Digitising Collections:

Leveraging Volunteers & Crowdsourcing to Accelerate Digitisation

AXIELL
ARCHIVES LIBRARIES MUSEUMS

INTRODUCTION

When it comes to managing a collection, such as a museum, archive or gallery, ensuring the preservation of the records, objects and data is critical.

Today's technology-driven, internet-enabled world has opened up new opportunities to conserve, record and share collections and transform how they are managed.

This report by Axiell and MCN, explores the current digitisation strategies of today's collections management organisations, drawing on the views of more than 100 professionals working in museums, archives, galleries and other collections institutions. 118 Respondents took part in a global survey launched in October 2016 and follow up interviews were conducted with a sample of the participants

In particular, we focus on how readily they are integrating volunteers into their approach and moving to crowdsourcing to increase their rate of digitisation. Examples from more than 10 institutions world-wide provide insight into how these strategies are working in practice.

CONTENT

Section 1:
Prioritising Digitisation
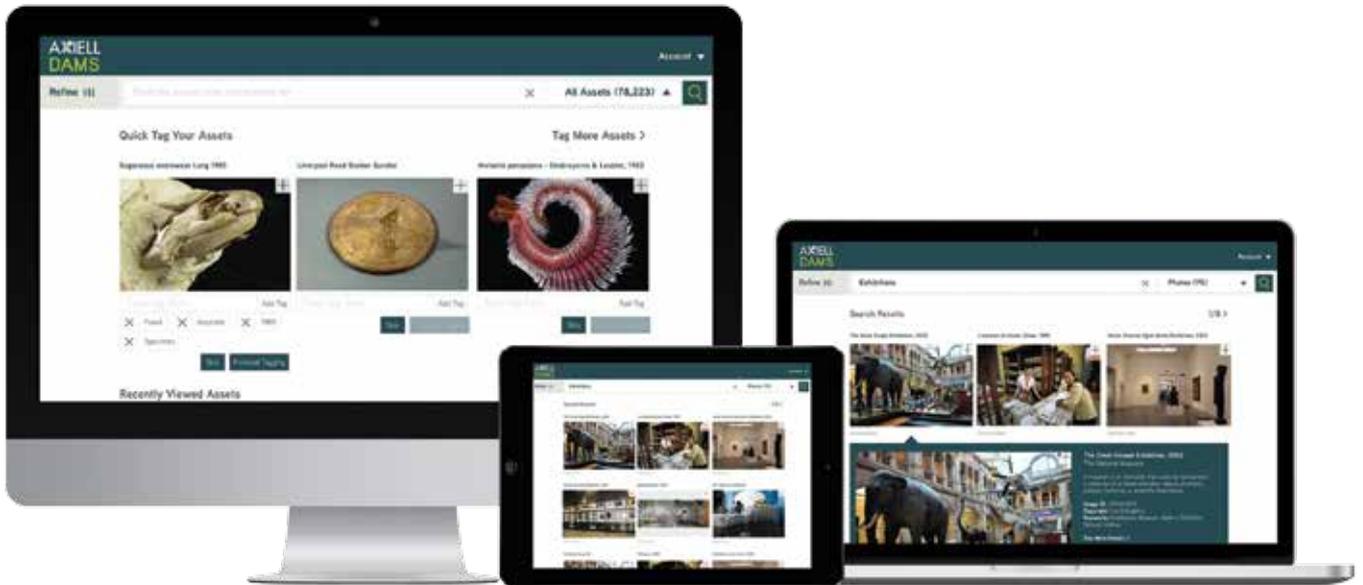
PRIORITISING DIGITISATION

According to a report by the Collections Trust, the number of devices connected to the internet already outnumbered the human population by 2008. With the internet becoming a day-to-day resource for consumers, it is no wonder that digitisation has become a core focus for today's museums.

Digitisation also goes far beyond making a collection available for public consumption. In addition to bringing collections to life to share with the world at large, creating digital replicas of objects and adding rich information and data also serves to preserve cultural and natural heritage and enables sharing with like-minded institutions worldwide.

DIGITAL DATA
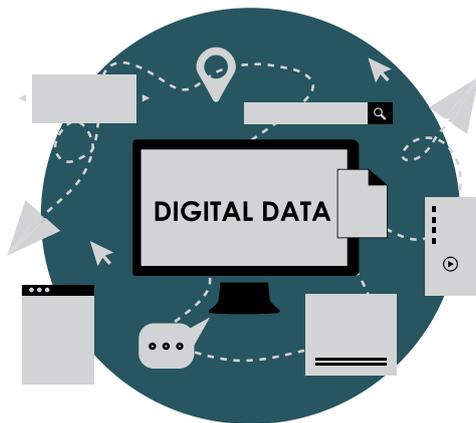
THE DIGITISATION JOURNEY

The article Digitising Museums, featured in Museums + Heritage Advisor, highlights how the number of items in a museum's collection is often staggering and digitisation projects can take years to roll out the infrastructure alone. Notably, once you start, the itemising and populating becomes an ongoing process rather than a defined, short-term project.

New research from Axiell and MCN (Museum Computer Network) carried out with collections professionals confirms that digitisation is a priority for the vast majority of collections institutions (86%). However, when comparing their goals for digitisation versus progress, we see that most still have a significant backlog to address, and based on their current rate of digitisation most are more likely to meet their goals in decades rather than years.

– TODAY, APPROXIMATELY WHATPERCENTAGE – OF YOUR COLLECTION IS DIGITIZED?

– WHAT PERCENTAGE OF YOUR COLLECTION ARE YOU ULTIMATELY LOOKING TO DIGITIZE ?

Today
Ultimately

	<10%	11-25%	26-50%	51-75%	75-99%	100%
Today	30%	22%	18%	19%	9%	1%
Ultimately	7%	12%	11%	8%	21%	40%
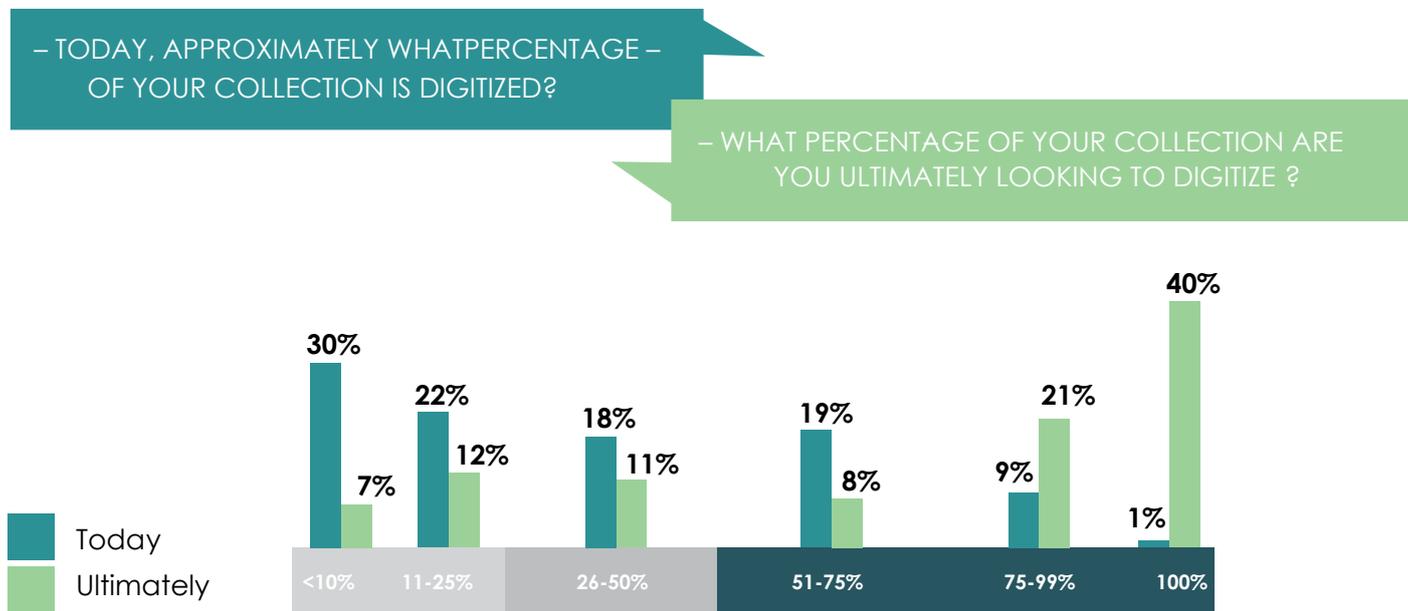
Two out of five (40%) of institutions want to ultimately digitise their entire collection and, overall, 69% are aiming for more than half. Yet today, given the current rate of progress, only 26% could confirm they will have met their goal within the next ten years and 16% can't even see the completion of their digitisation in their own lifetimes, predicting in excess of 50 years.

– AT THE CURRENT RATE, HOW MANY YEARS WOULD – IT TAKE YOU TO MEET YOUR DIGITIZATION GOAL?

16%
28%
8%
4%

| 1-5 years | 5-10 years | 10-20 years | 20-50 years | 50-100 years | 100-200 years | 200+ years |

27%
12%
4%
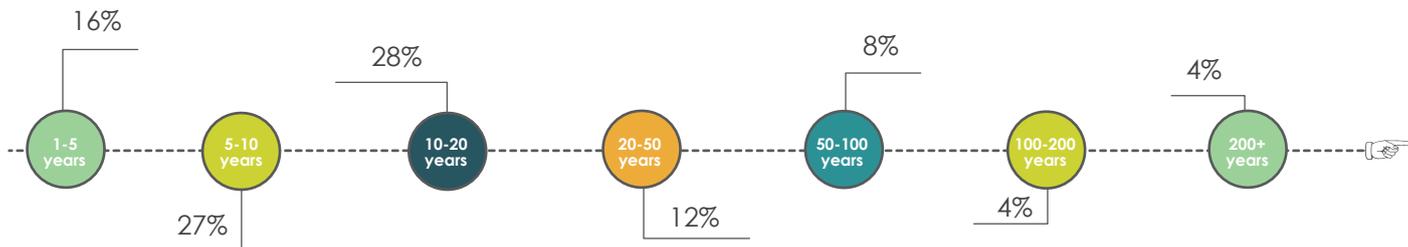
These numbers are unsurprising when we look at the scale of the task at various institutions.

At The Met there were about 60,000 art works on view at any given time and I would say that 99% of them had images, and of course they rotated some. There were around 260,000 different objects in the collection database and I would say we had images for 75% of them and about 95% of the more interesting ones. But we only had the best images for 45/50% of the objects. So the other ones would have been taken with big cameras or scanned in using professional photography.

Matt Mogan
President of Concrete Computing

We have a blend of collection types ranging from archival objects to museum objects, as well as documentation from the archives and a blend of subject matter such as history, art, science; a little bit of everything. The Smithsonian has 19 galleries, libraries, archives, museums, nine research centres and a zoo.

We have a feature at Smithsonian called Dashboard.si.ed that keeps running institutional statistics. We have over 138 million objects and specimens, over 153,000 archival cubic feet of collections at the Smithsonian and we have nearly two million museum objects and specimens represented by digital images. It's an ongoing process that will take time.

Meghan Ferriter
Project Coordinator at
The Smithsonian Transcription Center

No one knows exactly how many specimens we have collected, but my guess would be somewhere around 80 million. Just take a sub-group like entomology, the science of insects. It is the biggest in the world and immense on its own.

Adrian Hine
Data Manager at NHM

SHIFTING GOAL POSTS

On top of this backlog to address, institutions face the challenges of a changing landscape. With technology and consumer expectations rapidly changing, the goal posts for digitisation are constantly shifting. What constitutes good data and high quality imagery one year can easily fall below standards as new technology, higher-quality photographic equipment and 3D-imaging technology becomes available.

The mass adoption of mobile has seen a large shift in how consumers and potential audiences expect to consume information. By 2021 the average smartphone user is predicted to consume an average of 8.9GB of data per month, and changes like this in consumer expectations have an impact on the format, quality and searchability of data being digitised now.

Add to this new collections, specimens and objects being included over the years, and it is no wonder that collections managers can feel like they're on a never-ending conveyer belt of digitisation.

ENGAGING NEW AUDIENCES

Despite the challenges, engaging in the process of digitisation can create a significant opportunity to engage with audiences in new ways to increase the reach of your collections.

Kendal Museum, funded by the Heritage Lottery Fund (HLF), digitised its herbarium and two mineral collections as they have significant importance to Cumbria. Over 6,000 images have been taken to make these previously unseen collections accessible to everyone and volunteers and students were a key part of the digitisation process.

The Missouri Botanical Garden used crowdsourcing for a project called Art of Life, with the goal of increasing access to the natural history illustrations found within the Biodiversity Heritage Library (BHL). By 2015, over one million biodiversity images had been uploaded to Flickr from the books and journals of the BHL.

A lot of success that we've had with the crowdsourcing project is in bringing new audiences to BHL. People who didn't even know that we existed or what we did were finding out about us through these other platforms where we were doing the crowdsourcing and that brings an obvious benefit.

Trish Rose-Sandler
Missouri Botanical Garden

Unintended benefits are in some ways even more compelling – different engagement and the types of questions you get from the public. There's value in working with the public and breaking down the barriers. It matches many other forms of engagement with museums and archives in the industry, the idea of people wanting to be involved in learning. They want to pursue personal learning goals as part of their engagement with the museum.

Meghan Ferriter
**Smithsonian Transcription Center**

ENRICHING YOUR COLLECTIONS DATA

As well as creating new collections data records, enriching collection data is a key driver for digitisation. Only 10% of institution professionals said enriching data was not a priority, but there are variations in both how much this strategy is integrated into the overall institutional approach and the stage of maturity.

– OUR STRATEGY FOR ENRICHING COLLECTION DATA AT MY INSTITUTION IS... –
Respondents were asked to select the closest match

Still in development /to come

Distinct and related to the institutional strategy

Lying forgotten and ignored in some drawer

41% **37%** **26%** **23%** **10%** **5%**

Integrated into our institutional strategy

Referenced regularly by collections/documentation management and other practitioners on staff

Not a priority for the institution
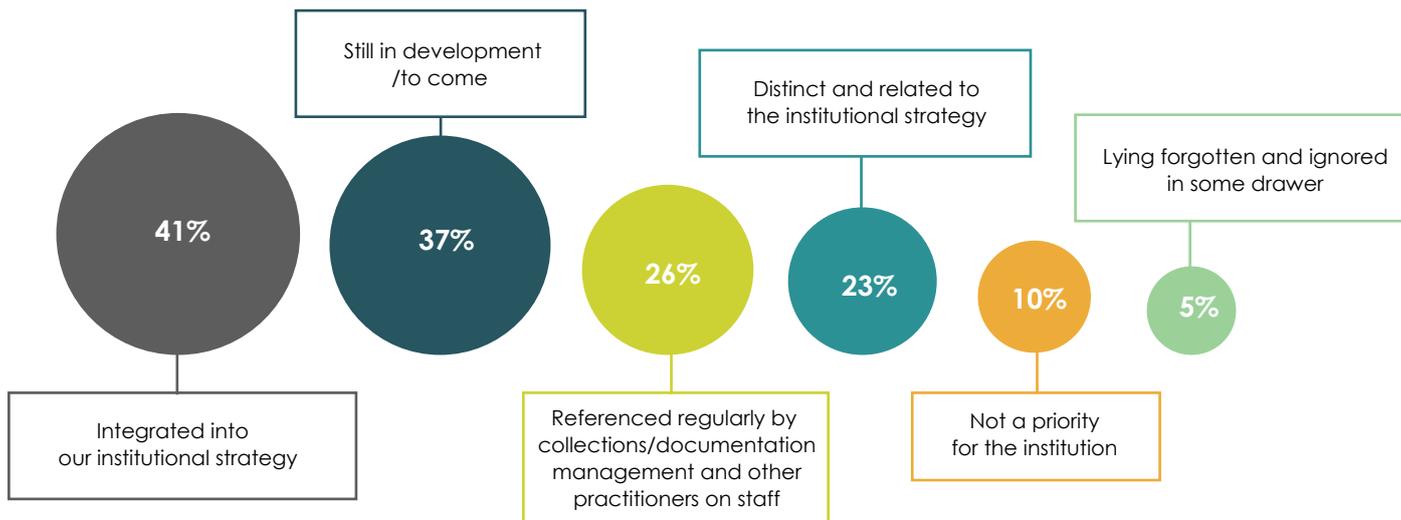
Of the respondents surveyed, 23% said that their strategy for enriching data collection at their institution is distinct, yet related to the institutional strategy. One of these respondents is Matt Morgan, President, Concrete Computing, who worked in a digitisation project at the Met, which has 60,000 art works on view at any given time.

"

At the Met I was involved in building a website that could house the entire collection online. Rather than data, it was the sheer amount of images coming from the photography department and the image library department that presented a challenge, with 99% of our collection having images. While some images were better quality than others, we took the approach to digitise everything. To fulfil on our digitisation strategy we felt that any image captured was better than none.

Matt Mogan, President of Concrete Computing

Section 2:
# Strategies for Digitisation

STRATEGIES FOR DIGITISATION

With such a monumental task before collections institutions, it is no wonder that the drive to accelerate digitisation has fostered much debate and experimentation in recent years.

With no one, clear solution to the digitisation challenge, and significantly varying needs between different institutions and collections types, it is no wonder that priorities vary widely from institution to institution:

– WHAT TACTICS ARE YOU USING / CONSIDERING USING TO INCREASE YOUR RATE OF DIGITISATION? –

	Leveraging volunteers for data entry	Crowdsourcing data	Conveyer belt rapid digitisation	Hiring additional staff for digitisation
Already doing this	70%	23%	10%	24%
Seriously considering this option	4%	15%	3%	11%
Considering this among other options	10%	22%	15%	21%
We have not yet considered this	9%	31%	51%	12%
We have decided against this	6%	10%	21%	33%
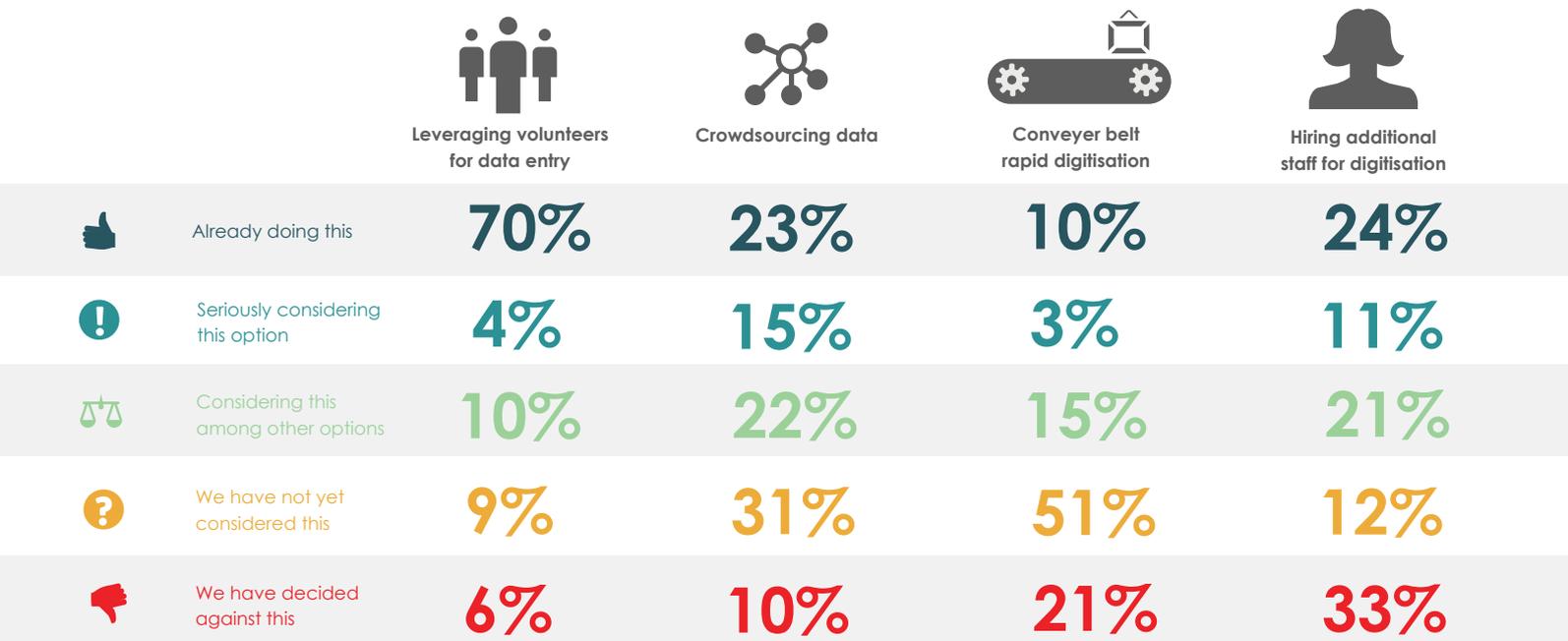
Using volunteers is a clear strategy, with 70% of institutions confirming they are already doing this and a mere 6% ruling it out as an option. Crowdsourcing, while not commonplace, is also seen as a clear opportunity, with 60% considering or already doing this. In contrast, only 24% of institutions have chosen to hire additional staff and less than 50% have not even considered conveyor belt rapid digitisation.

This is a likely an indication of the financial climate that museums and archives are currently working in and the associated cost of conveyer belt rapid digitisation and hiring additional staff.

Both leveraging volunteers and crowdsourcing offer museums and archives the opportunity to increase their digitisation rates without increasing their headcount or investing in expensive technology, so these strategies are often tempting to implement. This can be particularly appealing for smaller museums, where budgets can be tight.

Section 3:

Leveraging Volunteers & Crowdsourcing: What's Stopping You?

LEVERAGING VOLUNTEERS & CROWDSOURCING: WHAT'S STOPPING YOU?

Despite the obvious benefits of using non-paid staff to support digitisation, neither using volunteers or crowdsourcing for digitisation come without challenges - many of which cause real roadblocks for institutions looking to implement these types of strategies.

In this section of the report, we identify some of the key challenges that institutions face and provide examples of successful projects and key advice to help readers navigate the obstacles.

A. SOURCING VOLUNTEERS OR PARTICIPANTS FOR CROWDSOURCING

Over one-third of survey respondents state that the time investment needed to find volunteers was a challenge and 45% state that raising awareness of projects / sourcing contributors for crowdsourcing caused challenges.

THE CHALLENGE:

Sourcing contributors can be an ongoing task. Unlike paid staff there is no guarantee of how long each participant will continue to support your project and if participant recruitment isn't a constant ongoing process, projects can ebb and flow with the tide of participants or stop and stall altogether when people move on.

> *You are receiving people's labour for free if you're asking them to volunteer so there are responsibilities and things we need to think through around rewarding and acknowledging that participation and making it valuable for both the museum and participants. Offhand, 60% of peoples volunteering is for two weeks or less. If this type of transcription crowdsourcing should become more common, it will be a lot of effort to continue to recruit and more effort to sustain participation.*
>
> **Meghan Ferriter: Smithsonian Transcription Center**

WHAT'S WORKING FOR OTHERS

CREATE SPECIFIC AND TARGETED PARTICIPANT OPPORTUNITIES TO ATTRACT RELEVANT CANDIDATES

Volunteers can play a variety of roles at an institution, many of which require vastly different skills and personality traits just like paid staff. If you're looking to attract participants then create more targeted opportunities that list specific skills. Not only will you attract more volunteers, but they'll be better suited to the tasks that are set and also more engaged.

– WHAT ROLES DO VOLUNTEERS PERFORM AT YOUR INSTITUTION? –
Respondents were asked to select the closest match

59% Tour Guides

51% Conservation & Preservation

39% Data Cleaning

36% Other

59% Audience Engagement

81% Data Entry

48% Supporting Exhibitions
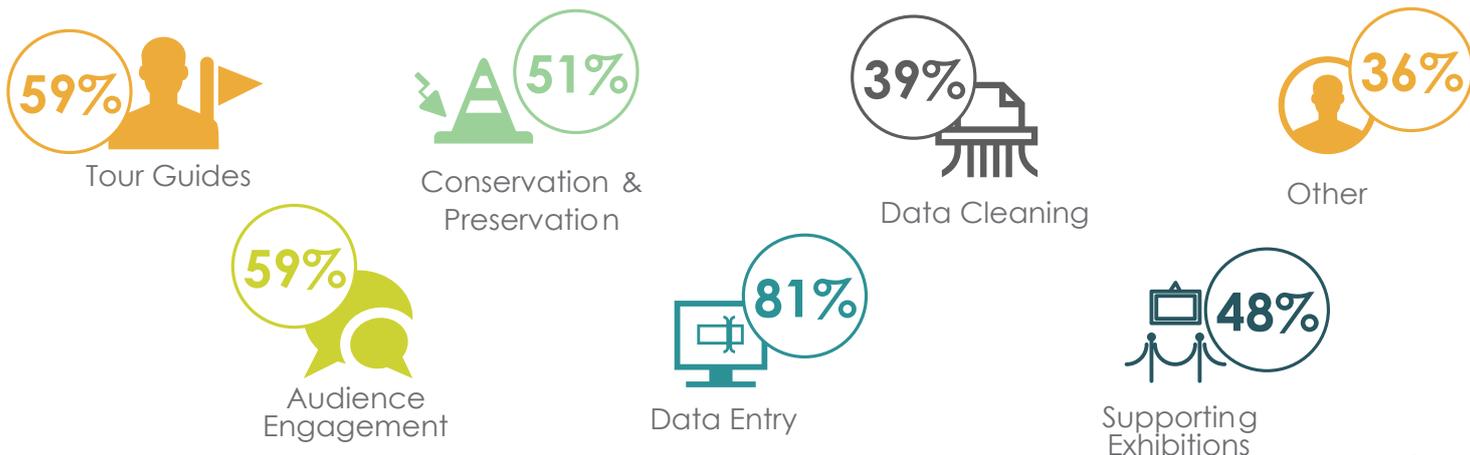
More than a thousand volunteers help the American Natural History Museum meets its goals by donating in excess of 115,000 hours per year to the museum. They have volunteers who work with the public by conducting tours, staffing information desks and even teaching other volunteers. However, many also help behind the scenes of the museum in the administrative and scientific departments, working alongside marketing, HR and IT.

The Harvard Museum of Natural History has a range of opportunities for volunteers, such as Nature Story Time Program Readers and Public Program Ambassadors, but does not allow volunteers to work behind the scenes with the research departments or curatorial staff.

Similarly, The Horniman Museum has non-specialist volunteering opportunities in collections management, but not in conservation roles, where they only use specialists.

The Royal Academy of Music has created a volunteer role profile specifically for digitisation. This role includes creating, editing and making accessible digital images from their museum collections to increase public access and awareness through the digital archive and online collections catalogue.

TARGETED RECRUITMENT

Recruitment can be a time-consuming process, especially if you're casting a very wide net. While it may sound counter-intuitive, narrowing your pool of potential recruits can actually deliver significantly more results if you focus all your energy on getting lots of engagement with smaller pools of more relevant candidates. Universities and educational institutions can be a good place to start, as well as relevant local adult groups whose interests match your objectives.

The South Western Federation of Museums and Art Galleries (SWFed) looked to streamline the process of recruitment and training by identifying roles and activities that people with little to no experience could do, such as data entry, basic research and visitor surveys.

SWFed also targets its recruitment drive to academic institutions, offering volunteering opportunities to the Guild of Students at Exeter University. This includes those studying archaeology but also all students who may have an interest in helping the institution. Interestingly, one of the important areas for training was in supporting existing staff to work with an influx of volunteers.

Avoncroft Museum recruits volunteers to 'help bring the past to life'. They dress in historicaly-accurate costumes and are given training on how to bring the stories behind the museum's collections to life through interaction with visitors and historical re-enactments.

The Street Art Museum in Amsterdam is a small not-for-profit that only has one full time employee and a few interns. The museum has been working closely with the local Reinwardt Academy (Academy for Museology) since 2015 and the academy has not only been the source of many of the mueum's volunteers and interns, but also provided invaluable recommendations and expertise to the start-up on its journey towards becoming an official museum and achieving ANBI status.

LOWER THE BARRIERS TO ENTRY

The easier you can make it for people to get started and contribute, the more response you'll get from the strategies above. Have a look through things like your website, your volunteer descriptions and your expectations for your participants. Are things easy to find and easy to understand? Are you forcing them to call in and speak to someone or come into the building? Can you make some easy enquiry forms online to make it simpler? In essence, are you making them do all the work?

Ultimately, they're providing you with their time for free despite having other priorities and interests in their lives. Making the process of getting started as easy as possible will reduce the number of participants who drop off throughout the recruitment process.

SMITHSONIAN TRANSCRIPTION CENTER

In 2014, the Smithsonian launched its Transcription Center, a web-site designed to leverage volunteers and crowdsourcing to help the institution unlock the content inside thousands of digitised images of documents. These included handwritten Civil War journals, personal letters from famous artists, 100-year-old botany specimen labels and examples of early American currency.

Meghan Ferriter, Project Coordinator at The Smithsonian Transcription Center, comments:

ON THEIR WEBSITE – GET STARTED EASILY:

JOIN US!

LEARN HOW TO TRANSCRIBE

Become a Smithsonian Digital Volunteer and help us make historical documents and biodiversity data more accessble.

Join 7,218 volunteers to add more to the total 218,848 pages of field notes, diaries, ledgers, logbooks, currency proof sheets, photo albums, manuscripts, biodiversity specimens labels that have been collaboratively transcribed and reviewed since June 2013 - Get started Now!

Follow us on Twitter and learn more about projects: @TranscribeSI

Partner on projects and ask your #volunpeers for best tips and tricks.

You'll also find updates on Facebook and behind-the-scenes shots on Instagram.

Our digitisation strategy is an ongoing process that will take time, but by combining more than 7,000 digital volunteers with 120 volunteers working in our transcription center we are making significant progress.

To make a project using volunteers a success we have realised the importance of making sure the barrier to entry is really low and that it's easy for volunteers to get started. We have had different projects that required existing knowledge of the objects, so providing as much guidance as possible and being readily available to respond to any volunteer questions is important.

B. OVERCOMING EQUIPMENT AND TECHNOLOGY BARRIERS

Nearly half of respondents said that equipment and technology in their institutions was a barrier to implementing volunteering and crowdsourcing strategies, with 47% saying they struggled to provide enough PCs with access to their collections management systems to support both volunteers and staff simultaneously and 48% unable to allow volunteers to carry out data entry remotely.

THE CHALLENGE:

Institutions often only have limited space, desks and PCs in their buildings, and much of this is taken up by existing staff. The physical question of "where do I put these participants" is one that can impede your ability to leverage large numbers of volunteers and participants. If your volunteers are specifically working with data entry into your collections management system this can be exacerbated, as the number of PCs with installed hardware can be even fewer.

Institutions that require participants to be based on location also risk significantly lowering their pool of potential candidates as geography, travel and the associated expense become factors for the participants.

WHAT'S WORKING FOR OTHERS

ENABLE REMOTE DATA ENTRY ONLINE

Enabling data entry to be performed remotely not only significantly reduces the demand for equipment and hardware in your institution, but also allows you to source participants from anywhere around the globe. It can also offer a greater degree of flexibility for your staff to enter data from home or from other sites and locations - or even just from other areas of the building such as storage rooms or exhibitions.

You can implement technology that allows both participants and staff to enter and review data through a web browser direct into your CMS. This means that participants can use their own devices such as PCs, laptops, tablets and smartphones to support your digitisation efforts.

To help in digitising and cataloguing its assets, The San Diego National History Museum's Research Library has been working with volunteers and interns over the past nine years. The institution continues to digitise its large collection of historic images, photos, field notes and diaries that recount research expeditions, collecting trips and conservation work, to make them available online.

A related project was launched in 2013 by UC Berkeley's Essig Museum of Entomology. Code-named Calbug, any volunteer with an internet connection could help to transcribe field notes on around a million bugs and insects. The specimens came from California's eight major entomological collections, which included the San Diego Natural History Museum and the California State Collection of Arthropods.

"

We created algorithms to help us find out which pages had illustrations and then we would push them out to platforms like Flickr for the public to help us describe them. Volunteers didn't have to be onsite and could easily help us from home. Our literature goes back to the 1400s, and using the public helped us to describe images that were not previously searchable.

Trish Rose-Sandler
Data Analyst
Missouri Botanical Garden

The Street Art Museum in Amsterdam chose web-based collections management, Axiell Collections. **Anna Stolyarova Museum Director of the Street Art Museum Amsterdam** says:

"

We are a small, non-profit museum and we can't afford to have our own in house IT support like a big museum can. We only have one full time employee and a few interns and if anything happens to our system or our equipment everything is safely saved in the cloud and can be accessed from another computer.

"

It's a huge benefit that we can work remotely. We are not dependent on the hardware that is stored in the museum. I can also give other people access, such as volunteers and students of the Reinwardt Academy, who will help out with upcoming collection management tasks. One intern recently had an accident and could only work from home. With Axiell Collections she can still do her job thanks to the online service.

"

Axiell Collections is very agile and turns a rigid process such as entering data into a flexible job that can be done from any computer as long as you have the login credentials.

ENABLE REMOTE ACCESS TO MATERIALS USING ONLINE TOOLS

It's not only important to allow data entry online, but if you have large stores of digital material already that you need to capture data about, it's important to make these as accessible as possible.

The Smithsonian believes that transcription by humans is the only way to make the text of its items searchable, which will open them up for endless opportunities for research and discovery. Many of the digitised documents in its collections are handwritten or have text that computers cannot easily decipher. This is exactly what its Transcription Center was designed for.

There are 7,082 participants mentioned on the website that are specifically digital volunteers, which represents anyone who has contributed any transcription ever to the Smithsonian Museum. That is distinct from approximately 6,300 onsite volunteers at the Smithsonian who may or may not overlap with the digitisation process.

Participants can view text and transcribe data online using their full service, user friendly platform.

C. TRAINING & MANAGING
PARTICIPANTS / VOLUNTEERS & VOLUNTEERS EFFICIENTLY

A huge 85% of survey participants said that the time investment required to manage and support volunteers was a challenge and a massive 75% cited that the time spent giving new users training on systems caused challenges for implementing effective volunteer strategies – making managing and training voluntters the single biggest concern.

THE CHALLENGE:

With a high turnover of participants, often with little to no experience, training and management will inevitably be an issue for institutions. When you factor in that collections management is a highly-skilled profession, which often draws on very sophisticated collections management technology, the time to competence for a new starter can be prohibitive.

Once someone acquires enough skill to largely self-manage, they're likely to move on and the whole process starts again. This puts pressure on full time, skilled staff and often leaves them feeling like they'd be quicker doing a job themselves than using volunteers.

WHAT WORKS

PREPARE, AUTOMATE AND SYSTEMISE

If you're using volunteers and crowdsourcing for digitisation, the chances are that much of the work that participants are doing is actually quite simple and highly repetitive. What takes your staff time is repeating the same training, providing the same information and answering the same questions over and over again.

But if you put in the time early to automate and document the most typical parts of the process, then participants can will be significantly less reliant on skilled staff to support them.

CREATE SIMPLE DATA-ENTRY WORKFLOWS

While your collections team may be making use of all the advanced functionality of your collections management system, a wealth of features and choices can be very overwhelming for an inexperienced user.

If you have the capability to customise or adapt your CMS you can create very prescriptive workflows for data entry. The key to getting this right is creating a fool-proof, step-by-step process thatrequires the least amount of decision making from the user.

GOLDEN RULES

1. Think about a process that will work seamlessly for 80-90% of cases and work to that standard,then provide help for the cases that don't fit the mold separately

2. Create a step-by-step workflow that participants can follow over and over again

3. Don't make them guess what to do next – end each step with one clear call to action

4. Hide any functions they don't need to use – these can be overwhelming

5. Think about mandatory fields and alerts – make it very difficult (if not impossible) to enter data incorrectly

6. Where possible have a few shorter pages – rather than one very long one – to make information more digestible

7. Lock down the options – where possible use drop-downs or checkboxes rather than free text

8. Test with a non-user before you start and see if they can get through the process without intervention

REMOVE TECHNOLOGY SKILL BARRIERS

If you're expecting volunteers to work with technology, make sure the technology is accessible and usable to those without in depth technical skills. Ensure the systems you are using are intuitive and easy to grasp.

> One part of making sure that the barrier to entry is really low and that it's easy to get started is ensuring that the interface is providing as much guidance as possible, because you really can't expect people to read instructions.
>
> **Meghan Ferriter**
> **Smithsonian Transcription Center**

> Our museum doesn't have professional staff employees; we work with students and volunteers who rotate regularly.
>
> We have to share access to the system and Axiell Collections supports this and makes it easier for us to manage our collection. A special demo account was set up and that helped me to better understand Axiell Collections and how it looks. You can play with it and it's not that hard to learn. I thought it would be very difficult to start working with, but it's totally the opposite.
>
> **Anna Stolyarova**
> **Museum Director of the Street Art Museum**
> **Amsterdam**

> Providing adequate training is a challenge, and also making sure that you are able to respond within a reasonable amount of time to questions. For this reason the Smithsonian Transcription Center project is available online, 24 hours a day. People usually want answers to their questions while they are working on it, so we try to follow up as soon as possible. We also take the kinds of questions they are asking and look at ways to revise our instructions so we how can we better support and communicate with volunteers.
>
> **Meghan Ferriter**
> **Smithsonian Transcription Center**

D. ENSURING ACCURACY OF DATA / RECONCILING DATA

It is interesting to see from our research that data is a key area where collections institutions are making use of volunteers, notably in data entry (81%) and data cleaning (39%). Many respondents also cited a number of other data-related tasks, including digital imaging, collection photography for digitisation and mass digitisation.

However, the quality of that data is a concern for those working with both volunteers and crowd-sourcing. The biggest challenge that respondents listed as a challenge for crowdsourcing was ensuring the accuracy and quality of the data captured – with 75% agreeing that this was a challenge.

THE CHALLENGE:

Allowing non-qualified participants to input data direct into the CMS can be a terrifying thought and this could be rife with inaccuracies. If you have followed the advice in section C effectively then the chances are this problem will be lessened, but it is always necessary to have a system of checks in place to protect the integrity of your collections data.

WHAT'S WORKING FOR OTHERS

STANDARDISE VOCABULARY

We've found that providing a controlled vocabulary list improves the quality of the data you receive and we've now put in adequate review processes. Our peer review means anyone can get started with transcribing, but only those who have created accounts can review someone else's work before it is passed to the Smithsonian staff for final approval.

Meghan Ferriter
Project Coordinator at the Smithsonian Transcription Center

Most museums in Canada use the CHIN (Canadian Heritage Information Network) data dictionary as their metadata standard. Since the 1970s, many museums across Canada used a common collections management system based on this data dictionary and that is why there is much commonality in metadata across Canada.

HOW TO KEEP YOUR DATA CLEAN USING YOUR COLLECTIONS MANAGEMENT SYSTEM

- **Minimise opportunities to enter dirty data**
  - Consider suppressing the display of any unused fields. Fewer on-screen fields translate to less opportunity to create dirty data

- **Personalise help messages and documentation**
  - Adjust your help messages to reflect the styles and practices in use at your organisation

- **Control vocabulary**
  - Configure pick lists to dynamically display terms based on a value entered in a related field

- **Use established standards**
  - Standardise vocabularies as an authority for negotiating terminology

- **Be vigilant**
  - Consider running vocabulary lists on key fields on a quarterly basis, auditing a percentage of records at regular intervals and nominating records for peer review

Read more advice on how to clean your data at http://alm.axiell.com/making-clean-get-away

Trish Rose-Sandler, Data Analyst at Missouri Botanical Garden, comments:

“

When you get a lot of metadata from the public, and you have to assume a lot of them are not experts in the area that they are adding information to, you have to proceed with caution. We do try to do some analysis of what's coming in assess the quality. People who have come to our sites are pretty committed to the idea of research and libraries and archives and what we are trying to accomplish, so we haven't really seen a lot of input we felt we couldn't take in.

“

It's interesting because in some cases we definitely want to have a consensus on the metadata. So with the OCR we were looking for a certain amount of agreement on what a word actually was, because then we could have a certain level of confidence that a word was not accurate in the system. But with something like an illustration everybody kind of brings their own unique perspective to what that image is and often the metadata is not going to be the same. In the context of images, I believe that's ok because it becomes an access point for other users. Having that variety of metadata is desirable, whereas you wouldn't want a variety of responses when you're trying to transcribe text.

Top tips for using volunteers and crowdsourcing to meet digitisation goals

The research shows a clear desire among collections institutions to progress digitisation strategies and enrich their collection data. The majority are already making use of volunteers to achieve this and, while crowdsourcing strategies are not as advanced, most institutions understand the potential opportunity and are considering this approach.

Follow these summary guidelines to ensure your institution can accelerate its digitisation approach and make the most of volunteer and crowdsourcing opportunities:

1. Set a strategy and stick to it

Trying to implement a broad strategy to generally use volunteers or crowdsourcing in your organisation rarely takes off. The institutions that see the most success use volunteers and crowdsourcing to achieve specific aims and objectives and base their strategies around those specific projects. Before you start implementing your plan make sure you are clear about the results you want to see and tailor your approach to meet those goals.

2. Find the right pool of volunteers

Seek relationships with groups or organisations who can provide you with a steady pool of relevant volunteers. Local universities with relevant courses or local interest groups who are interested in your topics are a good place to start. You'll always experience a certain amount of volunteer churn and so building an ongoing relationship prevents your project from stopping and starting when the initial volunteers drop off. Identifying this group should be easy if you have set clear objectives at step 1.

3. Allow web-based data entry

Allow volunteers and relevant groups to input data direct to your collection management system through a web browser. This approach will allow you to leverage a wider pool of volunteers as they can work remotely, there is less duplication of work for your staff and there is less demand on providing hardware within your organisation.

4. Make it easy to get started

It is not unlikely for a museum or archive to have a high turnover of volunteers. Having an easy to use systems and easy to learn processes will significantly reduce the level of training required by your staff and enable any new user to start work quickly and effectively. Set up clean and clear workflows, ensure the technology is easy to understand, and make help resources easy and accessible.

5. Control the vocabulary and terminology

Where possible, provide simple drop down lists or defined terminology for users to reduce the chance of them inputting unclean or inconsistent data. Sites with multi-disciplinary collections can have shared fields that require very different terminology depending on what is being catalogued, so in this case consider forms that display different information for different collections types. Select a solution that gives you this type of granular control for accurate data management.

6. Personalise help messages and documentation

Tailor your help messages and documentation to offer field definitions and data entry examples that reflect the style and practices in your institution. Examples and help functions are always useful to those entering data for the first time and help to accelerate how quickly they can get up and running on your system.

7. Don't over complicate instructions

Rather than risk users digging through hundreds of detailed steps, put in place a concise summary with a limited number of short, clear instructions. Keeping things simple is often enough to help guide users through a task.

SOURCES

http://326gtd123dbk1xdkdm489u1q.wpengine.netdna-cdn.com/wp-content/uploads/2016/12/Edited_-_Digitisation_-_A_Simple_Guide_24.11.15.pdf

http://advisor.museumsandheritage.com/features/digitising-collections-breaking-museum-walls-opening-collections-world/

http://www.kendalmuseum.org.uk/whats-on/exhibitions-and-events/test/

http://hmnh.harvard.edu/volunteer

http://www.horniman.ac.uk/get_involved/volunteering-in-the-galleries

http://lynnblackadder.com/wp-content/uploads/2016/05/Working-with-volunteers-in-collections-care.pdf

https://www.ram.ac.uk/public/uploads/documents/60f4f0_volunteer-digitisation-role-profile.pdf

https://www.swfed.org.uk/case_studies/investing-in-volunteers/

http://www.avoncroft.org.uk/wanted-volunteers-bring-life-avoncroft-museum/

http://alm.axiell.com/interview-anna-stolyarova-museum-director-street-art-museum-amsterdam

https://transcription.si.edu/

http://www.sdnhm.org/science/research-library/projects/digitization-projects/

http://canada.pch.gc.ca/eng/1443538179085

INTERVIEWS

Meghan Ferriter, Project Coordinator at The Smithsonian Transcription Center
(interviewed on 7 December 2016)

Matt Mogan, President of Concrete Computing
(interviewed on 7 December 2016)

Trish Rose-Sandler, Missouri Botanical Garden
(interviewed on 8 December 2016)

CONNECT WITH US

Thank you for downloading this report and taking the time to it.
We sincerely hope that you nd it interesting. **Please feel free to share it!**
If you are interested in museum content and collection management
specifcally, then please connect with us:

 facebook.com/axiellalm   @axiellalm   Axiell ALM

 sales-alm@axiell.com

 www.alm.axiell.com

THANKS

to MCN for collaborating with us on this report.
We sincerely appreciate the great cooperation.

mcn

AXIELL
ARCHIVES LIBRARIES MUSEUMS