HARNESSING THE POWER OF NATURAL SCIENCE COLLECTIONS

A BLUEPRINT FOR THE UK

DISSC UK

SUMMARY

Harnessing the full potential of the UK's natural science collections, by bringing them together in a distributed, interoperable research infrastructure, will make them physically and digitally open, accessible, and usable for all forms of research and innovation.

Between September 2021 and March 2022, a consortium across the twelve major regions of the UK, led by the Natural History Museum, London (NHM), participated in a study to develop the business case and plan to support a national programme of natural science collections digitisation. This work, funded by the Arts and Humanities Research Council (AHRC) sought to understand the size and diversity of UK natural science collections; establish the readiness of UK institutions to undertake digitisation; and develop training materials to pilot regional digitisation, building on prior investments in digitisation and informatics by the NHM.

This report sets out the key findings from this work and is presented as a "blueprint" to scale up digitisation and cooperation through a national infrastructure programme: DiSSCo UK is part of the UK's contribution to DiSSCo, the Distributed System of Scientific Collections, a partnership of institutions who share a vision to empower science and society, in balance with nature, to make reliable knowledge and evidence about the natural world available to all.

DISSCO UK WILL:

EMPOWER THE UK NETWORK OF COLLECTIONS THROUGH DIGITISATION

Ensure that UK institutions are equipped and supported to digitise, share and improve their collections, regardless of size or location.

ENHANCE UK BIODIVERSITY AND HUMANITIES INFORMATION INFRASTRUCTURE

Provide leadership, expertise and tools to support the integration of UK natural science collections information as an interconnected digital knowledgebase.

IMPROVE DATA QUALITY

Ensure that all data within the DiSSCo UK network are of the highest possible quality and associated with clear indicators, enabling users to assess data origin, relevance and usefulness for any application.

DELIVER RELEVANT DATA

Ensure that the DiSSCo UK network delivers data in the form and completeness required to meet the highest-priority needs of science and, through science, to wider society.

UK NATURAL SCIENCE COLLECTIONS

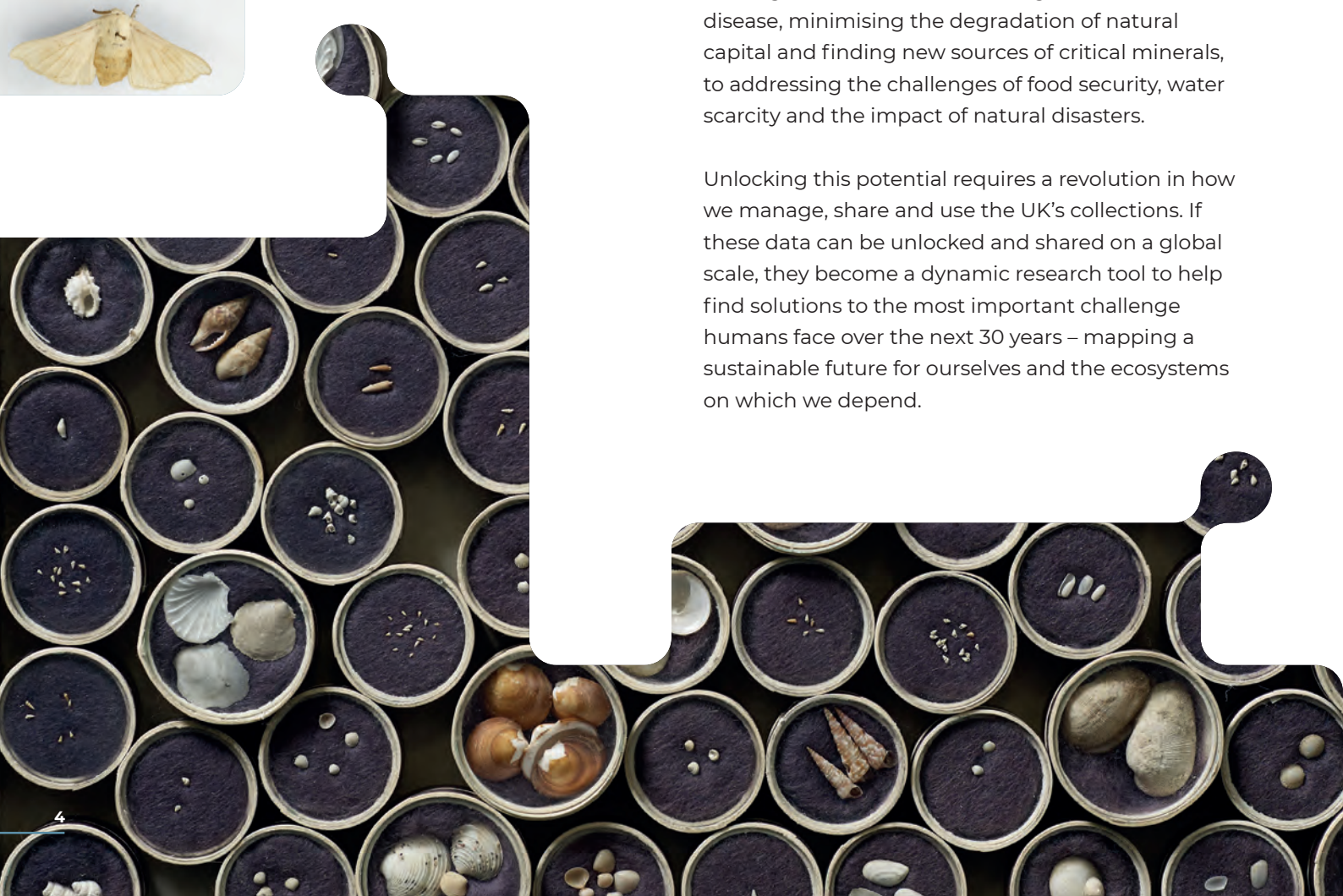
The UK holds some of the world's most important natural science collections, with more than 130 million specimens collected from across the globe and stored in more than 90 institutions throughout the country.

HARNESSING THE POWER OF UK COLLECTIONS

Developed over the last 250 years of human exploration, natural science collections tell us how the Earth and its natural systems formed over 4.56 billion years, and the impact of human life on the natural world over the past few thousand years. They are the gateway to inspiring a lifelong love of nature, the source of educational material for generations of schoolchildren, and a treasure trove of scientific information that underpins most of our understanding of the natural world.

Biodiversity provides the support systems for all life on Earth. Yet the natural world is in peril, and we face biodiversity and climate emergencies. Solutions to these problems can be found in data associated with UK natural science collections, from slowing extinction rates, reducing risk from zoonotic disease, minimising the degradation of natural capital and finding new sources of critical minerals, to addressing the challenges of food security, water scarcity and the impact of natural disasters.

Unlocking this potential requires a revolution in how we manage, share and use the UK's collections. If these data can be unlocked and shared on a global scale, they become a dynamic research tool to help find solutions to the most important challenge humans face over the next 30 years – mapping a sustainable future for ourselves and the ecosystems on which we depend.



WHAT ARE NATURAL SCIENCE COLLECTIONS?

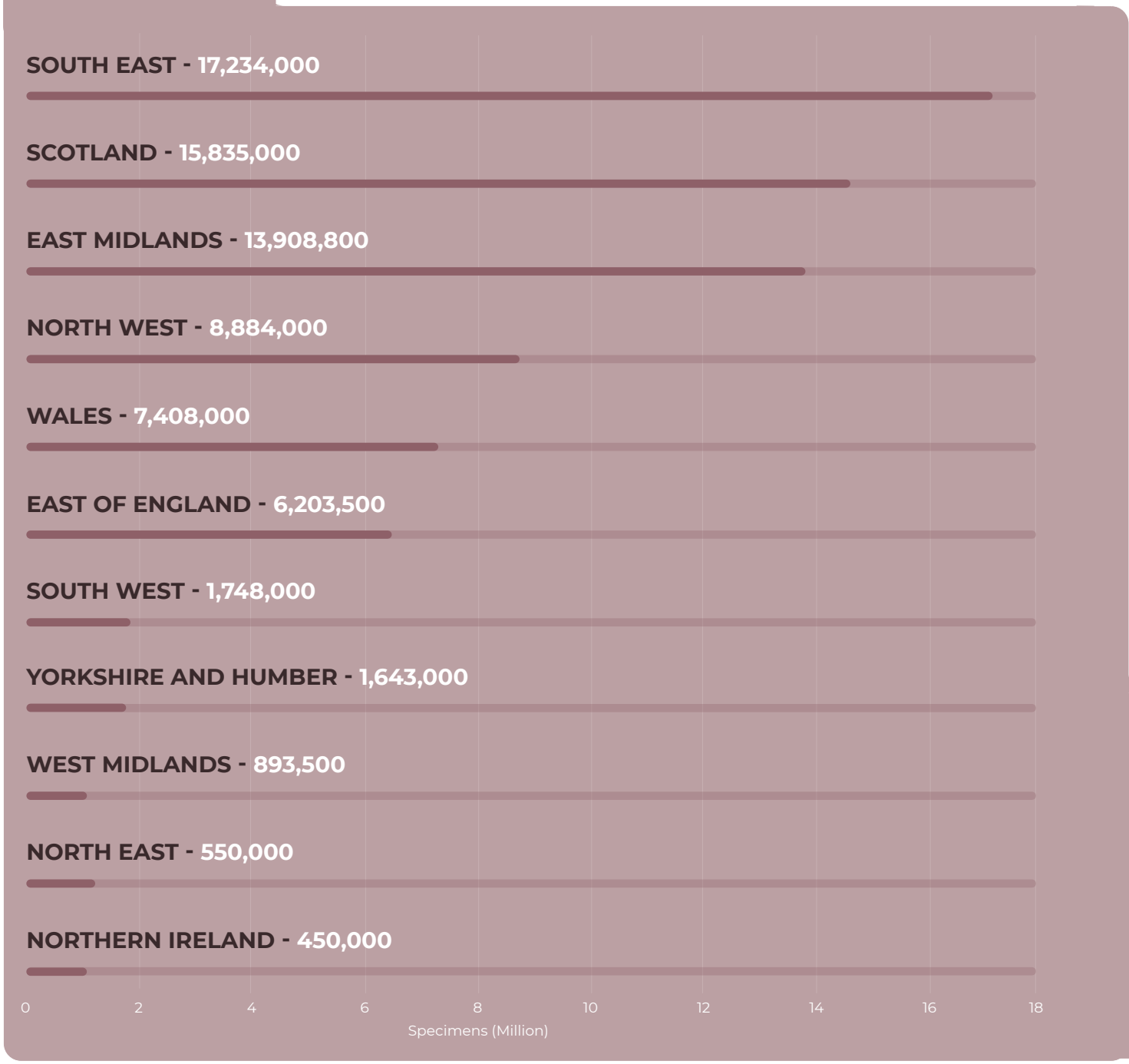
These consist of both organisms such as plants and insects; their associated biological material, such as preserved tissue and DNA; and geological material, including meteorites, rocks, minerals and fossils. All are usually accompanied by extensive digital and analogue publications and notes. Each specimen harbours many kinds of data and can be cared for by a range of organisations, most often museums or botanic gardens.

Collections contain diverse physical specimens, including seed banks, geological cores, living, frozen, fossilised or dried tissue collections and DNA banks, including material from endangered or extinct species from some of the most inaccessible parts of the Earth, and in some instances objects that originate from outer space.

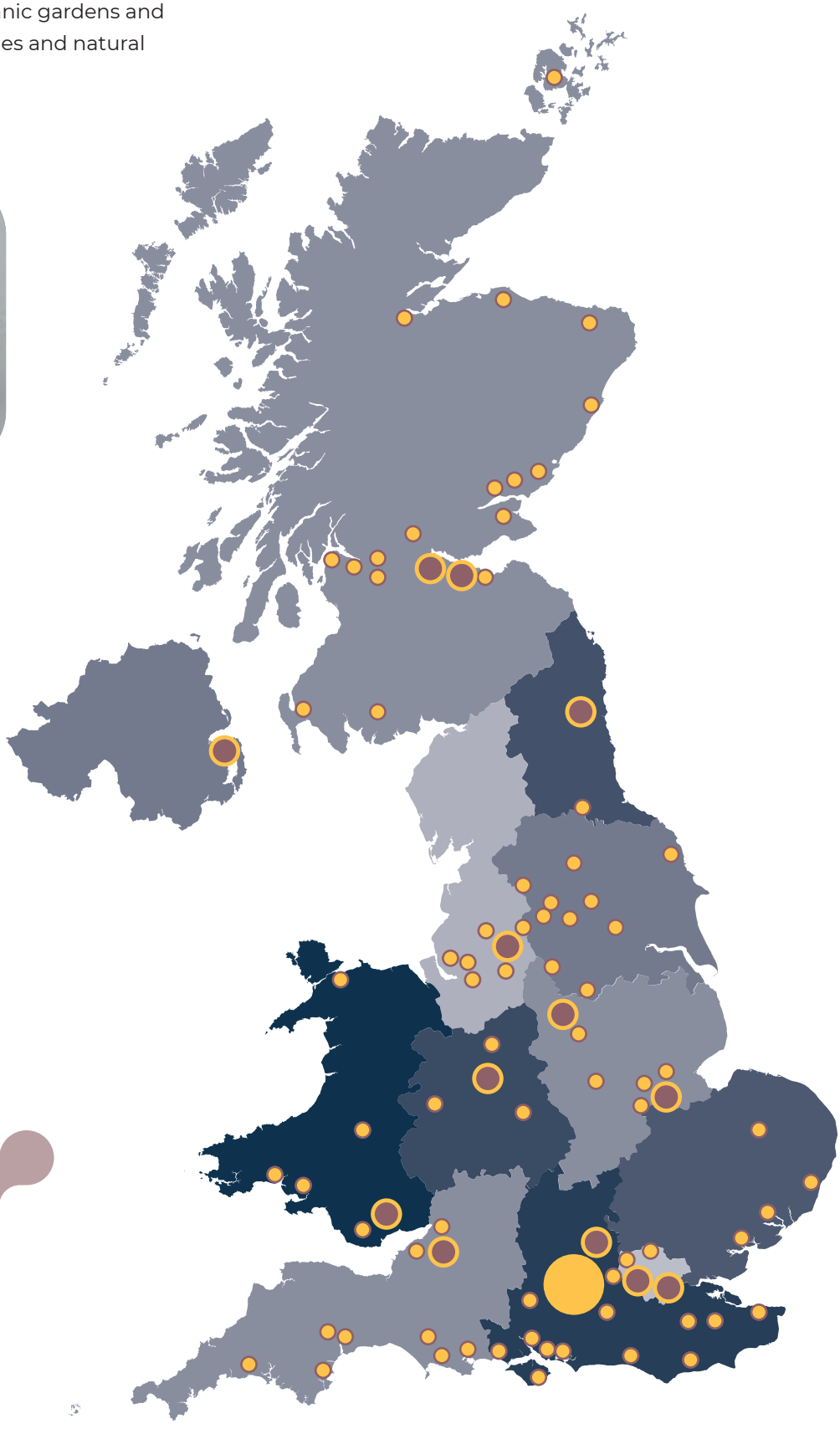
UK COLLECTIONS BY THE NUMBERS

There are an estimated 650 natural science collections in the UK and 94 of these contributed to our survey to understand the size, diversity, origin and current levels of digitisation.

GREATER LONDON
80,313,000
SPECIMENS



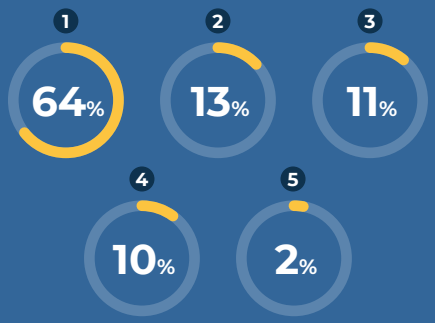
The UK collections community is a patchwork of smaller and larger institutions, spread between museums, universities, botanic gardens and herbaria, research repositories and natural history societies.



KEY

- Central hub
- Regional hubs
- Nodes/collections

TOTAL SPECIMENS SURVEYED



137,800,000 IN 83 INSTITUTIONS

COLLECTIONS PERCENTAGES THAT ARE:

MIDS refers to the Minimum Information about a Digital Specimen, a data standard describing different levels of digitisation.

1. **Undigitised:**
   There is no digital record of the specimen.
2. **MIDS-0:**
   A bare record of a specimen exists that reflects the early digitisation process of creating a database record that is useful in highlighting the existence of the specimen in a digital format.
3. **MIDS-1:**
   A digital record exists with basic information attached and a unique identifier, enabling all information to be unambiguously associated with a physical specimen.
4. **MIDS-2:**
   The specimen's digital record contains key information essential for most scientific purposes, including scientific name and geographic coordinates.
5. **MIDS-3:**
   The record contains further information known about the specimen, such as links to third-party sources and field notes.

59%
of the collections are invertebrates.

12%
are palaeontological.

4%
belong to the other disciplines (anthropology, extraterrestrial, microorganisms, zoology vertebrates).

16%
are botanical.

9%
are geological.

JUST **23%**
have any form of digital record.

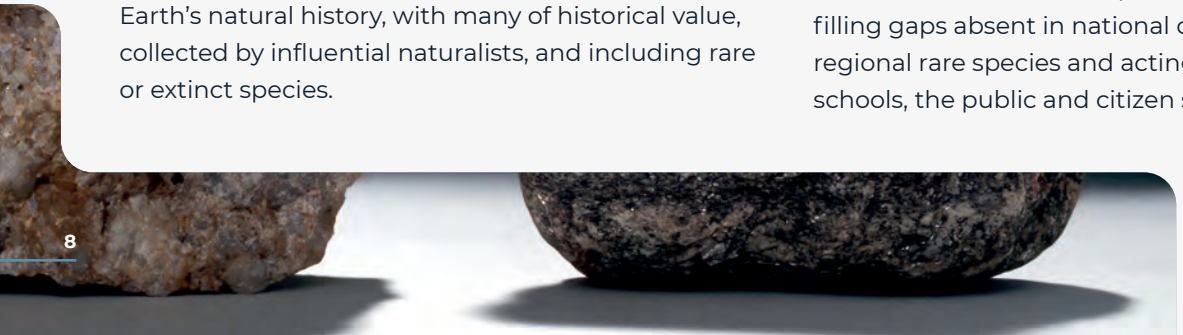
JUST **2%**
are considered research-ready.

KEY POINTS

The UK holds more than 137 million specimens recorded from 83 institutions and collected from across the world, providing a unique insight into Earth's natural history, with many of historical value, collected by influential naturalists, and including rare or extinct species.

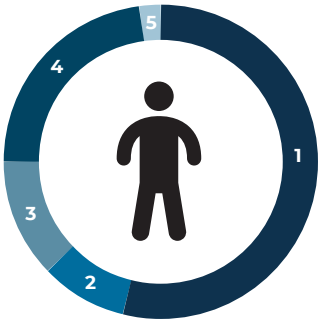
The vast majority of these specimens lack any digital record, effectively rendering them invisible to scientists and the wider public.

Local, smaller collections provide critical insight, filling gaps absent in national collections, housing regional rare species and acting as a bridge to schools, the public and citizen scientists.



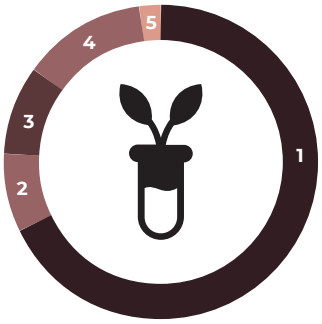
WHAT DO UK COLLECTIONS HOLD?

ANTHROPOLOGY



SPECIMEN COUNT **580,000**

BOTANY



SPECIMEN COUNT **21,586,000**

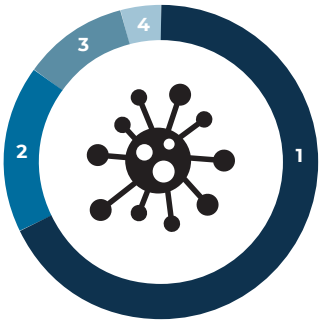
EXTRATERRESTRIAL

SPECIMEN COUNT **13,000**

GEOLOGY

SPECIMEN COUNT **12,697,000**

MICROORGANISMS



SPECIMEN COUNT **40,000**

PALAEONTOLOGY

SPECIMEN COUNT **16,673,000**

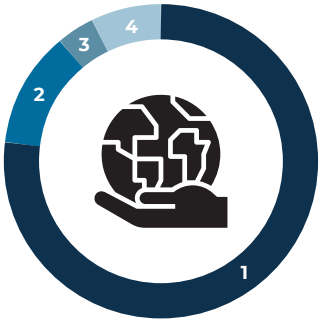
ZOOLOGY INVERTEBRATES

SPECIMEN COUNT **81,452,000**

ZOOLOGY VERTEBRATES

SPECIMEN COUNT **4,638,000**

OTHER GEO/BIODIVERSITY



SPECIMEN COUNT **92,000**

THE IMPORTANCE OF LOCAL COLLECTIONS

Local, often smaller, collections contribute unique specimens that bridge critical gaps in our taxonomic, geographic, and temporal understanding of global bio- and geodiversity. The impact of these collections often combines global significance with local embeddedness and activation, building on strong and equitable relationships with source communities and geographies. These collections are uniquely placed to connect with schools, volunteers and citizen science programmes, building a more creative and resilient natural science collection community, as well as enhancing our collective research capacity. They often hold multiple types of collections, making them well-placed to see inter-disciplinary potential and convene scientists to work with others, widening impact potential.

NHM SCIENCE & DIGITISATION CENTRE

The Natural History Museum, London, is developing a new research and digitisation centre outside of London, which will use natural science collections to deliver innovative solutions to global challenges. This centre will focus on expanding the museum's digital outputs through a dedicated digitisation centre, developing new technologies that accelerate the digitisation process. When complete in 2026, this new facility will act as the central hub for DiSSCo UK, working with the regional hubs to share digitisation resources and expertise and widen access to vital collections data.

RECENT UK RESEARCH ON COLLECTIONS INCLUDE:

1. The development of fungal-based biological solutions to enhance crop yields without the need for harmful chemical pesticides or fertilisers.

2. Work to catalogue global historic bat specimens and make them available for genomic studies to understand the origins of Covid-19 and its jump from animals to humans.

3. Research on lithium extraction to seek lower cost and more secure UK sources for the minerals needed to support the transition to a low carbon economy.

A VISION FOR THE NEXT DECADE

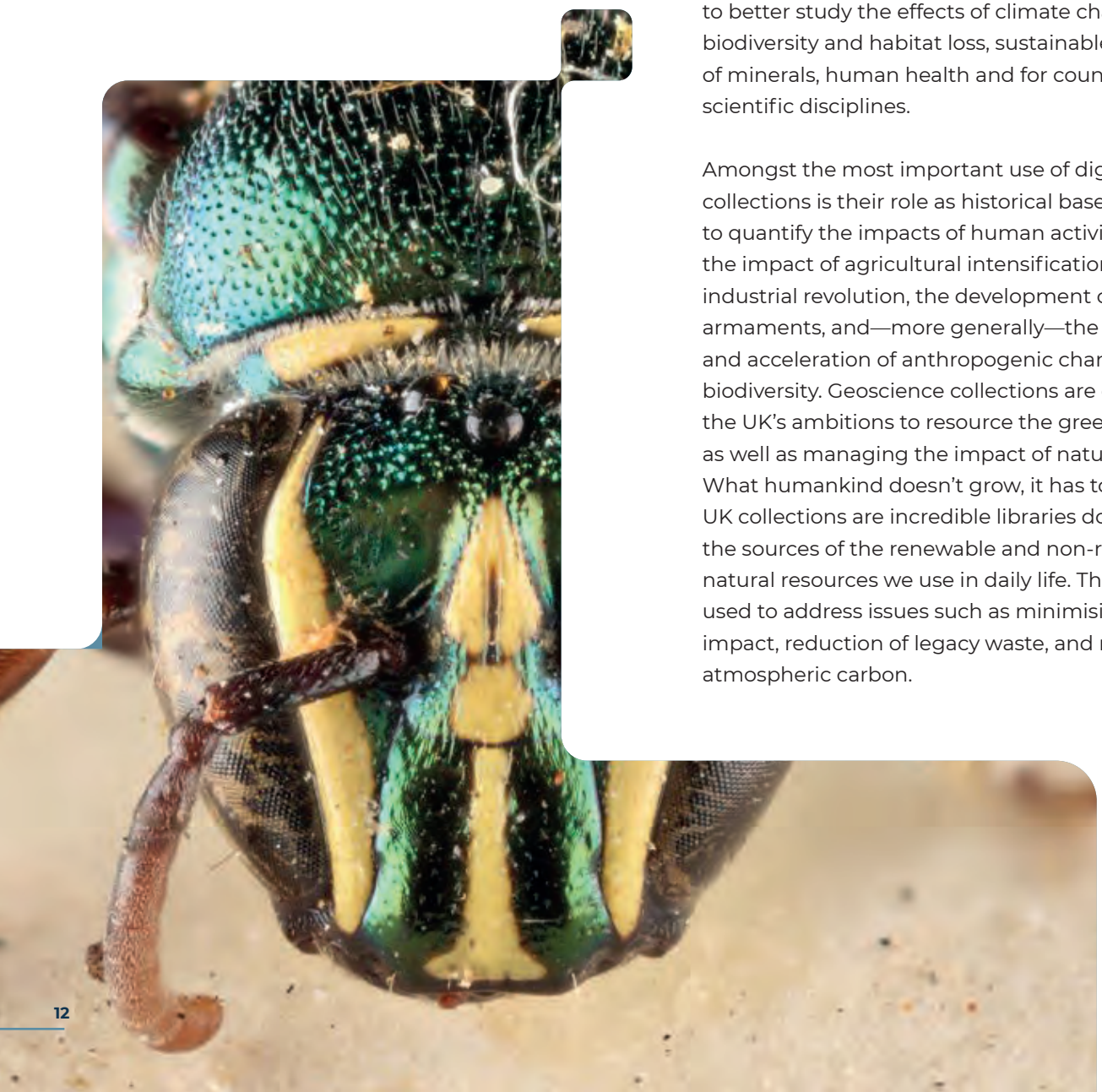
Our overarching objective is the compilation, integration and release of knowledge about the UK's natural science collections. This covers all species through the digitisation of the UK's unique collections of plants, animals, microbes, fungi, geology and meteorite collections.

The UK network is exceptionally well-placed and driven to achieve this, adopting a hub and spoke model where regional collections build their capabilities to act as centres of excellence for local, often smaller, collections. This network will be connected to the Natural History Museum, London's new Science and Digitisation Centre, scheduled to open in 2026.

UK collections are global in extent and encompass the period of time over which humans have transformed life on Earth. They provide perspectives on biodiversity change that are unrivalled by contemporary data, while our global breadth from ancient to modern helps us understand how and why biodiversity is changing and provides insights that will help us design a more sustainable future for people and nature. Smaller collections provide critical local insights and historical connections to the region, while larger national collections provide the capabilities and digital infrastructure to synthesise and integrate this information.

BENEFITS OF UK NATURAL SCIENCE COLLECTIONS

The UK's specimens are a hugely important scientific resource and infrastructure, as well as a major public and cultural experience and a source of economic and societal benefits.



ADVANCING SCIENCE, INNOVATION AND UK POLICY

A vast data-rich scientific repository
UK collections are amongst the world's richest in terms of their taxonomic, geographic and temporal diversity, reflecting the long and complex history of the United Kingdom. This treasure trove of collections, a priceless and unique scientific resource, will be unlocked through digitisation to better study the effects of climate change, biodiversity and habitat loss, sustainable sources of minerals, human health and for countless other scientific disciplines.

Amongst the most important use of digitised collections is their role as historical baselines to quantify the impacts of human activity: the impact of agricultural intensification, the industrial revolution, the development of nuclear armaments, and—more generally—the influence and acceleration of anthropogenic change on biodiversity. Geoscience collections are critical to the UK's ambitions to resource the green economy as well as managing the impact of natural hazards. What humankind doesn't grow, it has to mine, and UK collections are incredible libraries documenting the sources of the renewable and non-renewable natural resources we use in daily life. They can be used to address issues such as minimising mining impact, reduction of legacy waste, and reducing atmospheric carbon.

Supporting humanities research

The discovery and study of items in the UK's natural science collections has directly shaped our culture, our view of ourselves and our place in the world, adding social, aesthetic and historical value to the public understanding of science. Many items are part of the nation's cultural heritage as iconic, historically significant collections in their own right, such as those of Mary Anning, Charles Darwin and the polar explorers Scott and Shackleton. DiSSCo UK provides an opportunity to digitally aggregate these collections to support humanities research in ways that would be physically impossible.

DiSSCo UK will also play a critical and culturally sensitive role in addressing issues of inclusivity and relevance. Digital collections provide a unique opportunity for society to add content, interpretation, and representation around collections. With many specimens not yet sampled, documented, or investigated, digitisation can provide an opportunity to restore this "lost" heritage to communities, improving knowledge about them and boosting the integrity of the systems and institutions responsible for looking after them.

Underpinning local, regional and national policymaking

Protecting nature and biodiversity, for a future in which both people and planet thrive, is not only a moral imperative but is increasingly part of policy and legal commitments by local, regional, national and international authorities. Natural science collection data inform a range of these commitments. These data are central to monitoring change in the environment, but also for testing predictions of future change, to ensure UK policies are 'nature-positive' and evidence-led. Restoration and enrichment of UK ecosystems at any scale depends on information about past and present-day location of species, much of which comes from data held in natural science collections and the communities of specialists and naturalists recording this information. This information is also critical to ensuring the UK is meeting its shared commitments under a number of international agreements and UN agencies.

Natural science collection data underpin UK international policy and obligations in relation to:

- The Convention on Biological Diversity (CBD)
- The Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES)
- The United Nations (UN) Sustainable Development Goals
- The UN Framework Convention on Climate Change (including the Paris Agreement and the Glasgow Climate Pact)
- The International Union for the Conservation of Nature (IUCN)
- The World Health Organization (WHO)
- The Food and Agricultural Organisation (FAO)
- The Convention on Migratory Species
- The Ramsar Convention on Wetlands
- The Convention on the International Trade in Endangered Species
- IOC-UNESCO IODE Ocean Biodiversity Information System (OBIS)

ADVANCING SCIENCE, INNOVATION AND UK POLICY

Lifelong learning

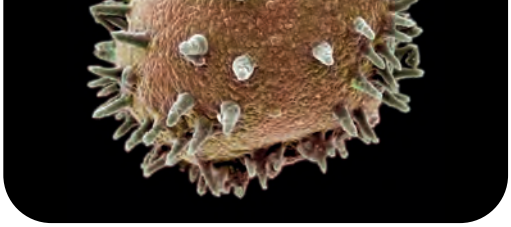
Managing natural science collections data requires skills and competencies that align with next-generation science standards for all age groups. This applies in both formal educational settings, such as science curricula, as well as informal learning opportunities. The digital data and specimens central to DiSSCo UK can be seamlessly incorporated into existing courses that include topics on evolution, biodiversity, systematics, conservation, climate change and ecology. Specimen-based data makes science accessible through the specimen itself, which is a tangible, place-based, engaging object, as well as through aggregated specimen data that is verifiable, relevant, and a logical gateway to data literacy.

DiSSCo UK cannot reach its full potential without the strong involvement of the citizen science community. The UK has a long history of amateur natural scientists, and there are many UK based citizen science projects already structured around monitoring biodiversity, such as the biological recording schemes and digital platforms like iSpot, iRecord and iNaturalist.

Related digital projects will play a key role in involving the public with DiSSCo UK through supporting involvement with collections-based science and databases. Contributions include transcription, taking measurements and annotating features such as flowering stages. UK collections, both small and large, are in a unique position to attract volunteers to help with the digitisation process, providing valuable training that is inclusive and engages participants from a wide range of ages, interests and backgrounds.

Growing UK collection capabilities

UK natural science collections in recent years have undergone some rationalisation, with some collections physically merged. Local collections are vulnerable to lack of support due to changes in institutional priorities. Most local and regional institutions lack a specialist natural science curator, or share a peripatetic curator. All involve curators in a range of tasks focused not only on collections management but also on visitor engagement, research, item loans, exhibitions and more. DiSSCo UK is an unparalleled opportunity to reinvest in UK collections and the expertise of the staff that care for them.



With collections becoming digital, specimens held at different sites can be managed online as a single entity via the national data portal, and searched in a unified way. Taxonomic expertise can be better coordinated, improving the resolution and spread of taxonomic skills in the UK. The voluntary sector, with its core of expert amateur naturalists, is an important repository of taxonomic expertise and by opening up UK collections digitally, DiSSCo UK has the potential to inspire a new generation of taxonomists amongst the volunteers who monitor changes in their local fauna and flora.

DiSSCo UK's technological innovation will create a powerful set of tools for UK researchers. While the state of collections in the UK is not susceptible to a "quick technological fix", technology is a driver that can be used to support the huge demand for taxonomic information created by biodiversity loss and global change, while also addressing the expectation amongst scientists and the public for information sources that are both accessible and easy to use.

THE ECONOMIC VALUE OF DIGITISING COLLECTIONS

In 2021, the Natural History Museum, London, collaborated with economic consultants Frontier Economics to explore the economic and societal value of digitising natural science collections. They concluded that digitisation has the potential to see a tenfold return on investment, creating benefits in excess of £2 billion over 30 years. The report examined the impact of collection data in five sectors: biodiversity conservation, invasive species, medicine discovery, agricultural research and development, and mineral exploration.

For more information, see https://doi.org/10.3897/rio.7.e78844.

BIODIVERSITY CONSERVATION

£0.7BN TO £1BN

Digitisation enhances taxonomic knowledge which improves detection of threatened species. This enables conservation efforts and maintains balance in the ecosystem.

INVASIVE SPECIES

£0.7BN TO £1.1BN

Digitisation enhances taxonomic knowledge which improves detection of invasive species. Reducing the frequency of losses from threats leads to significant economic benefits.

MEDICINES DISCOVERY

£0.8BN TO £2.8BN

Digitisation can increase the availability of samples which can be tested for the purposes of drug discovery.

AGRICULTURAL R&D

£20M TO £70M

Digitisation can help in the discovery and/or improve the understanding of Crop Wild Relatives (CWR) with regards to their genetic traits. This will benefit breeding of better crops for agriculture.

MINERAL EXPLORATION

£0.7BN TO £1BN

Digitisation can improve the accuracy of existing data and provide more geonomics data. This can accelerate the discovery process and minimise costs by de-risking it.

CASE STUDY: TYPE SPECIMENS

Digitising the Mollusca type specimens from the UK and Ireland (2016-2023)

Nomenclatural 'types' - the specimens from which a species was originally described - are the treasures of any natural science collection and are constantly sought out by researchers. Many museum curators are spread across several natural science disciplines and lack the skills or resources to attend to type research and curation, and such specimens are at risk of being 'lost' to the international scientific community.

Specialist malacological curators from Amgueddfa Cymru - National Museum Wales (ACNMW) and the Natural History Museum, London, funded by the John Ellerman Foundation are developing a jointly held, universally accessible resource connecting the Mollusca type specimens of national and regional UK museums for the first time. Staff at seven UK partner museums, lacking a malacological curator, were trained to recognise, research, database and interpret the molluscan type specimens in their collections. The website was launched with over 700 type specimens fully researched and imaged (https://gbmolluscatypes.ac.uk/). A further 450 types from the ACNMW collections and 500 records from three partner museums were added, exceeding the goal to cover primary types alone. A second phase of funding has allowed the inclusion of Ireland, and the site renaming as Mollusca Types in Britain & Ireland, involving 10 further collections including those in Belfast and Dublin, and adding further biographical details on collectors. This work is ongoing and is due to be completed in 2023.

CASE STUDY: 3D DIGITISATION

GB3D – an open resource for education, research and the public

All species – fossil or living – are defined by "Type Specimens" in museum collections. Many of the fossil species in the UK were described in the nineteenth century, and the current location of their types may now be poorly known. The GB3D Type Fossils Online project was funded by Jisc to address this. The British Geological Survey worked with 22 partner museums to take new high-resolution images of the type fossils in their collections – totalling over 17,500 – including many stereo pairs which were processed into anaglyphs. Two thousand of the best specimens were laser scanned to produce 3D-digital models. All the results are now freely available for viewing and download from the website: http://www.3d-fossils.ac.uk/

CASE STUDY: CITIZEN SCIENCE

Moths and the Edward Pelham - Clinton collection

With over 2.5 million insect specimens in our natural science collection, National Museum Scotland (NMS) has been fortunate to receive many important collections over the years, donated by prominent naturalists. One of great significance to us is Edward (Ted) Pelham-Clinton, who devoted his spare time to collecting and recording, compiling one of the most comprehensive collections of British and Irish moths (35,600 specimens) in 1,518 localities (https://pelham-clinton-points.netlify.app/).

He kept meticulous notes of his findings in 64 field diaries, documenting 2,169 species and approximately 160,000 observations. Together with the collection, this is probably the most comprehensive resource on British moths from 1930 to the 1980s.

This collection has become an invaluable reference for experts, biological recorders, and volunteers, who visit the collection regularly to extract data or to work with the specimens. The Covid-19 pandemic has made digitisation of his field diaries an urgent priority and we now have over 17,000 pages scanned and transcription is underway.

Ted not only made a substantial contribution to British Entomology through his writings but also inspired the current Biological Recording movement. NMS volunteers and associates continue Ted's work recording biodiversity in Scotland, re-visiting some of the same sites that Ted visited and thereby providing society with critically important environmental monitoring information.

The website is well used and frequently cited by academics and popular with fossil enthusiasts. It was launched to the public at the Lyme Regis Fossil Festival in 2013 and the online 3D experience has proved popular with schoolchildren, particularly during the recent periods of lockdown, including schools as far afield as Athens - http://gb3dtypefossils.blogspot.com/2015/04/.

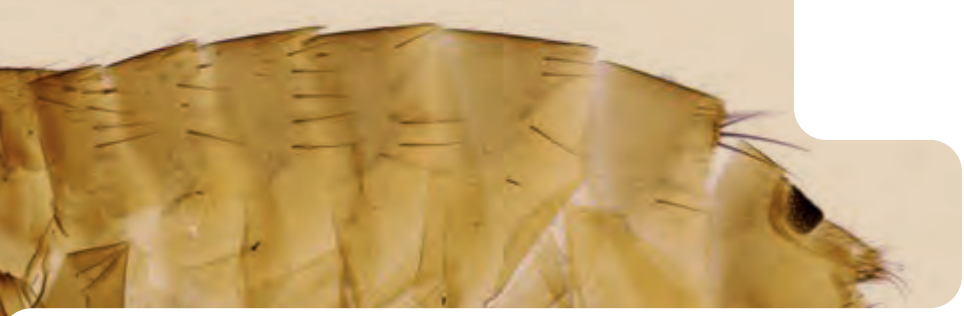
DIGITISATION

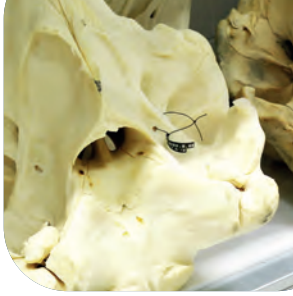
There is a vast and growing demand for the UK's digital natural science collections, with organisations keen to mobilise their data, but constrained by funding, training and a lack of digital infrastructure.

THE CURRENT STATE OF DIGITISATION

In December 2021, 90 natural science collections responded to our survey about their digitisation strategies, progress, resources, barriers and needs. While there is tremendous enthusiasm for digitisation, with 57% of organisations already doing some, most are unable to resource this activity consistently or to mobilise data to maximise impact.

The Natural History Museum, London, has been digitising systematically since 2014 and recently released its 5 millionth digital specimen. While this is only just over 6% of their 80 million items, demand from researchers is enormous, with more than 3 billion records downloaded in over 470,000 download events since February 2015 and more than 1,600 scientific papers citing this digital data.





BARRIERS TO DIGITISATION

78% of organisations need more funding to prioritise NSC digitisation.

FUNDING

13% of organisations have a full-time digitiser.

52% of organisations have a part-time digitiser.

STAFF

There is interest in training for all steps of the digitisation process. The topics with highest interest were on publishing specimen data and imaging workflows.

TRAINING

RESOURCES

51% of institutions have access to a macro photography set up.

34% of institutions are digitising their non-natural science collections at a higher priority compared to their natural science collections.

NATURAL SCIENCES LOW PRIORITY

37% of institutions have access to a flatbed scanner.

25% further equipment is available to less than 25% of institutions.

24% of institutions cite low leadership interest as a key barrier to natural science digitisation.

SCIENTIFIC EXPERTISE

Some institutions lack natural science specialists, restricting use and digitisation of scientific objects as there is limited knowledge on the contents of the collection.

WHAT IS DIGITISATION?

Digitisation is the process of converting physical information into a digital format. This can involve scanning or photographing the specimen, but sometimes just involves capturing fields of data, usually from the specimen labels, generating vast datasets that underpin scientific research on collections.

Different levels of digitisation yield different types of data for different purposes. At a minimum, digitised specimens require a unique identifier (usually a matrix barcode) linking the physical object to the digital record.

This stub is then augmented with further information, including the scientific name, geographic location, collector, date and various images of the specimen and their labels. More advanced information includes 3D models, DNA sequences, chemical analysis, field notes, and other collection details. The level of digitisation corresponds to a standard (the Minimum Information about a Digital Specimen, or MIDS for short), and the data are recorded in accordance with the Digital Extended Specimen data model, supporting a diverse range of data services to share, improve and exploit each record.

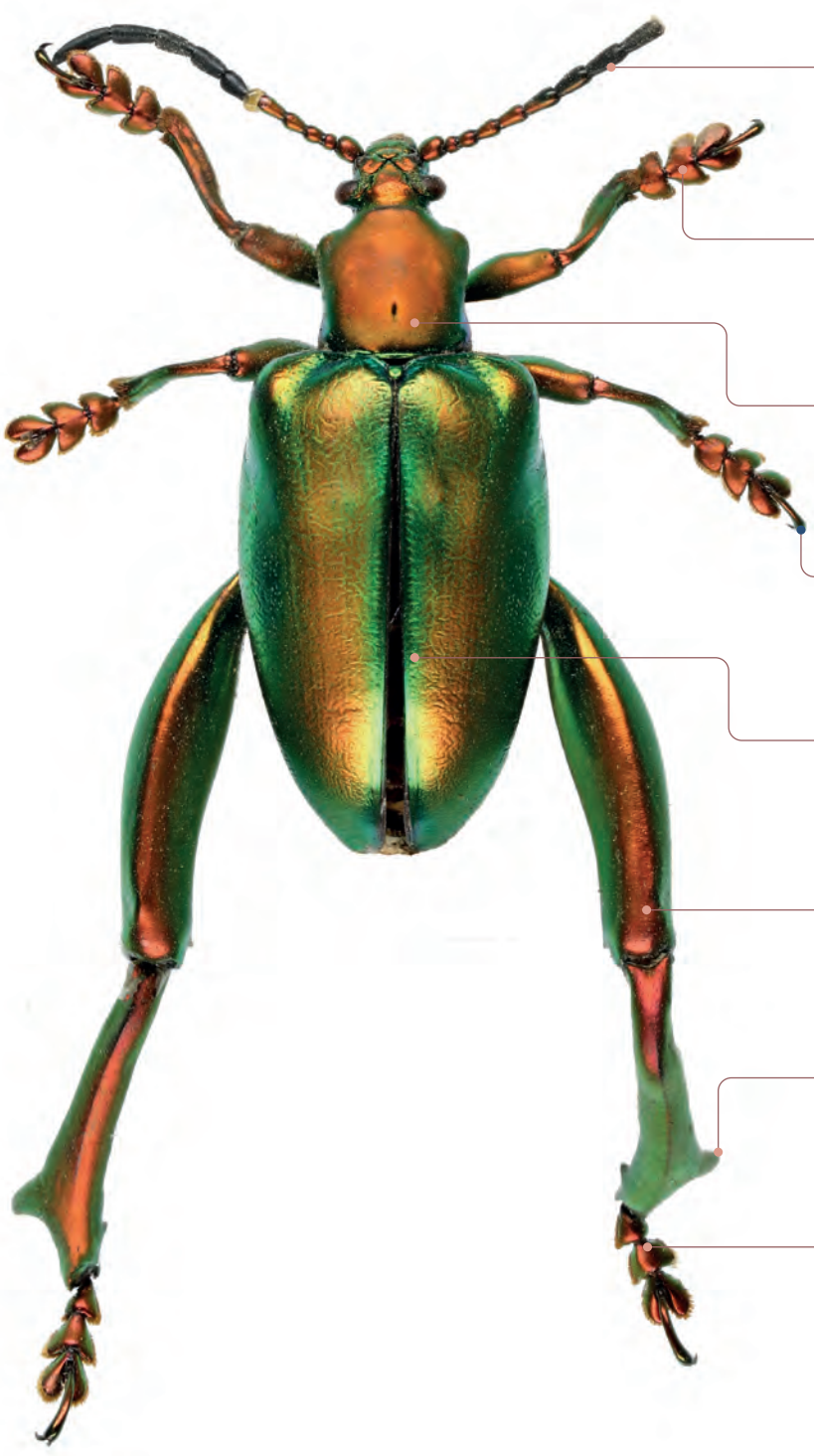
THE EXTENDED DIGITAL SPECIMEN

Until recently, science, industry, and society have relied on physical specimens housed in UK natural science collections. Ongoing advances in data generation and analysis being pioneered by DiSSCo are transforming biological and Earth science collections from physical specimens to dynamic suites of interconnected resources enriched through study over time.

The concept of the digital extended specimen is central to DiSSCo and describes the constellation of data types, preparations and processes centred around a natural science specimen, such as the specimen's physical, genetic and ecological properties, recorded in space and time.

The facts are represented as a digital twin of each physical specimen stored within a database, which is FAIR (findable, accessible, interoperable and re-usable), extensible as further information becomes available, and linked to other specimens through their shared properties. The resulting connected network of digital specimens supports analytical approaches that are simply not possible with physical objects, linking data to support coordinated co-analysis of specimens and richer, denser and more robust datasets.

WHAT'S IN A MUSEUM SPECIMEN



- Genomic data
- Chemical data
- Morphological data
- Geographical data
- Taxonomic information
- Species Interactions data
- Ecological data
- Geological data

WORKFLOWS, AUTOMATION AND NEW TECHNOLOGIES

The UK is one of the world leaders in developing and innovating processes to digitise and mobilise data from natural science collections, although we lag behind many countries in terms of the volume of collections digitised. Good imaging workflows at high speed and scale exist for herbarium sheets (flat pressed plants); insect specimens on pins; and microscope slides. Similar mass imaging workflows are being developed for other smaller specimens and preservation types, such as small dry invertebrates (fossil or extant). Larger, more complex, or fragile collection items may require a more bespoke approach.

Natural science collections represent a vast variety of object types, sizes and methods of preservation that can make them challenging to digitise but this diversity is also a catalyst for innovation. Robotics and various branches of Artificial Intelligence like computer vision and machine learning are central to this endeavour and have already been deployed in some digitisation workflows. Whereas human transcription of data, for example from specimen labels, is time-consuming, optical character recognition and handwriting recognition can extract more of these data in a fraction of the time. Robots may offer round the clock ways to pick-and-place specimens e.g., in imaging equipment for increased output. Automation can also be deployed to clean and enhance data, identifying possible errors in clusters of information. Many of these efficiencies are products of digitisation at scale, with some necessitating centralisation of specialist equipment.

CASE STUDY: EDUCATION

Kernels of Wisdom

In December 2021, Manchester Museum visited two local primary schools to deliver a series of workshops about Indigenous gardening, led by Border Crossings (a charity that celebrates the world's First Nations) with support from the National Lottery Heritage Fund. The Botany Bay project set out to "make use of the migration histories of plants and crops, and their Indigenous cultural heritage in relation to ecology and reciprocity, as a way to stimulate young people to explore new ways of living". In a school hall, inside Manchester Museum's Inflatable Museum, Alexandra P. Alberda, the Curator of Indigenous Practices, seamlessly wove together object handling with personal experience and storytelling.

Throughout the workshops, pupils reflected upon where our common food plants come from, how these plants were, and are, used by Indigenous peoples of the Americas, and how colonisation impacted upon these practices.

Using colourful corn, beautiful butterflies, and antique prints of medicinal plants, the group explored different ways of knowing, based on kinship and care, and the event was less about teaching and perhaps better described as gifting. The pupils started to make connections and comparisons with their own lived experiences, and to develop plans for their own sustainable school garden. It was sobering to see how discussing the museum objects revealed a genuine concern about climate change and its impacts for many of the pupils.

DIGITAL INFRASTRUCTURE

DiSSCo UK will develop and deliver biodiversity information infrastructure, providing the expertise and tools to support the integration of all collections information across the UK network as an interconnected digital knowledgebase.

REGIONAL DIGITISATION HUBS

Regional hubs will be centrally coordinated and supported by the Natural History Museum's experienced digitisation team, via a variety of self-service and direct options including support with recruitment, training, and data delivery to the national data portal. This team will provide coordination and support, adopting the Train-the-Trainer model, whereby partner organisations will be trained in digitisation and associated digital skills, and will then act as leads, passing on these skills and expertise to local organisations. The central hub will also provide a stable and persistent data infrastructure to support digitisation and research, implementing and promoting models and tools for persistent sharing and use of open data.

A NATIONAL DATA PORTAL

Digitised data will be made free, open and accessible through a national natural science collections portal that will be developed in coordination with the National Biodiversity Network (NBN). This Portal is likely to be based on the Living Atlas Software, which has become the international standard for making biological observations and collections data available. As part of DiSSCo UK we will explore the potential to expand the capabilities of the Living Atlas Software to Earth science collections. The central portal will aggregate data from digitising partners, and link out to the Global Biodiversity Information Facility (GBIF) and GeoCASe which are respectively the international aggregators of biodiversity and Earth science collections. GBIF and GeoCASe provide DiSSCo UK with the capability to track scientific exploitation and use of collections data, enabling us to see precisely which specimens and collections are being used, by whom, and for what purposes, at both the institutional and national level.

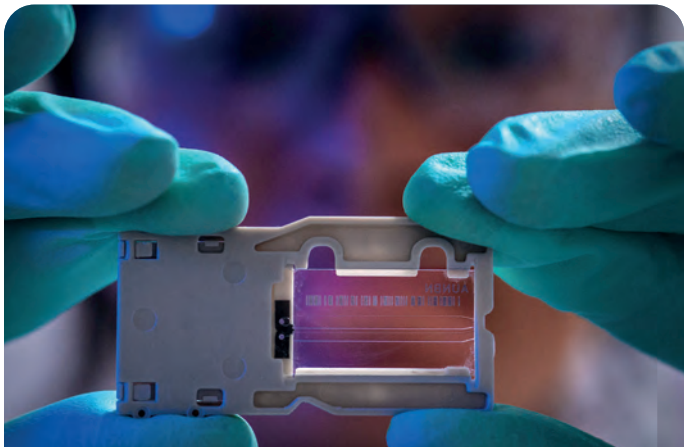
OPEN DATA STANDARDS

DiSSCo UK is fortunate to be able to adopt the relatively mature data standards and technologies that have already been developed within the global natural science collections community. Central to this is the concept of the Digital Extended Specimen, which is the basis of every digitised record in the wider DiSSCo network. Adapted from the mature Darwin Core standard, the Digital Extended Specimen provides a shared technical basis that enables us to digitise to a common standard, address quality issues associated with the data, and provide the flexibility to support the diverse outcomes associated with the aggregated information. This data standard is applied throughout the digitisation process and within our data infrastructures such as our collections management systems and data portal.

All data generated through the network are both open and FAIR (Findable, Interoperable, Accessible and Reproducible). This is a condition of participation within the wider DiSSCo network, with all data made available under the Creative Commons Zero (CC-Zero) waver and images available through a Creative Commons Attribution (CC-BY) licence. Limited exceptions will be possible where data meet certain criteria, such as disclosing the locations of threatened species.

COMPLEMENTARY ACTIVITIES

DiSSCo UK is embedded within a network of related organisations that aid the digitisation and dissemination of UK collection data. We are partnering with the National Biodiversity Network (NBN), created to exchange biodiversity information, and are aiding in the development of the national natural sciences collections portal. NBN is also the UK node for GBIF, the Global Biodiversity Facility, which is the global aggregator for biological collection data and observations. DiSSCo UK is also working closely with GeoCASe, which is the sister organisation to GBIF for Earth Science collections. Biodiversity information standards are being developed within the TDWG community, who promote the development and adoption of a comprehensive domain model and vocabularies for biodiversity and Earth science information. NatSCA (the Natural Sciences Collections Association) is the UK's subject specialist network for natural science and is a vital outreach and networking partner for DiSSCo UK. Lastly, DiSSCo UK is also the UK node of DiSSCo whose international coordination team is developing and promoting a shared technical architecture to enable biodiversity knowledge to be integrated and managed as a linked digital resource.



SUSTAINABILITY & COLLABORATION

Building a long-term partnership for DiSSCo UK

CHALLENGES

DiSSCo UK seeks to become a one-stop, dynamic, openly available and easy-to-use e-science infrastructure for the UK natural science collections, optimising collection access, curation and management practices by digitising and permanently linking collections to digital representations of these specimens.

The principal challenges relate to:

- Data and technologies, including fitness for purpose, transparency and completeness of the data we collect, as well as the application of AI and robotics to the digitisation process.
- DiSSCo UK's organisation and community, including gaps in participation from some regions, and the functioning of the network of partners.
- Funding, noting the challenges of maintaining sufficient funds to address the scale of the digitisation and core activities of the DiSSCo UK central hub.
- Management of expectations, addressing the risk that stakeholders may have unrealistic expectations of what DiSSCo UK is able to achieve given the level of resources and effort required.

ORGANISATIONAL MODELS

Operationally, DiSSCo UK will adopt a hub and spoke model enabling mutual learning across the network to enhance interdisciplinary impact. Centres of digitisation, some specialising in certain themes, will be developed across the UK's 12 major regions, supported by a primary DiSSCo UK hub operated by the Natural History Museum, London. The wider network of local UK collections will work with their closest regional hub or the central NHM facility, providing digitisation support, access to IT infrastructure, training and logistics for moving collections. Most collections will be digitised locally, coordinated by staff and equipment from the hub(s), although in some instances collections may be moved if appropriate to a digitisation centre. Standard operating procedures, digitisation hardware and IT infrastructure will be deployed to ensure consistency of processes and data quality assurance.

Digitisation priorities are likely to take a thematic approach, aligned with the different digitisation workflows required for different types of collections. Some regional digitisation hubs may specialise in particular workflows, becoming national centres of excellence, while others will cater for multiple workflows. Considerations include the variety of collections, the impact and volume of digital records created through different workflows, the regional distribution of the hubs, and the capability gains (e.g., number of staff/volunteers trained). Local, smaller collections will be able to bid for funding to digitise their entire holdings, while others will focus on thematic groups, often relating to how these collections are currently organised and used.

FINANCIAL NEEDS

Completing the digitisation of the vast majority of the UK's natural science collections will cost in the region of £150m, equating to an average of roughly £1 per specimen, but with considerable variance due to the diversity of UK collections, from dinosaur leg bones and whale skulls the size of a small room, to slide-mounted wasps smaller than the full stop at the end of this sentence. The balance of funds covers the operation of the DiSSCo UK data portal, coordination and support activities, and the associated digital infrastructure. It is estimated that the digitisation process will take 6-10 years, with a minimum return on investment of £2 billion over the next 30 years.

Participating institutions will provide a significant in-kind contribution, but a proposal of this magnitude that reaches every region of the UK, from small volunteer-led and local authority-run museums to the major collections of the UK's four nations, can only be addressed with significant national investment. These costs are grounded in experiences from equivalent national activities in Europe, the US and Australia and the development of a 'cost-book' organised by the wider DiSSCo network.

DiSSCo UK's funding would be distributed through a central fund supporting digitisation hubs and their digitisation needs, alongside a centrally administered grant scheme, enabling local, smaller institutions to bid for funding and support, to ensure that local and regional collections can join the initiative at a time and place of their choosing. Funding will require multiple 'tracks' to ensure inclusion for all, from the largest collections (which are already 'digitisation ready') to the regional hubs who will need to recruit and train teams and the small local collections who will need to use those resources for a short period of time.

NEXT STEPS

Large-scale problems require large-scale solutions. DiSSCo UK is an ambitious initiative to provide a dynamic and high-quality digital gateway to the UK's natural science collection. Despite leading the globe in digitisation innovation, the UK has slipped behind global peers on digitisation rates, with the US, Germany, France, Switzerland, Spain and the Netherlands all committing tens to hundreds of millions of dollars or euros to collection digitisation over the past five years. If we act now, DiSSCo UK provides an opportunity to bring together the UK's world-leading centres of natural science collections, halting biodiversity loss and ensuring that the UK does not miss out on the scientific, commercial, and societal benefits of the emerging biodiversity data revolution.

PARTNER LIST

Birmingham Museums, Bristol Museum & Art Gallery, British Geological Survey, Great North Museum, Horniman Museum & Gardens, Manchester Museum, National Museum Wales, National Museums NI, National Museums Scotland, Natural History Museum London, Natural Sciences Collections Association, Oxford University Museum of Natural History, Royal Botanic Gardens Edinburgh, Royal Botanic Gardens Kew, University of Cambridge Museums.

ACKNOWLEDGEMENTS

This work was led by the Natural History Museum (NHM) with the assistance of the partner institutes. We would like to thank the Arts and Humanities Research Council (AHRC) for their support (Grant number AH/W006715/1) and the 94 UK natural science collections who generously provided their time and data to develop the concept of DiSSCo UK.

UKRI Arts and Humanities Research Council

HOW TO CITE THIS BOOKLET

Smith, V.S., Hardy, H., Wainwright, T., Livermore, L., Fraser, N., Horak, J., Aspinall, J. and Howe, M. (2022) Harnessing the power of natural science collections: a blueprint for the UK, pp 1-28. Natural History Museum. doi: https://doi.org/10.5281/zenodo.6472239

Please recycle this brochure after use.

The Museum strives for environmental and financial sustainability. This invitation has been designed with this in mind and is printed on carbon balanced, FSC© recycled post-consumer waste pulp.

WEB
dissco-uk.org

CONTACT
dissco-uk@nhm.ac.uk

DISSCO UK