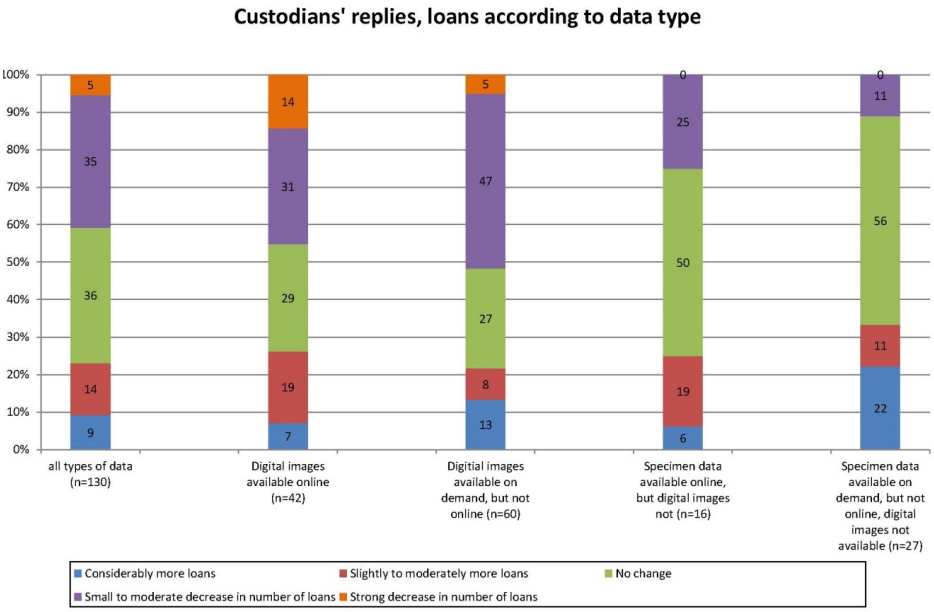


Figure 1a. Custodians' replies, loans according to data type (from Enghoff, 2017).

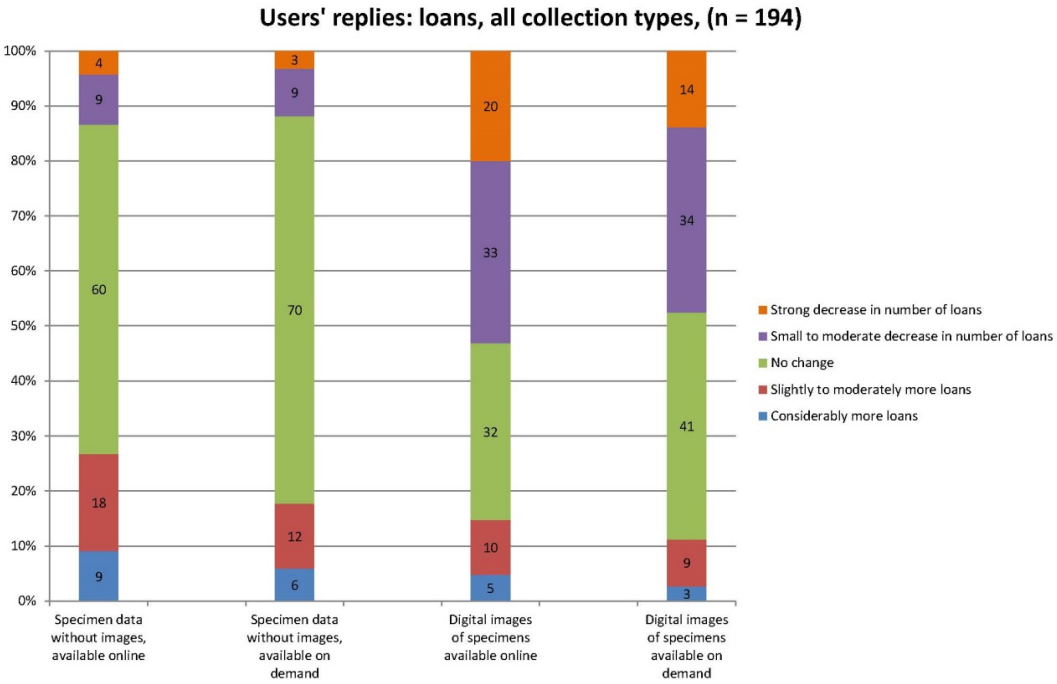


Figure 1b. Users' replies, loans according to data type (from Enghoff, 2017).

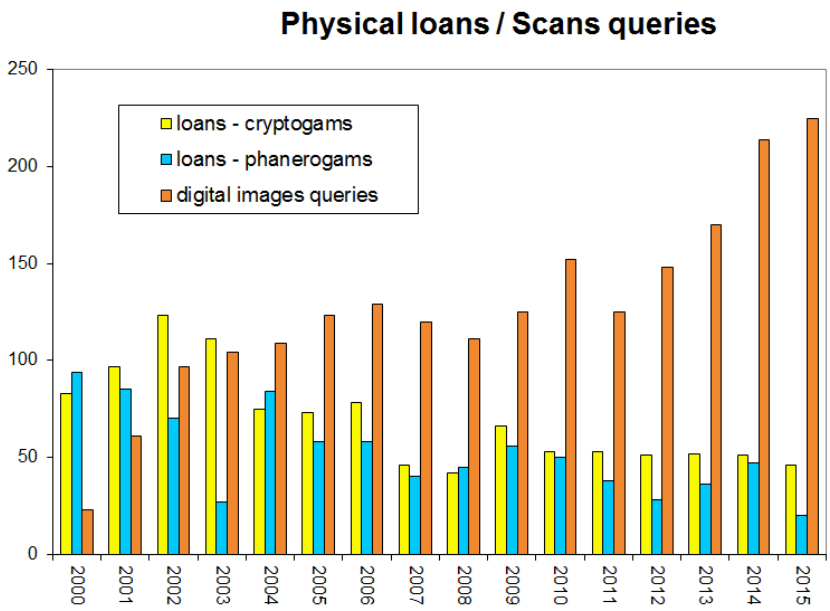


Figure 2. Loan statistics from the Herbarium, Conservatoire et Jardin botaniques, Geneve. The graph was kindly provided by Laurent Gautier (CJB) (from Enghoff, 2017).

2.3. Current status of digitisation capacity in EU Institutes

Digitisation, in the sense of imaging and databasing a specimen, has been increasing in the natural history sector since the early 2000s, with herbaria being some of the first to adopt it widely. Possibly due to the easy format of most specimens to image, and the success of the Andrew W. Mellon Foundation Global Plants programme, herbaria are one of the first disciplines to have large numbers of specimens imaged.

Three surveys have recently been carried out which provide some data on the current status of digitisation capacity within Natural History Collections.

2.3.1. Survey 1: Review of Current Digitisation Workflows and Equipment (Phillips et al., 2014)

A review of current digitisation workflows and equipment was carried out as part of SYNTHESYS3 NA3 (Deliverable 3.3) (Phillips *et al.*, 2014). The report was based on a survey with 14 responses.

This report, in 2014, found that all collection types have made some progress towards digitising their collections, with between <1% and 20% for entomological collections, <1% for geological collections, and an average of <1-10% digitised for paleontological and botanical collections, although two institutes have digitised significantly larger percentages of their botanical collections, Naturalis (up to 90%) and Paris (6 million). The percentage digitised for zoological collections varies widely, depending on the type of collections held.

This survey included questions on workflows and equipment. They found that approximately two thirds of institutes carried out some level of curation, specimen conservation or collections management steps prior to digitisation as part of the digitisation workflow. Equipment used for

digitisation of collections was found to be diverse, but was not viewed as a limiting factor for digitisation. Funding and suitable digitisation infrastructure were seen to be a greater impediment. Most institutes were still digitising their collections in-house.

2.3.2. Survey 2: Accelerating the discovery of biocollections data (Krishtalka et al., 2015)

A survey carried out by in 2015 by the GBIF Task Force on Accelerating the Discovery of Biocollections Data (Krishtalka *et al*., 2015) elicited 617 responses from institutes around the world. Of these, only 1% are not currently digitising their collections and have no plans to digitise. A range of barriers to digitisation were identified which were found to fall within six categories: funding, effort, size of task, not an institutional priority, perceived data quality issues, lack of skills/training. This survey focussed more on the data capture element of digitisation rather than imaging. However, they found that more than 80% of 513 respondents indicated that their institute intends to digitise their entire collection(s), and more than 30% responded that they plan to prioritise digitisation in response to research needs/requests.

2.3.3. Survey 3: Survey of existing digitisation capabilities in Natural History Collections (RBGE, 2017)

The survey carried out by RBGE in 2017 for this report, asked European institutes about the current status of digitisation. The questionnaire for the survey was built using SurveyMonkey and distributed in July 2017 to the SYNTHESYS Consortium as well as to institutes represented in the CETAF Digitisation Working Group. There was a total of 31 respondents to the survey from institutes across Europe (Table 1). Responses from several individuals from a single institute were combined. This left a total of 21 individual institutes which responded. Respondents were from a variety of backgrounds within their institutes.

Name of Institute	Country
Botanic Garden and Botanical Museum Berlin	Germany
Botanic Garden Meise	Belgium
Conservatoire et Jardin botaniques de la Ville de Genève	Switzerland
Finnish Museum of Natural History	Finland
Hellenic Centre for Marine Research	Crete
Hungarian Natural History Museum	Hungary
Museo Nacional de Ciencias Naturales	Spain
Museum für Naturkunde Berlin	Germany
Muséum National d'histoire Naturelle à Paris	France
National Museum Prague	Czech Republic
Natural History Museum London	United Kingdom
Natural History Museum of Denmark	Denmark
Naturhistorisches Museum Wien	Austria
Naturhistoriska Riksmuseet	Sweden

Royal Belgian Institute of Natural Sciences	Belgium
Royal Botanic Garden Edinburgh	United Kingdom
Royal Botanic Garden Madrid	Spain
Royal Botanic Gardens Kew	United Kingdom
Royal Museum for Central Africa	Belgium
Senckenberg Gesellschaft für Naturforschung	Germany
Swedish Museum of Natural History	Sweden

Table 1. Institutes which responded to the 2017 survey of the existing digitisation capabilities in Natural History Collections Across Europe.

From the results of the survey, we found that 20 of the 21 institutes have partial to extensive digitisation capability. These institutes cover a range of collection types and show that there is capacity for digitising all collection types across these institutes (Fig. 3). Two institutes noted that workflows and processes are developed to different levels depending on the collection or project. These differences can be seen in the following section on percentage of collections digitised. Two responses suggest that the digitisation programme is well developed and that they are well placed to complete the digitisation of certain collections.
In addition to the collection types listed, institutes also reported the capacity to digitise Micro CT scanned collections, cryptograms, prehistory specimens and map collections. One reported that they can digitise microscope slides individually, but not in bulk. Another reported having a digital microscope to digitise lichens and slime moulds.

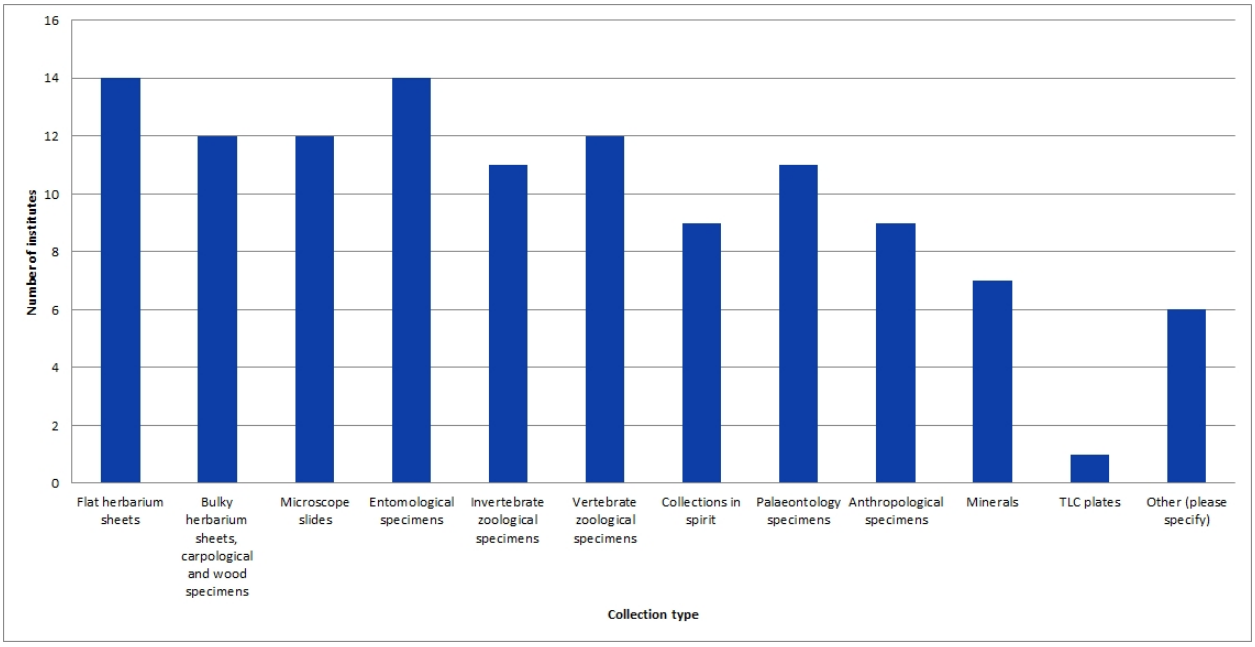


Figure 3. Digitisation capacity of institutes holding European Natural History collections resulting from the survey carried out in July 2017.

The survey asked for an estimate of the percentage of each type of collection that had been digitised for each institute (Fig. 4). In order to interpret the results of this question more clearly we defined the levels of digitisation as follows:

Percentage digitised	***Assumption made regarding status of digitisation capacity based on the percentage digitised***
0%	Digitisation hasn't started, there is no equipment and protocols have not been developed.
<1%	Very early stages, pilot studies undertaken to start exploring appropriate equipment and developing protocols.
1-10%	Protocols, equipment and workflows have been established.
11-50%	Significant progress has been made.
51-100%	Equipment, protocols and IT infrastructure are in place to allow for the digitisation of the entire collection.

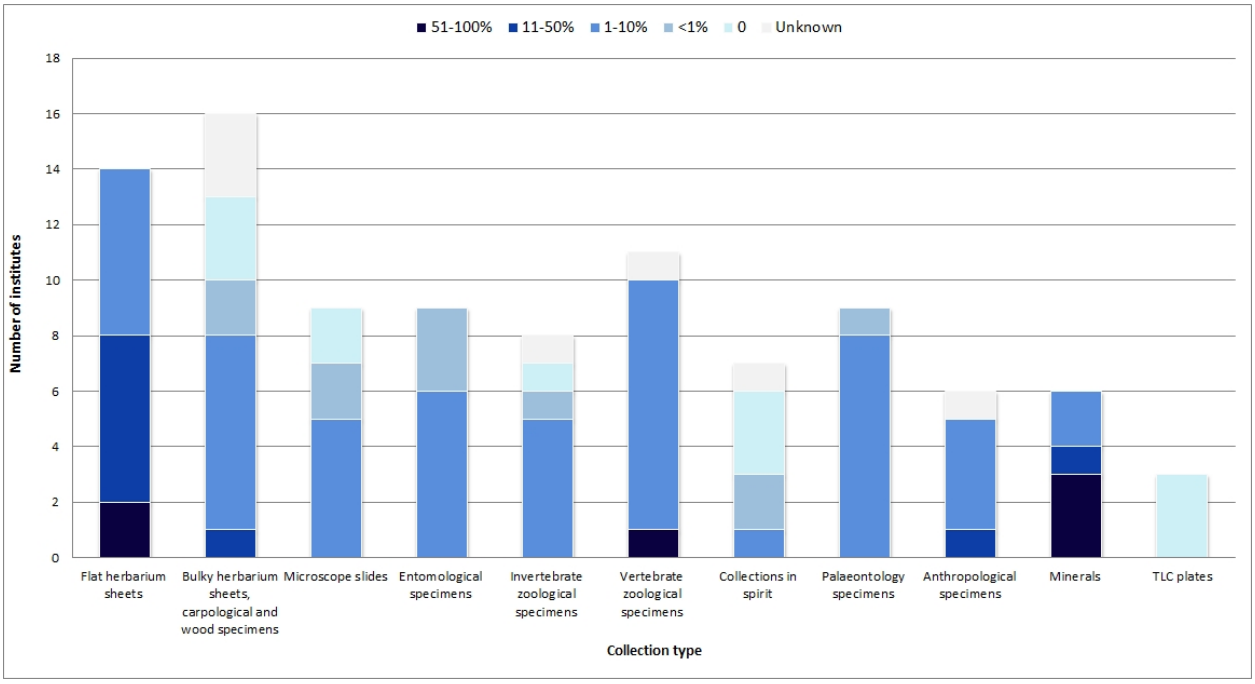


Figure 4. Percentage of collections digitised in institutes holding Natural History collections resulting from the survey carried out in July 2017.

From the responses relating to the percentage of collections digitised it is possible to see some grouping by collection types. The digitisation of herbarium and mineral collections appears to be most advanced. All respondents had digitised more than 1% of their collections, with at least half having digitised over 10%. Whilst more than half of respondents had digitised more than half of their mineral collections, it is possible that these respondents were considering digitisation to be solely data entry and so this result needs to be verified. Based on these responses it could be said that the equipment, protocols and workflows are in place for most institutes for the digitisation of these collections. Some institutes also have the ICT infrastructure in place to handle the large numbers of images and data. The digitisation of bulky herbarium sheets, carpological and wood specimens does

not appear to have progressed at the same rate, although some institutes were not able to separate this figure from the percentage of flat sheets digitised. Institutes still appear to be working on developing equipment, protocols and workflows for these collections.

The digitisation of the microscope slides, entomological, invertebrate, vertebrate, palaeontological and anthropological collections have all made some progress, with more than half of institutes who responded for these collections having digitised more than 1%. Whilst one institute had digitised more than 10% of their anthropological collections and one more than 50% of their vertebrate collections, two institutes had not started digitising their microscope slide collections and one institute had not started digitising their invertebrate specimens, showing that institutes appear to have been working on digitisation capability by collection type and some collection types appear to be more complex than others to digitise.

An important point about the entomological collections is the method of digitising whole drawers rather than individual specimens, followed by the segmentation of images using Inselect (Hudson *et al*. 2015). Whilst this was not specifically included in the survey, one respondent noted that they have made more progress with drawers than individual specimens, with 30% of these imaged.

There are two collection types which show markedly less progress: collections in spirit and TLC plates. Whilst the lack of digitisation of TLC plates may be due to institutes not considering them as a priority, the small percentage of collections in spirit that have been digitised is more likely to be caused by the lack of equipment and protocols and the inherent difficulty of the material. Interestingly, following discussion with researchers at RBGE, TLC plates are now being prioritised given their fragility, vulnerability, importance and particularly due to the rapid alteration of the colour over time. They are also easily digitised using the same equipment, protocols and workflows as the flat herbarium sheets.

2.4. Existing experience with a Digitisation on Demand service

The provision of images and data as the result of direct requests from users of collections (Digitisation on Demand) is becoming more common as digitisation becomes more firmly established within natural history collections. As users become more used to gaining access to specimens in a digital form there is an increasing acceptance of this as a way of working. In addition, as mentioned in earlier sections, changes to import regulations, limited funding and other effects are changing how researchers are accessing collections. Digitised specimens can allow for label information to be easily accessed and the range and quality of material held by collections can be assessed by the user prior to making a decision to visit or loan material. For some taxa and research purposes the digitised specimens are sufficient to provide taxonomic and other research data.

The main section of the survey carried out by RBGE for this task in the SYNTHESYS3 JRA, aimed to discover the current status and experience of DoD in Natural History Collections, including what metrics are used, and how images and data are made available following requests, in terms of format and access.

The majority of institutes (17 of 21) accept digitisation requests, with the frequency and systems used varying both between and within institutes. Two institutes reported that they have a 'formal' system to handle these using email, with a further one saying the management for this is not yet in place. Four institutes reported that they carry out digitisation on an occasional or ad hoc basis. In institutes with multiple collections DoD varied depending on the department or individual curators.

Some institutes provide DoD if the number of specimens requested is low, or aligned with the institute's research interests. One reported sending an agreement form to the user.
Further exploration of this question would be of great benefit to developing a DoD network, as there are different practices within the institutes surveyed which could impact on how the network could work and whether an institute would be able to contribute to it. Some of the areas that would be beneficial to investigate would be why some institutes currently limit the number of specimens in a request and how it could be managed within an institution
Institutes use a range of different services to make images and data available, with emails, transfer services and online collection catalogues being the most widely used (Fig. 5). In addition to the services included in the survey, institutes also listed hard drives, institute download servers and other online portals as methods for making images and data available. Several institutes reported multiple methods for sharing data, with the one being used for a particular request often dictated by the types of images and data to be transferred.

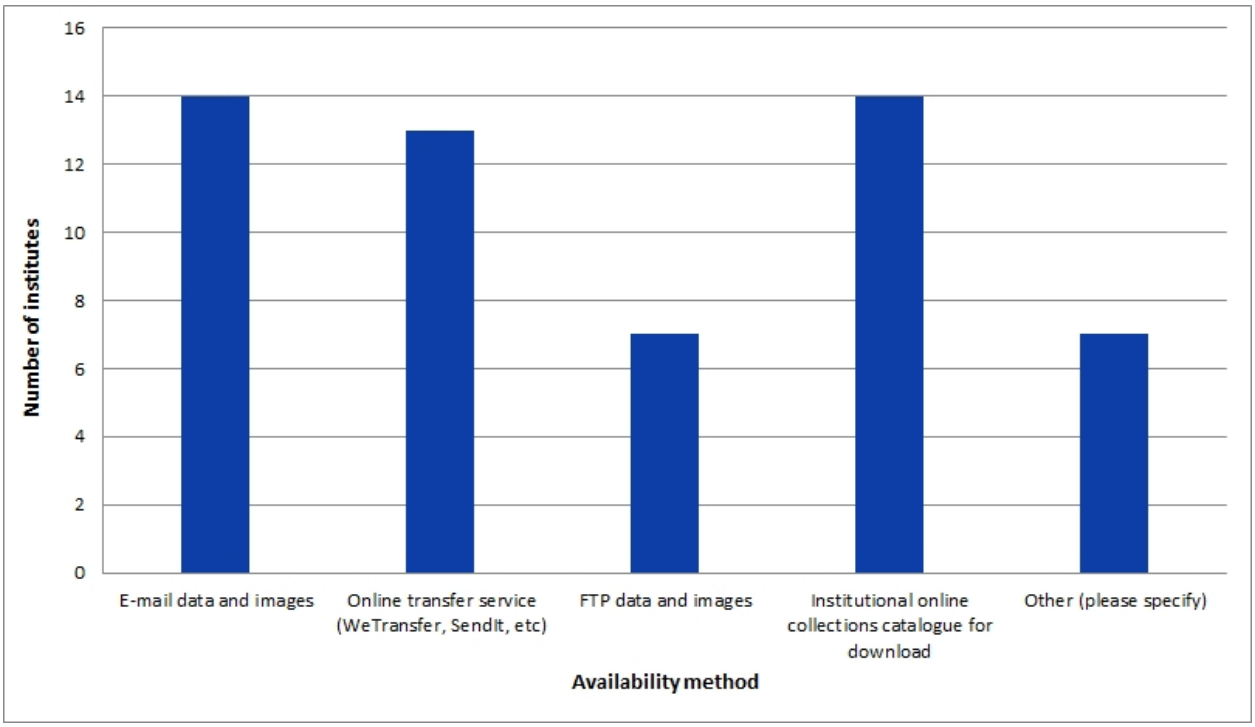


Figure 5. Methods used by institutes holding Natural History collections to make requested specimens available following digitisation, resulting from the survey carried out in July 2017.

The survey asked institutes about the image format and resolution that are made available to users requesting images (Fig. 6). High resolution images (both jpgs and tiffs) were the most common formats available. This is consistent with the reasoning behind the decision at the beginning of the Andrew W. Mellon Foundation Global Plants digitisation programme to digitise specimens at a standard resolution of 600dpi. This equates approximately with examining a specimens through a x10 hand lens, and was considered to be an acceptable standard for research purposes for most specimens. It therefore makes sense to make these high resolution images available to researchers.

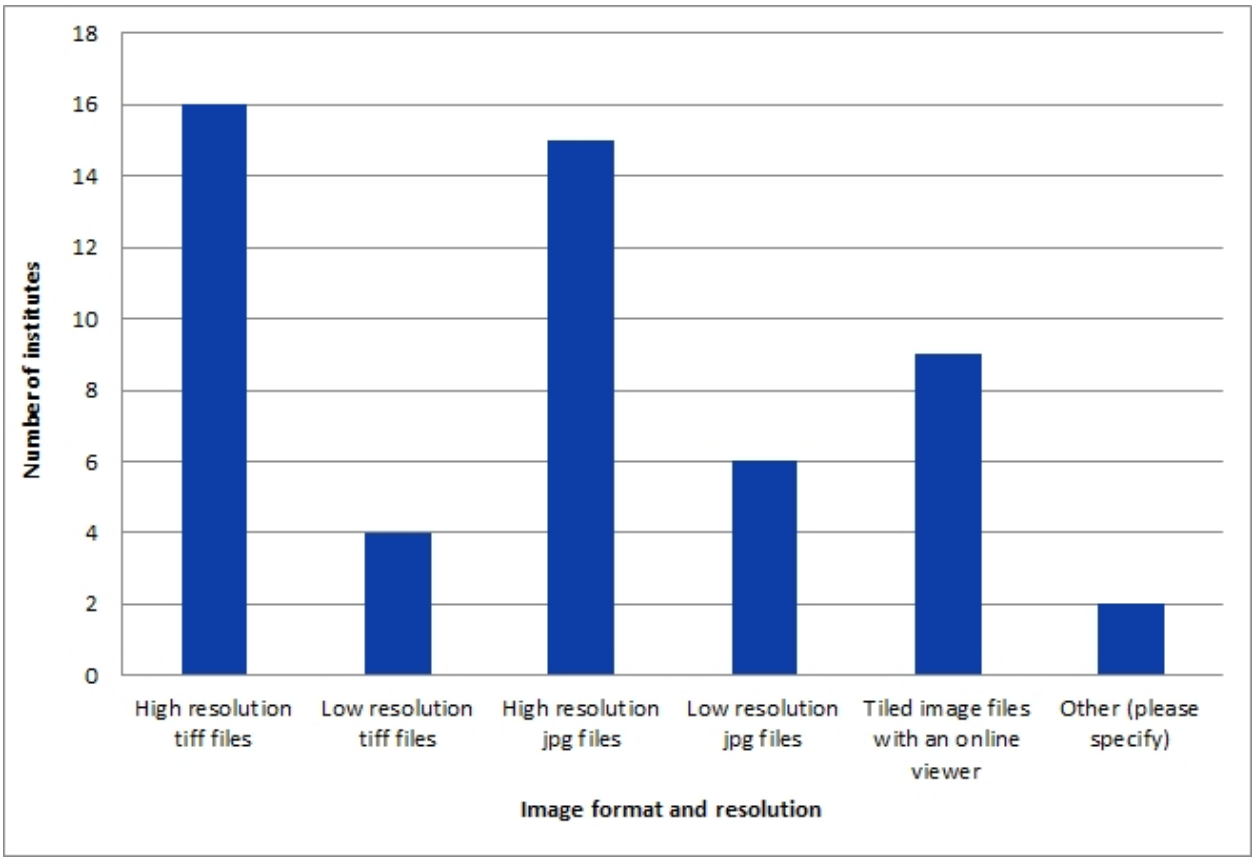


Figure 6. Image format and resolution made available by institutes holding Natural History collections resulting from the survey in July 2017.

Two institutes reported that the format and quality can depend on the purpose of the request and the individual's need. One institute normally sends the high resolution tiff files but will also send jpg files when requested. One institute will only send tiff files as a special request.

Nine institutes now make tiled images available with an online viewer which would enable the user to zoom into details without having to download or transfer the files. This survey did not ask for further information about the viewers used but this is an area that is expected to expand in the future. Work is now being undertaken to explore resources which would enable viewers to work between institutes (Crane, 2017). This would allow users to view specimens from different institutes together on a single screen without having to download or transfer the files onto their own computer or server.

An API (Application Programming Interface) is a set of subroutine definitions, protocols, and tools that can be used when writing software and building application. It can be used to allow developers to access data for use in the software or application they are writing. The ability to allow data to be accessed in a machine readable format has been considered to be an important part of the development of collections access. It is also inherently linked with the development and enhancement of the Stable URI system of identifiers which has now been adopted by the Consortium of European Taxonomic Facilities (CETAF) as a standard. The system is now being used in 13 CETAF institutes.

Most institutes either have no API, or the implementation of an API for their online collections catalogue was not known by the respondent (Fig. 7). One respondent reported that some of their

systems had APIs available, and another reported that they provide an API for a separate online portal.

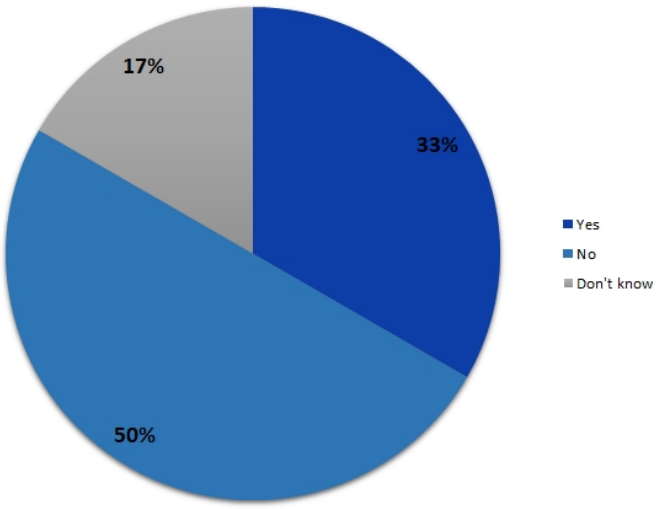


Figure 7. Institutes which currently have an Application Programming Interface (API) for their online collections catalogue, resulting from the survey in July 2017.

The range of metrics available varies between organisations and within the types of collections held (Fig. 8). The ability of an institute to gather some of these metrics may be dependent on how their digitisation workflow is set-up. It may be that with a more formalised system that some of these figures would be easier to provide.

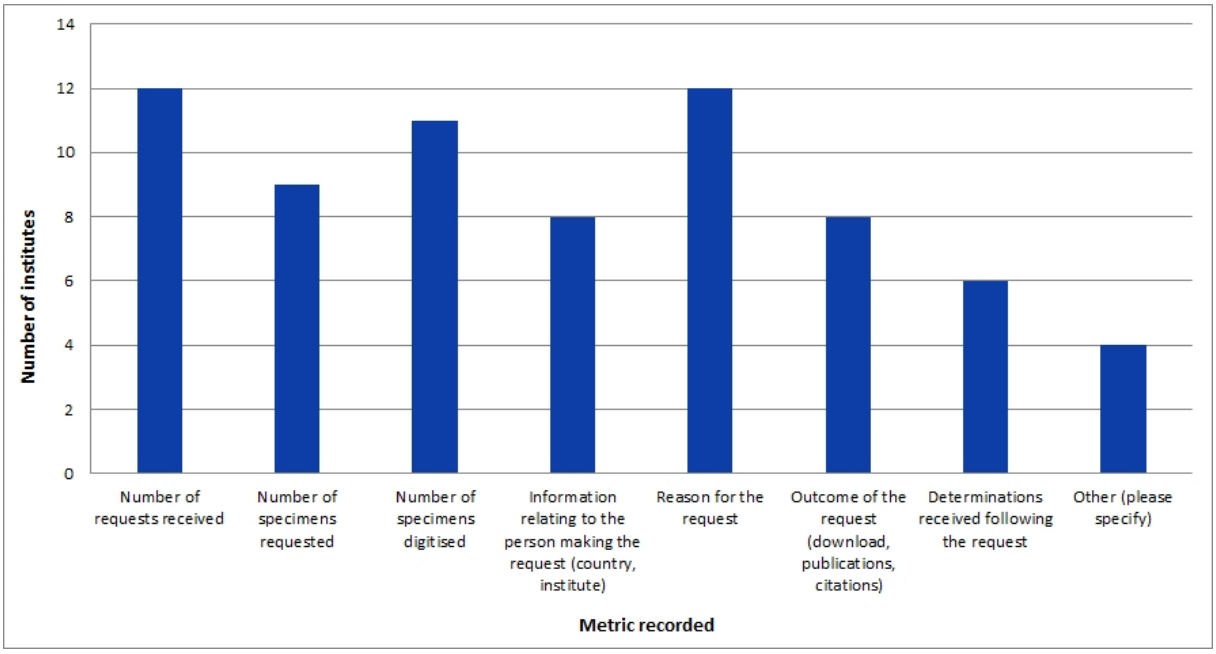


Figure 8. Metrics relating to digitisation requested recorded by institutes holding Natural History collections resulting from the survey in July 2017.

Just over half of the institutes which responded have carried out a costing of at least part of their digitisation processes (Fig. 9). However, where costing information was provided it is clear that figures from different institutes vary significantly in which costs are included and how the figures are

calculated. Based on the responses received, a range of factors have a major impact on the cost of digitisation and these are summarised below (Table 2).

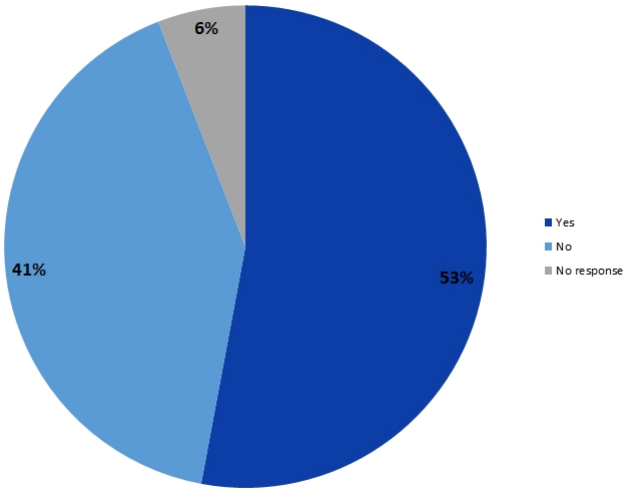


Figure 9. Percentage of institutes holding Natural History collections which have carried out costing of the digitisation processes, resulting from the survey in 2017.

Factor	Options
Digitisation equipment	dependent on material being digitised choice made by institute for other reasons
Who is doing the digitisation	institute staff or person requesting digitised material
Scope of digitisation	level of data capture done by staff whole drawer vs individual specimen
Post image processing	manual vs automatic cropping processing from raw file to tiff or other format file naming workflows
Number of images per specimen	overview image close up images of features, labels capsule contents
Level of pre-digitisation curation	cleaning and repairs nomenclature
Workflow being used	whole cabinets, collections vs pulling individual specimens level of automation in workflow
Location of digitisation	internal or external freezing and translocation of specimens
Inclusion of overheads	whether full economic costing is being used
Format and resolution of image	2D or 3D 300, 450, 600, 1200 etc ppi
Inclusion of digital preservation/curation costs	inclusion of digital storage

Table 2. Factors affecting the costs of digitisation in institutes holding Natural History collection.

2.5. Example of building a Digitisation on Demand system within an institute

For several years, the Royal Botanic Garden Edinburgh (RBGE) has been developing a Digitisation on Demand (DoD) service. The development of this service is described here.

In line with most institutes, RBGE has sent specimens on loan to researchers and has a well-developed loans policy and protocol. The loans data have been recorded electronically since 1994 when the *BG-BASE*™ herbarium collections database was installed. These data show a gradual decrease in the number of loans over time (Fig. 10). Given that these figures show a decrease in number of loans starting prior to the start of digitisation, there are clearly other factors impacting loans. One aspect of RBGE loans policy is to limit the number of specimens sent out for any one taxon in a single loan due to the risk of specimen damage or loss. This requires the curators to select the specimens for loan, and this is generally done without direct input from the researcher. One result is that the researcher may often be unaware of the full extent of the collections and may not always receive the most important specimens for their research. Researchers require specimens for many different kinds of research and different research questions will require different levels of information from the specimens. Access to digital versions of the specimens' data may sometimes be sufficient, whilst the addition of a digital image of the specimen may also decrease the number of physical specimens required by the researcher.

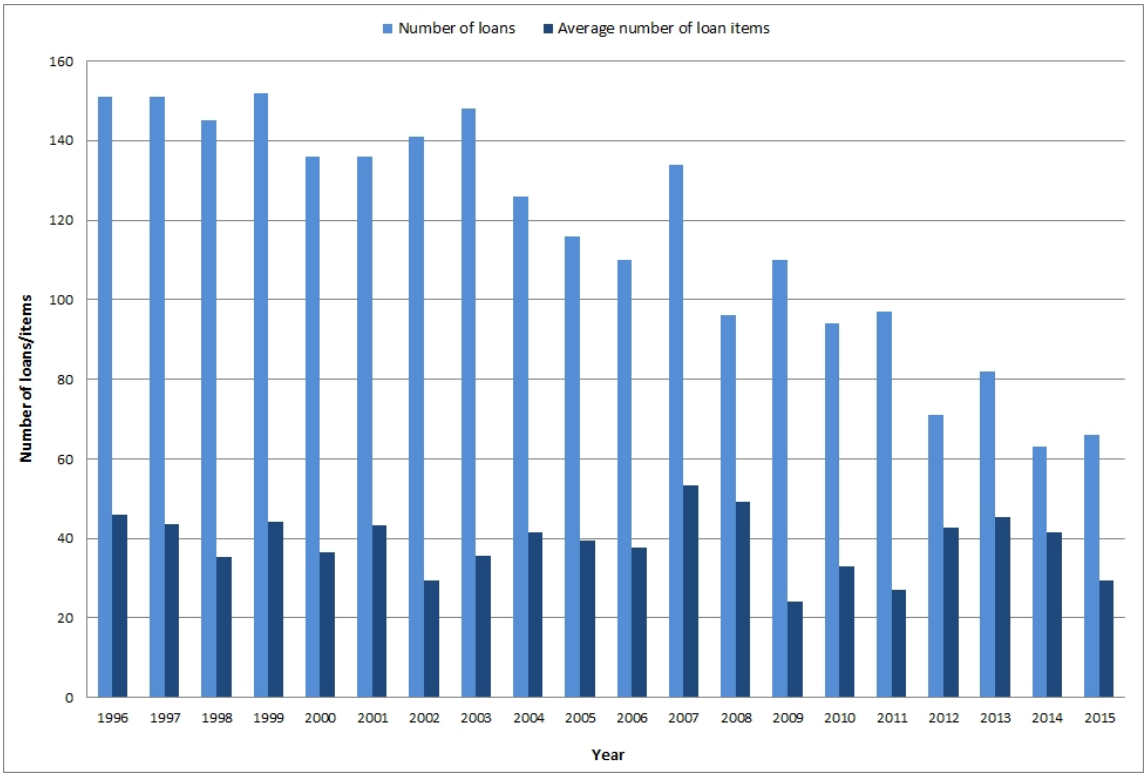


Figure 10. Number of loans and average loan size from data recorded at the Royal Botanic Garden Edinburgh (RBGE) from 1996-2015.

RBGE started to digitise (database and image) specimens in 2004 as part of the Mellon African Plants Initiative. When this programme expanded to become the Latin American Plants Initiative and finally the Global Plants Initiative, RBGE were able to digitise all the known types in the collections (c. 56,000). In addition, the Andrew W. Mellon Foundation funded the imaging of all the angiosperm specimens from Chile, Argentina and Paraguay, which had been previously databased and were filed under a single geographical region in the herbarium. Funding from the Scottish Government enabled the digitisation of a large proportion of specimens from SW Asia and the Middle East. The Flora of Turkey was produced at RBGE and the herbarium holds the most important collections for this area. Both of these major projects enabled RBGE to install digitisation equipment and develop a large scale digitisation methodology and capacity. More recently, RBGE appointed permanent digitisation staff to continue a large scale digitisation programme and it was decided to incorporate a DoD element to this programme.

The structure of the DoD service was influenced by both the current loans policy and the digitisation protocol as well as by current staff capacity. An important factor that had to be considered was the practice of not marking specimens which have been imaged at RBGE. This was a decision taken at an early stage and was based on the aim to speed up the digitisation process as much as possible and on the policy of maintaining to a minimum the amount of non-essential marks on the specimens. One result of this is that it is not clear whether a specimen has been imaged or not without checking electronically. Based on trials to speed up the digitisation process, we developed processes to rapidly capture minimal data working at a batch level, only capturing data that were shared by all specimens within a folder. This protocol enables one person to create minimal data records for up to 1,200 specimens per day.

In addition to increasing efficiency, another area of digitisation development at RBGE has been to bring all digitisation project workflows into a single workflow which is integrated with the curation of the Herbarium. The workflow is shown here (Fig. 11), and includes a large automated element.

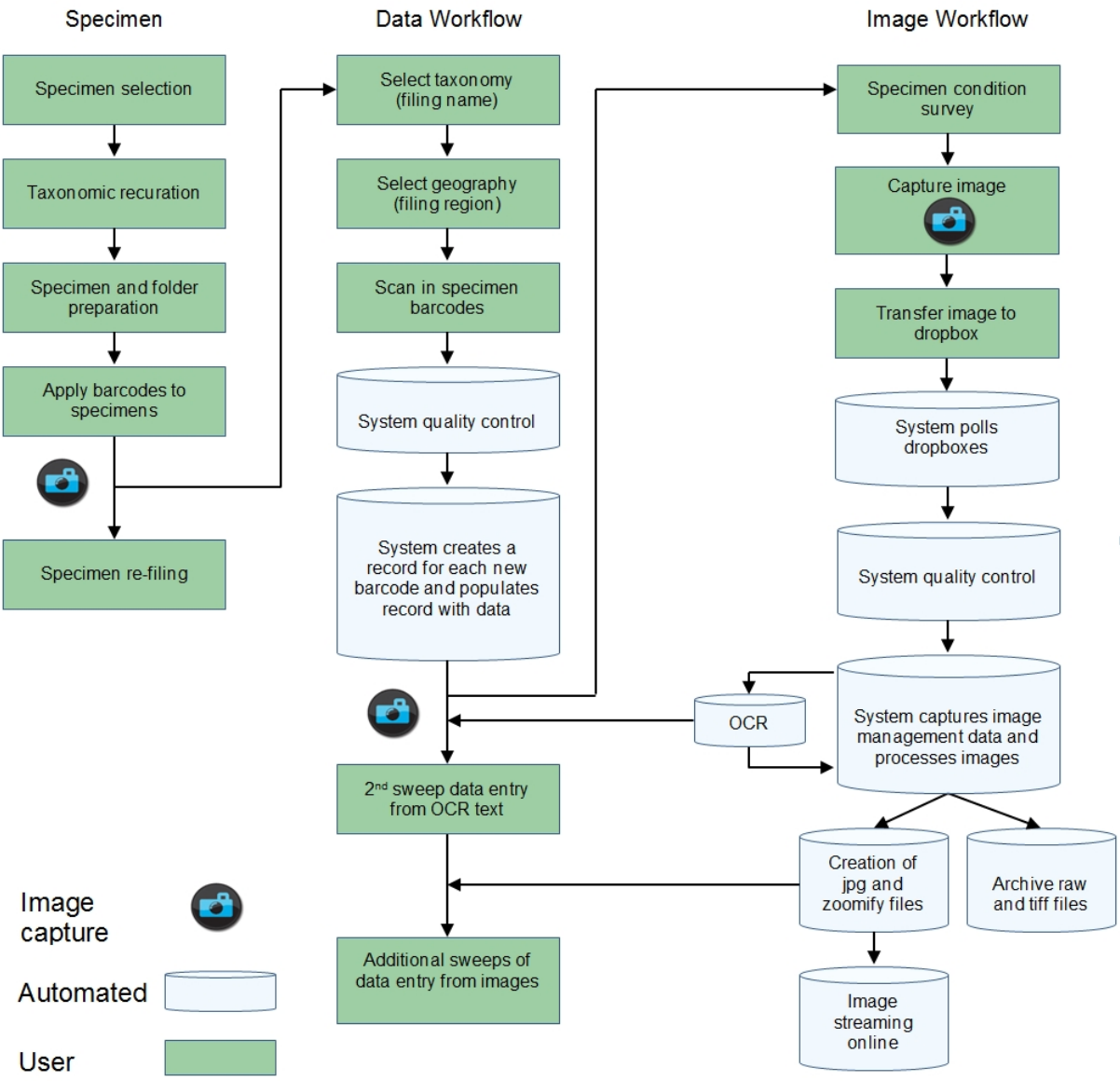


Figure 11. Diagrammatic overview of the digitisation workflow at the Royal Botanic Garden Edinburgh (RBGE) (Haston *et al*., 2012)

The DoD service is undertaken within this workflow, using the same efficient processes to enable us to digitise larger numbers (Fig. 12). We therefore use a system of minimal manageable units for requests. If a researcher requests only a small number of individual specimens within a taxon, we will check the cabinets and increase the number of specimens to a manageable unit of up to 1-2 cabinets (500-1,000) specimens. In some cases, we will digitise a complete family if it falls within this limit. If this unit is too large, we may limit the digitisation to a complete genus, or a particular filing region or, occasionally, we will limit it to one species where the number of specimens is large. The decision-making protocol is included here, but flexibility is important to allow for staff capacity and prioritisation.

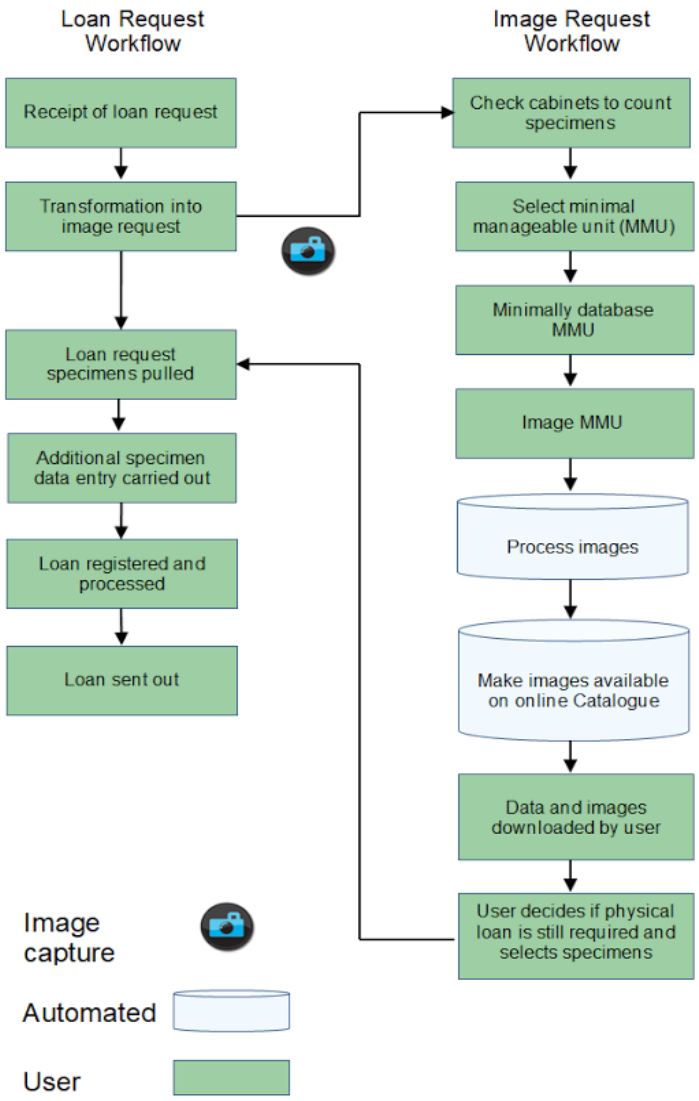


Figure 12. Diagrammatic overview of the loan and image request workflows a the Royal Botanic Garden Edinburgh (RBGE).

The benefits of this approach, where a larger number of specimens are generally digitised, is that the researcher is able to access more material and can then make an informed selection of specimens for a physical loan if it is still required. We are finding that for a number of researchers the digital versions are sufficient for their work and a loan is no longer required.

In order to manage the requests being received, a request management system is being developed. We are currently transitioning from a spreadsheet recording system to a MySQL database with a web-based user interface. This system records the person making the request as well as the taxa and geographical regions of the specimens which will be digitised for the request (Fig. 13).

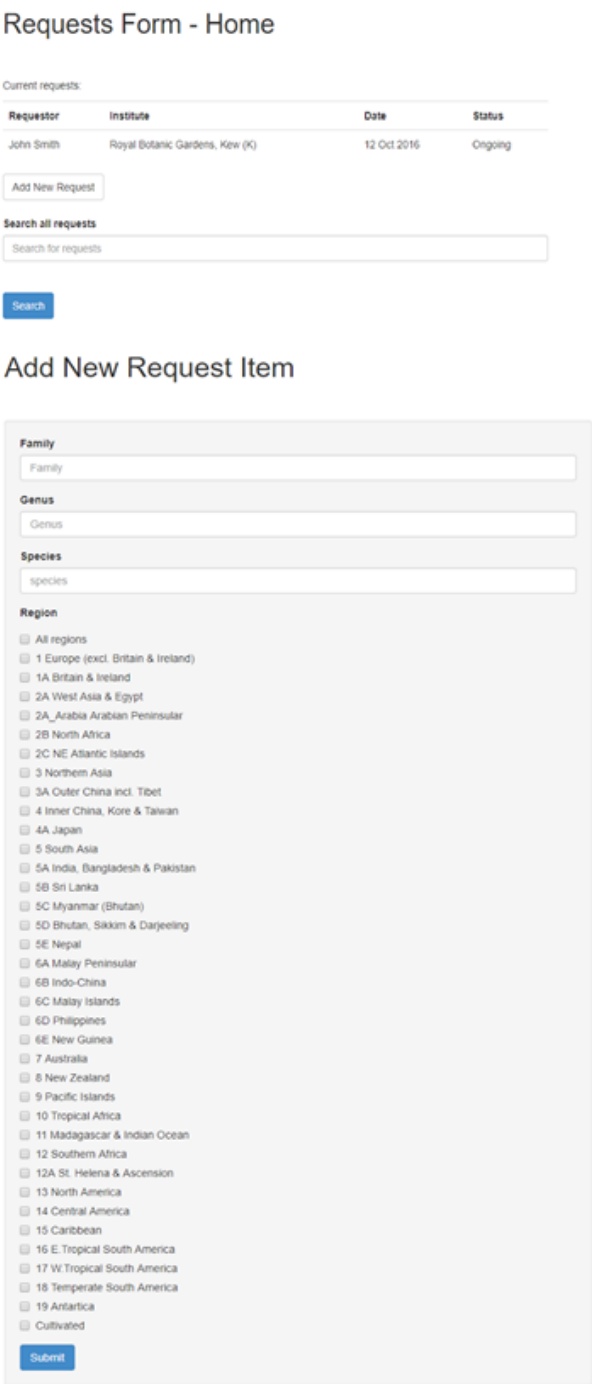


Figure 13. Screenshot of the web-based user interface of the request management system at the Royal Botanic Garden Edinburgh (RBGE) currently under development.

For many institutes which have partially digitised their collections, this is a particularly difficult transition phase. The complexity of managing a partially digitised collection has a major impact on several key curation tasks, including accessioning new specimens, laying away returned loans, and re-curating families. To help staff quickly know if a specimen needs to be imaged prior to laying it away, we have been developing web-based query tools. The first of these is a website in which the specimen barcode can be scanned in and the specimen record will be returned showing the image if digitised and the information required to lay the specimen away, including the family and family number, the genus number within the family, and the filing species name. However, whilst this tool

lets the user know if the specimen has been imaged it does not indicate whether it should be imaged, ie if all the other specimens in the folder have been imaged. A second tool is being developed which quickly lets the user know if the specimen should be imaged prior to being laid away. Both tools rely upon the collection staff having a web-based device (phone, tablet, computer) close at hand within the collection.

The DoD service at RBGE has not been openly advertised. We have been receiving an average of 8 requests per month, but there is variation in the number and size of requests throughout the year, possibly related to PhD programmes and researchers' teaching workloads (Fig. 13). The number is manageable at present with current staffing and we achieve most requests within one week. A significant increase in the number and size of requests would make it difficult to maintain the same level of service without additional staff. To date, a total number of 73,231 have been digitised through the DoD service at RBGE (Fig. 14). This is 18% of the total number of specimens digitised (408,718). The average number of specimens imaged for requests is 1,627 per month.

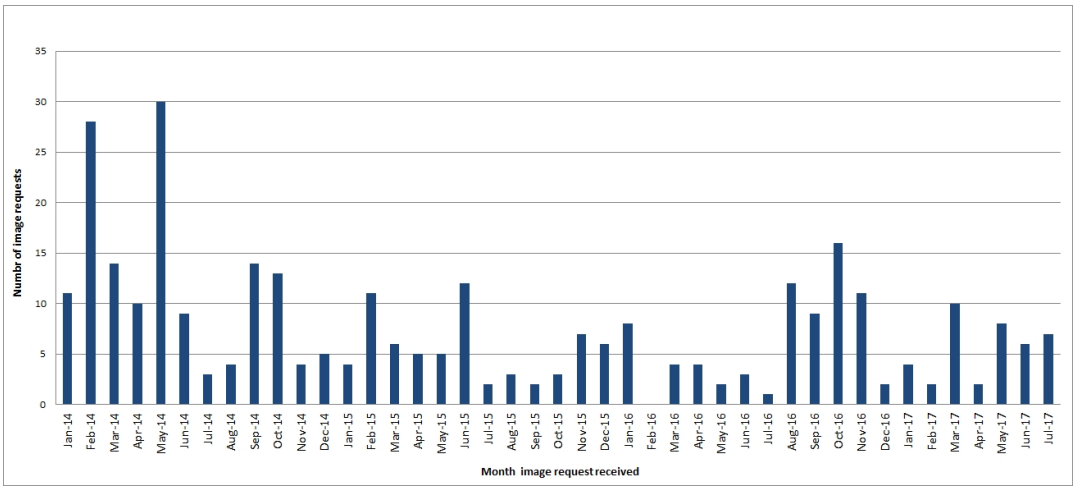


Figure 14. The number of image requests completed at RBGE from January 2014 to July 2017. Total number of requests was 360; number with no request date recorded is 31.

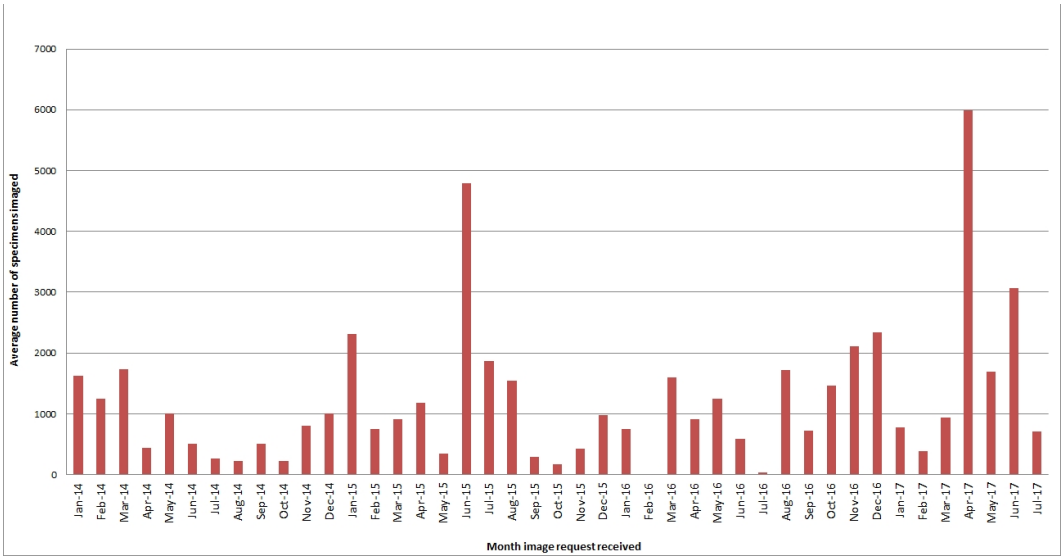


Figure 15. The number of specimens imaged through the Digitisation on Demand (DoD) service at RBGE. Total number imaged through the DoD service is 73,231; number with no request date is 20,318.

3. Building a network for Digitisation on Demand (DoD)

Elspeth Haston & Robyn Drinkwater (Royal Botanic Garden Edinburgh)

3.1. Identifying barriers to adopting a DoD approach across participants

The report produced for SYNTHESYS3 D3.3 highlighted some of the factors that limit current digitisation rates and could be barriers to institutions being able to provide DoD (Phillips *et al*., 2014). The main factors identified in this report were lack of human resources and funding, both of which will potentially impact on an institutions ability to participate in a European wide DoD network. Other limitations identified in this report included lack of an adequate data storage solution, which would likely become an increasing problem if institutes actively engaged in the DoD network and received regular requests. Physical workspace, collection handling, equipment and technology were also limiting factors which could act as barriers to DoD.

From our survey we found that 80% of respondents already accept digitisation requests, although the frequency of these requests and the formality of this process between, and in some cases within, institutions varies. However many of the responses suggested that these requests are currently driven by loans and research interests or are limited to small numbers of specimens. The institutional reasons behind these current practices would need to be examined and where necessary overcome to allow for the DoD network to be rolled out and successful.

Whilst not a barrier in itself, the way in which collections are used and how this use is affected by material becoming available digitally needs to be considered, as it has an impact on metrics which are used to show the 'value' of a collection. The report on the impact of digitisation found that both curators and users of collections found that having specimens available digitally reduced the number of loans, and users also found less need to visit collections (Enghoff, 2017). If these figures are reduced by digitisation, then robust metrics need to be developed for digital access to allow institutions to be able to show how their collections are being used. Whilst some metrics are already being collected by institutions, these may not be easily accessible and more robust measures could potentially be developed to show collection use.

3.2. Report on market validation

3.2.1. Gauging demand

The NA3 Subtask 3.2.1 report gauged user and custodian opinions on digitisation (Enghoff, 2017). One custodian noted that "our digitisation efforts are a response to demand for data, rather than something that generates demand", suggesting that there is a 'market' for a DoD service. A digitisation network would also make collections in overlooked institutes more accessible, as one user of dry zoological collections noted that "without such images in most cases I would not even know about these specimens" (Enghoff, 2017).

Preliminary data relating to the number of data records and images downloaded from institutional websites and aggregators strongly support the demand seen from institutional loan and image requests (Cubey, 2016).

3.2.2. Institutional buy-in and capacity

Users noted other advantages of having collections available online, which would be a direct consequence of others requesting specimens (Enghoff, 2017). Each DoD request would make more of the collection available to the public, helping to raise the profile of the collection in the public domain. Having images available also allows researchers to better allocate their limited resources, by targeting collections which have suitable material for visits or to better target their loan requests, which can be limited to certain specimens or even derivatives of specimens.

The results of the survey carried out for this report suggest that Natural History Collections in Europe are generally well-placed to undertake DoD in a future network. There is also a knowledge base within the network to help institutes develop appropriate workflows for the collections they have yet to start digitising.

Some institutes would find servicing DoD requests easier than others, as a high proportion of their collections have already been digitised and it would just be a case of making these discoverable through the network. For other institutes, there would be a risk that the number and size of the digitisation requests coming through a DoD service would not be manageable and this could have a negative impact on the reputation of these institutes. This aspect of a DoD service would need to be carefully monitored and managed. Managing expectations from users would be particularly important for collection types which have been identified as complex and expensive to digitise.

A wide variety of methods are currently being used to make images available to users of the collections. All of these would potentially work well with a DoD network, as in the preliminary approach proposed the individual institute would be responsible for the method used to provide the requestor with the images. Depending on how the network develops it may be that more standard formats would need to be agreed, although this could be dependent on the institution's capability.

Most institutions don't currently have APIs (Application Programming Interface) for their collection catalogue. With the initial approach to forming the network it is not necessary for there to be an API, however this may change as the network develops in the future.

Some institutes calculate the cost of digitisation as a daily rate, others calculate a per specimen rate. The variation in equipment and methodologies between institutes will inherently result in different cost rates for digitisation. This should not be a problem in terms of calculating the cost of digitising a request, but there should be a clear structure to be sure that all institutes were potentially including standard elements, such as an agreed level of pre-curation, data capture, image format and resolution, etc.

Data quality needs to be considered for a DoD service - as one user noted "the number of blatant errors in the available data often gives low trust in the information that appears credible" (Enghoff, 2017). Depending on the institution responding to the request, the type of data made available may vary, as institutional protocols will differ. Differences in the data provided already exist, as individual organisations completing requests will follow their own protocols, so this issue does not need to be addressed within a DoD network.

3.3. Recommendation for a technical Digitisation on Demand infrastructure

Most institutes holding natural history collections have existing protocols in place to handle requests for loans and images of the specimens. This currently involves a fragmented approach with each researcher contacting institutes separately to make a request. This approach has benefits and

drawbacks. The benefits include the assurance that in most cases the researcher has carried out preliminary investigations to ensure that the material in that institute would be present and relevant. The drawbacks include the amount of time required by the researcher to make a list of institutes, find out contact information for each collection, and write a number of individual e-mails, as well as the frequent missing out of collections which have important, relevant material. The exclusion of material means that not only are these data not included in the research which renders the research less robust, but these specimens are then not being studied, determined and annotated, actions which make the specimens more useful for future research.

The current system is fragmented and inefficient. There is therefore a need to build a more efficient system which would enable researchers to effectively discover and contact multiple relevant institutes, but which would still allow direct communication. At this early stage of system development this communication is necessary for processing requests.

There are several key elements to the process of requesting material from institutes which would need to be part of any future development.

3.3.1. Discoverability

The researcher needs to be able to discover which institutes hold relevant collections. There are online resources listing institutes which hold biological collections and these include the Global Registry of Biological Repositories (GRBio) which incorporates Index Herbariorum, the Biodiversity Collections Index and Biorepositories.org. The data held in GRBio is generally high level and does not include detailed content relating to the collections. At a European level, the Consortium of European Taxonomic Facilities has been developing an information portal with more detailed collection information for the 60 member institutes. However, neither of these resources currently enables a researcher to determine whether these institutes hold collections for a particular taxon or geographical region. The Global Biodiversity Facility (GBIF) enables a researcher to search the collections but will only include the collections that have already been catalogued and databased. There is therefore a need for a resource which can help a researcher make contact with the institutes who may hold relevant material.

3.3.2. Searchability

Having discovered the relevant institutes to contact the researcher now needs to find out what specimens they hold for the taxon or geographical region of interest. Whilst the aim would be to have an automated system based on some level of inventory produced by the institutes, most institutes do not have the capacity for this exercise. In addition, there is no clear cost benefit to carry out an inventory to the level required to provide a resource for a DoD service. Given very limited resources, priority is being given to preserving, curating and digitising the collections at present. The primary method of searching the collections currently is by contacting the curation staff who then manually check the collections for a particular request. In the short term it is feasible to continue using this method until more of the collections can be catalogued and made searchable.

3.3.3. Request

Having discovered which institutes hold the relevant collections, the researcher then needs to contact the institutes directly to request access to the specimens. This would normally be in the

form of a loan, a visit or a digitisation request. Although a small number of requests are still received as paper documents through the post, most requests now arrive by e-mail, often sent to a central collection e-mail address rather than to an individual.

3.3.4. Response

Each incoming request requires assessment and consideration. When a request is received the curator responsible for requests will either assess the collections and make a decision directly or they may forward the request on to another member of staff for additional input. There are several levels of decisions to be made for incoming requests. A loan request may be accepted (fully or partially) or refused, it may be transformed into a digitisation request and/or it may be transformed into a recommendation for a visit.

3.3.5. Access

If the request is accepted, the researcher will receive access to the specimens, either physically in the form of a loan, and/or virtually in the form of digitised specimens. The digital specimen files may be sent to the researcher directly or they may be made accessible online for the researcher to download as desired.

3.3.6. Feedback

Following access to the specimens, a certain level of feedback is generally requested. This includes notification of any taxonomic changes and annotations as well as notification of any published research/citations along with references. This feedback is often not received, and it then falls to the institute to try to discover any citations and references.

3.3.7. Impact

The impact resulting from access to the specimens is generally accepted as one of the most difficult elements to discover and monitor. It is also one of the most important and desirable metrics for demonstrating the importance of Natural History Collections to funding agencies.

A technical infrastructure for a DoD service needs to provide a system which includes all of the above elements to some extent. At a simple level this could be a listserv with a filter system so that a researcher requesting images for a particular taxon would tick a box identifying the type of collection and the e-mail would then only be sent to institutes which had listed that collection type. Therefore a plant request would only go to herbaria, and an insect request would only to institutes with entomology collections. On receiving the request it would then be entirely up to the institutes to respond directly to the researcher. That discussion would include a discussion about the size of the collection and whether the institute would be able to digitise it or whether they ask the researcher to arrange a visit.
Institutes signing up to the system would indicate which major taxon groups they hold and would be prepared to receive requests for to allow the filters to work. A structural overview of a potential DoD network is presented below (Fig. 15).

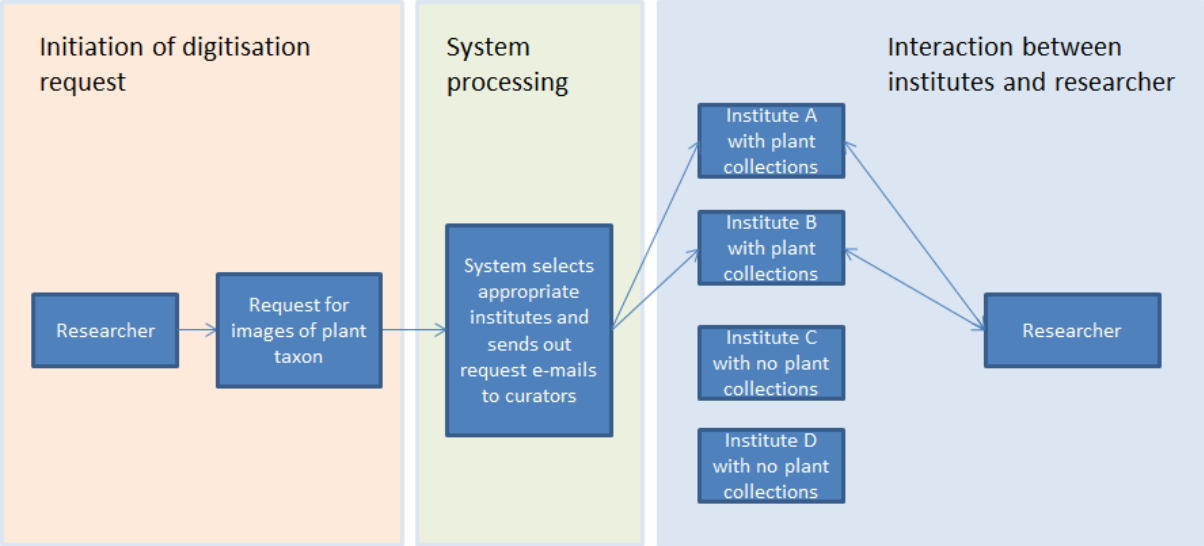


Figure 15. Diagrammatic overview of a potential European Digitisation on Demand network of institutes holding Natural History collections.

3.4. Plan for deployment of a European networked Digitisation on Demand (DoD) service

We recommend that the deployment of a networked DoD service for European institutes be carried out in a staged process. Designing a system that is simple to implement would result in a high uptake of institutes joining the network. This will be critical for the service to fully succeed and develop further.

An example of a deployment model can be seen in the development of the Stable URI system. This was implemented by a small number of institutes who became early adopters. Following the acceptance of the system as a CETAF standard, more institutes signed up and there are now 13 institutes implementing the Stable URI system for specimens. This model, whereby the network comprises only two or three institutes at the start, would result in an inefficient service and would discourage use.

The deployment plan would therefore involve focussing on building the institutional network with a simple system. Starting with a largely manual system, the only automated element in the system would be the filtering of appropriate institutes for inclusion in the e-mail list. The text of the e-mail would be generated by the researcher, who would also tick a checkbox for the type of collections. This would be used to filter the institutes to which the request would be sent, including only those who hold the selected type of collection, eg plants, fungi, invertebrates, etc. The individual curators would then respond directly to the researcher and all subsequent communication would be off list, following existing institutional protocols. One model for this is the Plant Search of Botanic Garden Conservation International (BGCI) where individual Botanic Gardens provide details of their plant collections to BGCI, and BGCI provide the search tools (https://www.bgci.org/plant_search.php) to search the aggregated data. The user searches the site for a species that they would like to access or find out more about. If the species is held in one or more collections, an email query is then sent to all collection holders with material who then decide how to respond.

As stated, the key aim at this stage would be to build the institutional network, encouraging institutes to collaborate and making entry to the network easy. In future it would be possible to develop a more automated system as more institutes carry out surveys of their collections. It may also be possible to incorporate costs to feed into a crowdfunding programme.

4. Can we assign priorities for digitisation based on crowdfunding?

Jennifer Pullar and Laurence Livermore (Natural History Museum London))

4.1. Report Background

One of the SYNTHESYS3 Joint Research Activity objectives (4.1) is to undertake research on the feasibility for "digitise on demand" services for European natural history collections. We have looked at the potential of using Crowdfunding as tool to cover the costs of a specific digitisation initiative, as an alternative to pay-as-you-go or other funding models.

In this report we present: (1) an overview of the current crowdfunding landscape, (2) a summary of the experience of the SYNTHESYS partners with crowdfunding, and (3) a brief assessment of the feasibility of crowdfunding as an alternative funding method for digitisation.

4.2. Introduction to Crowdfunding

Crowdfunding is described in the Oxford Dictionary as 'the practice of funding a project or venture by raising money from a large number of people who each contribute a relatively small amount, typically via the Internet'. Crowdfunding is closely related to and historically preceded by *crowdsourcing,* described as 'The practice of obtaining information or input into a task or project by enlisting the services of a large number of people, either paid or unpaid, typically via the Internet.'

4.3. Crowdfunding models and platforms

There are four broad models for crowdfunding, some of which (especially donation and rewards-based) can overlap within the same platform and sometimes occur in the same project:

Donation-based – entirely philanthropic

Rewards based – non-financial incentive to funders e.g. t-shirt, posters, digital rewards, event invites.

Lending-based – repayment of funds to funders with interest. (e.g. Lending Club, Zopa)

Equity-based – share profit with funders (e.g. Crowdcube, Pepins, Seedrs)

Figure 1.

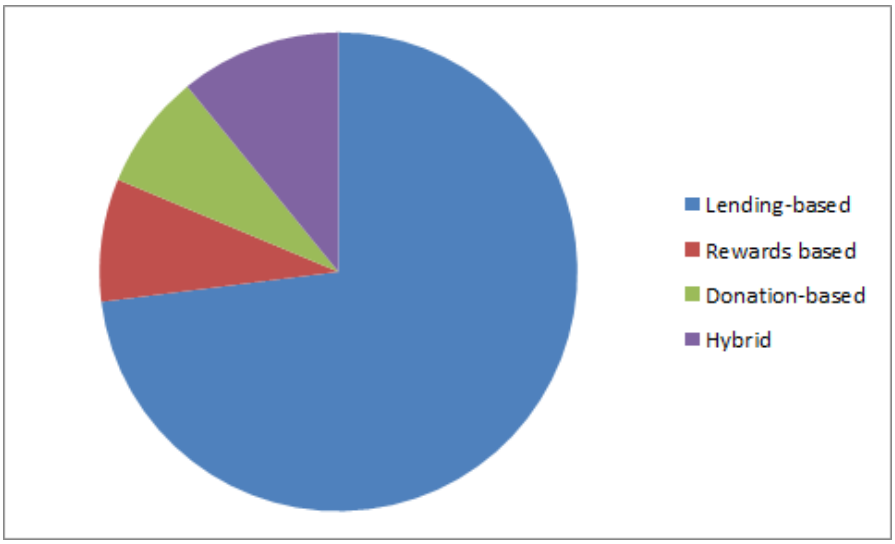


Figure 1. Shows the estimated use of crowdfunding models by the 2015-16 Crowdfunding Industry report. The report estimates the total raised through crowdfunding during 2015 to be $34 billion. Of this 73% of crowdfunding donations worldwide were lending based; 8.3% donations based; 7.8% rewards based, 7.5% equity and the rest hybrid. Traditionally the donations and rewards based systems have been used well within non-profit projects and organisations, so we would potentially be interested in looking at the combination of the donation based and rewards based models for crowdsourcing for european natural history institutions projects. 'The lending and equity models are essentially forms of investment, subject to additional regulation and therefore not suitable in the context of this report.'

4.3.1. Popular crowdfunding platforms

These are platforms that are known to SYNTHESYS partners and referred to in crowdfunding literature. These examples focus on the donation and reward-based platforms. All crowdfunding platforms take commision and some have rules about the types of funding supported by their platform. 'All or nothing' funding options return the funds to donors when the project's financial goals are not reached whereas with a partial funding option, you may be able to collect the money that was raised. Partial funding options usually come with a higher level of commision.

Platforms	**Commission**	**Type of funding permitted**	**Other**
Indiegogo 'Go fund yourself.'	4% if reach funding goal plus 3% credit card fee	All or nothing and partial funding options	If unsuccessful 9% to keep funds or no fee when funds are returned. Additional $25 fee for non US campaigns
Kickstarter 'Fund and follow creativity.'	5% Commission and various paypal/ credit card fees	All or nothing	Boasts a 44% success rate Monthly traffic highest of all platforms 5.5m
JustGiving JustGiving is the #1 platform for online giving in the world	5% commission and various paypal/ credit card fees	All or nothing and partial funding options	Mainly used for charity and personal causes. Currently used by The Natural History Museum, London and many UK museums and educational institutions for non-project specific donations.

Crowdfunder 'the UK's #1 crowdfunding website, where ideas happen'	6% each pledge plus credit card and paypal fees	All or nothing and partial funding options	UK site, have 'crowdfunding for museums' section – useful for public outreach.
Experiment 'help fund the next wave of scientific research'	5% commission and 3-5% credit card fee	All or nothing	Can mark projects with Open Access badge to indicate that results of project will be released in openly accessible way.
Rockethub 'The leading global community for entrepreneurs'	4% commission and 4% credit card fee	All or nothing and partial funding options	If unsuccessful 8% to keep funds or no fee when funds are returned
DigVentures 'Archaeology / In Your Hands' DigVentures supports and delivers exciting crowdfunded and crowdsourced archaeology projects around the world	4% commission and credit card fees. If unsuccessful 9% fee.	'keep it all' model, rather than 'all or nothing'	Only supports rewards-based projects - stating that 'each funding level must have deliverable benefits for purchasing.'
Crowd.Science Crowdfund science with us	5% commission and 3% credit card fees	All or nothing and partial funding options	Only supports educational rewards based crowdfunding and only accepts science projects.
Ulule 'Make good things happen'	5% commission and 3% credit card fees	All or nothing	Based in Paris, Montréal, Barcelona, Rome and Antwerp and boasts success rate of 68%.
World of Crowdfunding 'Crowdfunding for everyone'	6% commission and various credit card transaction fees	All or nothing and partial funding options	Project pages must contain at least one image or video and must have collected at least $ 100 online donations before they appear in site's search results.

Growfunding	10% commission	All or nothing	Only supports projects in Brussels and requires these to be rewards-based.

4.4. Relevant examples of Crowdfunded Museum projects

We looked at crowdfunding case studies in related project areas that we thought would be comparable areas of interest for SYNTHESYS partners. We looked at REBOOT THE SUIT, a conservation and digitisation project at the Smithsonian; Rebuilding Eric, a project to rebuild the UK's first robot at the Science Museum and Defying Dementia, a current crowdfunding campaign to fund Alzheimer's drug research at Lancaster University.

4.4.1. Science Museum - Rebuilding Eric - the UK's first robot

'Rebuilding Eric' was a secondary project alongside the Science Museum's major exhibition 'Robots.' this meant that the project was timely; had relevant promotional opportunities, there was a ready made community to publicise to and also if the project didn't receive funding there wasn't a plan to go ahead without it. This meant that the funders knew that the project was only happening because of their donations.

Rebuilding Eric used the platform Kickstarter and had a goal of £35,000. The Museum added stretch goals to the project when it looked like the full amount would be pledged. A total of £51,813 was raised with 861 backers. The campaign used the rewards model of crowdfunding in which the most popular amounts pledged were £5, £20, £25, £30. The reward for up to a £5 donation was a digital thank-you, but the other options rewarded funders with a specifically designed t-shirt or tote bag with the slogan "I brought Eric back to life" and mini versions of Eric the robot.

Fundraising period: 10/07/2017 - 16/6/2017 (37 days)

Project URL: https://www.kickstarter.com/projects/sciencemuseum/rebuild-eric-the-uks-first-robot

4.4.2. REBOOT THE SUIT Smithsonian

REBOOT THE SUIT was a Smithsonian project to raise funds to conserve, digitise and display Neil Armstrong's spacesuit. This used the Kickstarter platform and had an initial goal of $500,000 which was exceeded by over $200,000 and raised $719,779 with 9,477 backers.

Rewards included virtual gifts for the lower level rewards, physical items for the medium level and unique experiences for donations above $2,000. The higher amounts tended to include all the rewards available below this amount too. The four most popular rewards received more than triple the amount of backers as the perks neighbouring them – as such, we can take these numbers as direct evidence for the popularity of the rewards.

Most popular Pledges
$1 thank-you email (1,015 backers)
$20 cling on decal boot print (1,985 backers)
$46 3D scan data for space suit glove (1,317 backers)
$55 exclusive mission patch designed by Star Trek and NASA designer (1,594 backers)

Fundraising period: 20/07/2015 - 19/08/2015 (30 days)

Project URL: https://www.kickstarter.com/projects/smithsonian/reboot-the-suit-bring-back-neil-armstrongs-spacesu

4.4.3. Defying Dementia

Defying Dementia is an ongoing Crowdfunding project run by Lancaster University. Defying Dementia are raising money to to fund the safety testing to get an extremely promising Alzheimer's drug into human clinical trials. In the UK, crowdfunding is common in the arts and community project fields, but not in basic medical research (Pugh, R. 2016). Defying dementia have a long time scale for reaching their project goal and have been actively crowdfunding for two years.

The team are using an entirely philanthropic donation-based approach to fund this work. They are gathering funds from a mix of platforms: Just Giving, posted cheques online donations. This campaign is currently at £125,000 of the £165,000 target (June 2017). Interestingly in comparison to the reward orientated campaigns, the most common donation amount was not the lowest possible; fifteen donors on just giving donated £5 and £100 also received the same amount of donations. £10 and £20 were by far the most popular pledge amounts and several of these were accompanied by personal reasons that the donor was pledging the amount: "for x who is fondly remembered".

Fundraising period: ongoing

Project URL: http://www.lancaster.ac.uk/defyingdementia/

4.5. Review of the Academic Literature

There has been a surge in the literature around crowdfunding over the last five years. For this project we looked at a variety of literature and began noticing some obvious trends to draw out that might be of use to SYNTHESYS partners. Mollick, in *The Dynamics of Crowdfunding* (2014) discusses patterns associated with successful fundraising. These are: project quality, social networks, appropriate goals and careful planning to ensure that when a project is ready to execute once funded. *Crowdfunding in Museums* (Bump, 2014) discusses the significance of crowdfunding for museums and best practice standards to running a crowdfunding campaign. Wheat et al. in *Raising money for scientific research through crowdfunding* conclude their report by stating that 'completing a crowdfunding project only marks the beginning of the relationship between scientists and the 'crowd'' (p.71. 2013) they also posit that the true potential of crowdfunding is not just the money

that can be raised, but the opportunities for education and outreach. Crowdfunding has the potential to shift perception and increase public involvement to foster future positive relationships between scientists and nonscientists.

The key aspects that were emphasised in the academic research were the *audience* that engage with crowdfunding; the *importance of campaigns* in spreading the word, communicating and engaging crowds with the project both during the fundraising and through the execution of the project that was fundraised for and the the *types of projects* suitable for crowdfunding. We will take a further look at each of these areas.

4.5.1. Audience

Bump states that *'crowdfunding is a key tool to engage Generation X and Millennial donors'* (p20. 2014) she backs this up with figures from the 2013 Millennial Impact report in which 70% of all annual giving came from the 'Baby Boomers' (those born 1946-64) or previous generations. In contrast, crowdfunding is most likely to engage Generation X (born 1965-1980) and millennials (1981-1995) who may give less money, but want to track the progress. *'Both generation X and Millennial also want to know that their donations make an impact and desire to see proof throughout the success of the campaign.'* (p. 20) This provides a challenge for museums to provide proof of impact to a wide audience who may not individually contribute much. However there are possibilities here to engage your audience to become ambassadors for the project through their networks.

4.5.2. Campaigns

Successful crowdfunding develops a crowd, so outreach is essential. Wheat et al. state that *'the probability of reaching a fundraising goal depends upon a successful outreach campaign, with larger crowds translating into more money raised.'* (p.71, 2013.)Thorpe in 'Crowdfunding for social good' states that the success of a crowdfunding campaign *'has much more to do with your marketing than your cause'* (p19. 2013) Developments in digital technologies mean that we are more able to interact with audiences who might never be able to visit the physical museum. Mollick found that *'social network ties have been found to be important in crowdfunding'* (p. 8, 2013.)

Crowdfunding platforms encourage two main ways to pitch a project, the narrative and story of a project and the media that goes with this. The narrative is the reason and passion that ties the whole project together. Kickstarter highly recommend using video to hook your audience into your campaign. 'Rebuilding Eric' was the able to really engage the public well with video. Combining archival footage and new video content with a current curator. The project was able to convey the passion behind the project and engage backers to be part of the story. Mollick used crowdfunding data to analyse what contributes towards campaign success. His findings suggest that *'social network size predicts success'* (p8, 2014) and demonstrates that *'videos and frequent updates are associated with greater success, and spelling errors reduce the chance of success.'* (p8) from his findings Mollick concludes that *'Success is therefore linked to the quality of products and not just the social network of founders'* (p8.) Engaging donors and non- donors in meaningful ways can not only encourage the audience to give to a campaign, but could also encourage them to become an

ambassador for the campaign supporting and validating the message on social media. Clarity of key messages become even more important when asking the online community to share the project further and tell others that they are supporting the campaign.

The literature suggests that contributors to a project generally expect a reward and these are key factors for success. All but one of the crowdfunding case studies featured creatively tiered digital and physical rewards. Although in the case studies, the physical rewards of t-shirts, tote bags and physical models were popular. Digital rewards; thank-yous, invites to special events and digital downloads were also popular. These rewards don't cost the organisation a huge amount, but do give out more campaign information and empower the donor with resources to share the project they support with their peers.

4.5.3. Project

As with any project, its advisable to set goals and assess the suitability of your project and your potential audience before embarking on the project. This will help you assess whether crowdfunding is the right platform for your idea, provide measures of success and also provide learning opportunities if the project doesn't meet its target.

If appealing to the Generation X and millennial target audience it may be possibly to call upon your audience to champion and share the project with peers. This audience is also primed to assess the success rate on your project. As contributors to the project, will want to know that their contributions are going into realising the project. The audience want to know that the links they may have shared with their peers were worthwhile and a positive or negative experience in their perception of project success could affect their view of the institution and if they would give to similar projects in the future. To address this, Mollick suggests that crowdfunding projects should have appropriate goals that allow a project to be delivered on time and careful planning to set these goals to *'rapidly execute a promised venture.'* (p14, 2014)

Setting realistic goals include setting an amount of money to be raised by a certain date, reaching a target audience, producing a product implementation a projects or advancing an idea or cause. Crowdfunding campaigns must be responsible and accountable to every donor and clearly articulate what the museum will do with the donors' contributions. Regular reports to supporters and the wider community should be included in the planning stages.
Selecting an idea or project that will suit the crowdfunding format and target audience is key. Thorpe (p 13-14, 2013) in *Crowdfunding for social good* identifies six aspects to consider when planning a crowdfunding project.

- **Face** - an image, product or message selected to best convey the story and point of the campaign
- **Urgency** - a sudden problem or threat which emotionally connects the audience and underlines why they must contribute now.
- **Geography** - local projects that serve a local community are more likely to achieve local success.

- **Community** - projects addressing a community concern or needs and those that create connections to a community is a positive factor in project success.
- **Event** - a event or activity that has natural associating to the project and can be a catalyst for promotion and engagement.
- **Politics** - Project should not be politically divisive as can alienate audience.

These aspects take into account the museum's online and local community and the potential interest in a promotional opportunities for the project. By evaluating how the suggested project meets these criteria a museum can determine whether the project is suitable for crowdfunding.

As Wheat et al.(2013) states that reaching the funding goal is only the beginning of the relationship, Mollick asks whether successful crowdfunding leads to successful delivery of the end goal of the project and found that only *'23.4% of projects delivered on time'* (p12, 2014.) If a project has been funding by several hundreds of donors, this could be letting down the exact audience that you have worked so hard to get an emotional reaction and attachment to the project with. Mollick suggests that projects should be selected on quality. Those communicating about project campaigns should look for ways to signal preparedness to potential funders to gain their trust.

From reviewing the literature and case studies around crowdfunding, it appears that best practice for a crowdfunding campaign to be successful involve having the right project, setting goals, developing a concise message and marketing strategy including good visuals, accessing a social media community and creating lasting relationship of trust with donors including carrying your project out as soon as possible once a funding goal has been realised. We wanted to use this background information to find out awareness of and use of crowdfunding within the SYNTHESYS community.

4.6. Survey of SYNTHESYS Partners

A short survey was sent out to x SYNTHESYS partners and gained a response rate of x. Overall there was an awareness of Crowdfunding and the popular platforms. It possible that there were confusion over the term *crowdfunding* and *crowdsourcing* with a few partners saying they would use crowdfunding for 'data entry' or 'transcription'. Others had much more awareness of crowdfunding but hadn't found a suitable project yet. One partner had run a successful crowdfunding project. This is detailed below, as well as some of the responses to the survey.

Figure 1

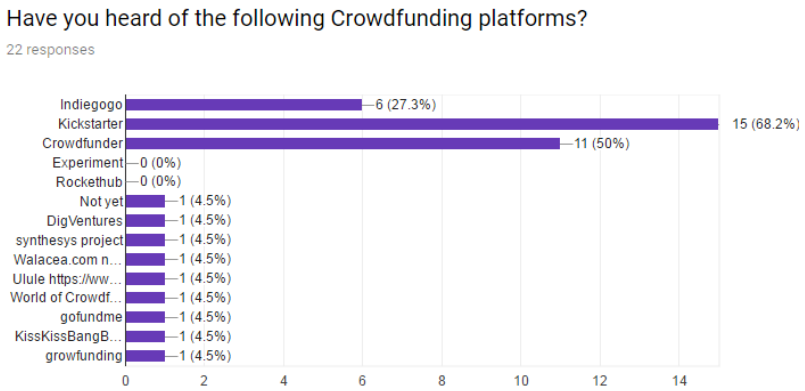


Figure 2

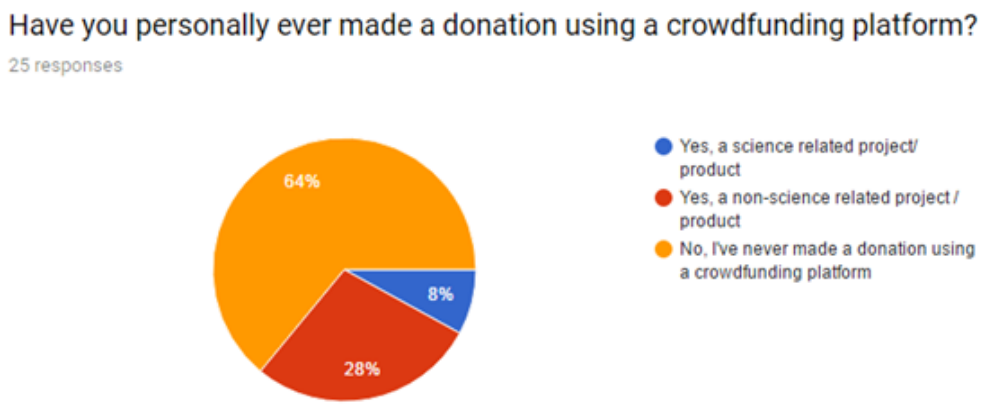


Figure 3

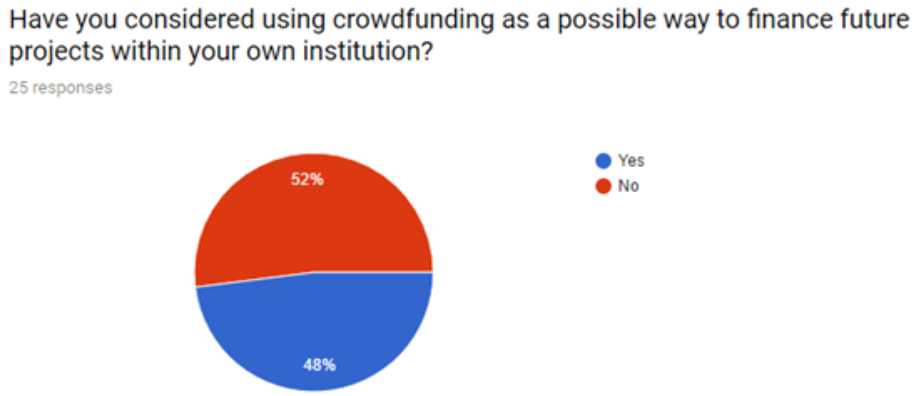
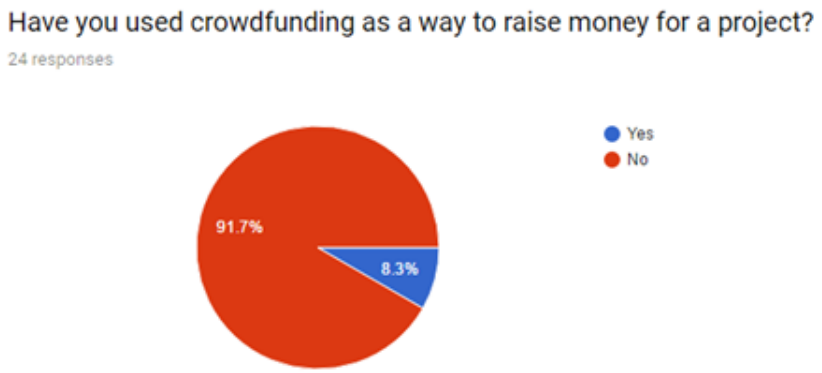


Figure 2 & 3 show us that although 64% of SYNTHESYS Partners have not given to a crowdfunding platform, nearly half of all partners said that they had considered using crowdfunding as a possible way to finance future projects and 71% have heard of had heard of Kickstarter as a Crowdfunding platform. Perhaps this could indicate that the partners believe they are targeting an audience other than their peers via crowdsourcing.

Partners suggested that crowdfunding could be suitable for projects that furthered science and also engaged the public:

"DNA sequencing project with a clear, interesting outcome."
"Taxonomy project with some conservation aspects"
"to develop digital galleries with 3D models of scanned Museum specimens"

Figure 4.



While nearly half of partners have considered using crowdfunding, most hadn't found a suitable project yet. Two partners were aware of funding campaigns within their institution although not involved in the projects. Partners from the Museum of Natural Sciences in Brussels provided details on a crowdfunding campaign they had been involved in and shared their experience.

4.6.1. Meet Ben the Plaeosaurus - Museum of Natural Sciences, Brussels

The Palaeo team at the Museum of Natural Sciences wanted to raise €25,000 to fund the curation and display of a new Plateosaur for their Dinosaur gallery. They reached 198% of their goal raising a total of €49,510 using the growfunding plaform and a rewards system.

Their feedback is below:

" This project was very exciting and stimulating. It had a quick and impressive success (almost 200%.)

- But the success requires
  - a good story to tell to which the citizen can identify
  - a good network and communication plan
  - a lot of work and implication before, during and after the action"

Fundraising period: 23/2/2017- 21/5/2017 (87 days)

Project URL: https://www.growfunding.be/plateoteam

We asked SYNTHESYS partners to rate the following aspects in determining success in crowdfunding projects:

- The size of an institution's social network
- A compelling story/ reason
- Compelling media
- A well planned communication campaign

- Building relationship/ trust with donors over the longer term

Figures 5 To 9 show the ratings of importance given by SYNTHESYS partners to aspects that are key in determining success in crowdfunding. Compelling media such a videos, images or graphics was ranked highest, even higher than a compelling story or reason which was ranked second. Third ranked was a well planned communication campaign followed by the size of an institution's network. Finally building relationship and or trust with donors was ranked the least important aspect in determining success of a crowdfunding project. This could be because success in crowdfunding is seen as the success of raising the funds rather than an overview over the end product or project. It might be interesting to see what regular donors to crowdfunding campaigns might respond to this question.

Figure 5.

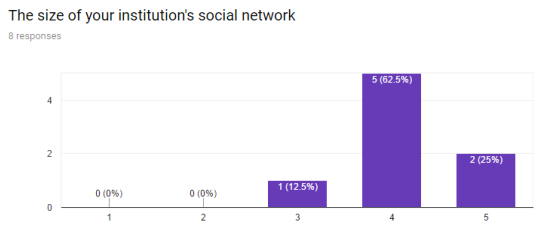


Figure 6.

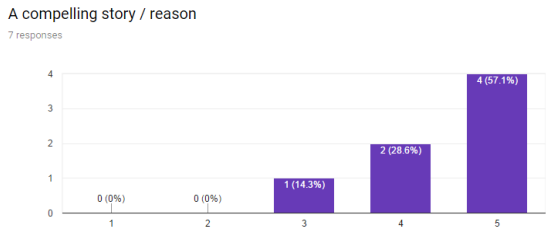


Figure 7.

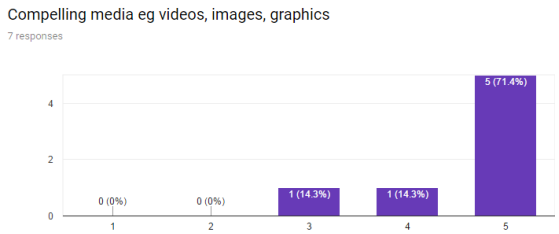


Figure 8.

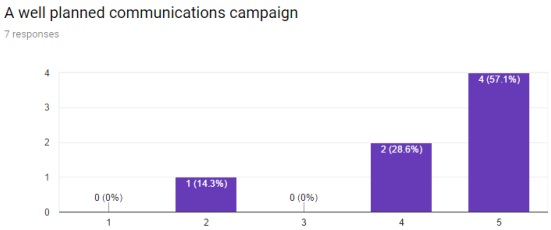
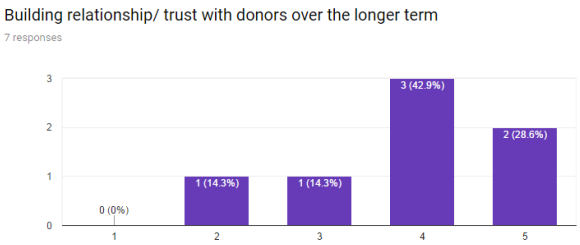


Figure 9.



4.7. Recommendations for Crowdfunding as part of Digitisation on Demand

The academic literature indicates that Crowdfunding is contributed to by Generation X and increasingly a Millennial audience that may only give a small amount of funding, but want to receive ongoing feedback to enable them to track project success. This audience have multiple social media accounts and the ability to become a digital ambassador for a project. The experience a donor has with a project will have knock on effects on how that donor feels about the institution attached and support projects in the future.

Results from our survey of SYNTHESYS partners tell us that there is a broad awareness of the most popular crowdfunding sites within the consortium. However, only 36% of partners have ever given to a crowdfunding project, while nearly 50% consider Crowdfunding as a potential viable route for future funding. This indicates that the audience intended to contribute to projects, falls outside the SYNTHESYS consortium and their peers. This report would therefore recommend that research into how best to adapt communications to engage a millennial audience would be a key consideration to SYNTHESYS partners considering a future crowdfunding project. This should include plans on how to provide feedback and enable the donors to track progress and share the project with peers.

The case studies and literature show that planning the marketing campaign and choosing the right project is key to crowdfunding success.Thorpe (2013) identifies key aspects to consider when planning a project as indicators of success. These aspects are face, urgency, geography, community, event and politics. Considering how a project meets these factors not only allow project managers to find out whether the project is suitable for the crowdfunding platform, but also start to think about how to integrate the marketing campaign into the project planning.

The Science Museum's 'Rebuilding Eric', the Smithsonian's 'REBOOT THE SUIT' and the Museum of Natural Sciences, Brussels' 'Meet Ben the Plaeosaurus' put careful thought into aligning projects with timely and relevant marketing opportunities. For example the Science museum launched their project at the same time as marketing material was going out about their major exhibition, Robots. This provided a relevant and interested community who may wish to take their engagement to the next level and support the museum by becoming a donor.

These projects also embedded creative tiered rewards in their marketing campaign to acknowledge donors and show appreciation for contributions. Donations and rewards based systems have been used well within non-profit projects and organisations and an 'all or nothing' approach is recommended most strongly within the literature. It is worth considering which approach will most suit the project, institution and organisation from the beginning of the project. There may be aspects of the project that may lend themselves well to 'rewarding' the donor with relatively low resource needs on the institution. Many projects will be reporting to key stakeholders on the success of the project with project updates. Why not tailor these updates to your donors and consider them a stakeholder? Many crowdfunding platforms offer a customisable feature to generate a thank-you at the time of donation and most projects could have some web presence other than the crowdfunding platform, it would be relatively easy to feature donors names on a webpage as a thank-you.

Crowdfunding can be seen as an effective tool for raising funds for special projects. However, due to the relatively small amounts being given by each donor, it takes a lot of people to form 'the crowd.' This means that a lot of energy and resources needs to be embedded in the communication side,

telling the story and translating passion to the potential donors. It's not a suitable platform for all projects and therefore not a substitute from traditional funding. If a crowdfunding project doesn't reach its original financial goals, it should be used as an opportunity to learn and adapt future projects to be more successful. Crowdfunding is still a relatively recent development in the fundraising landscape and institutions will need to monitor and track their progress so that they can tailor approach to their donors and non-donors in order to make best use of Crowdfunding as a potential alternative funding method for digitisation.

5. A case study on Digitisation on Demand

Steffen Kiel, Johannes Lundberg, Stefan Daume, Irene Bisang (NRM, Stockholm) and Margaret Gold (NHM London)

5.1. Introduction

Natural History (NH) collections represent an essential research infrastructure for a range of disciplines, including systematics, ecology, and environmental sciences. Swedish collections, with their 33 million specimens, represent an outsize proportion of the global holdings, both in terms of size and importance, and the research use of Swedish collections is extensive. Until recently, the use of collections has involved personal visits and loans of specimens but we are now facing transformative changes involving big-data and e-science approaches to digital representations of the collections. This will enhance existing use and open venues for new innovative uses of NH collection data for a variety of purposes, including addressing societal challenges. The development is driven by initiatives like the Global Biodiversity Information Facility (GBIF), whose objective it is to make digital biodiversity information from a multitude of sources available to researchers, authorities, and the interested general public.

This case study on 'digitisation on demand' at the Swedish Museum of Natural History (NRM) in the scope of SYNTHESYS WP4 was designed within this context. It aimed to explore workflows for rapid digitisation and label transcription that can contribute to the increasing demand for digital collections. The main project objectives were:

1. Test imaging and transcription workflows to inform future 'digitisation on demand' projects;
2. Gain experiences in building and promoting highly specialized crowdsourcing or citizen science challenges crossing into the area of cultural history that eventually can support the digitisation efforts at NRM;
3. Support future digitisation proposals with a partial quality assessment of crowdsourced information;
4. The project will also lead to making available previously not digitised material which will now support ongoing and future research projects.

The project consisted of two subprojects – Skåne brachiopods and Münchenberg herbarium – which will be presented below. Both subprojects were presented and advertised to the interested public at a Citizen Science Day at NRM, which is also reported upon.

5.2. Subproject I: Fossil brachiopods from Skåne, Sweden

The objectives of this subproject were to develop workflows and best practices for rapid imaging of fossil invertebrate specimens. We wanted to estimate time and effort needed, and the potential costs for digitisation on demand. We also wanted to gain experience (effort, time needed, data quality) with crowdsourcing of label transcriptions. The collection consists of an estimated 8,000 specimens ranging in diameter from < 1cm to ca. 8 cm, plus specimens embedded in rocks of even

larger size, up to 20 cm length. The majority of the specimens had already been sorted to species level. Because the Swedish government requires that all specimens included in the holdings of NRM are registered, we added numbers to previously unnumbered specimens.

The tasks we faced were:

(i) sorting the specimens of each species by locality and giving numbers to unnumbered specimens,
(ii) photographing all specimens,
(iii) connecting each image to its specimen, species and locality, both in the hard drive directory and in an Excel spreadsheet,
(iv) setting up and running the crowdsourcing project on Zooniverse.

5.2.1. The team

Project coordinator and senior curator Steffen Kiel was responsible for instructing and guiding the collections staff, and for most of the data handling. Sorting and photography was done part-time by a member of our collections staff who holds a PhD in paleontology and had been working in our department for about one year. The second person was a Syrian refugee with a degree in archeology who had previously worked at the Archeological Museum in Aleppo. She had no training in paleontology and zoology and was thus not familiar with animal systematics, but she had some experience with collections and collection management. She first came to the Paleobiology department at NRM through an internship program by the Swedish employment office (Arbetsförmedling), in January/February 2017, during which she gained expertise with paleontological collections management.

5.2.2. Work flow

(i) Sorting

The material was initially divided into three batches, the first two (ca. 50 drawers) were already sorted by species; the third batch consisted of unsorted and unidentified material and was left for a later part of the project. In batches 1 and 2, the specimens of each species were arranged first by species, within species alphabetically by locality, and within localities ascending by specimen number. Specimens without number were placed at the end of each locality section and were given new numbers once the sorting of the species in question was completed.

(ii) Photography

The specimens were photographed successively. Prior to photographing each set of specimens, a photo of a large yellow label indicating the species name and the locality of the following specimens (hereafter referred to as 'yellow label') was taken. This enabled us to quickly assign these data to the images. In the first batch, for each species, the camera was positioned so that the largest specimen, its label, and a 1 cm scale bar would fit into the frame. All specimens were photographed with this setup regardless of their size. Thus, we avoided adjusting the camera anew for each specimen to find an optimal frame size, which ensured efficient time management. Typically two images were taken of each specimen: the first was taken in a standard orientation for scientific purposes. A second image was taken when the specimen had a label attached to it, in an orientation to optimize the visibility of the label.

The second batch included many small-sized species (1 cm diameter or even smaller). In 5–10 % of the cases we took one image of all specimens from one tray, to save time. The specimen images on such a 'group photo' will later be isolated using the software *Inselect*. Photos were taken by one person at a time. We used a moveable camera stand that was positioned right next to the collection to avoid long transport of heavy drawers.

(iii) Data handling

Using the yellow labels with species and locality data, the images were sorted by species; there was one folder per species. Using a short computer script written in R by project coordinator Steffen Kiel, the images in each folder were renamed and re-numbered, and an Excel spreadsheet with the image names was produced for each folder. Again by using the yellow labels, the locality of each specimen image could then be easily added to the Excel spreadsheet. This procedure was easy to explain to the collections staff and was done by them. Finally, all Excel sheets were collated into one list. The filling in of locality data in the Excel spreadsheet from the yellow labels could potentially be done automatically using OCR (optical character recognition), but this would have involved more time for programming that the project coordinator could afford.

Afterwards, the images for the crowdsourcing project were selected: only images either with an individual number, or if no individual number was shown, a range of numbers, were included. Locality data for each image was extracted from the main Excel spreadsheet using another short computer script written in R by the project coordinator. The renaming of the images, the collection of their locality information, and the extraction of the images for the crowdsourcing project required three working days for the 7,271 images of the first batch.

(iv) Label transcription via crowdsourcing

For the crowdsourcing project for label transcriptions, we used the online platform Zooniverse, which is designed specifically for such tasks. Individual projects on Zooniverse are called 'expeditions', and we coined our expedition 'Exploring Tropical Sweden'. The images were split into two separate workflows, one with instructions in English, the other in Swedish, with the aim to see how the Swedish vs. the international public would perform. We asked the volunteers to transcribe specimen number(s), name of the collector, and the year of collection from the labels. Setting up the project on Zooniverse took most of the working time of project coordinator Steffen Kiel during ca. 0.5 PM. Significant input to the set up was provided by Margaret Gold, science community coordinator at the Natural History Museum in London and funded by SYNTHESYS. The first Exploring Tropical Sweden expedition was launched on Monday May 15th 2017 and included 1810 images. Focus after the launch was on responding to various details and inconsistencies reported by the volunteers. It included a moderate re-structuring of the setup, which considerably eased the workflow for the volunteers. Exploring Tropical Sweden was advertised via the NRM's website, NRM's Facebook site, SYNTHESYS' Facebook site, by specifically informing amateur geologist/paleontologist groups around Sweden, and by a Citizen Science Day described below (see Impact for details). The museum's IT, Communication, and Pedagogical departments strongly supported these efforts. The project closed on July 4th after each of the 1810 images had been transcribed by three different volunteers, and had attracted a total of 125 volunteers. A detailed analysis of the transcription data was in progress at the time of writing this report.

5.2.3. Results

(i) Work flow

The team found it most efficient when each task (sorting, photography, data handling) was done by one person at a time, because when two people sorted at the same time they tended to be in the way of each other. The most time consuming aspect of the sorting, at least initially, was to get used to the old hand writing on the labels,. It often required asking curators or collection staff with long experience with this type of handwriting, and with the geography of Sweden. After some initial sorting it was possible that two persons could work on the different tasks simultaneously, i.e., while one was sorting, the other was taking images of already sorted specimens. This avoided that one person had to do one task 8h/day and alone. Because the photo stand was positioned in the collections, they still could work in a team but without being in each other's way while working on the same task.

(ii) Crowdsourcing

Perhaps the most interesting result in this respect was that the English workflow was finished long before the Swedish workflow. Yet, most of the volunteers that had been in contact with us appeared to be Swedish citizens. Because most Swedish citizens have an excellent grasp of English, our interpretation is that it didn't matter to them on which of the two workflows they were working. They may simply have picked the first, which was the English one. Thus as long as the task does not specifically relate to the Swedish language, it appears that 'Swedish' crowdsourcing projects can be run on international platforms such as Zooniverse without the need for extensive explanations in Swedish.

(iii) Numbers

Sorting of the first two batches was completed after two weeks, 0.5 PM. 51 drawers of fossil invertebrate specimens were sorted by species, locality, and number. Batch 1 consisted of 33 drawers, batch 2 consisted of 18 drawers, and a total of 2548 new numbers were assigned to previously unnumbered specimens. A total of 12,360 images were taken. See Table 1 for detailed specimen/image data per batch. Note that in batch 2 there were typically fewer specimens per locality and many very small specimens (< 1 cm). These factors resulted in more numerous adjustments to the camera setup and likely explain the longer time it took to photograph the second batch.

At the time of writing this report (mid-July), label information (specimen number, collector, year of collection) from 1810 images have been transcribed via crowdsourcing and are currently being processed. Furthermore, all images were digitally archived including species name and locality. Once the data from the crowdsourcing project have been processed, these data will be imported into the collections database at the Paleobiology Department.

Table 1. Detailed specimen/image data per batch.

	Drawers	Specimens (est.)	Images	Days for photography	Images selected for crowdsourcing	Notes
Batch 1	33	4,000	7,271	18	1,810	7,271 is the number of registered images; ca. 250 images had to be repeated due to quality issues
Batch 2	18	3,000	5,089	12	To be completed Autumn 2017	

(iv) Cost estimates

Based on the estimate that we have 7,000 specimens and that the salary including social fees of a collections assistant is 35,000 SEK/month, the costs are the following:

- Sorting of specimens: 0.5 PM: 17,500 SEK / 7,000 = 2.5 SEK / specimen (~0.26 €)
- Imaging of specimens: 1.5 PM: 52,500 SEK / 7,000 = 7.5 SEK / specimen (~0.78 €)

Based on the same salary, the combined costs per image are: 52,500 SEK / 12360 = 4.25 SEK (~0.44 €)

In this particular collection, many boxes contained numerous specimens associated with a single label, which made photography fast because the label did not have to be moved from the box to the photo stand for each specimen. But this would be the case for example in type collections, where one specimen has one label. Photography of such one-specimen-one-label collection takes longer and hence costs more. No systematic assessment of this was made; however, examination of a small, random sample of specimens and images indicated the following numbers:

- 'multiple specimens' photography took on average 0.65 min / image,
- 'one-by-one' photography took on average 0.9-1.3 min / image

Thus 'one-by-one' photography can take up to twice as long as 'multiple specimens' photography, nearly doubling the cost estimates above. Evidently, the cost estimates provided above need to be adjusted depending on the salary costs of the involved personnel that depend on local conditions.

5.2.4. Discussion

The workflows developed specifically for this project resulted in a very reasonable cost for the imaging of an individual specimen of less than 1 Euro. This cost does not include the transcription of label information and data basing of the specimen. These would not necessarily need to be conducted when responding to a digitisation request. Nevertheless, they might be desirable to add to the digitisation task of the collection. If not performed on a batch basis, efforts for selecting individual specimen, and handling and sending the images, would need to be added.

For the current estimates of digitisation time and costs, the work of the project leader in introducing the team, guiding the actual sorting, imaging and labeling, including some programming for data handling is not considered. However, once the routines and processes are established at a NH collection institution, this would also not incur when responding to an individual Digitisation on Demand request.

In this particular project the project leader spent 149 hours (corresponding to 0,93 PM), which also included setting up the crowdsourcing project on the Zooniverse platform and responding to queries from volunteer transcribers. The costs of the project leader SK were only partly charged to the SYNTHESYS WP4 budget; the remainder constitutes a substantial contribution-in-kind by NRM to WP4. In addition, the efforts of several staff members at the NRM Paleobiology department in translating the instructions and explanations for the Zooniverse project into Swedish increased considerably the in kind contribution by NRM to WP4.

5.3. Subproject II: The Münchenberg herbarium

The 'Münchenberg herbarium' is the oldest known preserved herbarium from Sweden. The text is handwritten and a mixture of early 18th Century Swedish and Latin, complemented with plant names in German. The original aim of this subproject was to gain experiences in building and promoting highly specialized crowdsourcing or citizen science challenges crossing into the area of cultural history. The main task for the specialized crowd would be to transcribe the complete text. However, this much higher complexity of narrowly handwritten text covering a whole page, often with unclear relation to the associated plant specimen, was not straightforward to deal with. The Zooniverse platform does not provide a 'standard template' for such a task. The effort to create such a template was beyond the possibilities of this WP4-task and the efforts that the project coordinator Johannes Lundberg could devote to it. We therefore narrowed down the Münchenberg project from the transcription of whole pages to the following tasks:

(i) mark the presence of Scientific, Swedish and German names on the pages,
(ii) transcribe marked names.

This will enable to create a suitable template, and to construct tutorials. One of the reasons for asking for transcriptions of the names is that the herbarium contains unpublished Swedish (and possibly German) vernacular plant names. The Münchenberg herbarium subproject will be re-animated after the summer break beyond SYNTHESYS WP4.

5.4. The Citizen Science Day (CSD)

The CSD was held on 20 May 2017 in the main hall of NRM, five days after the launch of the Exploring Tropical Sweden expedition. The aims were to advertise the project and to engage an interested crowd in label transcription. Three computers where visitor could test the workflow were provided, a 'label reading challenge' that featured some difficult-to-read locality names was held, some of the actual fossil specimens were shown, and a table with drawing utensils and fossils for kids was available.

The Münchenberg herbarium subproject was also presented at the Citizen Science Day. We exhibited the original herbarium. On a large printout of one of its pages, visitors could transcribe individual

words using sticky labels. Everyone involved in the SYNTHESYS WP4 project at NRM was present at the event to explain the project to the public. The costs for other NRM personnel than the project team and -coordinators present at the CSD, as well as the efforts spent on promoting the project and the event, were not charged to the SYNTHESYS WP4 budget.

A very encouraging outcome for the Münchenberg herbarium subproject was that even rather young people (children, young teenagers) were able to read and transcribe part of the text. We also established contacts with the paleographic community which was represented at the CSD by a staff member of Riksarkivet (Swedish National Archive) and lecturer at the Department of Paleography at Stockholm University. The members of the paleographic communities and genealogical associations will be a crucial crowd to approach once the Münchenberg site will be launched in fall.

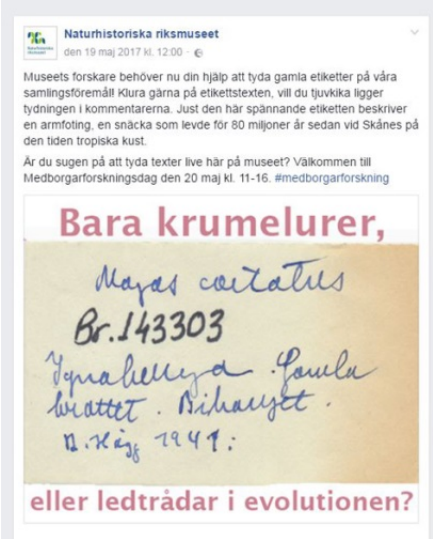
5.5. Impact, dissemination and outreach

This case study (Task 4.1.4) contributes significantly to Deliverable D4.5, "Digitisation on Demand" with time and cost estimates for imaging NH specimens, as well as to Deliverable D4.3 "Crowdsourcing for metadata enrichment" with insights and data on setting up, promoting and maintaining crowdsourced specimen label transcription projects. The former will facilitate the planning for Digitisation on Demand offers at NH institutions. The crowdsourcing experience gained in this task will be most valuable for designing future large-scale digitisation of NH collections.

The project and the related CSD on 20 May 2017 were extensively promoted on NRM's website and repeatedly posted on social media of both NRM and of SYNTHESYS .Through these actions, the SYNTHESYS project as a whole gained an additional wide exposure to the general public. NRM's Facebook posts received a social impact (sharing, linking, and interaction) above NRM's average.

Specifically, NRM performed the following dissemination actions

- Two editorial Facebook posts: 15 May, 19 May (see image)
- Facebook event announcing the CSD on 20 May (see above)
- Directed written invitation to CSD to Karin Borgkvist Ljung at the Swedish National Archives (Riksarkivet) who sent on the invitation to her networks
- Directed written invitation to CSD to Stockholm University, Department of Paleography (Director of Studies)
- Directed written invitation to CSD to the Swedish Genealogical Association



The case study and its results were presented at two meetings: in one oral presentation at the Nationella samlingsmöte [meeting of Swedish NH collections staff] on May 4, 2017, and in two oral presentations at the SYNTHESYS final meeting at NHM London on June 6, 2017 (http://synthesys3.myspecies.info/node/668).

At NRM, this project was the first step towards the initiation of a dedicated Citizen Science Platform on NRM's website. This site will assemble all NRM's Citizen Science activities (e.g., Reporting of animals in the context of environmental monitoring; Bird Ringing), including this case study on the Skåne brachiopods and the Münchenberg herbarium. It will feature a blog where results will be published regularly and contacts with the volunteers be maintained. The Citizen Science Days was well received among the staff including the management at NRM, and NRM is now aiming to hold such events more regularly (i.e., biannually).

Finally, the project allowed, with the significant support from the Arbetsförmedlingen (Swedish employment office) to employ a Syrian refugee. The financial provision by the Arbetsförmedlingen enabled us to hire the team member during five (mid-March to mid-August; see above Subproject I, Team) instead of the two months that would have possible with the SYNTHESYS WP4 funding, which significantly increased the output of imaged specimens.

6. References

Bump. M.R (2014) *Crowdfunding in Museums* available: http://scholarship.shu.edu/cgi/viewcontent.cgi?article=2978&context=dissertations accessed 20 June 2017

Crane, T (2017). IIIF (Interntional Image Interoperability Framework) for Natural History Collections. https://www.youtube.com/watch?v=waaxCqgr1V0

Crowdfunding industry statistics (2016) available: http://crowdexpert.com/crowdfunding-industry-statistics/ accessed 20 June 2017

Crowdfunding survey for SYNTHESYS partners (2016) https://docs.google.com/a/nhm.ac.uk/forms/d/1nLOYWIHsJOpbd9Vw7atBiE5-Uy0dBhTYofyynlJMgMw/edit?usp=drive_web

Cubey, R (2016). The usage of the RBGE Online catalogue & aggregators. SPNHC, Berlin.

Enghoff, H (2017). Does digitisation of natural history collections reduce the need for physical access and physical loans? SYNTHESYS3, NA3, Subtask 3.2.1.

Krishtalka, L, Dalcin, E, Ellis, S, Ganglo, JC, Hosoya, T, Nakae, M, Owens, I, Paul, D, Pignal, M & Thiers, B (2016) Accelerating the discovery of biocollections data. Copenhagen: GBIF Secretariat. Available online at: http://www.gbif.org/resource/83022.

Haston, E, Cubey, R, Pullan, M, Atkins, H, Harris, DJ (2012). Developing integrated workflows for the digitisation of herbarium specimens using a modular and scalable approach. Zookeys, 209: 93-102. doi: 10.3897/zookeys.209.3121

Hudson, LN, Blagoderov V, Heaton A, Holtzhausen P, Livermore L, Price BW, van der Walt S and Smith VS. 2015. Inselect: automating the ation of natural history collections. PLOS ONE. 10 (11), e0143402. 10.1371/journal.pone.0143402.

Mollick, E. (2012) *The dynamics of crowdfunding: determinants of success and failure.* Available: http://www.sciencedirect.com/science/article/pii/S088390261300058X accessed 20 June 2017

Phillips, S, Green, L & Weech, MH (2014). Review of Current Digitisation Workflows and Equipment, SYNTHESYS3, NA3, Task 1.3.Deliverable 3.3.

Pugh, R (2016) *Researchers turn to crowdfunding to develop Alzheimer's drug* available https://www.theguardian.com/healthcare-network/2016/apr/05/research-crowdfunding-alzheimers-drug accessed 20 June 2017

Stanley. W (2016) *How Can a Museum Use Kickstarter to Crowdfund?* available:https://www.museumnext.com/insight/how-can-a-museum-use-kickstarter-to-crowdfund/ accessed 20 June 2017

Thorpe, Devin. D (2013) *Crowdfunding for social good* via Bump. M.R (2014) *Crowdfunding in Museums* available: http://scholarship.shu.edu/cgi/viewcontent.cgi?article=2978&context=dissertations accessed 20 June 2017

Wheat, R.E. et al (2013) *Raising money for scientific research through crowdfunding.* Available: http://www.jarrettbyrnes.info/pdfs/Wheat_et_al_2012.pdf accessed 20th June

7. Appendix 1: RBGE Digitisation on Demand Survey results

7.1.1. Q.4. Does your institute currently have digitisation capability?

Yes, partial to extensive capability	20
No capability or very minimal capability	1

Additional commments:

Conveyor belt driven set-ups for vascular plants and insects respectively; macro lens stacking labs
We've developed capacity for certain workflows, but not all types of collection. Most have been done within bespoke projects so they cannot necessarily be replicated without some effort.
The capability and extent of current and previous imaging differs strongly among collection departments
We currently have a running project with the objective of scanning the entire vascular plant collection which is databased up to 80%. We started in January 2017, with the annual target of scanning 175,000 databased specimens. With this approach we will need to keep our funding capability for the next 4.5 years to complete the whole objective of 1 million specimens digitised (fully databased and scanned)

7.1.2. Q.5. Do you currently have the capacity to digitise the following collections?

Flat herbarium sheets	15
Bulky herbarium sheets, carpological and wood specimens	12
Microscope slides	12
Entomological specimens	14
Invertebrate zoological specimens	11
Vertebrate zoological specimens	12
Collections in spirit	9
Palaeontology specimens	11
Anthropological specimens	9
Minerals	7
TLC plates	1
Other (please specify)	6

Additional commments:

we don't have any entomological specimens but we have a digital microscope to digitise lichens and slimemolds
Micro_CT scanned collections
Among the minerals only the holotypes (c 200 specimens) are imaged, as we considered that morphology is not relevant for mineralogical and geoscience research
Cryptogams specimens, Prehistory specimens
Geological map collections
We can digitise microscope slides individually only, not bulk

7.1.3. Q. 6. What percentage of your collections are currently digitised?

Collection type	0	<1%	1-10%	11-50%	51-100%	Unknown
Flat herbarium sheets	0	0	7	6	2	0
Bulky herbarium sheets, carpological and wood specimens	3	2	7	1	0	3
Microscope slides	2	2	5	0	0	0
Entomological specimens	0	3	6	0	0	0
Invertebrate zoological specimens	1	1	5	0	0	1
Vertebrate zoological specimens	0	0	9	0	1	1
Collections in spirit	3	2	1	0	0	1
Palaeontology specimens	0	1	8	0	0	0
Anthropological specimens	0	0	4	1	0	1
Minerals	0	0	2	1	3	0
TLC plates						

7.1.4. Q. 7. Do you currently accept digitisation requests (Digitisation on Demand)?

Yes	17
No	4

Additional commments:

We are digitising specimens on loan on a regular basis. We are also offering help if external digitisation requests are relevant for BGBM research foci.
predominantly handled by individual curators; Zoosphere
management not yet organized
But, this is just occasionally and very individual. Till now this was not our primary cause of digitisation. occasional service
We do support DoD but only through bespoke projects, with the exception of Herbarium sheets which are sometimes digitised for free on request, if the number requested is low.
different among departments, as not all types of specimens are suitable to study as an image. Not relevant for mineralogical collection
Ad Hoc
We receive some 200 requests for specimens imaging per year. The way for sending those requests is via email, and our average response is one week. These on demand scanning are carried out with the same size (Din A3) and quality standards (600 ppi) than the rest of our images.
logging of New specimens imaged through Loan management system. Maximum of c10 images per request. When sending images use an Image Supply agreement (ISA).
Ad Hoc
we would accept them, but had no requests recently
E-mail and spreadsheet system, developing a database system. Turnaround approximately 2 weeks or less.
limited to 50 specimens per request; generally much less

7.1.5. Q. 8. If you do accept digitisation requests are you able to provide any of the following metrics?

Number of requests received	13
Number of specimens requested	10
Number of specimens digitised	12
Information relating to the person making the request (country, institute)	9
Reason for the request	13
Outcome of the request (download, publications, citations)	9
Determinations received following the request	7
Other (please specify)	5

Additional commments:

This is very on individual bases and just a few cases per year. We have all documentation if needed. We are sure that this will change in future, when we will open our digital catalogue (index) of our collections online. Few cases only
plants: 1-6; palaeoentomological: 1-3, zoological: 1-7 IF researcher provides feedback on determination, otherwise 1-6
Unsure about some of accuracy of some of the metrics and not all easy to extract. We log all requests to the HerbCat email address but this is more than just image requests and image requests may come from other routes. If requests leads to digitisting specimens then this is logged in Loan management system and ISA sent. Outcome of the request only sometimes known. Reason for request may be vague - general for Research.
We have limited information on outcome and the number of specimens requested is an estimate.
The 4 last bullets will need a lot of work to get the info

7.1.6. Q. 9. How do you currently make the images and data available to the person requesting them?

E-mail data and images	15
Send data and images using online transfer service (WeTransfer, SendIt, etc)	14
FTP data and images	7
Make available on institute online collections catalogue for person to download data and images	14
Other (please specify)	7

Additional commments:

hard disks
Make micro-CT scannings available from the LifeWatch Greece web site
institution's download server
hard disk drive for some CT Scan
Some images are able to be downloaded via our online culture heritage portal www.esbirky.cz
Due to the size of the high quality images, we usually send them via WeTransfer
Sometimes people want higher res images than available through our catalogue. No real download process of images from online catalogue. For more than one/two images we will use something like Webtransfer.
We can use the other options on request.

7.1.7. Q. 10. What image resolution and format do you provide for the person requesting them?

High resolution tiff files	17
Low resolution tiff files	4
High resolution jpg files	16
Low resolution jpg files	6
Tiled image files with an online viewer	9
Other (please specify)	3

Additional commments:

The image quality and format depends individually on the request purpose. Image quality

can vary in individual cases
Usually we send the high resolution tiff files, but sometimes also jpg files when requested
Tiff only upon special request

7.1.8. Q. 11. Does your online collections catalogue currently have an API?

Yes	7
No	9
Don't know	3

Additional commments:

URL RESTful
But we are providing API for our cultural heritage portal www.esbirky.cz. Only for on-line portal
See http://data.nhm.ac.uk/about/download
Different systems, some with API some other not
I believe we can transfer data to GBIF or other portals

7.1.9. Q. 12. Have you carried out any costing of your digitisation processes?

Yes	9
No	8
No response	1

Additional commments:

never for scientific and educational purposes but occasionally for business usage
for the CT Scan
we are not able to carry this. The cost relates to general service costs only (SEM etc).
We have approximate figures for slides, herbarium sheets, and palaeontological material.
Since NRM is time recording, it is possible to retrieve the time and costs for the digitisation efforts at NRM; however, I think it is not possible to distinguish imaging. Also, we have not broken down this to a specimen level.

They are free of charge
Different cost models depending on depth of digitisation
we are working on that
(regarding "on-demand " digitisation)

Where costing information was received this is not published here but has been used to inform the crowdfunding section