Mapping the Species Data Pathway: Connecting species data flows in England

Cabinet Office – Geospatial Commission

May 2021

Geospatial Commission

OGL

4 City Road
London EC1Y 2AA

+44 (0) 20 7580 5383
eftec@eftec.co.uk
eftec.co.uk

This document has been prepared for the Cabinet Office by:

Economics for the Environment Consultancy Ltd (eftec)
4 City Road
London
EC1Y 2AA
www.eftec.co.uk

eCountability
Chancery Cottage,
Kentisbeare, Cullompton,
Devon, EX152DS

www.ecountability.co.uk

ALERC
C/O NEYEDC
10a Minster Gates,
York,
North Yorkshire
YO1 7HL
http://www.alerc.org.uk

NBN Trust
Unit F
14 – 18 St Mary's Gate
Lace Market
Nottingham
NG1 1PF
https://nbn.org.uk

BRC/CEH
Wallingford
Oxfordshire
OX10 8BB
UK
brc@ceh.ac.uk

Study team:

Ian Dickie (eftec)
Natalya Kharadi (eftec)
Sophie Neupauer (eftec)
Bill Butcher (eCountability)
Jo Treweek (eCountability)
Jo Judge (National Biodiversity Network Trust)
Steve Whitbread (ALERC)
Tom Hunt (ALERC)
David Roy (BRC)
Martin Harvey (BRC)

Reviewer

Rob Tinch and Ece Ozdemiroglu (eftec)

Acknowledgements

The authors are grateful for the many contributions of time, knowledge and thinking to this report by stakeholders within and outside the study team. The authors are responsible for any errors or omissions.

Disclaimer

Document evolution

Draft Report 1	20/12/2020	Reviewed by Ian Dickie
Draft Final Report 1	26/12/2020	Reviewed by Ian Dickie
Draft Report 2	03/02/2021	Reviewed by Ece Ozdemiroglu
Final Report (combined)	23/02/2021	Reviewed by Rob Tinch
Final Report (combined)	12/04/2021	Reviewed by Ian Dickie

This report is based on eftec's Version 2 – January 2020 report template.

Contents

Tables

Figures

Boxes

Abbreviations & Acronyms

ALA	Australian Living Atlas
ALERC	Association of Local Environmental Records Centres
API	Application Programming Interface
BCR	Benefit-cost ratio
BDF	Biodiversity Data framework
BRC	Biological Records Centre
BTO	British Trust for Ornithology
CBA	Cost-benefit analysis
Cefas	Centre for Environment, Fisheries and Aquaculture Science
CEH	Centre for Ecology and Hydrology
CNN	Convolutional Neural Networks
DAC	Data Archive Centre
DASHH	The Archive for Marine Species and Habitats Data
Defra	Department for Environment, Food and Rural Affairs
DOI	digital object identifier
eftec	Economics for the Environment Consultancy Ltd
ESAS	European Seabirds at Sea
EU	European Union
EurOBIS	European Ocean Biodiversity Information System
FAIR	Findable, Accessible, Interoperable, and Reusable
GBIF	Global Biodiversity Information Facility
GDPR	General Data Protection Regulation
GEMINI	GEo-spatial Metadata INteroperability Initiative
GIS	Geographical Information Systems
ICES	International Council for the Exploration of the Sea
INSPIRE	INfrastructure for SPatial InfoRmation in Europe
ISO	International Organization for Standardisation
JCDP	Joint Cetacean Database Programme
JNCC	Joint Nature Conservation Committee
LERC	Local Environmental Records Centre
MCS	Marine Conservation Society
MCZ	Marine Conservation Zone
MEDIN	Marine Environmental Data and Information Network
MEDIN	Marine Environmental Data and Information Network
MOAT	Marine Online Assessment Tool
MPA	Marine Protected Area
NBN	National Biodiversity Network
NGOs	Non-Governmental Organisations
NMBAQC	North-East Atlantic Marine Biological Analytical Quality Control
NPV	Net present value
NSS	National Recording Schemes
NWEBS	National Water Environment Benefits Survey

OBIS	Ocean Biodiversity Information System
OGL	Open Government Licence
ONS	Office for National Statistics
OSPAR	Convention for the Protection of the Marine Environment of the North-East Atlantic
PH	Priority Habitats
PVB	Present value of benefits
PVC	Present value of costs
QA	Quality Assurance
SBIF	Scottish Biological Information Forum
TDWG	Taxonomic Databases Working Group
UKCEH	UK Centre for Ecology and Hydrology
WFD	Water Framework Directive
WoRMS	World Registry of Marine Species

Executive Summary

The UK Geospatial Strategy sets out a commitment for the Geospatial Commission to identify how improved access to better location data can support environmental outcomes. This is part of Mission 2 in the strategy, to improve access to better location data. To support this mission, the Geospatial Commission's Data Improvement Programme has sponsored this study to look at the costs, benefits, and management of species data in England. It presents options to make species data more consistent, joined up and accessible for end-users by encouraging FAIR data principles.

High quality, current and accessible species data are essential to underpin environmental policy and land use planning. Implementing and evaluating outcomes from the Environment Bill, Biodiversity Net Gain and the new Environmental Land Management Scheme (ELMS) will depend on access to high quality species baseline and monitoring data. Species data tells us about the diversity of natural assets, which is associated with greater resilience (Dasgupta Review, 2021).

The current species data pathway in England enables a large amount of data to be recorded and shared. The current system utilises extensive volunteer input and is reliant upon continued good will and interest. The system has evolved ad-hoc over time, leading to complex species data flows and data inconsistencies. There are a confusing number of data flow-routes, incomplete species group and spatial coverage and a lack of clarity regarding roles, data quality, access, responsibilities, and processes (e.g., for verification).

Economic analysis was conducted of the costs and benefits of the role of the species data pathway in enhancing stewardship of and access to species data, and the decisions that use such data. A baseline (current operation) was compared to a 'no species data pathway' scenario in which data stewardship and use is severely impaired due to the absence of the data pathway. The cost-benefit analysis (CBA) found that the benefit of the current species data pathway strongly outweighs its cost. The benefit-cost ratio ranges between 14:1 and 28:1. The availability of resources is a risk to maintaining the current species data pathway and a funding gap of £6 million has been identified through this study.

This study outlines 14 recommendations that are grouped into four themes:

- Defining biodiversity data framework (BDF)
- Principles and standards
- Investment
- Data use and re-use

This study also reviewed the 2018 Scottish Biological Information Forum's recommendations on biological recording infrastructure - not all are relevant to England, but there are similarities in the priorities both this study and SBIF identify. These include needing to support the financial viability of, and benefits from the pathway, improving current practices (e.g., on data verification), and using FAIR data principles.

Technical Summary and Recommendations

This study is made up of three main parts:

1. An Investigation of the species data 'landscape' in England
2. An economic analysis of the costs and benefits of the current species data pathway to society.
3. Recommendations

The species data landscape in England

High quality, current and accessible species data are essential to underpin policy and land use planning towards sustainable outcomes for the environment and biodiversity. The species data landscape in England covers all types of marine, freshwater and terrestrial species. The National Biodiversity Network Data Flow Pathway (Figure S1) summarises the operations generally involved in creating, checking and sharing species data so that they can be put to use. Species data provided via this pathway are collected by numerous volunteers and professionals and then processed and stored across a wide range of organisations. Species data can be used for a wide range of purposes including developing and evaluating biodiversity and environmental policies, and decision making within local and national planning systems.

Figure S 1: National Biodiversity Network (NBN) data flow pathway

Looking forward, implementing and evaluating outcomes from the Environment Bill, Biodiversity Net Gain and the new Environmental Land Management (ELM) scheme will depend on access to high quality species baseline and monitoring data. Species diversity and abundance in the UK continue to decline and the UK failed to meet most of the CBD's 2020 Aichi targets for biodiversity. Monitoring of impacts on biodiversity has also not been systematic, but the *State of Nature 2019* (Hayhow et al., 2019) concluded that more species have shown strong or moderate decreases in abundance, and/or have decreased in distribution than have increased since 1970. Reliable data on species' distribution and abundance, shared through a robust and efficient biodiversity data framework, will be essential to plan actions towards nature recovery and,

importantly, to evaluate their effectiveness.

There are currently a number of limiting factors and risks associated with the maintenance of the biodiversity data framework that supports the species data pathway. These are:

- Availability of resources - Resources come from funding of organisation's core functions, cross-subsidising some core functions from the various revenue-generating services that utilise the data, and the large amounts of volunteer time.
- Poor data access - Approximately 50% of potentially useful data were identified as "currently inaccessible" Hassall et al. (2020).
- Variable data quality - not all records marked as 'verified' are correct.
- Incomplete spatial and species group coverage.
- A confusing number of data flow-routes - due to the organic nature in which the species data system has evolved since the Victorian era.

Economic analysis of the species data pathway

The economic analysis focused on the costs and benefits of the species data pathway, and its role in enhancing access to species data for decision making. The baseline is defined as the current operation of the pathway in England, through which species data is verified, put into a consistent context and shared. This is compared to a scenario in which the species data pathway does not exist, so while the same data is collected, its stewardship and use is severely impaired, as illustrated in Figure S2.

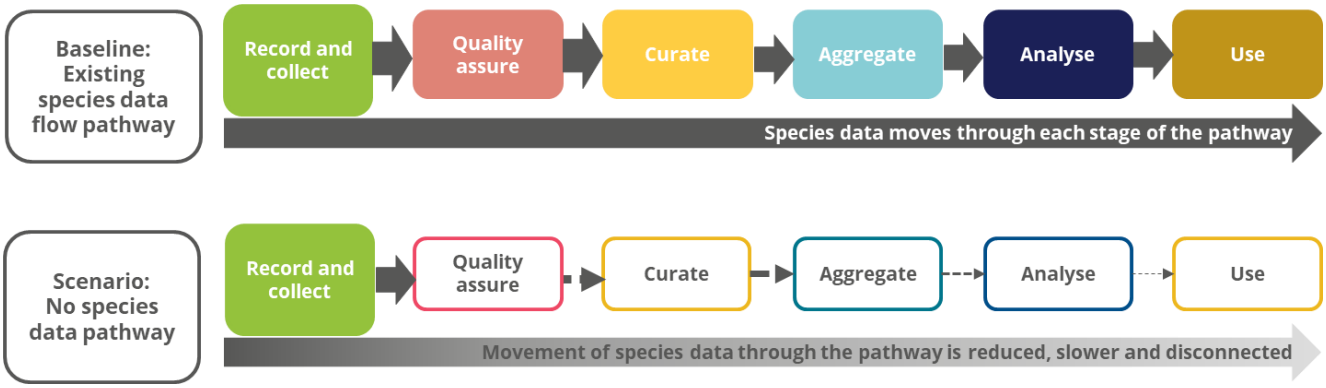


Figure S 2: Baseline and scenario of the species data flow pathway compared in economic analysis

The economic analysis compares this baseline and scenario in a cost benefit analysis (CBA) framework. The CBA focuses on the data provisioning service of the existing species data pathway. This service is critical, as for users, it is not just about having access to data, but also knowing that they have access to all relevant data.

In the 'no data pathway' scenario, data does not flow effectively along the pathway (i.e., from record and collect to end use in Figure S1.). As a result, while the processes that use species data

are the same, they operate with no or unstructured species data.

Results of the cost benefit analysis

Table S1 presents the results of the CBA. The 'minimum' results represent costs and benefits that have been estimated with the highest confidence indicators, whilst the 'best' result reflects a mixture of high and moderate confidence in the methods used. The 'maximum' reflects the maximum possible costs and benefits of the system including measures with 'poor' confidence. Upper and lower bound indicators of costs and benefits were tested in sensitivity analysis.

Table S 1: Cost-benefit analysis outline results: no species data pathway compared to the baseline of current pathway

Description	Minimum	Best	Maximum
Total present value costs (PVC) PV60 £m	331	829	2,483
Total present value of benefits (PVB) PV60 £m	5,643	22,939	34,003
Net present value (PVB – PVC) PV60 £m)	**5,311**	**22,110**	**31,520**
Benefit-cost ratio (PVB / PVC) PV60 £m)	**17.0**	**27.7**	**13.7**

Table notes:

The baseline year is 2020. As recommended by the Green Book (HM Treasury, 2020), present values (PV) have been estimated using the standard discount rate (3.5% and declining) over a 60-year appraisal period

While the monetised estimates account for a number of significant economic costs and benefits, it was not possible to quantify all potential impacts. Nevertheless, the conclusion is that the benefits of the species data pathway to a variety of private and public sector decision-making processes strongly outweigh the costs. Key benefits include achieving water industry and agri-environment objectives and avoiding delays in the land use planning system.

While there may be some omitted costs, the main costs of the pathway are believed to be captured and have been adjusted for optimism bias. The costs for the baseline include a small sum – approx. £100,000 per year – that is not currently funded. This is needed to shore up the system (e.g., to maintain IT hardware and software), and retain the existing large benefits outlined.

There may be significant omitted benefits, such as the role of species data in:

- Organising and motivating wildlife tourism activity and spending.
- Motivating volunteers to be physically active in a way that benefits their health; and
- Understanding several of the ecosystem services identified in the UK national ecosystem assessment (e.g., as an indicator of the quality of vegetation and soil's capacity to store and sequester Greenhouse Gases).

Conclusions

The species data pathway is of great value to the UK, both to private sector activities and public policy, providing services that are essential to a wide range of decision makers. It is fundamental to activities that interact, evaluate and change the environment. Moreover, it is key to informing strategy on how the environment is managed to benefit people and meet future challenges.

There are strengths of the current species data pathway: it is good at getting a large amount of data recorded and shared, and there is extensive volunteer input in terms of volume of records and expertise. However, biological recorders, many of whom are volunteers, will typically choose a data flow route that reflects their personal experience and the legacy experience of their peer networks. This often leads to errors, omissions and inefficiencies within the biodiversity data system, as well as resistance to new technology or methods of working. However, it has been historically underfunded, and increased funding could enable significant improvements in practices and the application of technology.

Novel data capture methods, together with rapid developments in technology, are creating opportunities to accelerate rates of recording and data sharing. The biodiversity data system will need to develop new ways to realise this potential. Economic analysis has identified significant benefits, but the scale of the investment required now, and in the future has not been evaluated.

Recommendations

This study's recommendations are based on the review of the existing species data landscape, the economic analysis reported above, the results of interviews and workshops held as part of the study, and on the expert opinions of the consortium members who carried out the study. The study also reviewed the applicability of the SBIF recommendations for Scotland to England. It then developed further recommendations for the future management of the species data pathway in England.

While this study looks at the current species data pathway in England, the SBIF review undertook an investment analysis of enhancing the pathway in Scotland. The SBIF review was based on an assumption of open data, which means free data in the sense that the costs to those recording and managing it cannot be recovered. This led to SBIF's approach to seek 100% government funding for species data pathway and infrastructure enhancements. The SBIF review suggested options for service provision and governance relevant to such a public funding model. This funding model is considered unlikely in England. Therefore, some of the SBIF recommendations do not apply to England. However, both reviews identify similar issues and priorities, and conclude that there are many uses of species data, the benefits of which significantly outweigh the costs of the pathway.

This study makes 14 recommendations in 4 areas, summarised below (further detail can be found in Section 8.3 of the main report).

Biodiversity data framework

The National Biodiversity Network (NBN) includes all the varied organisations and individuals involved in collecting and managing species data in the UK. The key organisations are the NBN Atlas, Local Environmental Record Centre (LERC) databases, National Recording Scheme (NSS) databases (including those hosted by Biological Records Centre). While marine data infrastructure is likely to remain substantially separate from terrestrial/freshwater systems, a level of interoperability around the coastal zone and species that use both environments is desirable. There is a diverse array of actors with an interest in species data. Many of them are supportive of existing system to an extent, but also recommend changes and improvements to deliver better quality, more joined up approaches, better data accessibility or faster response times. Clearer definition of organisations' roles would assist in improving data flows.

The general consensus between participants in this study and expert opinion is that local and taxon-specific databases are required to enable local and taxon-specific knowledge to be embedded into the data entry, data management and quality assurance processes. It is acknowledged that, in order to improve this system, there needs to be greater transparency of data flows and an adoption of the FAIR Data Principles (Wilkinson et al. 2016, for data to be Findable, Accessible, Interoperable, Reusable), which could be facilitated via a certification system (see recommendation 6). While there was some support for one centralised database for all species records, there are inherent risks in having a single system, such as over-reliance on a single IT framework and its managers and a loss of engagement from voluntary recorders.

Biodiversity Data Framework Recommendations

1. *Recognise the key components of the Biodiversity Data Framework as local data centres, National Schemes & Societies, UK-wide and marine data portals. Each should have a clearly defined role, enabling them to work together as a collaborative, connected community.*
2. *Promote to key sectors the public good, efficiency and conservation benefits of collecting, sharing and using datathrough a clearly defined and recognised BDF.*
3. *Maintain the local and taxon-specific biodiversity databases, with a greater emphasis on transparentdata flows and data sharing via a UK data portal.*

Principles and standards

Although the BDF comprises separate organisations, they should operate to shared principles and standards. The FAIR principles are relevant as they aim to ensure that maximum value can be gained from existing data products. They have attracted wider attention and support and received almost unanimous support during the stakeholder consultation for this study. They recognise the rapid acceleration of online data access and the need to support machine discovery and use of data as well as human operation.

Consultation for this study has identified that while the current system enables access to a significant amount of high-quality species data, there are weaknesses. The system has bottlenecks and constraints around data quality and accessibility. Wider adoption of standards, data-sharing protocols and consistent use of clear data flow pathways, which are currently limited due to funding shortages, would be one way to improve this. The need to communicate standards is most acute before the data collection stage, in both professional and volunteer recording. Additionally, training and agreed processes would give assurance to recorders, data contributors, funders and users that improvements were being made. Certification of processes, with a focus on advancement and improvement, is preferred to accreditation of organisations.

Data Principles and Standards Recommendations

4. *Base the generation, management, collation and sharing of species data on FAIR Data Principles – to make species data in England Findable, Accessible, Interoperable and Reusable throughout the species data network. These principles should be adopted in the BDF and should underpin all parts of the system*
5. *Develop and promote data standards throughout the BDF and the wider network of species data collectors and contributors. Particular emphasis should be on the data collection stage to ensure data meets its potential for use.*
6. *Develop a system of certification of BDF processes to drive high common standards across the network and the adoption of the FAIR principles. This will aid clear data flow pathways and efficient interoperability between BDF nodes.*

Investment

Across almost all stakeholder groups there was support for a mixed-funding model, as it offers a diverse, sustainable framework that promotes innovation opportunities inside and outside of government. The commercial users of species data consulted in this study recognised the fairness of contributing to the costs of data stewardship at the point of use. For the private sector it is often preferable to pay for data that is immediately accessible, known to be comprehensive and high quality, than to receive data of unknown quality and coverage for free.

The public sector could fund the core of the species data pathway, in return for free access to species data. This data is needed for delivery of policies in the government's 25-Year Environment Plan and to support planning, agricultural and environmental legislation. This approach recognises that Open Data is not synonymous with free data (but is more concerned with data accessibility) and aligns to the UK Geospatial Strategy approach to Open Data. The strategy aims to maximise economic, social and environmental value, and for the efficient and fair use of public money.

The SBIF review found that verification systems were under considerable strain and may be unsustainable in the face of rapidly rising demand (i.e., numbers of records requiring

verification). This finding is confirmed for England by consultation in this study: data needs to become accessible more rapidly. Technology is both a source of the problem (a conduit of increased records) and may offer part of the solution, through machine learning. Verification suffers from inconsistencies and unclear processes, which can be addressed by development and use of a protocol. It could cover criteria by which to assign the status of a record, and roles of human and machine verification.

Consultation in this study also revealed that high volumes of species data collected for specialised purposes are not reaching the species data pathway and do not become available for subsequent re-use. These data could contribute to knowledge and decision-making. In future this issue may be partially addressed through implementation of other recommendations in this report. A targeted programme of capture and mobilisation of such datasets would improve access to high-quality species data.

Investment Recommendations

7. *Adopt a Data Sharing approach across the pathway, in which many end-users contribute to reasonable costs of data stewardship within the BDF and, where practicable, species data is accessible online*
8. *Invest in the BDF, from central government budgets, as part of a mixed public/ private sector funding model, recognising the essential role of accessibility to high quality species data as a public good to deliver environmental legislation and policy*
9. *Develop a verification protocol with key stakeholders in the verification process, which aligns to current and future verification requirements and technologies. The resources required to support implementation of the protocol should also be identified.*
10. *Invest in building capacity for verification, through expert training and the use of new approaches such as automated assessment to support verification decisions.*
11. *Invest in processes to capture and mobilise species data generated by research and high volume, novel recording methods that can be used to supply the BDF.*

Improving data use and re-use

In the land use planning process improvements can be made to the sharing of primary data and the re-use of existing data. Many stakeholders expressed concern that, in a high proportion of planning applications affecting biodiversity, existing species data is not being accessed and that omission adversely affects the quality of decision making.

Species data collected by consultants to support assessments of projects that require regulatory approval do not normally reach the species data pathway and therefore are unavailable for re-use. This represents a wasted resource. In the marine environment the Marine Data Exchange provides an example of how this can work successfully for all parties. Established by The Crown Estate in 2013, this resource helps to make valuable data freely accessible, promote collaboration within the sector, reduce survey costs and ultimately de-risk investment offshore (Crown Estate, 2021).

For over a decade it has been a requirement of the Chartered Institute for Ecology and Environmental Management (CIEEM) consultant members to share data. A portal has been in place (Vogel, 2016, Smith et al 2016) to facilitate data re-use, but it is regarded as ineffective.

Although there is a requirement for data collected using certain public funds, such as government or National Lottery, to be shared, there is little or no follow-up to ensure data are shared. Data collection and sharing are often afterthoughts in a project and at project closure.

Data Use Recommendations

12. *Mandate the re-use of species data collected by consultants in transparent processes that support regulatory compliance, potentially through new regulation. This will reduce survey costs, improve professional standards and support environmental outcomes.*
13. *Require proponents of development to certify that best available species data through the BDF have been accessed in the preparation of applications where there is risk of impact on biodiversity, potentially through new regulation. This will help ensure that existing and newly collected species data is equally available to project proponents, regulators and evaluators and will support environmental outcomes.*
14. *Require organisations collecting data funded by public money to provide a plan for data collection and sharing, in accordance with FAIR data principles, before funds are received.*

Systems to support these recommendations and implement FAIR data principles require further investigation.

Future development of the Species Data Pathway

The following principles are suggested as a basis from which to manage the species data pathway in future:

1. Recognise the specific characteristics of the pathway, in particular the role played by volunteers. Data recording and collection (and possibly stewardship along subsequent parts of the pathway) are often not part of funded activities, meaning useful data is recorded but budgets are not allocated to stewardship of that data.
2. Organize the system to support FAIR principles to maximise the value of species data to society. In the data pathway, organisational structures, relationships and governance, staff resources and skills and funding should all aim to support the FAIR principles.
3. Enable those who support the species data pathway through stewardship of data (in line with FAIR principles) to cover the costs of efficiently doing so and be financially viable. This means being able to retain rights over data where needed to generate income.
4. Embed the interpretation of 'accessible' in FAIR as 'shared data' (rather than 'free data').

The challenge for managing the species data pathway is to balance the FAIR data principles with

the need for financial viability. This needs to be done in a way that supports stewardship of higher quality, shared species data in the long term. The funding of the biodiversity data system from public funds is justified based on the benefits it supports for wider society and taxpayers. Further revenue can be obtained from the sectors who need to access biodiversity data and realise value from doing so. The focus should be on sustaining revenue for data stewardship from those who get value from it.

There are a number of actions that can be implemented, and provide benefits, relatively quickly. This study has identified 'quick wins' in areas such as data verification, implementation of FAIR data principles, and engaging funders of the species data pathway to build awareness of good practice. The most immediate need is to shore up the system to stop collapse through investment in national IT support and additional staff to enhance resilience and capacity for development.

Investments to improve and simplify the species data pathway in England require consideration of what an enhanced biodiversity data system/pathway would look like, which is beyond the scope of this study. Mapping of the current data flows is needed to identify options for rationalisation, and whether organisations in the pathway have appropriate capabilities.

Planning of investment in England, and at UK scale, would also benefit from research into the needs of biodiversity (and species) data users. In developing investment options, it is recommended that:

- Full consultation of the organisations and stakeholders involved is carried out to help design and implement changes;
- Provision is made for training to develop the skills needed to realise the benefits and minimise the risks from any changes; and
- Key objectives are coordinated across the UK.

It should be noted that changes in organisations and/or process will also require the culture and behaviour for staff and volunteers to adapt. If the benefits of any change are not communicated to and accepted by those involved through lack of transition planning and funding, this increases the risk that expertise could be lost and data sharing reduced. Funding this culture change should be seen as an investment in retaining and growing the total contribution of volunteers to data recording and the long-term function of the species data pathway.

1.Introduction

This study for the Geospatial commission examines the costs, benefits and management of the species data pathway in England. It was developed by firstly investigating the current species data landscape in England, and then looking at its costs and benefits to society. This introduction describes the background to the study, the objectives, and the structure of this document.

The *UK Geospatial Strategy* under Mission 2 aims to improve access to better location data. To support this mission, the Geospatial Commission's Data Improvement Programme has sponsored this study to look at the costs, benefits, and management of species data in England. It presents options to make species data more consistent, joined up and accessible for end-users by encouraging FAIR (Findable, Accessible, Interoperable, Re-useable) data principles.

1.1 Study background

eftec, in partnership with eCountability ltd, the NBN Trust, the Association of Local Environmental Records Centres (ALERC) and the Biological Records Centre (BRC) of the UK Centre for Ecology & Hydrology (UKCEH) have undertaken this project to provide advice to the Cabinet Office's Geospatial Commission on the current species data landscape in England and to consider how species data could be made more consistent, joined up and accessible, as needed to support new biodiversity policy and other applications in England.

The value of data is realised when it is used to create insights that support decision-making and drive action. Its value to society, the economy and the environment can increase when:

- It is used by more people and businesses.
- It is used to solve problems beyond its original purpose.
- It is combined with other data.
- Its quality improves.

Objectivity in the consultation process was ensured through the use of a standard questionnaire across all interviews and a standard presentation for each workshop and the involvement. This also aided with the comparability of responses. The recommendations are developed and endorsed by the whole study team.

1.2 Objective of this study

This study provides an overview of the current species data landscape (data, tools, systems and challenges) and existing approaches to the collection, funding and provision of species data in England. All types of species data are considered, including marine, freshwater and terrestrial species. The report considers how biological recording works in England at present and provides an overview of issues and challenges identified by stakeholders. Recommendations for potential improvements are provided with a view to improving levels of access to high quality species data for a variety of applications in policy and decision-making. The main objectives are to:

- Provide an overview of the current species data landscape (data, tools and systems) and its efficacy.

- Provide an overview of current approaches to collection, funding and provision of species data.
- Identify issues and challenges.
- Provide recommendations to make high quality species data more consistent, joined up and accessible.

Building on these findings on the current species data landscape, this report covers two areas of analysis. Firstly, it presents an economic analysis of the current value of species data in England taking into account the Scottish Valuation Model in the SBIF Review (Wilson et al., 2018). Secondly, it considers whether the SBIF Review recommendations for Scotland can be applied in England based on the economic analysis for England and the current species data landscape in England. SBIF's recommendations reflect the outcomes of a comprehensive engagement with stakeholders involved in biodiversity data collection in the UK over a period of several years, but not all are necessarily directly transferable to the English context.

It should be noted that the economic analysis in this study differs from that in the SBIF review. The SBIF review considered the costs and benefits of investing in a recommended biological data management structure in Scotland (i.e., investment appraisal). **The economic analysis in this study is a cost-benefit analysis of the existing species data pathway in England as compared to a 'no data pathway' scenario, and it is aligned to the HM Treasury Green Book guidance (2020)**[1].

The study steering group's comments were taken into account in producing this report, particularly with regard to assumptions in the economic analysis and the review of relevance of the SBIF recommendations to England.

[1] HM Treasury guidance document which details how to appraise and evaluate policies, projects and programmes. https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-governent

2. The species data landscape in England

This section gives an overview of how species data flow from recorders to users along the species data pathway. It explains how species data are currently collected, managed and distributed in England and identifies the main organisations involved. This outline of the species data landscape is the basis for identifying its issues and challenges (in Section 3) through the consultation processes (Section 4), and for the economic analysis (in Section 5 and 6) and review of the SBIF recommendation's applicability to England, (in Section 7).

2.1 Data flows and pathways

Species data flows in the UK can be complex, even for a single species group. The National Biodiversity Network Data Flow Pathway provides a useful summary of operations generally involved in creating, checking and sharing data so that these can then be put to use. The data involved may be a biological record or a dataset that aggregates multiple records (Figure 2.1).

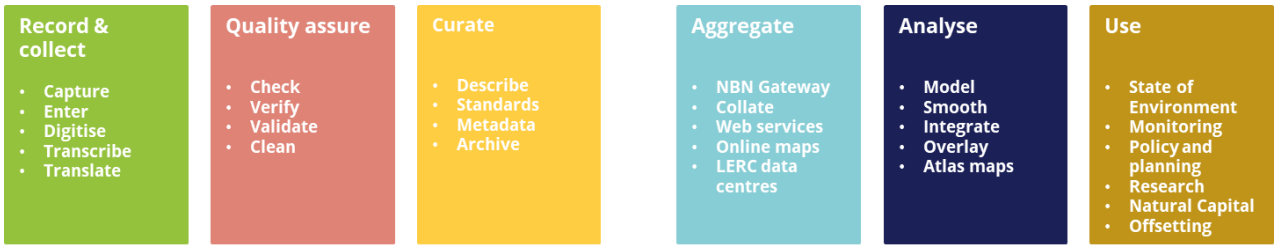


Figure 2.1: National Biodiversity Network (NBN) data flow pathway[2]

The main processes listed in each step of the pathway are briefly explained in Appendix C.

The six steps represented Figure 2.1 can be simplified into four groups as follows:

- Data Collection
- Data Stewardship
- Value Added Data Services
- Data Use

Data Collection includes both human observation and automated detection of species occurrence and the capture of the data into a record that can be communicated.

Data Stewardship relates to many of the processes in the Data Flow Pathway excluding original data collection and final data use. It includes quality assurance, collation, secure archiving, communication with data collectors and users and ensuring that data is managed, made accessible and used in line with the wishes of the originator.

Value Added Data Services that turn raw data into information and knowledge increase in importance towards the use end of the pathway. They enhance use, but may also overlap with data stewardship, for

[2] After National Biodiversity Network (2014). Note that the NBN Gateway has now been superseded by the NBN Atlas.

example in sharing metadata and verification information. Recorders may also be data users, so that communication of knowledge back to recorders can be regarded as a service.

Data Use includes the application of data, information and knowledge by individuals and organisations across the public, private and third sectors.

A comprehensive assessment of the complex species data flow landscape in England is beyond the scope of this study, but a useful characterisation of the main data flow routes was produced by the Field Studies Council BioLinks project[3]. Three data flow 'perspectives' are identified: via Local Environmental Records Centres, National Recording Schemes and Societies and via the online platform iRecord (Figure 2.2).

Faced with a range of data entry systems, a biological recorder will typically choose a data flow route that reflects their personal experience and peer networks:

- A recorder with a focus on a particular taxonomic group, (such as dragonflies or flowering plants) will typically adopt the data flow route recommended by the relevant National Recording Schemes (NSS), including their regional co-ordinators (i.e., Vice County Recorder).
- A recorder with an interest in local conservation and a motivation for their biological records to influence local decision-making may follow data flow routes recommended by their Local Environmental Records Centre (LERC).
- In many cases there are good arrangements to share data locally, for example via a Vice County recorder affiliated to a National Scheme and working closely with their LERC. BirdTrack is a good example of a system to facilitate data flow from recorders to the NBN Atlas (as in Figure 2.2)
- iRecord facilitates flow from a recorder to the NBN Atlas on behalf of several recording schemes.

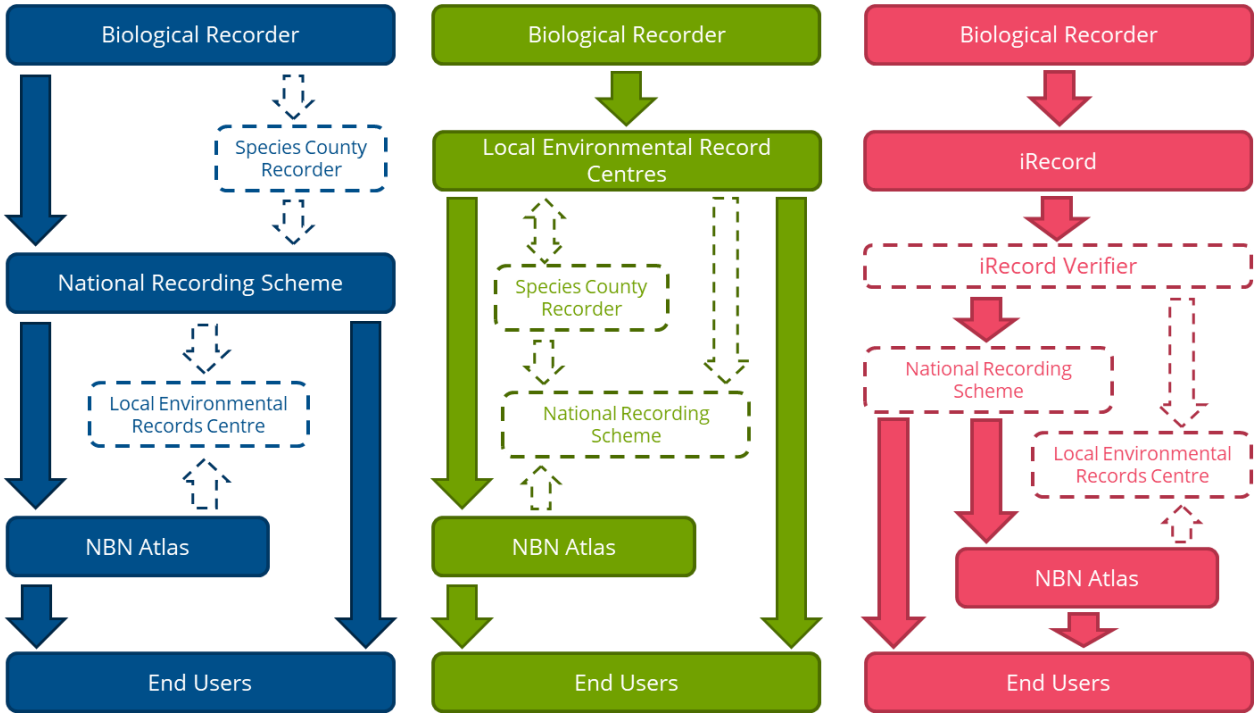


Figure 2.2: Three data flow examples

[3] https://www.fscbiodiversity.uk/dataflow

A more complex representation of current data flows was developed for the SBIF Review workshop on data flows (Figure 2.3). While this has a Scottish emphasis, the model can be considered broadly representative of the situation in England, but further review and engagement would be needed to develop a fully representative and/or ideal model. Development of a similar model is recommended for England.

The data providers to the NBN Atlas by classification and the proportion of records on the NBN Atlas and shared by each classification of data provider are shown in Figure 2.4 and Figure 2.5. Of those, 79% (118.9 million records) are supplied by the British Trust for Ornithology (BTO), with the remaining 21% (31 million records) supplied by 40 national schemes and societies. Around 75% of the species records shown on the NBN Atlas are supplied through National Recording Schemes and Societies (Figure 2.5). The large number of records provided by the BTO provides a heavy bias towards birds with over 60% of all records on the NBN Atlas relating to birds.

Around 10% of records on the Atlas, or around 20 million, are shown at high spatial resolution, 100m^2 or less. This is only a small proportion of high-resolution records; for example, it is estimated that LERCs curate over 75 million records at 100m^2 resolution or less (see Section 2.2.2). Many national schemes and societies, recording groups and environmental/conservation NGOs also hold high resolution records which are not shared via the NBN Atlas. It is not possible to accurately estimate how much data is held by these organisations without surveying them all.

The proportion of records on the Atlas by resolution and by data provider is shown in Figure 2.6. The reasons why many high-resolution records are not currently shared through the NBN Atlas are various and discussed further in Sections 3.1 and 3.2.2.

The variation in record spatial resolution is important as it is a key driver of multiple data flow pathways to data use. Spatially precise species data are essential to support most decision-making processes and some research but less important for national distribution analyses and Atlas production. Spatially precise data require more resource to collect – there are more spatial options – although this aspect is becoming less of a constraint with technological advance in survey methodology. The preferred approach is to record capture resolution, as precise as possible, and process data for use at varying output resolution.

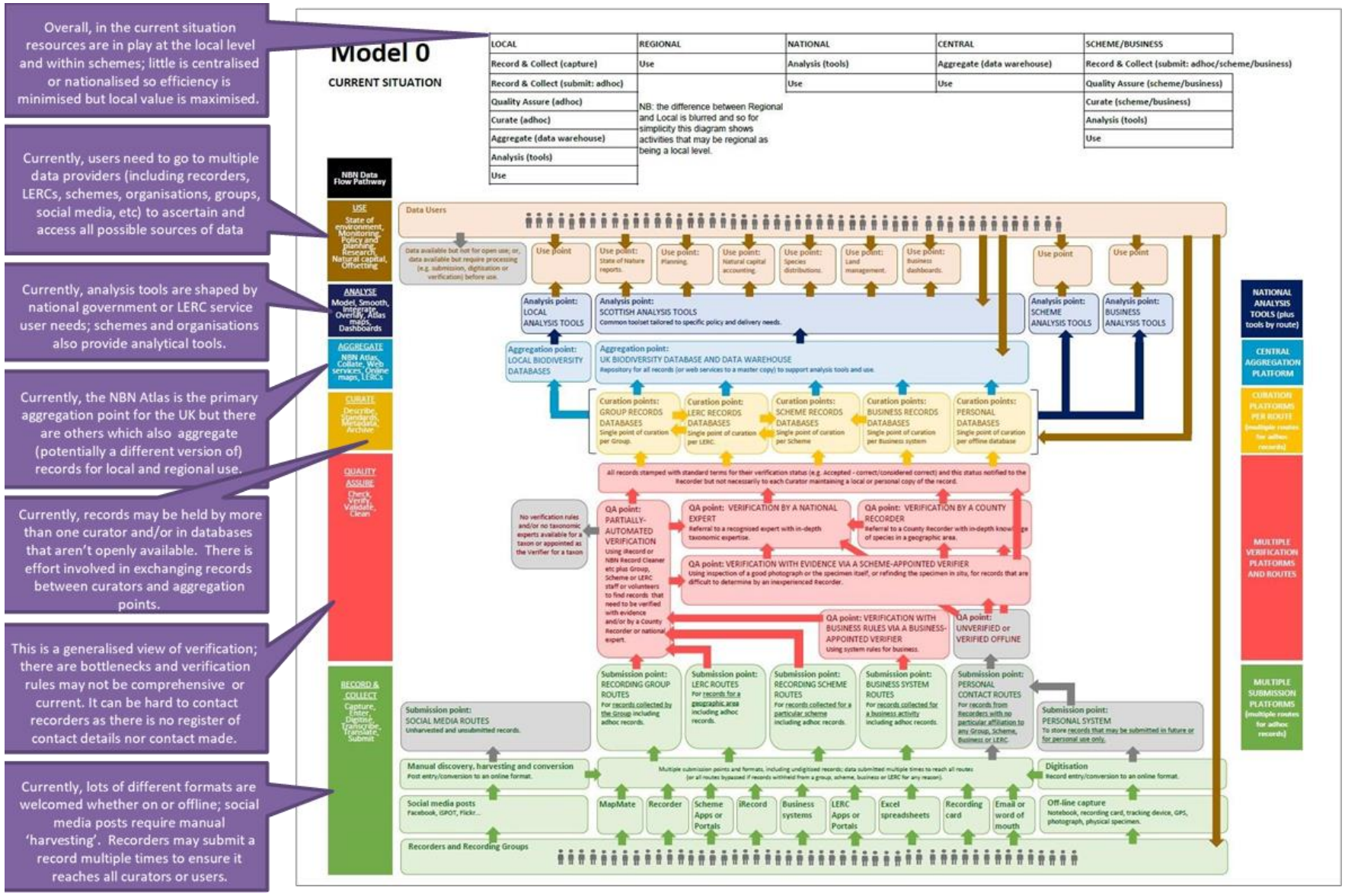


Figure 2.3: Diagrammatic representation of the current data flows in Scotland

Source: Outputs from the SBIF Review Workshop on Data Flows. See Appendix A for higher resolution.

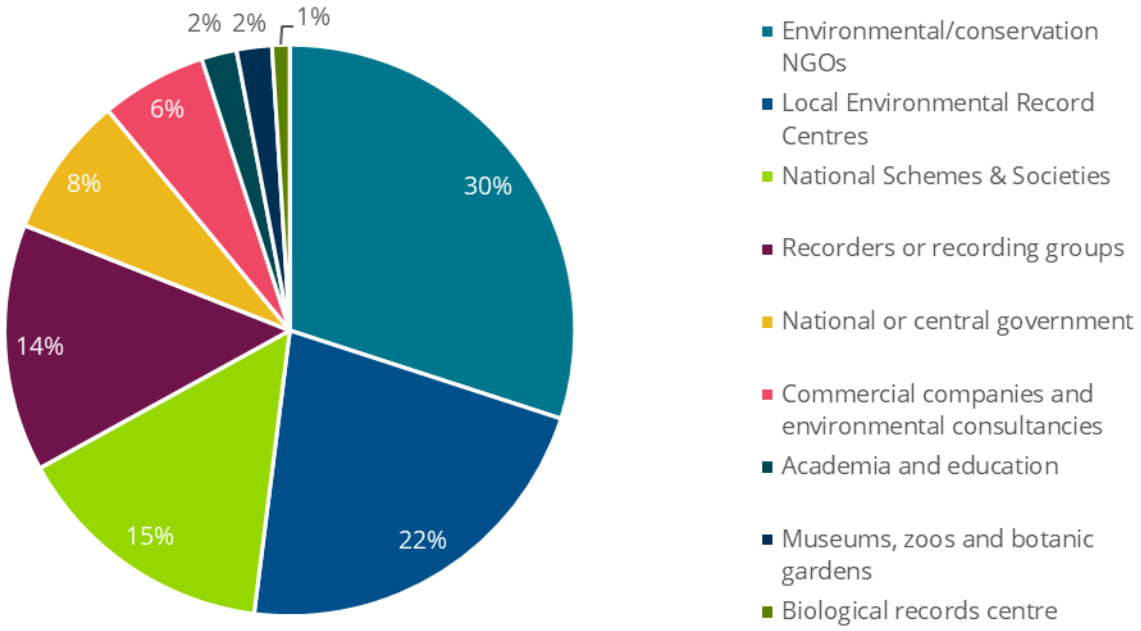


Figure 2.4: Proportion of data providers on the NBN Atlas by classification

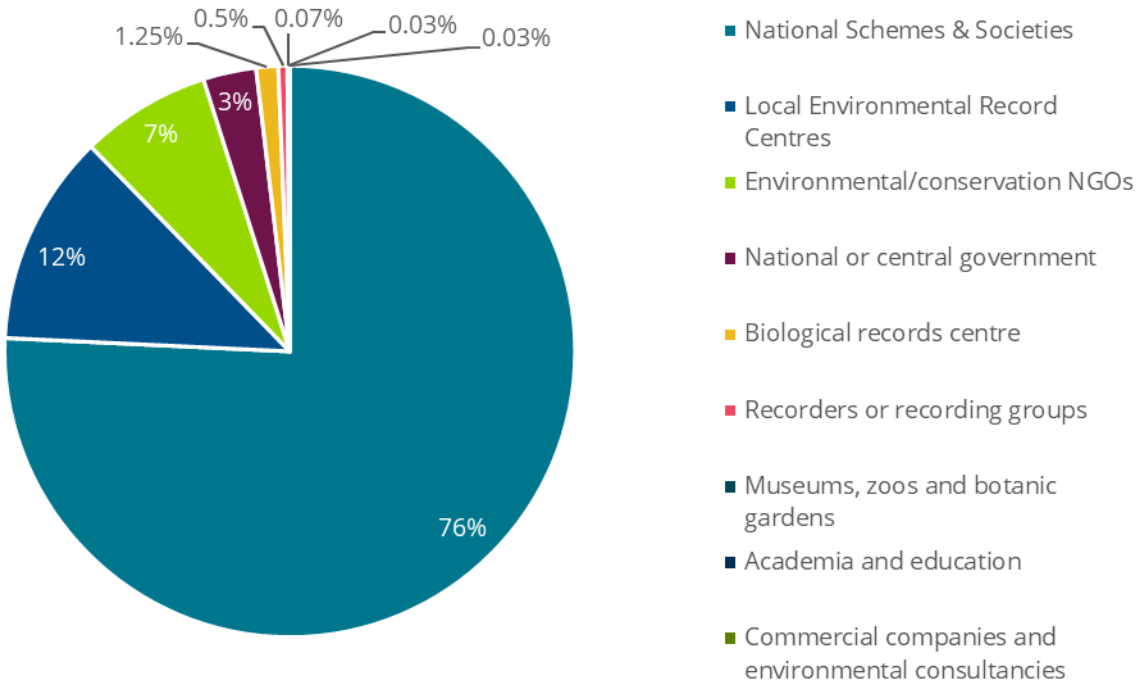


Figure 2.5: Proportion of records on the NBN Atlas by provider classification

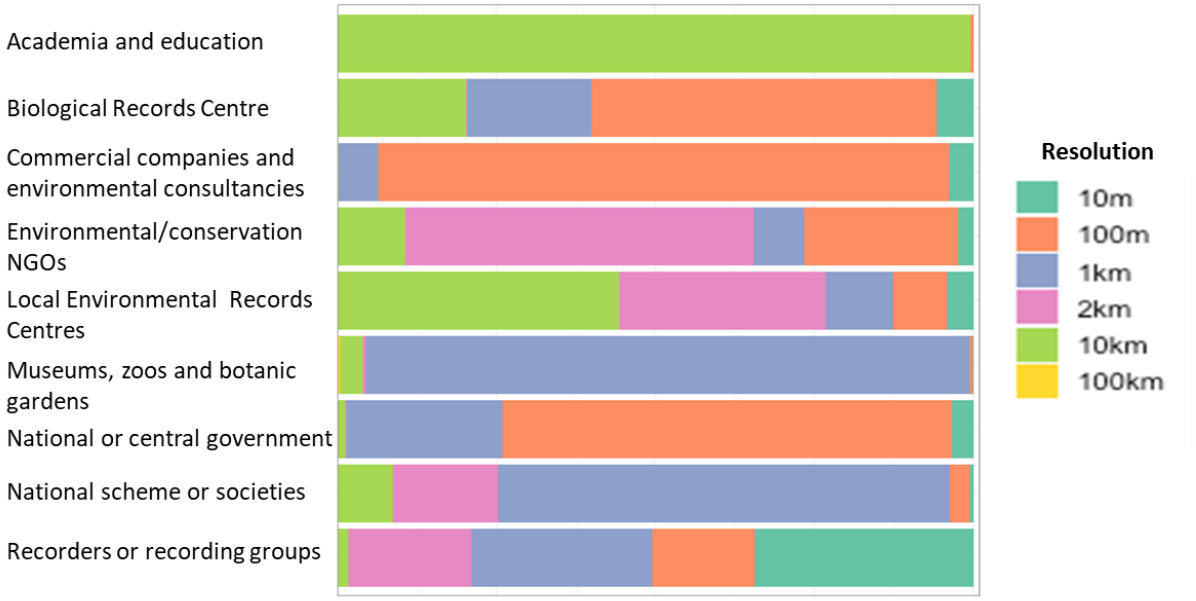


Figure 2.6: Proportion of records by resolution and by data provider classification on the NBN Atlas

The flow of marine data through DASSH to OBIS is shown in Figure 2.12.

2.2 Who is involved?

A great many organisations and individuals are involved in collecting species data in England (Box 2.1).

Box 2.1: Organisations and individuals involved in collecting species data

Recorders and local recording groups
National recording schemes and societies
Local Environmental Records Centres
Environmental/conservation NGOs
Biological Records Centre
Commercial Companies
Environmental Consultancies
Academia and Education
Museums, zoos and botanic gardens
National or central government agencies

There are five main organisations (or groups of organisations, or partnerships) responsible for managing species data: the NBN Trust, LERCs, the Biological Records Centre (BRC), MEDIN Data Archive Centres (see Section 2.6) for marine and fish, and the National Schemes and Societies (NSS). These organisations, partnerships and their people are a key part of the system in which species data flows along the 'data pathway'. Other key elements include their infrastructure (i.e., physical premises, IT capacity - hardware and software, and shared structures, including databases - some of which are shared across organisations) and the tools, applications, and practices they use. Collectively these enable stewardship of data along the pathway and its availability to users. These are referred to in this report as constituting the "Biodiversity

Data Framework" (BDF).

2.2.1 The NBN Trust

The National Biodiversity Network (NBN) Trust is an umbrella organisation, which has been championing the gathering, curation, sharing and use of biodiversity data in the UK, Channel Islands and Isle of Man for over 20 years. The Trust is the UK's largest partnership for nature working with over 200 partners under the banner of the National Biodiversity Network (the Network). The partnership includes volunteer-led taxon or region-specific recording groups, Local Environmental Records Centres, large conservation NGOs, museums, botanic gardens, universities and government agencies. The NBN Trust provides a facilitation, coordination and support role for the Network, working collaboratively with the Network partners to design, adopt and implement national and international standards for wildlife data and solve problems associated with data collection, sharing and use. The NBN Trust also hosts the UK's largest aggregation of multi-taxa biodiversity data, via the NBN Atlas[4].

The NBN Atlas is a digital data data-sharing infrastructure which provides a comprehensive, multi-taxon view of the UK's terrestrial and marine wildlife. It enables species occurrence records to be shared, visualised, analysed and viewed for multiple purposes. All data on the NBN Atlas are stored in the Darwin Core standard format and are publicly available to view and download under one of four licences, three open (Creative Commons licences CC-0 and CCY-BY and the Open Government Licence (OGL)) and one shared (CC-BY-NC) which does not permit use of the data for commercial purposes without the prior approval of the data owner. Figure 2.7 shows the number of registered users and records downloaded each month. The Atlas has information on 45,922 species. Figure 2.8 gives a breakdown of records by licence type and verification status.

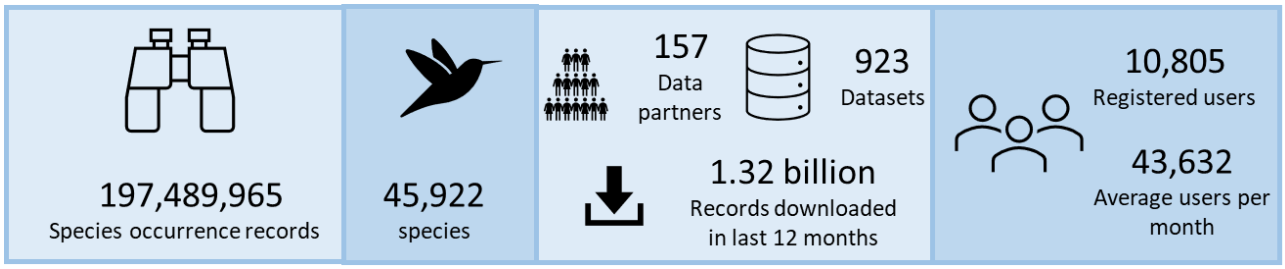


Figure 2.7: NBN Atlas in numbers

[4] Accessible at: https://nbnatlas.org/

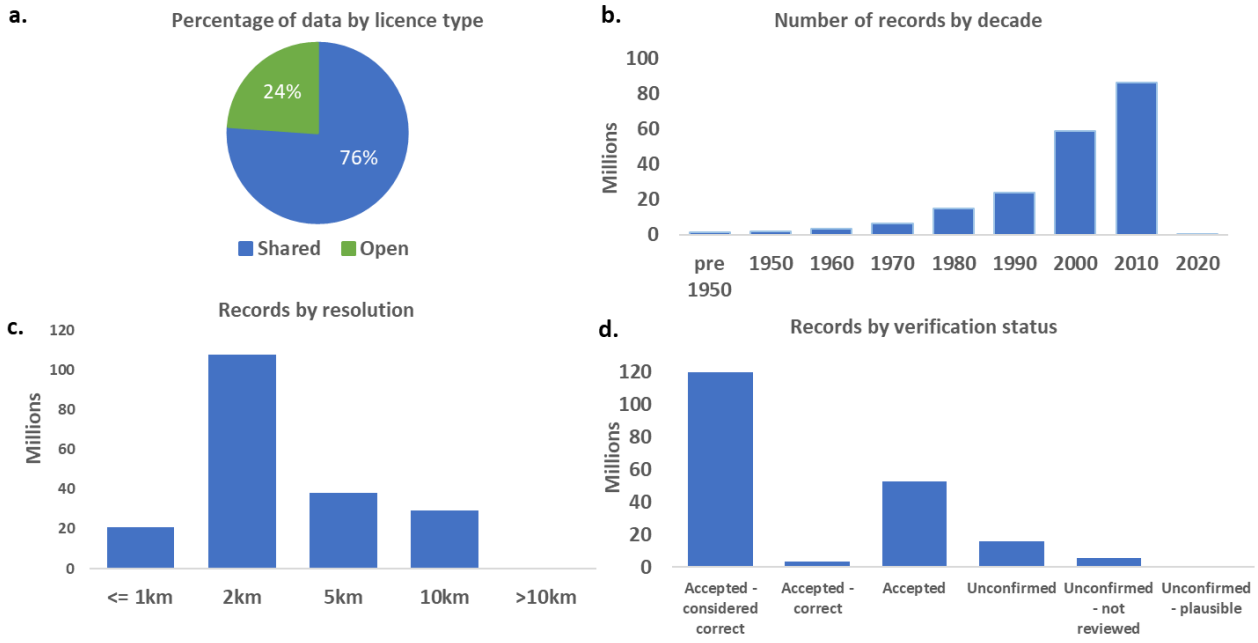


Figure 2.8: Breakdown of the records on the NBN Atlas by a) licence type, decade, resolution and verification status
Source: https://creativecommons.org/licenses/

2.2.2 *Local Environmental Record Centres (LERCs)*

The definition of LERCs accepted by Association of Local Environmental Record Centres (ALERC) members is "*... not-for-profit organisations that collect, collate and manage information on the natural environment for a defined geographic area. LERCs support and collaborate with networks of experts to ensure information is robust and make information products and services accessible to a range of audiences including decision-makers, the public, and researchers*." LERCs cover terrestrial and freshwater species and some, such as those in Devon, Cornwall and Northern Ireland, also include coverage of the marine environment.

LERCs are instrumental in the work of local recording communities. The majority have over 100 (and in some cases thousands of) individual active recorders at any one time and are continuously recruiting new recorders. They engage with the local recording community via social media, workshops and events, field recording events and newsletters. LERCs also provide many services for the local recording community (see below), the vast majority of which are provided free of charge. A summary of services provided by LERCs is provided in Box 2.2.

LERCs in England

Five of England's LERCs are independent legal entities, the remainder being hosted by Local Authorities or Wildlife Trusts (Figure 2.9). Collectively they are represented by the Association of Local Environmental Records Centres (ALERC), formed in 2009. The Association aims to provide a central voice for LERCs and to provide a support-based network of knowledge and advice for its members. ALERC also provides an accreditation scheme for those LERCs who demonstrate that they meet a set of minimum standards. Currently fifteen LERCs in England are accredited, but a requirement of ALERC membership is that LERCs not yet accredited are working towards it (ALERC, 2015). Currently there are 42 LERC members, of which 27 are not yet accredited.

Box 2.2: LERC services provided to recorders and the public and private sectors

Services to recorders include:

- Digitisation of paper records, including validation tasks
- Secure local storage and back-up of valuable data
- Training on IT systems including online recording sites and apps
- Forwarding of records to expert verifiers
- Events organisation, networking opportunities and news updates
- Physical and digital libraries and equipment storage
- Survey coordination, gaining landowner permission and gap analysis
- Small grants schemes
- Onward provision of records in support of biodiversity policies and decisions
- Feedback on use of their data

Services to the public sector include:

- Provision of verified species records in spreadsheet or GIS
- Individual reports and phone / email based advisory service
- Communication with recorders to address data gaps

Services to the private sector include:

- Individual site reports to support assessment of planning applications
- GIS layers of data for a wide area
- Digitisation of records in paper format

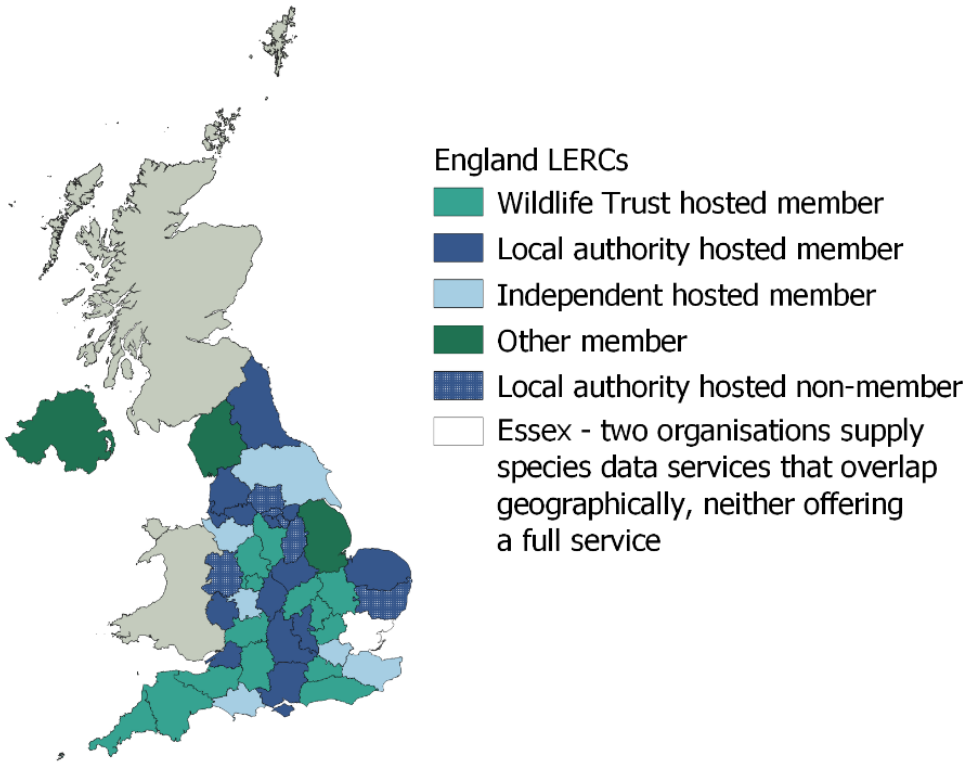


Figure 2.9: The English LERCs

Scotland

Scotland uses a mixed model to provide services and data management at the local level. Some areas are covered by established LERCs offering both community and commercial services, whilst other areas are covered by volunteer-managed databases with a reduced offer and who don't consider themselves to be LERCs.

Wales

Wales has comprehensive geographic coverage from four LERCs. These are independent not-for-profit legal entities that also work together through LERC-Wales for the purposes of sharing data with online access through a common standards data portal, Aderyn. This Biodiversity Information and Reporting Database aims to present informative data for the wildlife recording community and decision makers.

LERC Wales has a strategic partnership agreement with Natural Resources Wales (NRW) which provides NRW with online access to Aderyn and includes capturing and sharing NRW data on the NBN Atlas on an Open Government Licence (OGL).

Northern Ireland

Northern Ireland is comprehensively covered by one LERC, the Centre for Environmental Data and Recording (CEDaR), which is fully funded by the Northern Ireland Environment Agency (NIEA) and works in partnership with National Museums NI and the recording community. CEDaR covers coastal waters as well as terrestrial and freshwater habitats.

Services and data stewardship provided by LERCs

Key services of LERCS, amongst those identified in Box 2.2, include the following.

LERC Data Stewardship

A 2017 survey by ALERC estimated that over 129 million species records were curated in UK LERCs at that time. Assuming typical annual growth rates, the total in 2020 is likely to be in the region of 160 million records. It is believed that the majority of these are held at high spatial resolution.

LERC Services

LERCs deliver services to a range of users in the public, private and third sectors, all working towards the same aim, the collation and sharing of biodiversity data. While services vary according to need, charged-for services to public and private sector clients enable the LERC to provide free services to voluntary recorders and other non-commercial users.

Services to recorders

As described elsewhere in this report, much of the species data recording is performed by volunteers. Many services are therefore aimed specifically at volunteer recorders and amateur naturalists. These include annual conferences to learn about how data is being used and the latest technological developments in field recording, and workshops with a national expert to improve identification skills.

Services to the public sector

Services to the public sector are usually in the form of data provision to inform decision making. A typical

service of this kind would be a GIS layer of priority species records that a planning authority uses to inform local plan development. The Environment Agency combines data from all English LERC records into a dataset to inform their licensing activity. A public sector body might also request work to analyse the records for protected species hotspots or to target recorder effort on recorded sites.

Service to the private sector

These services include day to day work to supply data search reports to support assessment of planning applications through ecological consultants. The data requirement is normally for species occurrence records at high spatial resolution from the application site and its vicinity, together with related environmental information. An estimated 10,000 data search reports are provided to consultants per year across the England LERC network. Consultants may also require data for a large area for biodiversity audits, where the service is provided in GIS format under a year-long licence.

2.2.3 The Biological Records Centre (BRC)

The Biological Records Centre (BRC) within the UK Centre for Ecology & Hydrology (UKCEH) was established in 1964. BRC is a national focus for terrestrial and freshwater species recording in the United Kingdom and works with the voluntary recording community through support of national recording schemes, including those covering marine species. National recording schemes are independent of the BRC but are offered support appropriate to their needs and requests, while value is added by the BRC to the contribution of volunteers and their data by applying scientific expertise and a cross-taxa perspective. BRC is able to act as a conduit between the communities of recorders and potential users of the data, especially government agencies and academic researchers.

BRC supports a diverse range of outputs, such as atlases showing national distributions of over 12,000 species from over 40 taxonomic groups, and in some cases quantified trends in their distributions and abundance (now available for over 5,000 species). BRC pioneers the use of new technology for data capture (online portals and smartphone apps) and verification (including automated verification) through customisable, inter-operable database systems to facilitate efficient data flow, including to the NBN Atlas and to LERCs. (Pocock *et al*. 2015.)

BRC maintains the Indicia suite of software tools that have been used by many stakeholders to build interoperable web- and app-based species recording systems that have promoted the verification and sharing of data between NSS, LERCs and the NBN Atlas. The Indicia software is used for the iRecord[5] online wildlife recording system maintained by BRC.

2.2.4 National Schemes and Societies

National Schemes and Societies are organisations dedicated to the recording of particular taxonomic groups. Levels of recording and monitoring vary between taxa (Figure 2.10 and Figure 2.11). There are currently 85 recording schemes and societies, which come in many different shapes and sizes. The largest organisations collect and use records collected by hundreds or thousands of members of the more "popular" or easy to see species, for example the British Trust for Ornithology and birds.

[5] Accessible at: https://www.brc.ac.uk/irecord/

These larger organisations may employ staff to support their volunteers and recording work. At the other end of the spectrum are the recording schemes for the more specialist taxon groups, which are generally entirely led and run by expert volunteers, in some cases by a single volunteer. An example is the Caddisfly (Trichoptera) recording scheme, which relies on a single person to pull together and mobilise a dataset for the whole of the UK. A very large number of volunteers are involved in collecting species data and contributing to local and national recording schemes (estimated at 70,000 by Pocock *et al*. 2015) and this number is increasing through citizen science programmes.

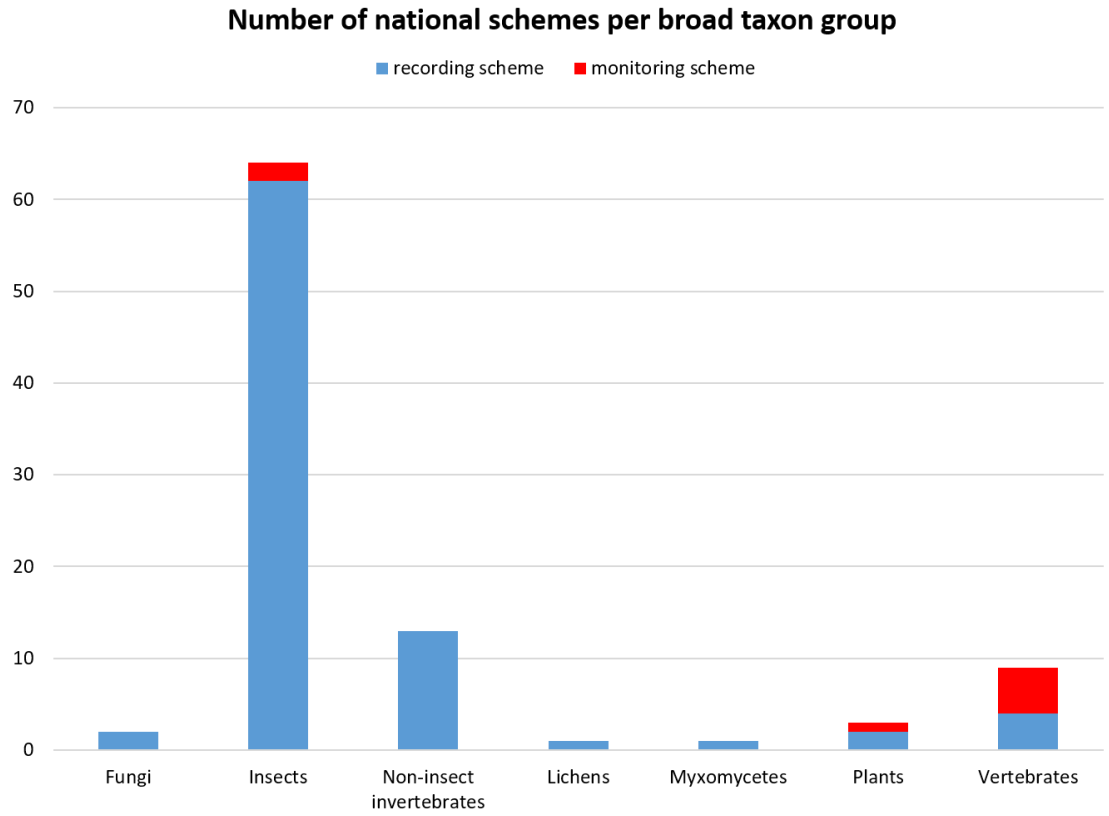


Figure 2.10: Number of recording and monitoring schemes per broad taxon group in England

There are also eight structured monitoring schemes that focus on collecting species records as part of defined monitoring protocols, e.g., the National Plant Monitoring Scheme and the UK Pollinator Monitoring Scheme. National Schemes and Societies work closely with the Biological Records Centre, who make their records available for scientific analysis, and support them in various ways, including data management, atlas publication and website hosting, depending on the requirements of each scheme. Schemes and societies have published over 120 distribution atlases, covering over 10,000 species. Data from schemes and societies have been used to analyse trends in the distribution of over 5,000 species of invertebrates, bryophytes and lichens, measured as changes in occupancy. These trends underpin government biodiversity indicators and State of Nature reports.

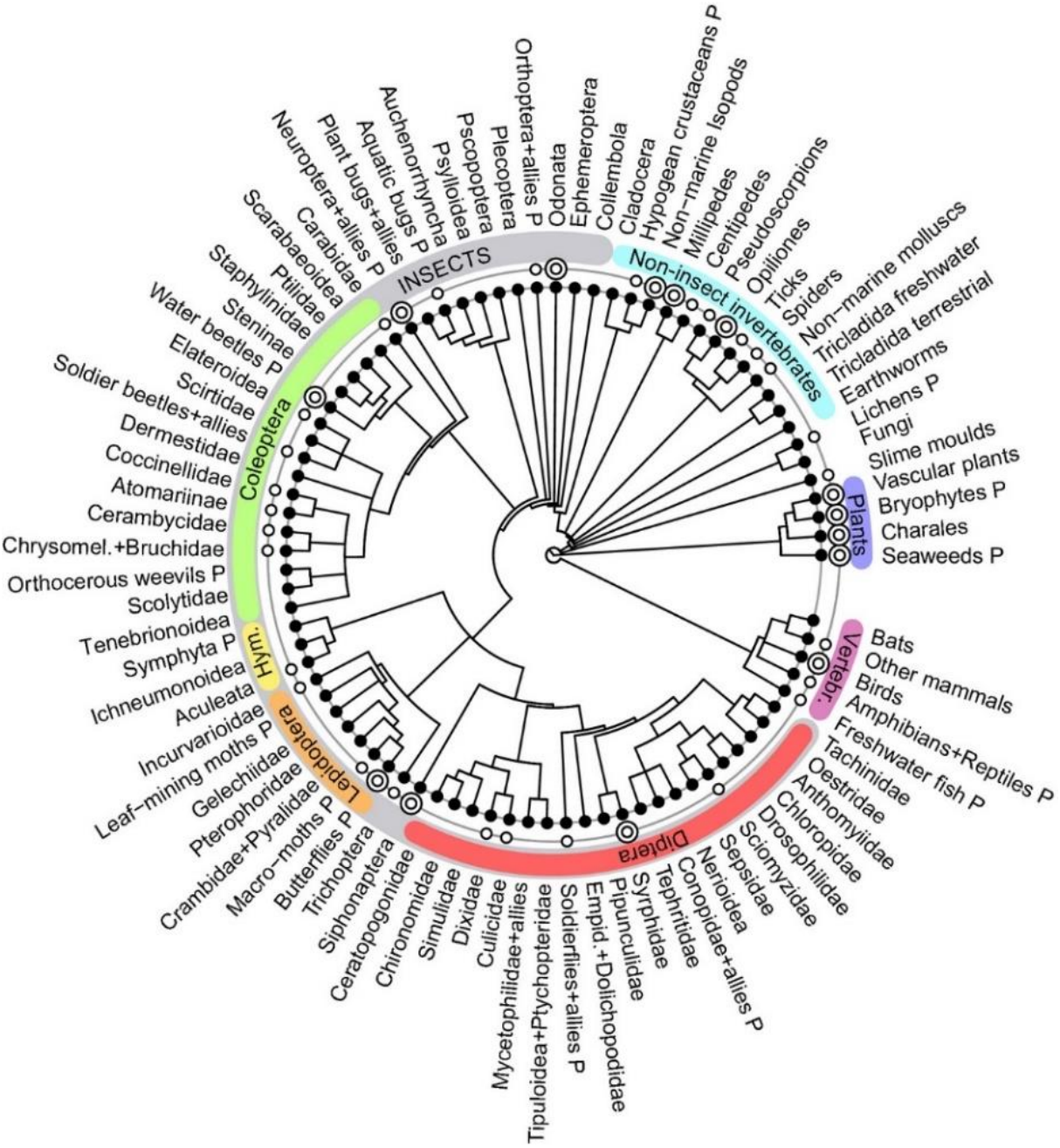


Figure 2.11: The diversity of taxa recorded through national recording schemes in the UK. Open circles indicate the taxonomic groups for which a single (single circles) or repeat (double circles) distribution atlases have been published.
Figure source: Pocock et al. 2015.

2.3 Data standards

The application of data standards provides for consistency, facilitating data management, aggregation and interoperability, saving time and providing users with a greater degree of confidence in data. A number of data standards are used in the species data flow pathway. Some are applicable only to a particular process or data type and there is variable adoption of standards across the data flow landscape. The use of standards-compliant recording tools can assist the process, e.g., facilitating format compliance, or ensuring that sample meta-data are captured. Resources to assess and fit collected data to required standards are limited. There is ongoing debate among professional and volunteer participants regarding the extent of the

benefits of applying data standards universally and more uniformly. However, there is general agreement that data standards should be in accordance with the FAIR principles (see Section 8.3.2).

Darwin Core

Darwin Core is an internationally recognised standard maintained by Biodiversity Information Standards (TDWG)[6]. It includes a glossary of terms (properties, elements, fields, columns, attributes or concepts) intended to facilitate the sharing of information about biological diversity by providing identifiers, labels and definitions. Darwin Core is primarily based on taxa, their occurrence in nature as documented by observations, specimens, samples and related information.

NBN Atlas

Occurrence records on the NBN Atlas use the Darwin Core data standard, making them interoperable with those of other countries and allowing an easier export of data to GBIF. All new datasets submitted to the NBN Atlas must be accompanied by a metadata file to describe the dataset. This includes a number of required standard attributes.

MEDIN

The MEDIN Partnership (including Cefas, JNCC and MBA/DASSH), support the Darwin Core standard for marine biodiversity data along with a number of data exchange formats. MEDIN provide a 'MEDIN Discovery Metadata Standard' for the collection of consistent discovery metadata – information that accompanies a dataset to allow other people to find out what the dataset contains, where it was collected and how they can get hold of it. The MEDIN Discovery Metadata Standard is a marine profile of the UK Government Standard GEMINI2 which also complies with other international conventions such as INSPIRE and ISO19115. The templates for this standard include mandatory fields for both 'general' and 'detailed' metadata. The mandatory fields are similar to those used for the NBN Atlas.

There are then 12 separate MEDIN Data Guidelines for the collection and curation of marine biodiversity data from different methods and approaches including benthic and pelagic data collection at sea in addition to guidance for submission of ad hoc data. These provide templates with pre-defined fields.

UK Species Inventory

The use of an agreed approach to the names of species and other taxa is essential for the sharing of biodiversity data. The UKSI provides a well-established and comprehensive set of taxon names that is widely used in the UK, and underpins the NBN Atlas, iRecord and associated websites and apps, the Recorder 6 database and other systems. The UKSI is managed by the Natural History Museum, with input from a wide range of accepted experts in particular taxonomic areas.

iRecord

iRecord uses the Indicia data model[7] which integrates with the UK Species Inventory and allows custom taxonomies to be used. Most data in iRecord are entered directly using online forms. All data inputs to

[6] https://dwc.tdwg.org/terms/ Biodiversity Information Standards (TDWG - formerly known as the Taxonomic Databases Working Group) is a not-for profit, scientific and educational association formed to establish international collaboration among creators, managers and users of biodiversity knowledge to promote the wider and more effective dissemination and sharing of knowledge about the world's heritage of biological organisms

[7] For more information see: https://indicia-docs.readthedocs.io/en/latest/developing/data-model.html

iRecord or any other Indicia based system must provide minimum required fields such as valid dates, valid grid references and controlled terms. iRecord uses controlled terms for verification status.

2.4 Public policy landscape

Species data are used in the development and implementation of a wide range of public policies in England. Only a short summary of the processes through which these policies use species data is possible here, under the broad headings of biodiversity policies, and other environmental policies.

2.4.1 Biodiversity policies

Information about species abundance, and changes in range and status have informed the drafting of UK species protection legislation and the schedules of species afforded different levels of protection, subject to licenses for controls, or identified as invasive non-natives. It has also been employed in the identification of Priority Species, nationally, and locally notable taxa. This has been reinforced by ecological understanding of species, their relationships with each other and their environment, and the ecosystem service role(s) they play. Accurate species records are essential for monitoring the condition of protected areas, the status of protected and priority species and the state of the habitats that support them. A review of the biodiversity information needs of the UK's environmental public bodies was published by Pocock (2018), but new policies could place new demands on biodiversity data in general, and species data in particular. Some examples are given below to highlight the importance of reliable species data and an effective BDF.

SSSI monitoring and assessment

The monitoring and assessment of SSSIs is an important aspect of Natural England's overall monitoring programme and provides evidence for the delivery of the UK Government's 25 Year Plan to Improve the Environment. Natural England has published a new strategy for SSSI monitoring (Natural England, 2019). This will guide monitoring of the extent and condition of Protected Sites. Data from the monitoring of SSSI features will contribute to national-level indicators on habitats, species and heritage features. It will also be used to monitor and evaluate delivery of the Nature Recovery Network, Net Gain, Local Natural Capital Plans and the Environmental Land Management System. Protected Sites monitoring is anticipated to have a greater focus on specific features moving forward, and this will include species. Data from SSSIs also contributes to the monitoring and assessment of internationally important sites and to the UK Government's international reporting requirements. It does not cover examples of non-designated priority habitats and species that may occur on a SSSI.

A reliance on data from third parties and indirect monitoring (proxies) is envisaged, and direct field surveys are envisaged only if these are insufficient or if validation of indirect assessments is needed. Recording for monitoring purposes may be done by Natural England staff, partners or volunteers, but with Natural England verifying final assessments of SSSI condition.

The standard also sets out the approach that Natural England will take in relation to data sharing and transparency. Where possible the intention is to use open data wherever possible and to make any new data openly available within 2 years at the level of detail originally captured. There is also a commitment to provide data from SSSI monitoring to the NBN Atlas (provided permission has been given to share data by

partners).

Monitoring and reporting on the condition of Ramsar Sites

The UK ratified the Ramsar Convention and designated its first Ramsar sites in 1976. The UK's ratification also extends to its Overseas Territories and Crown Dependencies, within which the first Ramsar site was designated in 1990. While the initial emphasis was on selecting sites of importance to waterbirds, other species are increasingly taken into account, both in the selection of new sites and when reviewing existing sites.

The designation of UK Ramsar sites has generally been underpinned through prior notification of these areas as Sites of Special Scientific Interest (SSSIs) (or Areas of Special Scientific Interest (ASSIs) in Northern Ireland). Accordingly, these receive statutory protection under the Wildlife & Countryside Act 1981 (as amended), the Nature Conservation (Scotland) Act 2004 (as amended) and the Nature Conservation and Amenity Lands (Northern Ireland) Order 1985 (as amended). Government and the devolved administrations have also issued policy statements relating to Ramsar sites which extend to them the same protection at a policy level as Special Areas of Conservation and Special Protection Areas.

As a party to the Ramsar Convention, the UK is required to submit a report to the Ramsar Secretariat every three years. Many sites are designated to support populations of particular species and play a key role in maintaining their distributions and abundance at national and international levels. Reliable species data therefore underpin both site-selection and reporting obligations.

The Environment Bill

The Environment Bill[8] introduces a series of far-reaching policies that will support action on targets under the Twenty-Five Year Environment Plan and assist progress towards the Prime Ministers' recent 30by30 commitment to help protect the world's oceans[9].

The introduction of Local Nature Recovery Strategies (LNRSs), to cover the whole of England is a key element of the Bill. These will contribute to a new, national Nature Recovery Network and provide a spatial basis for protecting and enhancing wildlife at local level, including the targeting of Biodiversity Net Gain and the new Environmental Land Management scheme, set to commence in 2023. They will also be critical to local plan preparation and decision making.

Biodiversity net gain will become a mandatory part of the planning process within the next two years. The proposed metric for calculating gains (Biodiversity Metric 3.0[10]) uses a habitat proxy to determine pre- and post-development biodiversity units and resulting gain obligations in a standardised way. More developers will be obliged to have regard to biodiversity – and at a much earlier stage in the planning process - than is the case at present.

A central provision of the Bill is the establishment of the Office for Environmental Protection, as an independent regulator to hold the government to account and provide scrutiny and advice to public bodies.

[8] Available at: https://services.parliament.uk/bills/2019-21/environment.html

[9] Of at least 30% protected area coverage at sea by 2030: https://www.gov.uk/government/news/uk-creates-global-alliance-to-help-protect-the-worlds-ocean

[10] The current metric version 2.0 can be found at: http://publications.naturalengland.org.uk/publication/5850908674228224

Biodiversity will be one of its areas of responsibility and will require species data to undertake at least some of its functions or is likely to expect others to use species data to meet their responsibilities.

Despite the recent inclusion of Species Conservation Strategies, the Environment Bill makes no explicit reference to species data, nor to the information infrastructure that may be needed for evidence-based development and evaluation of the policies, strategies, reports and outcomes that will flow from the Bill.

Some organisations (e.g., JNCC 2019, Natural England 2020) have evaluated (species) data needs in relation to their purposes and budgetary constraints and have produced evidence strategies accordingly. For many other organisations and for LNRS areas, equivalent strategies have yet to be developed. A holistic approach to integrate data and information systems needs of the Bill's initiatives with existing systems could enable the BDF to facilitate, add value to and benefit from the investment in these new initiatives.

In July 2020, announcing a £5 million pilot for a new Natural Capital and Ecosystem Assessment, the Secretary of State for the Environment expressed the intention to improve the baseline understanding of habitats and species abundance across the country. This would require action within every planning authority to reduce bureaucracy of processes, whilst improving the quality of data in order to make better decisions. It would also be predicated on front-loading of ecological considerations in the planning development process. This would necessarily need to consider the BDF and the benefits that better support for current arrangements could offer alongside their evolution in support of Environment Bill delivery. This would accord with the aspirations of many who contributed to this report.

Species information is only likely to become more necessary in the future. But national conservation agencies are operating within greatly reduced budgets and relying more on automated approaches, alongside a smaller structured monitoring programme. Therefore, there is a risk that a silo-based approach or a failure to understand the value and benefits of species data and the BDF might result in the opportunities offered by the Bill and the NCEA being missed.

2.4.2 Policy influences on species data collected

The use of species records and derived data in England is partly driven by the operation of policies at the UK, devolved nation or local level, as well as existing obligations and commitments at the European and global scale. These policies influence the creation of species records, the need for curation and analysis, and the extent to which sampling is coordinated and systematic. They have also influenced the flow of resources in support of these activities over the years.

Information about species' abundance, and changes in range and status have informed the drafting of UK species protection legislation and the schedules of species afforded different levels of protection or identified as invasive non-natives. It has also been employed in the identification of locally notable taxa and national Priority Species. This has been reinforced by ecological understanding of species, their relationships with each other and their environment, and the ecosystem service role(s) they play.

Species information has been used more widely in policymaking at local, England and UK level in relation to:

- Reporting on the state of the natural environment.
- Developing nature conservation strategies and biodiversity action plans.
- Designating protected sites of national and international importance, nature reserves and local wildlife sites.
- Deciding how species of local to international importance are dealt with in the planning system; and
- The targeting of resources by statutory and non-statutory bodies and environmental funders in relation to land use, species recovery and landscape scale conservation management.

A range of policy requirements necessitate or depend on the use and/or collection and management of certain species data, either as a mandatory obligation or to conform with recognised standards and best practice guidance. Different policies demand different types of supporting evidence and may also impose specific requirements with regard to sampling, data formats, curation, meta-data, analyses and dissemination.

In response to these policy requirements, species data provision is also driven by factors such as the facts that:

- Different uses have greater or lesser importance from meeting statutory requirements, enabling or constraining economic activity, protecting or increasing the value of natural capital, enabling responses to trends and raising awareness, depending on stakeholders' goals and perceptions.
- The value of an individual species-record or dataset will depend on its utility and the extent to which it can be used for different purposes or will add to the value of existing data in meeting information needs.
- Species data are more likely to be collected and available where there is an incentive to collect and share, and this isn't precluded by cost or difficulty.
- Where representative species datasets can be accessed at proportionately low or no cost to the user, they are more likely to be used.

It is for a combination of these reasons that effort and resources have been directed more into the recording of groups such as plants, birds, butterflies and, more recently, bats and whales, than other, more diverse, species groups, together with recent marine recording efforts linked to the creation of Marine Protection Zones. However, this has also funnelled limited resources, encouragement and effort away from less-well recorded groups or more holistic species recording support.

The 2019 State of Nature report (Hayhow et al., 2019) highlights a general, ongoing decline and loss of many priority and other species in terms of both distribution and abundance. Our ability to monitor trends is only possible for a small fraction of the UK's biota and we don't presently have reliable data even for all priority and threatened species. The lack of a suitably comprehensive species evidence base creates challenges for the targeting and monitoring of effective environmental biodiversity policies.

There are a number of key issues relevant to the development, effectiveness and viability of the BDF in relation to national policy and guidance, and the resulting local policies. One is a need to identify and quantify our species and other environmental data needs, the gaps in provision and these, how these might

be addressed, and the infrastructure required to address these. A 'State of Nature Data' report could provide a helpful first step.

Land Use Planning

In the planning system, ecological assessments focus on habitats and protected species. Other priority species are rarely mentioned despite being of material consideration and it is unusual for reports to assess habitat quality or environmental conditions based on other species occurrence even where data are available. Despite industry guidelines, Local Planning Authority (LPA) policies and the NBN Trust's creation of the Consultants' Portal, relatively few species observations collected in connection with development are shared with LERCs or the NBN Atlas. This highlights where a lack of obligation on the main actor is presently failing to support the BDF, even though the surveys are funded, the information collected for public decision making, and the purpose is financial gain.

At a local level, alongside the changes resulting from the Environment Bill, many councils' operations and policies will also be influenced by their recognition of the climate and biodiversity emergencies.

UK Marine Strategy

The UK Marine Strategy provides the framework for delivering marine policy at the UK level and sets out how we will achieve the vision of clean, healthy, safe, productive and biologically diverse oceans and seas. The UK Marine Strategy consists of a simple 3-stage framework for achieving good environmental status (GES) in our seas. Achieving GES is about protecting the marine environment, preventing its deterioration and restoring it where practical, while allowing sustainable use of marine resources. The strategy covers 11 elements (known as descriptors) including: biodiversity; non-indigenous species; commercial fish; food webs; eutrophication; sea-floor integrity; hydrographical conditions; contaminants; contaminants in seafood; marine litter; and underwater noise.

Monitoring and reporting on the status of Marine Protected Areas

There are national and international commitments for networks of Marine Protected Areas (MPAs) that are ecologically coherent and well managed (Agnesi et al., 2017). In addition to commitments under the UK Marine Strategy, the UK has committed to several international agreements on MPAs, including under the OSPAR Convention. Many habitats and species protected within the UK MPA network are listed by OSPAR as threatened and/or declining. Reliable data on species are therefore important for national and international reporting. Such data is essential for supporting the designation, monitoring and management of Marine Protected Areas in UK Secretary of State waters. Between 2011 and 2019 the partnership program coordinated by Cefas to collect seabed species and habitats evidence from MPAs has conducted more than 140 surveys from more than 90 MPAs. In excess of 60 MPA sites have had partial or 100% cover seafloor acoustic data collected, more than 7,500 sediment samples have been analysed in respect to physical nature of the substrate and biological communities present, more than 48,000 minutes (8000 hours) of video have been collected and analysed, and in excess of 90,000 still images have been collected and analysed to record epifaunal communities. Outputs from this work total over 100 reports produced to date. JNCC has developed a standardised catalogue of UK MPAs and their protected features (The UK MPA Stocktake (JNCC, 2020)) This was considered to be important at a UK level because of the different approaches used by each country in identifying MPAs and monitoring their features. It includes an MPA network features list, including OSPAR threatened and declining habitats and species as well as features

with status through the NERC Act (2006).

Data collected as part of the MPA programme is made publicly accessible via MEDIN Data Archive Centres

2.5 Funding Landscape

While the majority of terrestrial or ad hoc biological records in England may be collected by volunteers, there are costs incurred at every step of the process from observation to making data available for use. Costs here are referred to in the broadest sense, not just entailing direct financial costs, but the costs of time and in-kind resources:

- Recorders incur time and travel costs in collecting and submitting records.
- Organisations receiving records (recording groups, schemes and societies, LERCs etc.) incur costs in verifying species identities and validating records, along with the costs of maintaining data management systems.
- National aggregators/data sharing platforms (e.g., iRecord, the NBN Atlas and MEDIN) incur the costs of their own data validation processes and the infrastructure required to expose the data to verifiers/ users.

The State of Nature Report 2019 estimated that the financial value of volunteer effort to conservation in the UK was £20.5 million per annum, reflecting the effort of 18,700 volunteers contributing to structured monitoring schemes (Hayhow et al., 2019). An additional 70,000 volunteers also submit biological records to National Recording Schemes or Local Environmental Records Centres (Hayhow et al., 2019).

Public funding supports species recording through the government agencies, local authorities and research bodies. However, the wider data stewardship aspects of the BDF are under-supported and not all funded recording results in species data becoming widely available. Some recording schemes receive no funding at all and while others may receive funding for data collection or monitoring schemes, most bear the full costs of data management. The Biological Records Centre, which supports schemes and societies and hosts and manages iRecord, receives funding from JNCC and UKCEH (through National Capability funding via the Natural Environment Research Council (NERC)).

The general list of major funding sources for LERCs includes:

- National government agencies
  - The Environment Agency
- Local Planning Authorities and local county councils
  - Borough councils
  - County councils and combined authorities (often providing an ecological service to a region)
  - National Park Authorities
- Private sector services
  - Individual site reports for ecological consultants supporting planning applications
  - Service Level Agreements with businesses with ecological impacts, typically water companies

Staff costs are the largest component of any overall budget. Other costs, such as financing premises, can be variable and often depend on arrangements with a host organisation (e.g., County Council or NGO). Sources of funds for recording activities vary between LERCs, an important factor being the number of local planning authorities in their area and the willingness of those local authorities to engage with the LERC. Funding sources are changing, with reduced levels of funds coming from national government agencies and more from private sectors.

On top of the data management costs, recording groups, schemes and societies, LERCs etc also provide training, support, engagement and sometimes equipment for recorders. This is constrained by finance.

The NBN Trust and NBN Atlas currently receive 43% of their funding from the Country Nature Conservation Bodies and JNCC, the rest being made up of grant funding and membership fees (total income in 2019/20 was £439,508). Funding from government for the NBN Atlas covers hosting costs and basic maintenance; there has been no government funding for development or improvements to the NBN Atlas since its launch in 2017. Limitations on resources not only affect how much work can be done on the NBN Atlas, but also the amount of support the NBN Trust is able to give the NBN Network in adopting standards, assisting with data management and solving the issues associated with data sharing.

The costs (and benefits) of the BDF are explored in more detail in Report 2.

2.6 Marine species data

Marine data collection and science face unique challenges due to the complexity of the marine environment and high cost of data collection. At the same time, the importance of marine species data has increased with expanding human activity in the marine environment, requiring a stronger focus on managing human impacts to reduce the pressures on marine ecosystems. Therefore, any action that makes collation, analysis and access to data easier and cheaper will continue to deliver benefits for all users of the marine environment and wider society. The Marine Environmental Data and Information Network (MEDIN) was established in April 2008 for this purpose. Its objectives are to improve access to, and the management of, UK marine data and information through:

- A Partnership Network of 50 organisations working collaboratively with shared aims
- Accredited Data Archive Centres (DACs) for long term curation of data.
- A web portal to make it easy to find UK marine data
- Data Guidelines
- Training provision and support.
- A Discovery Metadata Standard to provide information about datasets.

MEDIN implements a partnership approach to meet these objectives with funding from a consortium of fourteen sponsors (mainly from the government sector), active participation in Working Groups and partnership agreements with various government departments, agencies, research councils and private organisations to assist with ongoing development, implementation and dissemination of MEDIN. MEDIN reports to the Marine Science Coordination Committee (MSCC) and coordinates the UK Marine Sector

response to the EC INSPIRE directive, for sharing environmental information across Europe, and to data.gov.uk, the UK government's portal for public data.

MEDIN works through a number of themed Data Archive Centres (DAC), including DASSH with its biodiversity data scope. The MEDIN partners including MBA/DASSH, Cefas and JNCC play a major role in supporting primary data collection in the marine environment and networking with partners and data stewardship organisations.

Organisations involved in collecting marine environmental data in UK waters are encouraged to submit data to a relevant DAC and/or metadata to MEDIN. This includes both public and private sector organisations. Data and metadata can be contributed through completion of simple forms. Other organisations can then find data online through the metadata discovery portal or through searches on DAC portals. Search results can be exported, and data can be downloaded where available and many datasets are also made accessible using web mapping services, increasingly API's and with Digital Object Identifiers for citation purposes, for example in peer review publication. In general, over 85% of the data described in the MEDIN portal is downloadable from the MEDIN DACs.

The Archive for Marine Species and Habitats Data (DASSH), an initiative of the Marine Biological Association, is the UK Data Archive Centre for marine biodiversity data for both species and habitats. Accredited through the MEDIN partnership, and core-funded by the Department for the Environment, Food and Rural Affairs (Defra) and the Scottish Government, DASSH provides tools and services for the long-term curation, management and publication of marine species and habitats data, within the UK and internationally. DASSH is committed, with its partners and data providers, to the FAIR Data Principles, to make marine biodiversity data Findable, Accessible, Interoperable and Reusable. Through partnerships with other UK, European and international data centres DASSH contributes to data portals including the NBN Atlas, EMODnet, EurOBIS and GBIF.

Seasearch is a project co-ordinated by the Marine Conservation Society for volunteer scuba divers and snorkellers who have an interest in what they're seeing under water, want to learn more and want to help protect the marine environment around the coasts of Britain and Ireland. The main aim is to map sea-bed types and species distributions, establishing the richest sites for marine life, the sites where there are problems and the sites which need protection.

CEDaR, the LERC for Northern Ireland, has a significant role in marine data for Northern Irish coastal waters. A number of coastal LERCs and Wildlife Trusts have marine survey and data operations.

2.6.1 Key elements of marine species data flow

While there is a degree of overlap in taxa between marine and terrestrial/freshwater environments, this accounts for a small minority of species. Marine species data requirements and flows therefore need additional consideration.

Marine data flow pathways, although moderately complex (see Figure 2.12) are more structured and understandable than for terrestrial/freshwater data. This may be a consequence of the more recent

development of systems, fewer organisations involved and higher proportion of professional recorders.

Marine species data systems use international database reference systems, such as The Register of Marine Species WoRMS, and database applications specialising in the marine environment, such as Marine Recorder. They are already in the process of applying FAIR data principles and are progressing towards more advancement in terms of machine-discoverable datasets that comply with Darwin core data standards. There is also a clearer linkage between marine species data systems and wider environmental data e.g., the role of DASSH as the marine biodiversity data archive centre, linking with MEDIN.

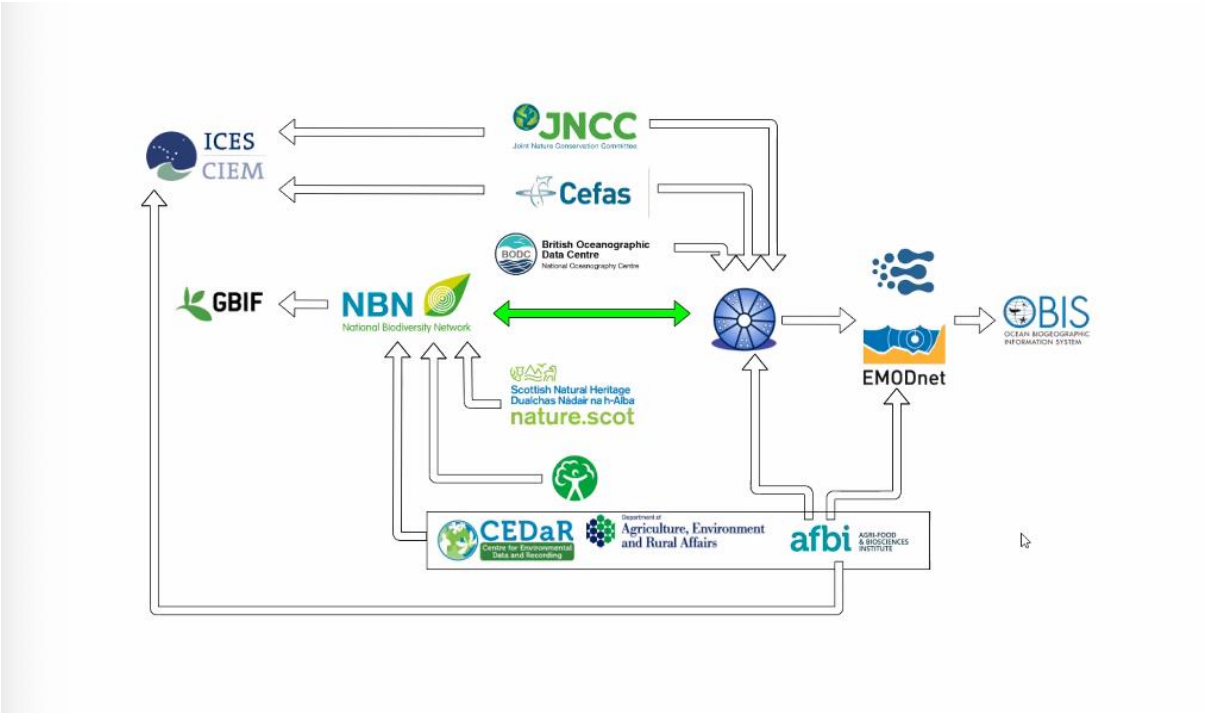


Figure 2.12: Marine data flow between DASSH (shown as blue shell) and UK and international data portals.

Relatively more marine data are professionally collected than by volunteers, except in coastal and nearshore environments. This is at least partly a consequence of the high cost of sending vessels offshore, advancements in more automated data collection and data storage infrastructure costs.

The proportion of the BDF funded by the public sector appears to be higher in the marine environment. Much of the primary data collected is funded by the public sector and made available through Open Government Licencing. Consequently, the private sector contributes very little to data stewardship costs. Nevertheless, marine data stewardship remains under-resourced (Agnesi et al., 2017). There are concerns about the sustainability of the Marine Recorder and other databases and the resourcing of work to capture, verify and make accessible datasets supplied in non-standard format.

3. Issues and challenges

This chapter provides an overview of some of the main issues and challenges associated with the species data pathway (described in Section 2) in England, identified by stakeholders (through the processes described in Section 4) or the study team. The opinions expressed in stakeholder interviews carried out for this study align, at least to some extent, with conclusions of other studies, including SBIF and the review carried out by JNCC of biodiversity data use in the Country Nature Conservation Bodies (Hassall *et al.,* 2020), but there are some significant points of departure.

The species data pathway has grown organically over many decades in response to volunteer enthusiasm for nature and a plethora of government, NGO and private sector influences and requirements. Over a hundred legal entities are involved in various ways - many more than in Scotland – each with its range of interests, expertise and goals. This diversity has produced both successes, as expressed by the quantity and quality of species data aggregated through the NBN Atlas, LERCs and NSS and the positive views of many data users but also significant frustration at its deficiencies and concerns for sustainability and capacity to deliver escalating species data demands arising from the ecological emergency.

Limitations to current biodiversity data workflows identified by Hassall et al. (2020) include:

1. **Data access**. Approximately 50% of potentially useful data were identified as "currently inaccessible". Reluctance to share data to the National Biodiversity Network (NBN) Atlas and other repositories was identified as an issue, attributed to concerns around commercial exploitation of data. The study also identified the existence of some key datasets (for example data from environmental impact assessments and historic datasets) that are not readily accessible and could be mobilised for use.
2. **Data flows**. There are too many submission portals and routes, causing confusion around "who is collecting what from where". Data flows are subject to significant time lags.
3. **Data coverage**. There are important taxonomic and spatial gaps in data, especially outside protected areas.
4. **Data quality**. Some potentially useful data (for example data collected through citizen science projects) need verification to improve confidence in their use in decision-making.

Stakeholders in this project echoed these concerns and amplified some aspects (Table 3.1).

Many stakeholders expressed concern that existing species data are poorly used in planning applications affecting biodiversity. Development proponents and their specialist advisers fail to use existing species data and this omission adversely affects the quality of decision making. One planning authority reported that over 90 % of biodiversity-sensitive applications failed to use best available data and a 2016 report for the Greater London Authority found that while 18% of planning applications potentially impacted on biodiversity, data searches were made for only 1% (GLA, 2016). Similarly, the data generated by environmental assessments and other planning processes are poorly shared.

Issues identified by stakeholders and the extent of agreement are summarised by theme in Table 3.1.

Table 3.1: Summary of stakeholder views

Biodiversity information needs

Species data coverage

- Species data needs have not been clearly identified.
- Data needs are being met for some "headline" indicator species, birds, cetaceans, bees & bats), but there should be more use of a wider range of indicator species, with less focus just on protected species.
- Data needs are well-defined for a small number of taxa (birds, protected species) but not for biodiversity as a whole
- Infrastructure exists but is complex and needs to be better organised and more consistently used.
- Future needs are as yet unspecified in terms of infrastructure.

Environment Bill and policy review

- Recording will be very important to support effective monitoring of policy implementation and outcomes. A focus solely on broad habitat outcomes (as in Biodiversity Net Gain policy) will not be sufficient.
- Citizen Science will become increasingly important.

Recording effort and methodology

Obstacles

- It is not always clear how to get involved.
- Fragmented systems are confusing for recorders
- Recorders have many motivations and not all are able or willing to be steered, e.g., to 'fill gaps.
- There are gaps and bias in terms of species and locations covered by recording and there is also recorder-bias, requiring improved validation methods.
- Improved ID skills are needed.
- Demographic of volunteers: typically, the older generation.

Measures to address obstacles

- To inspire recorders to take part, there needs to be much more dynamic feedback showing how records are used, with 'live' updating.
- There will never be enough recording to fill all the gaps for all taxa, and using models based on available data is the way forward.
- Better training and improved access to training is needed.
- More structured recording/citizen science projects.
- Include social aspects.
- Administrative/coordinating roles should be paid, but not necessarily volunteer time.
- Further development of handheld apps.

Novel sampling methods

- Able to use less skilled volunteers.
- Do not necessarily deliver higher quality data.
- Need consistency in survey methods.
- Good for more elusive species.
- Cannot replace traditional recording but can support and complement it.

Data flow

- Verification is a bottleneck: a huge problem in onward sharing of data. Not all stakeholders are convinced that automated verification can work effectively.
- Data are not being shared widely enough by some NSS, LERCs, schemes etc.
- External funding & support needed to speed up verification
- Data standards could ease data sharing.
- Consultants – make it a requirement/legislate to share their data.
- Make apps interoperable with rest of infrastructure e.g., LERCs or NBN Atlas.
- Quality assurance – sub verifiers would help as well as increased used of automated verification.
- Should be a legal requirement to share data collected in planning system.
- Should be an obligation for academics receiving public funds to share data that they generate.

- Verification is an essential process that adds value and enables data to be trusted for use.
- Authoritative sources of data should be known and publicised
- Variable data quality causes problems – not all records marked as 'verified' are correct and provenance should be captured and available for all data.
- Data from some consultancies and research projects cannot always be used as the quality is too poor.
- There will always be multiple systems, enabling flow between them is critical.
- Data flow between national and local levels is still poor, despite being a longstanding problem.

Administration

- Funding – hybrid model might be the best option, but opinions differ regarding preference for a focus on local databases, a single centralised database or a mixed approach.
- Need to make good use of local knowledge through effective and well-funded LERCs.
- Should adopt FAIR data principles.
- Different views regarding focus on local databases (Q20b) vs Mixed approach (Q20d) vs Single centralised database (Q20a).
- Views differ regarding making all new data open by default.

Data Use

- NSS are valued for expertise, atlas production and supply of species distribution data through NBN atlas.
- NBN Atlas is a valuable resource for research and signposting to detailed datasets.
- LERCs are valued for local knowledge and networking and supply of detailed data for decision-making.
- Users want verified data and more clarity on fitness-for-use information through use of metadata.
- Data gaps in NBN Atlas need to be filled to make it even more useful.
- LERC Services need to be broadened and improved in many parts of England.
- Consistency and joined-up working across LERC boundaries would improve accessibility of data for decision making.
- Faster response times, including direct access to online data through APIs is desired by many users.
- Data relied upon for peer-review publication should be made accessible using Digital Object Identifier links by default, especially where publicly funded (public and academic sector)
- Where data is made available under license outside of OGL, greater clarification is needed on permissible re-use.

3.1 Efficiency of data flow

As explained in Section 3.2, the current species data landscape has complex data flows and pathways and suffers from a lack of clarity about submission routes for new data. A key finding of the SBIF Review was that poor data flow hampers access to/ use of biodiversity data collected in Scotland and the evidence collected for this report shows that the same is true in England.

Ongoing diversification of routes for data submission has increased the volume of species data collected but has also increased the complexity of flow pathways. More effective integration of data pathways would improve the potential of species data for a range of uses. Although the current diversity of systems reflects the rich history of biological recording in Britain and Ireland, it can be bewildering for those who want to contribute wildlife observations to support conservation and research.

There are several reasons for the current complexity of data flows:

- The long history of biological recording in the UK and the large number of organisations that play a role in it.
- The reputational need for those organisations to be recognised as being data providers and/or

verifiers, and the temptation to judge success by how many records can be claimed under an organisational banner. These perceptions may be perceived as having financial consequences.

- There can also be a reputational need for individuals or groups to be recognised as 'the expert' in a particular context, which can lead to them being unwilling to share the data that supports their expertise.
- Although all species records share some core commonalities, the additional attributes that need to be captured vary depending on the purpose for which the data is being collected, and different contexts enforce different priorities on what is important about a record.
- The tension between wanting data to be shared and used, while at the same time needing to control who can access it, (for example concerns over data being 'free' for commercial use, or data on sensitive species being misused). Many recorders object in principle to their data collected voluntarily, often for conservation motivations, being freely available for commercial use, perceived often as potentially harmful to nature. While data shared under non-commercial licence should allay these concerns there needs to be an effective compliance system in place to prevent abuse. Concerns may also exist for potential abuse of data relating to sensitive species; such records are often shared at less precise spatial resolution than the original capture resolution.
- Potential conflict over who decides which data can be used and how – e.g., different views over data sharing may be held by the original recorder, the record collator and the verifier.
- The difficulty of tracking verification decisions, and of bringing such decisions together (e.g., if a local and national verifier have each played a role).
- Technical innovation across many years, combined with often insufficient funding to fully support the development, integration and maintenance of such innovations, and an unwillingness among stakeholders to adopt new systems (often for good reasons).
- The difficulty of maintaining trust and common purpose among many different organisations, with different priorities, different funding sources for professional work and the complication of combining this with volunteers, especially the limited number of expert volunteers. This is especially difficult with data quality issues, where there is often mistrust over the standards being applied by various parts of the system (some of this being based on well-documented examples of lapses in quality).
- The need for constituent parts of the network to be able to guarantee that they can provide data within a known timescale for their particular work, which can mitigate against the time it takes to ensure that data are shared and peer reviewed.
- The tendency for funding bodies to prioritise new, stand-alone projects over the continued development and support of existing ones (perhaps less of a problem now than it used to be).

Many of these aspects of biological or species recording were recognised by the SBIF review. The SBIF recommendations to transform data flows in Scotland are reviewed in Report 2 and assessed for their feasibility and potential application in England. Report 2 will also consider economic aspects of the SBIF recommendations.

The recommended data flow model for Scotland is based on the concept of a central data warehouse and curation tools (Figure 3.1), which is fed by data from recorders and recording groups as well as harvesting of records from social media. Verifiers have access to the central data source to provide quality assurance alongside automated approaches. Data users access the central data source through dataset services, giving a range of data access points.

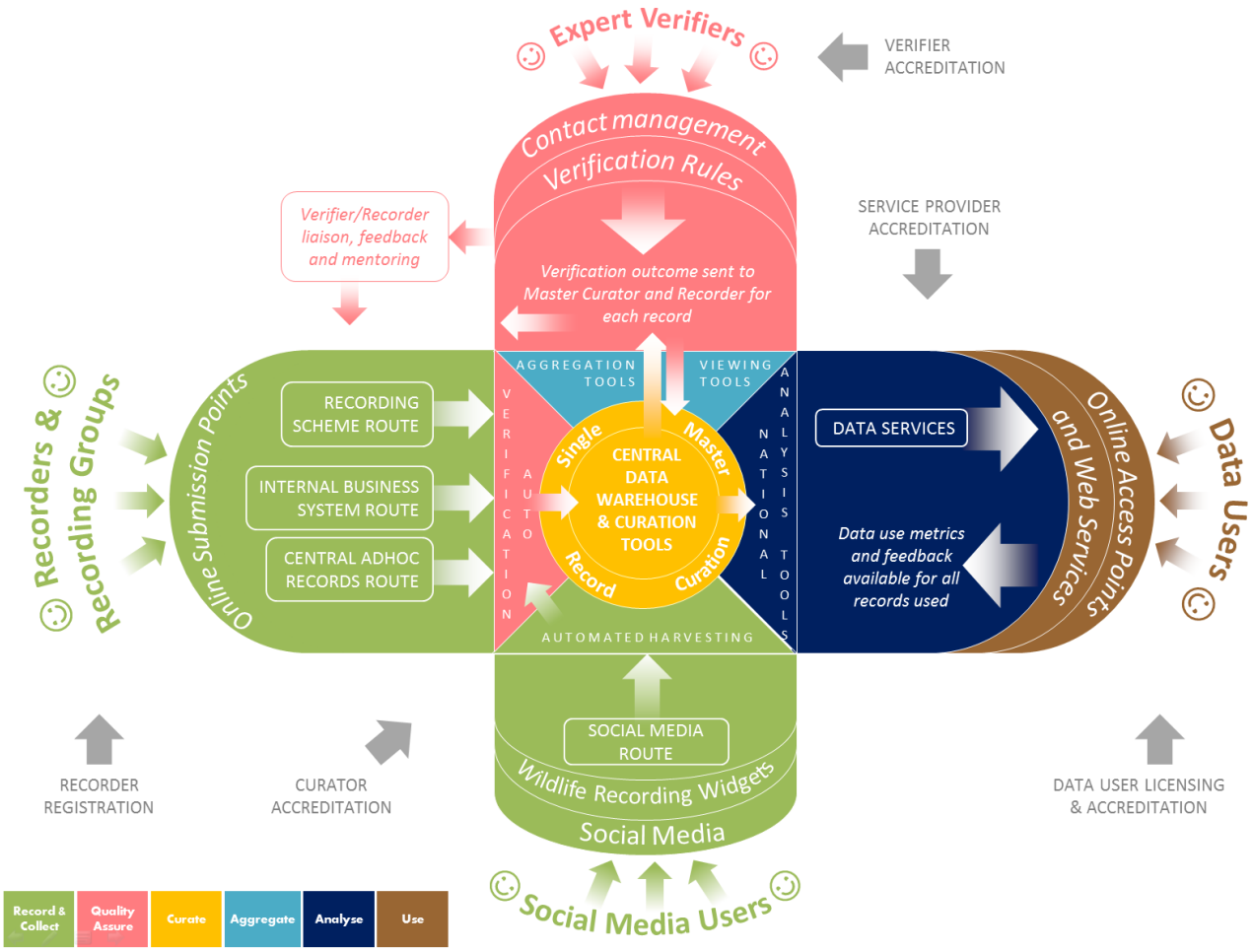


Figure 3.1: Potential data flow recommend for Scotland
Source: Wilson et al., 2018

Although appealing, this idealised data flow model is far removed from the current approach in England, where data capture, collation, verification, and use are highly fragmented. Much of the *ad hoc* biological recording which is currently undertaken by volunteers has the potential to continue in current ways, but better coordination is needed. In addition, there is an unknown number of organizations and individuals holding data and not sharing them openly (Groom *et al.*, 2017). Recommendation 3 relates to improvements in Data Flow efficiency.

3.2 Barriers to data sharing

The main barriers to more efficient data sharing are:

- Funding
- Data ownership
- Verification

3.2.1 Funding

Most organisations who received data from recorders receive it in multiple formats such as emails, bespoke spreadsheets and even paper records, they therefore incur costs capturing data into their data management systems. Once received they undertake data validation and verification and other data management tasks. The vast majority do not receive any funding for these tasks, and many rely on the provision of chargeable data information services to bring in income to enable them to continue. This means there is a cross-subsidy between these services to the provision of tasks that are necessary to ensure data is available in a usable format. Therefore, as reported by consultees in this study, some providers of chargeable information services do not want to share high-resolution data openly as this could impact severely on their income streams that support data stewardship.

Recommendations 8 and 9 relate to improvements in funding.

3.2.2 Data ownership

In simple terms, all data collected is owned by the person making the observation. There are some exceptions to this, such as when the collection of data is paid for by public funds or a commercial organisation. However, for the vast majority of data submitted to LERCs and national schemes (and others) ownership is retained by the recorder. Therefore, the resolution at, and licence under which, data can be made available is the recorder's prerogative.

"Some individual recorders will not allow their data to be shared openly, although there is a wide range of opinions on this (Fox et al. 2019). Some recorders won't agree to data being made available on a shared (CC-BY-NC) licence, some will agree to sharing data in a restricted manner: at a low resolution, after a period of time since collection, or with recorders' names removed. There are a minority of recorders who won't release data at all because of the risk of misuse by developers and/or the risk of trespass. Retrospectively, gaining permission to share data can be a huge task. As it may be necessary to get permission of all contributors to all datasets, the number of people who have to be contacted can number in the thousands. Conversely, if data sharing is made a pre-requisite for submission in order to reduce this communication task, the willingness to submit data may be suppressed for some contributors.

3.2.3 Verification

Many organisations will not share data unless it has been verified. There are also issues with verification capacity for many taxa, some groups having very few individuals with sufficient expertise to verify species. Since the great majority of verifiers currently work in a voluntary capacity there are limitations on the ability to insist on timely delivery. Consequently, there may be significant delays between data being submitted to an BDF node and it being shared; sometimes it may not get released at all. Building sufficient knowledge and expertise to verify species identifications is not straightforward. A need for increased resources for ID-training has been identified by stakeholders in this study but it takes time to build the necessary knowledge.

Verifiers are not normally accredited although some organisations may have processes in place. Recruitment and training of new verifiers is highly variable across organisations and taxa.

In response to verification delays in parts of the system some databases include unverified data while flagging verification status, leaving the data user to assess fitness-for-purpose. This in turn can supress data submission as expert recorders may object to their verified data being presented alongside unverified data that contains errors. There is also concern that many users ignore verification status flags and that this may adversely affect the quality of decision-making.

Recommendation 10 relates to investment in verification systems.

3.3 Novel methods

Technology has always played an important role in the collection, review and use of species data (August *et al*. 2015). There are rapid developments in computational and communication technology taking place now that are revolutionising approaches to species data capture and sharing. Data capture via websites and smartphones is making it easier to collect and share data, with the potential to collect greater volumes and diversities of data types. It is also improving geospatial referencing and the efficiency of communication between recorders and verifiers. In parallel with technology to support traditional biological recording, new environmental sensors have emerged for collection of high-resolution images, and video, sounds and molecular data. Once collated these data can be analysed using high-performance computing and novel statistical models, offering new insights into the state of the environment. Interactive visualizations and computer-generated text can also engage new audiences and engage them in data collection.

Realising the potential offered by novel data capture methods carries certain risks. Increased data volume and diversification of sources through wider participation, however, can disrupt existing data management structures and affect data quality. It is important to anticipate these risks and the societal challenges that new technologies may bring.

This issue was raised in JNCC's review (Hassall *et al.*, 2020) on biodiversity data use in the Country Nature Conservation Bodies. The potential to make more use of new technological advances in remotely sensed imagery (for example automatic species identification and counting large populations of seals and seabirds from high-resolution drone or satellite imagery) and eDNA techniques was recognised, as well as the need for development of systems and capacity to verify and interpret resulting data.

3.3.1 Overview of three new approaches

New technologies have the potential to increase the amount of data and the efficiency of data flow, from collection to review and quality assessment to sharing to publishing outputs. New analytical approaches can improve integration of multiple species datasets with other types of environmental information and can potentially provide products that are then used to stimulate more data collection. However, there are risks as well. An increased reliance on technological approaches could discourage people from engaging more directly with species recording and the environment, and lead to a reduction in trust in, and understanding of, the data being collected. Careful design of technological interfaces is needed, and there is a need to ensure that adequate data systems are in place to allow the capture of these new types of data, along with any associated metadata that can help integrate the data obtained using more traditional approaches.

Workshop participants in this study highlighted the potential benefits of using automated data capture and taxon identification approaches for species recording, but also cautioned that more needs to be done to understand what level of confidence can be placed in the results, and to ensure that the methods used are documented so that it is clear how any individual species record was arrived at. A good understanding of the capabilities and limitations of new technologies is needed to successfully integrate their results with existing traditional records. Some of the main emerging technologies are described below.

DNA based technology

DNA-based techniques for species detection and measuring biodiversity have recently moved from academic proof-of-concepts to operational tools being used by environmental and regulatory monitoring agencies worldwide. Recent reviews by Bush *et al.* (2019), Piper *et al.* (2019) provide up-to-date summaries and references describing the application of DNA techniques for both terrestrial and freshwater monitoring. There are also potential applications in the marine environment (e.g., Eble et al., 2020).

The ability to make use of DNA for species recording has greatly increased in recent years but is still a long way from being routinely applicable across multiple taxon groups. A recent study (Macadam *et al*. 2020) found that of over 13,000 UK invertebrate species in the UK, only 168 were represented in the Barcode of Life Database (BOLD) by high confidence barcodes that were based on verified UK specimens, while a further 3,025 species were assessed as having medium confidence. Macadam *et al*. 2020 provide recommendations for making better links between traditional species recording approaches and DNA approaches, including the development of standard protocols for data recording to ensure that the provenance of barcodes is clearer and confidence in their use will increase.

Acoustic monitoring

Following rapid developments in the capabilities of hardware/software and assessment methods over the last 20+ years, the last 12-24 months have seen the publication of a number of useful 'state of the art' reviews of acoustic survey/monitoring methods. These include general reviews such as Gibb *et al* 2018, and Sugai *et al* 2019; together with the more targeted reviews on animal communication (Teixeira *et al* 2019); bird bioacoustics (Darras *et al* 2018); freshwater habitats (Greenhalgh *et al* 2020); and the use of microphone arrays for localising individual animals (Rhinehart *et al* 2020) and for bird survey (Machado *et al* 2017; Eldridge *et al* 2018). Recent guidelines for the use of acoustic indices have been produced (Bradfer-Lawrence *et al* 2020), with recommendations also available for how to use these in landscape ecology (Villanueva-Rivera *et al* 2011) and agri-environmental contexts (Doohan *et al* 2019). Xie *et al* (2020) produced an analysis of global research trends in bioacoustics and ecoacoustics over the last 30 years. Fingas (2019) provides a review of developments in remote sensing in the marine environment.

Image recognition and machine learning

Automated identification of species has made considerable progress through development of deep learning and Convolutional Neural Networks (CNN) in particular (Wäldchen & Mäder, 2018). Goëau *et al.* (2018) demonstrated an accuracy of 88% across 10,000 plant species', while Van Horn *et al* 2018 achieved 81% spanning over 5,000 categories of plants, animals, and fungi. Furthermore, in Bonnet *et al* 2018 it was shown that software-classifiers were able to provide more accurate identifications than five out of nine specialists in the French flora. UK work (Terry *et al* 2020) demonstrates that these approaches can be further improved by building country-specific ecological knowledge into machine learning algorithms. These

demonstrations of image classifiers capabilities, not just for plants, build off rapid developments in the computer sciences. At the cutting edge of plant classification, we are now seeing real-time classification (as in the iNaturalist 'seek' app), and classification of plant organs using Mask-R-CNN (Mora-Fallas *et al* 2019).

Recommendation 11 relates in part to novel data methodologies.

4. Approach to stakeholder engagement

This Section describes how the main issues and challenges (identified in Section 3), associated with the species data pathway in England were identified through review of published literature and reports and engagement with stakeholders. The species data pathway (described in Section 2) involves the collection, curation, aggregation, dissemination, analysis, quality and use of species data. Views of the study team and other stakeholders regarding the efficiency of the current system were considered in relation to findings of other initiatives and used to develop the recommendations presented in this report.

4.1 Review of published literature and other initiatives

The findings and recommendations of other recent initiatives relevant to improvement of biological data flows were reviewed and considered, including the following:

- The Scottish Biodiversity Information Forum (SBIF) published "A Review of the Biological Recording Infrastructure in Scotland" in 2019. The applicability of SBIF's recommendations in an English context was reviewed. The review was informed by expert opinion and the results of interviews and workshops carried out as part of this study.
- JNCC's review of biodiversity data use in the Country Nature Conservation Bodies (Hassall *et al.*, 2020) "… and other recent JNCC reports that are relevant to this topic (Pocock 2018, August *et al*. 2019, Plummer *et al*. 2019)".
- The UK's Geospatial Strategy 2020 to 2025: Unlocking the power of location Geospatial Commission.
- The Open Data Initiative materials including https://theodi.org/about-the-odi/the-data-spectrum/

4.2 Study team perspectives

The study team comprises representatives of organisations directly involved in the collection, curation, aggregation, dissemination, analysis, coordination or use of species data in England. Internal meetings and workshops were held to ensure that the expertise and experience of the team was fully reflected in this report's findings and recommendations. Organisations represented are:

National Biodiversity Network NBN Trust

The NBN Trust is the UK's largest partnership for nature, bringing together organisations involved in the collection, sharing and use of information on the UK's wildlife species, under the banner of the National Biodiversity Network. The Trust works with approximately 200 partners in the network to make reliable, high quality, biodiversity data available and promote its wise use. The Trust has a co-ordination and facilitation role, working collaboratively with its stakeholders to design, adopt and implement national and international data standards, assist organisations with sharing their data and making it accessible for use, as well as aiding those using the data for multiple purposes. The Trust provides the only UK wide aggregation of multi-taxa biodiversity data via the NBN Atlas, which also provides tools for visualisation and analysis of species information.

Association of Local Environmental Records Centres (ALERC)

ALERC represents and supports Local Environmental Records Centres across the UK and facilitates

networking and collaboration between them. It promotes and develops quality assured standards for the collation, management, dissemination and analysis of biological and geological records, develops tools to support the work of the sector and advocates on behalf of LERCs at regional, country and UK level to data providers, information users and policy makers.

UK Centre for Ecology & Hydrology UKCEH – Biological Records Centre
The UK Centre for Ecology & Hydrology is an independent, not-for-profit research institute carrying out excellent environmental science with impact. Their 500 scientists work to understand the environment, how it sustains life, and the human impact on it. They provide the data and insights that governments, businesses and researchers need to create a productive, resilient and healthy environment. The Biological Records Centre (BRC) is a group within UKCEH, established in 1964, as a national focus point for terrestrial and freshwater species recording. BRC works closely with the voluntary recording community, principally by supporting national recording schemes and societies. BRC generates the data and insights that researchers, businesses and governments need to solve complex environmental challenges.

eCountability ltd
eCountability is a consultancy specialising in development of tools for environmental assessment and planning, working for public and private sector clients on establishment of data management systems, interpretation of datasets and development of applications to apply evidence-based approaches to various aspects of planning and development. This includes aspects of Environmental Impact Assessment, a potential source of species data generated by the private sector

4.3 Stakeholder engagement

A key component of this study was engagement with stakeholders involved in the collection, curation, aggregation, dissemination, analysis and use of species data, to gain insight to the current data landscape in England and the opinions of stakeholders regarding the efficiency and effectiveness of current systems. Stakeholder organisations and individuals (listed in Appendix B) were approached to participate in semi-structured online interviews and online workshops. Organisations were asked to provide a representative in most cases; if initial contacts failed to elicit a response, then known individuals in that organisation were invited directly. COVID-19 restrictions prevented any in-person meetings from taking place

4.3.1 Stakeholder consultation

Consultations were held with 27 individuals, selected to represent the breadth of activities, taxonomies, and organisational types within the species data landscape.

Selecting participants
Stakeholders were assigned to four groups for purposes of consultation:

Recording specialists - people or organisations primarily dedicated to gathering or collating species data, publishing them, carrying out quality assurance or engaging in other tasks to ensure that species records are reliable and reusable. This group may also be involved in providing support and capacity for generating species records, for example by loaning specialist equipment or providing training or mentorship to amateur or volunteer recorders.

Environment specialists – people or organisations dedicated to reporting on, managing and improving the natural environment. These people and organisations may overlap to some extent with recording specialists, for example if they conduct field surveys. However, the important distinction is that if they do this, they do it to facilitate their wider environmental work and they are not dedicated to recording nature as an activity in its own right. Organisations included in this category are a mix of professional bodies (such as LERCs) and voluntary groups (such as many of the national recording schemes).

Decision makers - users of species data, for example to inform policy or planning decisions. This group includes environmental regulators.

Marine specialists - identified as a separate group due to the different approaches used to collect, organise and analyse species data for the marine environment.

Within these groups, consultees were selected at random, with some additional organisations or individuals being added to achieve representation across the range of stakeholder organisations/ individuals in the species recording network (Table 4.1).

Table 4.1: Stakeholder groups examples

Groups	Example	Constituent stakeholder organisations or individuals within the species recording network
Recording specialists	A volunteer recording scheme whose members collect records of particular species, verify them as correct and publish them in an atlas.	Recorders or Recording Groups
		National Recording Schemes & Societies
		Local Environmental Records Centres
		Cross-sectoral partnership or secretariat organisations
Environment specialists	A museum with a natural history collection, whose work may influence knowledge of taxonomy and how species are categorised.	Environmental / conservation Non-Governmental Organisations
		Museums, zoos and botanic gardens
		Academia and education
Decision makers	An environmental consultancy who uses records of protected species as evidence that a site proposed for development needs detailed surveys.	National or central government departments, agencies or public bodies
		Commercial companies and environmental consultancies
		Landowners and farmers
		Member of the general public
		Local authorities and national park authorities
Marine specialists	An NGO concerned with promoting scientific work in the marine environment.	Marine conservation organisations.
		Local authorities with coastline

Relationships between stakeholders involved in species data in England can be complex, resulting in duplicate data pathways. Not all stakeholders are dedicated primarily to species data sharing and this can introduce bottlenecks and inefficiencies. Questions were designed to draw out feedback on such issues and the views of stakeholders regarding challenges and also potential opportunities for improvement.

Workshops

For local authorities and LERCs there were a relatively large number of bodies each having relatively similar interests, and a subset of these was chosen at random to be invited to four workshops. For other more diverse groups, a subset was chosen to represent as wide a range as possible of the diversity of relevant organisations. For the national recording schemes, this diversity includes larger schemes with staff and smaller schemes that are entirely run by volunteers. Consideration was also given to taxonomic diversity. For local recording groups and societies, it was recognised that not all such groups are involved with species recording and data collation, so a subset was chosen from those known to be active in this area. Inevitably, not all invitees were able to participate on the dates available, but responses were gathered from a wide and representative range of perspectives across the different sectors.

The workshops were structured to address a range of issues within the scope of this study, as follows:

- Project Background
- The Species Data Flow Pathway
- Participants' mapping to the Data Flow Pathway
- Participants' roles and experiences of using species data
- New policy demands for species data
- Experiences of using novel methods for collecting biodiversity species data and their likely impact and implications for species data flow
- Potential structural options for funding the biodiversity data framework
- Value and Benefits of access to species data

After an introduction, participants were requested to provide their responses, under Chatham House rules (Chatham House, 2021). Some adjustments were made for the different audiences, for example the marine workshop included a section on "Mapping the marine species data landscape – fundamental differences from terrestrial recording."

Semi-structured interviews

An interview protocol was designed and used for semi-structured interviews (Appendix D). A summary of the main themes and discussion topics is in Table 4.2. Interviews included quantitative questions with numeric answers and qualitative questions designed to capture subjective thoughts and feelings on certain subjects. Interviews were conducted on a one-to-one basis online, using Zoom or Microsoft Teams, with a third person taking notes of responses. Interviewees were assured that their responses would not be individually identified, either personally or attributed to their organisation. Interviewees were sent a copy of the questionnaire in advance. Questions not relevant to the roles of the participant were not asked, for example database questions for a participant taking part predominantly as a data user.

Scheduled for one hour, the interviews often took longer. If unable to cover all questions in the time available, the interviewer and participant were able to agree that the answers to remaining questions would be completed on the document and sent to the interviewer later.

Table 4.2: Themes addressed in interviews

Theme	Description
Biodiversity information needs	Questions about what the needs of biodiversity information and species records, whether they are being met and what measures need to be introduced to ensure all needs are met.
Recording effort and methodology	Questions about changes in technology and the way data can be collected in the field and the effect these may have on species records.
Data flow	Questions on how data move from place to place, how they are stored and accessed how their quality is assured.
Administration	Questions around how data should be administered, including important principles such as the FAIR principles, and how this should be funded.

4.3.2 Collation and analysis of responses

Responses to interview questions with numeric options on degree of agreement with statements were entered into a combined spreadsheet table and analysed by range and average response per question. The results were used by the study team to inform discussion and recommendations. Any significant variation in results by stakeholder group is noted in this report.

Responses to open questions in interviews and comments made by stakeholders in the workshops were compiled in workbooks and documents on the study data portal, organised by subject matter. This material was accessed by the study team to develop analysis and recommendations. From the topics discussed in the workshops, we drew out common themes that produced consensus from a wide range of participants and areas where consensus was lacking, with differing participants expressing conflicting views or diverging suggestions for best practice. From the questionnaire responses, the numeric answers were collated and tabulated, and the qualitative answers were grouped into the themes shown in Table 4.2; both sets of responses were used to explore areas of consensus or divergence. In further developing findings and recommendations, including interactions with Report 2 and based on reviewer feedback, this record has been re-visited to make best use of relevant information and detail.

5. Approach to cost-benefit analysis

This Section describes the economic analysis of the costs and benefits of the species data pathway (described in Section 2). It lays out the principles, the data collection process, the cost-benefit analysis (CBA) steps followed, and the assumptions (e.g., defining the counterfactual) and parameters (e.g., discount rates) used. The results of the CBA are reported in Section 6. Along with the findings from Section 2 and 3, they are used to inform the review of the applicability of the SBIF recommendations to England (in Section 7), and this study's conclusions and recommendations (Section 8).

5.1 Cost-benefit analysis

The objective of cost-benefit analysis (CBA) is to identify the positive and negative impacts that an intervention has over time and compare them against its costs. This is done with respect to a baseline or comparator scenario, i.e., what would have been the case had the intervention not occurred. The purpose of this is to establish whether the intervention results in a net benefit to society, meaning that the beneficial outcomes outweigh the costs, thereby justifying the intervention (a positive Net Present Value – NPV). The approach can also involve option appraisal, whereby different ways of delivering the intervention can be compared in terms of their NPV or benefit cost ratio (BCR).

The structure of the CBA (see Section 5.3) is intended to create a systematic and transparent approach to defining the boundaries and assumptions used, stating any key non-monetary costs and benefits. The distribution of costs and benefits over specific groups and time are identified, particularly where this is helpful for policy making (e.g., identifying costs to businesses). In drafting results, we are mindful of potential sources of error in the way the objectives and scope of the CBA are defined; and costs and benefits are identified and measured (such as omissions; back casting and forecasting; data used and valuation methods).

Like any other analysis, the outputs of a CBA are only as good as the inputs that feed into it, including both the data and the assumptions. The results must therefore be interpreted with an understanding of the reliability of the inputs so that the degree of confidence in the results is understood. Sensitivity analyses around various data and assumptions provide greater understanding and confidence in the findings (see Section 6).

5.2 Data collection

As part of the initial research for the CBA, a number of documents were reviewed in order to scope the potential costs and benefits that should be included in the analysis:

- Scottish Biodiversity Information Forum (SBIF) review's economic analysis (Wilson et al., 2018);
- Alluvium (2016) impact assessment of the Atlas of Living Australia (ALA);
- eftec and ABPmer (2019) cost-benefit analysis of Marine Environmental Data and Information Network (MEDIN);
- Griffin et al. (2017) review of studies using cost-benefit analysis of marine spatial data infrastructures; and

- JNCC (2018) review of terrestrial evidence programmes identifies benefits of species data.

The above is not an exhaustive list but they informed our understanding of the potential costs and benefits provided by data systems in general (e.g., all data types) and, more specifically, for species data.

Further research was undertaken to collect data for the quantification of costs and benefits. Additional sources have been reviewed to provide statistics (e.g., MHCLG live tables for planning developments), help form assumptions and develop an understanding of the importance of the biodiversity data framework. This also includes research into recent evidence and approaches to valuing the impacts of invasive species, time savings and volunteer effort.

This study benefitted from evidence provided by a few interviewees in follow-up to the Report 1 consultations (e.g., Mark Diamond (Environment Agency), Karen Kramer-Wilson (Natural England)), as well as expert opinion from those involved in organising species data collection processes to inform environmental policies and implementation (e.g., Tom McKenna (NatureScot), Lisa Norton (Centre for Ecology and Hydrology)). A range of other experts, for example on legal, planning system, water industry and property development issues, have also been consulted to gather data for the analysis.

5.3 Methodology applied

This section describes the structure of the cost benefit analysis used for this study:

1. Define scope and baseline
2. Identify costs and benefits
3. Quantification and valuation of costs and benefits
4. Compare costs and benefits
5. Sensitivity analysis
6. Reporting and interpretation

5.3.1 Scope and baseline

The scope of the analysis has built on the findings reported in Section 2, establishing the structure for the cost benefit analysis using the following key parameters:

- Baseline year: 2020
- Boundary: As stated in Section 2, the boundary is the species data landscape in England and covers all types of flora and fauna species data, including marine, freshwater and terrestrial species.
- The baseline against which to assess the value of species data (see Table 5.1).
- Time scale: The CBA follows HMT Green Book guidance (2020) and uses a 60-year time horizon.
- Discount rate: Following HMT Green Book guidance (2020) the standard discount rate (3.5% and declining) is used.

The CBA focuses on the data provisioning service of the existing biodiversity data framework. As the scale of the analysis is national (i.e., for England) the assessment also factors in the impacts of giving any user the ability to access species data. This distinction is critical, as for many decisions it is not just about having

access to the data users think they need, but also knowing the entire set of data that is available to a potential user.

Defining the scenario that is appraised is a crucial aspect of a CBA, as it determines the scope of the additional costs and benefits of an intervention. To assess the value of the species data pathway, the CBA examines how the value of activities that use species data might change. The assumed baseline and 'no species data pathway' scenario which it is compared to are laid out in Table 5.1 and illustrated in Figure 5.1.

Table 5.1: Baseline and 'no species data pathway' scenario descriptions

Baseline	No species data pathway scenario
Current situation with the existence and operation of the species data pathway in England, as described in Section 2. Key aspects of the system continue at current levels in the future, for example: • Activity at the different stages of the data flow pathway, from recording effort and processes to data analysis and presentation of the data (e.g., in monitoring programmes) • Level and sources of funding • Volume and range of species data handled In conclusion, it is assumed that the baseline scenario delivers benefits from both the use of species data and the access to that data.	The species data flow pathway does not exist: • The volume and range of species data recorded is the same as the baseline, but the subsequent movement of the data through the species data flow pathway does not happen • The decision-making and knowledge systems that use species data are the same as the baseline, but they operate without organised data inputs from the species data pathway • Instead, secondary data is used, obtained through bespoke search activities and/or primary research In conclusion, it is assumed that this scenario delivers some of the benefits of the use of species data, but not the full benefits of the current access to that data that the species data pathway enables.

An economic analysis has investigated the costs and benefits of the species data pathway in England. This focusses on the role of the pathway in enhancing access to (and therefore decision-making that uses) species data, rather than the total benefits from all species data. The baseline is defined as the current situation, with the existence and operation of the existing species data pathway in England. This is compared to a scenario in which the species data pathway does not exist, so while the same data is collected, its stewardship and use is severely impaired, as illustrated in Figure 5.1. As a result, a reduced volume of the species data recorded moves through the pathway at a slower rate than in the baseline. This is compounded by the fact that each stage of the pathway is disconnected, the implications of this are, follow a similar vein to the limitations of having an unstructured data system as described in JNCC (2018).

An alternative scenario could be one in which there is no access to current or pre-existing species data *at all*. In this scenario, any activity requiring species data would have to obtain that data by primary research or locate existing data and purchase it from data holders (e.g., rival commercial developers) and there would be minimal evidence on trends and contextual factors that would come from historical records of species data. However, as the focus of the CBA is the additional costs and benefits of the species data *pathway*, this alternative scenario is not preferred.

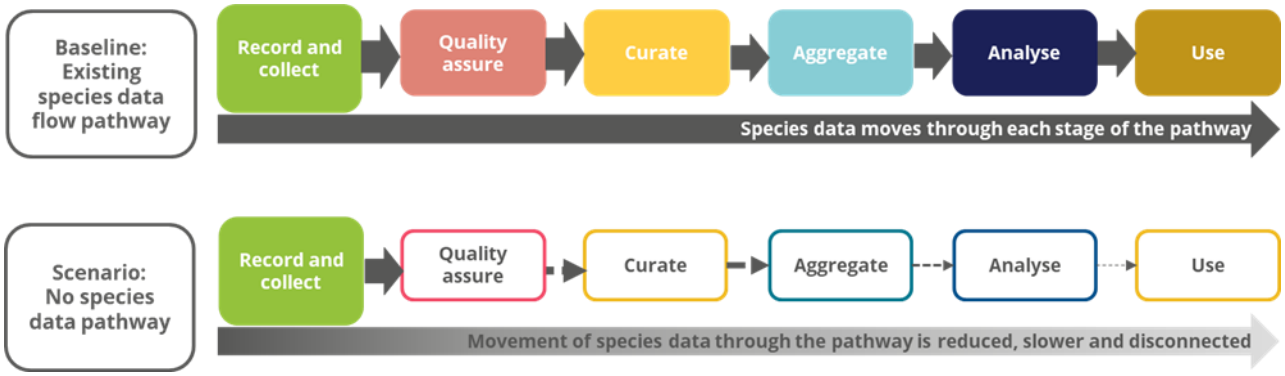


Figure 5.1: Baseline and 'no species data pathway' scenario
Adapted from Figure 2.1.

The definition of the baseline and scenario feed into the analysis of specific costs and benefits through two important assumptions. Firstly, in that the scenario looks at the absence of the species data pathway, but all else remains the same (Ceteris Paribus). This assumes society would try to pursue the same environmental outcomes through the same systems but do them less well without systematic species data (i.e., inability to target interventions, outcomes would be achieved at random). This ignores potential alternative approaches, which are assumed to be less beneficial (otherwise they would be done now) and probably substantially different. In reality, policies would need to be designed differently (e.g., laws would be written differently) but that is a complex issue and is an alternative scenario that could not be developed for this study. In addition, it is also likely that, in the absence of the species data pathway biological recording would be different – for example, there might be less motivation for data to be recorded at all.

Secondly, is the importance of species data in overall natural environment management. Without species data, biodiversity decisions and management would be reliant on data on habitat types and structures, while wider environmental management could also use measures of ecosystem services that were not dependent on species data. Habitat type alone is rather crude (e.g., identifying improved/unimproved grassland). Identification of specific habitat types (e.g., chalk grassland) and/or assessment of the condition or quality of habitats, requires data on species (i.e., on specific species or representative species groups[11]).

Therefore, as soon as you get into management options in policies (such as what is higher distinctiveness in the biodiversity metric for biodiversity net gain (BNG), or what is an appropriate pollinator mix for Common Agricultural Policy (CAP), you need species data. In the scenario with no species data pathway, you wouldn't have the evidence to assess habitat condition/quality in sufficient detail, and so policy delivery as envisaged would be largely ineffective (e.g., unable to be targeted based on habitat detail). Without species data it would be impossible to judge improvements in biodiversity from policies such as biodiversity net gain. It is also impossible by definition to re-construct species data baselines from the past. Existing species data robustly verified and archived and findable now and at any point in the future (the basis of FAIR data principles) cannot be re-created. Those data are essential to assess the outcomes of policies.

These assumptions justify attributing a significant value to the role of species data in several areas, but the

[11] In general, the more detailed the categorisation of habitat, the greater the reliance on species to identify the habitat. So for example, in UKHab, the highly generic level 2 of the primary habitat hierarchy, "Ecosystems" e.g. Woodland/Grassland/Heathland, can be identified largely from vegetation structure, while at the much more detailed Level 4, which includes "Priority Habitats" and is the basis of BNG metrics, the identification of species becomes essential to differentiate, for example, types of woodland.

analysis still tries to distinguish the additional value of outcomes as a result of having species data through the pathway. In some ways this may underestimate the value of the species data pathway: even if relying on primary data collection, distribution patterns and trends established through analysis of data from the pathway informs that primary collection (e.g., which time of year to conduct surveys).

5.3.2 Identify costs and benefits

The NBN data flow pathway underpins the basis of the existing species data landscape in England for the analysis, as illustrated in the baseline in Figure 5.1. As described in Section 2.1, the pathway is a coherent summary of the current operations generally involved in creating, checking, and sharing data so that these can then be put to use, further details on the processes in each stage of the pathway are outlined in Appendix C.

The relevant costs and benefits of the existing biodiversity data framework are identified for each stage of the data flow pathway. Not all of those costs and benefits can be quantified and monetised. Those that are considered material to the results of the CBA (i.e., they can change the net cost and benefit positions) are listed as 'key non-monetised effects'. It should be noted that the cost and benefit typology in the SBIF Review's economic analysis primarily focuses on benefits within the data recording process. A comparison between the typologies is provided in Appendix E.

Benefits

The many benefits along the data flow pathway can be categorised under three headings: (1) input benefits (e.g., savings of time and cost due to easier access to data through the pathway); (2) output benefits due to the use of data to support decision making and (3) outcome benefits due to maintaining the pathway and the decision it supports. Table 5.2 provides an overview of the study benefit typology with examples of what could be included in each category.

The majority of the benefits are associated with 'Use', which is as expected. It can be argued that these benefits arise along the whole pathway, however the identified material benefits are in line with the benefit descriptions in Table 5.2. For example, facilitating innovation would result in benefits along the entire pathway system (e.g., machine learning and artificial intelligence for quality assurance), however any quantified evidence on this has not been identified. The analysis focuses on the outcomes of the 'Record & Collect' and 'Use' stages of the pathway and includes specific policy spend as well as physical health outcomes for volunteer recorders. All benefits are valued in terms of additional or change from the no species data pathway described in Section 5.3.1.

Table 5.2: Benefit typology

Benefit category	Benefit description (examples)
Input benefits from easier access to species data	
Easier access	Data search – time savings: Avoided time that would have been spent accessing data under the no species data pathway scenario
Data generation	Avoided cost of (repeated) primary research
Output benefits for decision-making	
Supporting better decisions based on better evidence	Biodiversity management itself, management of associated activities (e.g. wildlife tourism) and activities for which biodiversity acts as an indicator (e.g. drought resilience of soils, water quality risks); Cost of achieving biodiversity targets; Local, national and international; Biodiversity net gain; 25 Year Environment Plan; WFD status; WiNEP driven by WFD; Agri-environment payments directed; Value for understanding impact of pressures and conservation action; Value for informing management; Value for reporting biodiversity status and meeting legislative requirements; Other strategic objectives
Planning system	Planning fees and applications; Planning for biodiversity; Avoided time-delays
Facilitating innovation	Use of machine learning & AI; Development of tools (e.g., NEVO)
Species management	Monitoring change in biodiversity; Use of biodiversity data; Controlling invasive species; Preventing extinction; Conservation efforts; Resilience of ecosystem services
Outcome benefits for decision-making	
Supporting research	Research publications and projects; Understanding of existence value
Avoided loss of opportunities related to future use and re-use of data	See technology developments discussed in Section 3.
Education	Courses; Training/Skills enhancement; Public engagement
Improving public health and well-being	Health & well-being benefits of volunteers
Cultural services	Opportunity cost of land use (i.e., maintaining priority habitats); Tourism and recreation activities driven by species
Social cohesion and integration	Networking; Development of partnerships; Facilitating the conversation; Development of standards

Costs

Costs along the data flow pathway accrue to the data providers of the pathway: (1) capital costs like replacing computers over time; (2) operating costs (e.g., operations and maintenance); (3) investments / savings grants and to the data users (4) costs for system use. Table 5.3 provides an overview of this study's cost typology with examples of what could be included in each category.

Unlike for the benefits, costs are more likely to arise in each stage of the data pathway due to variations in staff and volunteer requirements as well as data management infrastructure. Costs for system users (e.g., purchasing data) is mainly reflected under 'Use', however similar costs can arise while accessing analysis tools and data repositories.

The relevant 'Use' is determined on the basis of whether species data is needed as a majority input into an activity (e.g., planning applications, policy-making) or a minority input, and will be merged with other data types (e.g., habitats) for similar purposes. In the latter case, the associated cost is not reflected within this analysis. This is because these activities will still occur regardless of the existence of the species data pathway.

Table 5.3: Cost typology

Cost category	Cost description
Costs to data providers	
Data pathway capital costs	Data management infrastructure (e.g., communication, processing and storing); hardware; software
Data pathway operating costs	Staff time; Volunteer time; In-kind contributions; Advice and services; Other resource contributions
Investments/Savings	Statutory funding; Other funding sources
Costs to data users	
Costs for system users	Data searches/requests, data access, and use; Public sector (e.g., policy) and private sector (e.g., EIA)

5.3.3 Quantification and valuation of costs and benefits

In this step we collected the relevant data for the costs and benefits identified in Section 5.3.2. This section provides a summary of the indicators used in the analysis. Although order of magnitude figures for benefits, as identified in the SBIF report (Wilson et al., 2018), are useful to inform policy, the focus here is on measuring impacts relative to the baseline defined in Section 5.3.1. We have prioritised analytical effort in proportion to the significance (expected size and importance within the analysis) of impacts. Further details on the methods used to quantify costs and benefits are in Appendix F.

Benefits

Table 5.4 provides a list of the indicators and identifies who the benefits arise to (e.g., planners, wider society) as well as whether it has been monetised and what that figure represents[12]. As the focus is on the impacts, most values represent an additional benefit that is associated with the existing species data pathway. Where it is not possible to directly estimate the magnitude or value of a benefit, proxies have been used – for example, the financial resources directed in setting and delivering biodiversity targets, as a result of having access to species data. A summary of key variables and assumptions is presented in Table A.3. Unless stated otherwise, monetary unit values are assumed to remain constant over the appraisal period and benefits are assumed to start in Year 0 (i.e., 2020).

Attribution to the species data pathway is underpinned by a ceteris paribus assumption and the importance of the species data pathway to the natural environmental management benefit (as described in Section 5.3.1).

Non-monetised benefits are also identified and discussed alongside the results.

[12] For more detail on benefit methodologies see Table A.3 and Section F.1 .

Table 5.4: List of benefit indicators included in the CBA

Benefit category	Benefit indicator	Link to species data	Benefit(s) to...	Monetised?
Supporting better decisions based on better evidence base	Biodiversity Net Gain market value	Species data is a key component in assessing habitat condition.	Wider society	Y
	Improvements in Water Framework Directive status from water company actions	The species data pathway allows for targeted interventions to be implemented on rivers to improve species status (biological quality).	Wider society	Y
	Common Agricultural Policy payments	Species data allows funding to be directed to Pillar 2 of the Common Agricultural Policy (CAP).	Wider society	Y
	Biodiversity targets	Species data is a key component in assessing habitat condition.	Decision-making	Y – resources directed
Species management	Avoided damage cost from late detection of invasive species	Access to species data facilitates monitoring of INNS and enables early eradication.	Species management	Y
Planning system	Avoided legal fees	Developers can face legal action if they do not do due diligence on species data.	Planning system	Y
	Avoided interest payment on development loans	Developers that face legal challenges are also subject to delays in their construction process.	Planning system	Y
	Avoided loss of ecosystem services delivered by maintaining priority habitats	Species data is a key component in assessing habitat condition and identifying priority habitat types.	Planning system	Y
Easier access	Time savings for existing data search	Reduced time searching for existing species data due to the availability and accessibility of data through data repositories (e.g., data portals)	All users	Y
Cultural services	Cultural opportunity cost of farming	Landholdings that maintain a systematic record of species is motivated by species data.	Wider society	Y
	Domestic tourism expenditure for 'Watching wildlife, bird watching, other nature' activities	Wildlife tourism is an activity partly driven by the visitors understanding of species existence.	Wider society	Y
Improved health and wellbeing	Physical health benefits for volunteers	In addition to improving the general welfare of volunteers, if people are active during their volunteer time, volunteering can also have measurable physical health benefits.	Wider society	Y
	Opportunity cost of volunteer time	Volunteers' personal enjoyment of partaking in recording and collecting species data.	Volunteers	Y

Costs

Building on the categories in Table 5.3, a series of indicators have been developed to quantify and monetise the impacts[13].

Table 5.5 provides a list of the indicators and whether the indicator has been monetised. The majority of the costs of operating the pathway (i.e., infrastructure and operation) have been monetised. The costs for species data activity are partial as cost estimates are based on a sub-sample of non-governmental organisations and does not currently reflect national recording and monitoring schemes. The capital costs of the biodiversity data framework (i.e., hardware and software) have been estimated for the NBN Atlas, MEDIN and LERCs, and future capital replacement costs have been estimated assuming linear depreciation. These capital costs have been adjusted for optimism bias, as per Green Book guidance on assessing costs for 'data/equipment' (HM Treasury, 2020). A summary of key variables and assumptions is presented in Table A.4. Unless stated otherwise, monetary unit values are assumed to remain constant over the appraisal period and costs are assumed to start in Year 0 (i.e., 2020).

Non-monetised costs are also identified and discussed alongside the results.

Table 5.5: List of cost indicators included in the CBA

Cost category	Cost indicator	Link to species data	Monetised?
Data pathway capital costs	Unmet capital replacement costs	Capital and replacement costs for LERCs, NBN Atlas and MEDIN portals, i.e., cost of the biodiversity data framework	Y
Data pathway operating costs	Species data pathway operating costs	Operating costs of species data pathway for LERCs, NBN Atlas and MEDIN portals	Y
	Species data activity	Expenditure by select non-government organisations on species data activities.	Y
	Public sectors spend on data generation	Public sector spend on generating new data records.	Y
	Volunteer effort (data stewardship)	Species data recording and collecting is primarily undertaken by volunteers.	Y
Investments / Savings	Funding from UK public sector bodies attributable to species data	Funding received by NBN Atlas and MEDIN is used to support their operations	Y
	Funding from grants	Funding received by NBN Atlas and MEDIN is used to support their operations	Y
Costs for system users	LERC chargeable data requests	Access to species data via LERCs	Y
	Data requests from public sector bodies	Public sector bodies such as Natural England and the Environment Agency purchase data through the species data pathway. However, it is difficult to ascertain the total spend on data requests by these organisations.	Y
	Environmental Impact Assessment costs	Environmental impact assessments are a statutory requirement in some planning decisions. Part of the content generated for an EIA requires species data.	Y

[13] For more detail on cost methodologies see Table A.4and Section F.2 .

5.3.4 Compare costs and benefits

The estimated values of costs and benefits are collated and summed to give an overall assessment of impacts of the species data landscape. The calculations consistently use the parameters (e.g., discount rate, time scales) identified in Section 5.3.1. The time profile of costs and benefits assumes that all annual value estimated for the baseline year remain constant over the appraisal period.

To compare the costs and benefits of the species data pathway, the annual values for the appraisal period must first be converted to the same price year (2020 prices) using the latest HM Treasury GDP deflators (HM Treasury, 2021) and then adjusted for social time preference (HM Treasury, 2020). To do so, in line with UK government policy, the Green Book standard discount rate (3.5% and declining) is applied to future values. The individual costs and benefits are then aggregated into present value terms and directly compared to arrive at a net present value (NPV) – the present value of the benefits minus the present value of the costs, and the benefit-cost ratio (BCR) – the present value of the benefits divided by the present value of the costs. The formal calculations of the NPV are given in Appendix G and results are in Section 6.

5.3.5 Sensitivity analysis

Sensitivity analysis was performed on key assumptions identified during the analysis based on the most significant costs and benefits, assumptions with greater uncertainty and policy choices. This includes:

- **Sensitivity 1 – Choice of discount rate:** The CBA follows HM Treasury Green Book guidance (2020) on discounting, making use of the standard discount rate (3.5% declining) for the baseline assessment. HM Treasury (2020) recommends performing sensitivity analysis for longer appraisal periods. Therefore, a set of results will be estimated using a where the reduced standard discount rate (3.0% declining) has been applied to all main results 'best'. This rate reflects the equal weight that society puts on the welfare of future generations.
- **Sensitivity 2 – Time horizon:** The CBA follows HM Treasury Green Book guidance (2020) on time horizons, which is 60 years in the main results. However, the SBIF Review (Wilson et al., 2018) used a 30-year appraisal period, as recommended by previous versions of the Green Book for non-infrastructure projects. Therefore, to facilitate some comparison with the SBIF Review BCRs, a set of results estimating present values for 30 years is adopted.
- **Sensitivity 3 – Treatment of volunteer costs and benefits:** In the main results, volunteer effort is treated as a cost to reflect the contribution of volunteers to operation of the species data pathway, as well as the private benefit received by volunteers from the activity as a whole (i.e., enjoyment from what they are doing). Both components are only factored in to the 'maximum' results of the CBA. A set of results are produced showing the impact of excluding one or the other from the analysis.

  However, it can be argued that the cost of volunteer time should be treated as a disbenefit to society rather than a financial cost. Since a key benefit of the existence of the data pathway is time savings, anything that moves in the opposite direction would be a disbenefit (or negative benefit). A set of results are produced where the estimated value of volunteer effort is subtracted from the total present value of benefits. Note, that although the treatment of time spent as a disbenefit aligns with the notion that automation will result in time-savings, it is unlikely that in the species data pathway that automation would replace the need for volunteers to record and collect data. Time savings from technological developments would arise further along the species data pathway (e.g., verification using machine learning). It is also important to highlight that volunteers underpin

the ultimate functioning of the species data pathway, with more experienced volunteers being treated as equal to full-time employees. For example, RSPB note that 85% of their workforce is comprised of volunteers (RSPB, 2020). Therefore, treating this time spent would not acknowledge the important role that volunteers have in the existing pathway.

- **Sensitivity 4 – Including high values:** A handful of benefit and cost indicators are estimated with a range of values (low and high). In the main results, the lower values are included in the CBA. As part of a stress-test of these results, where relevant the CBA results are produced including the high values. This allows for the identification of indicators that are driving the results.

5.3.6 Reporting and interpretation

The results from the cost benefit analysis are presented in Section 6, with conclusions and interpretation in Section 8. The main CBA results are aggregated to produce minimum, best and maximum estimates of key metrics (i.e., total PV costs and benefits, net present value and benefit-cost ratio). The 'minimum' main results represent costs and benefits that have been estimated with the highest confidence and the lower bound of the quantified indicators, whilst 'best' reflects a mixture of high and moderate confidence in the methods used. The 'maximum' reflects the maximum possible costs and benefits of the system including measures with 'poor' confidence as well as the upper bound of the quantified indicators.

6. Results of the cost-benefit analysis

This section presents the results of the cost-benefit analysis (CBA) of the species data pathway. The objective of the CBA is to analyse and compare the key monetised and non-monetised costs and benefits of the existing species data pathway. It does this identifying the positive and negative impacts of the pathway over time, as compared to a 'no data pathway' scenario (described in Section 5.3.1). Each monetised benefit and cost indicator is given a confidence rating, using the ratings described in Table 6.1.

Table 6.1: Description of confidence rating

Level of confidence	Description
Poor	Evidence is partial and significant assumptions are made to attribute value to the species data pathway. The data provides only order of magnitude estimates of value to inform the cost-benefit analysis.
Moderate	Science-based assumptions and published data are used but there is some uncertainty in combining them and attributing values to the species data pathway. There is reasonable confidence in using the data to inform the cost-benefit analysis.
Good	Evidence is peer reviewed or based on published guidance, and values are clearly attributable to the species data pathway. There is good confidence in using the data to inform the cost-benefit analysis.

6.1 Benefits

Table 6.2 presents the estimated annual and present values of the benefit indicators identified in Section 5.3.3[14]. Overall, there is moderate confidence in the estimates. Where confidence is poor, this is primarily due to the difficulty of attributing a benefit to the provision of species data. Indicators with good confidence are primarily from approaches and estimates published by government with strong attribution to species data (e.g., physical health benefits, and Biodiversity Net Gain (BNG) market value, respectively).

The main beneficiary of the existing species data pathway is wider society through a range of outcomes. There are also benefits to the private sector, for example to developers in the planning system and farmers who receive agri-environment payments. The species data pathway has direct benefits that arise to those that collecting and using the data (e.g., physical health and well-being), as well as indirect benefits through time savings by those that interact with data collectors and users.

Minimum, best, and maximum estimates are produced as part of the cost-benefit analysis results. Where relevant, the low value of indicators is used[15]. The main results 'minimum' estimates aggregate indicators with good confidence and low ranges of indicators with moderate confidence: this is because there is greater certainty in the methods applied as well as the link to the species data pathway. This implies that the avoided eradication costs from late detection of invasive species is included in the minimum estimates. The main results 'best' estimates reflects both indicators with good and moderate confidence ratings. The value of volunteer time, despite moderate confidence is only included in the 'maximum' main results. The 'maximum' main results estimate is the 'best' estimates with the inclusion of indicators with a 'poor' confidence.

[14] Further details on methods can be found in Section F.1 .

[15] Sensitivity 4 tests the inclusion of the high values of these indicators. Results are in Section 6.4.

Table 6.2: Identified and monetised benefits of the England 'no species data pathway' scenario compared to baseline

Benefit category	Benefit indicator		Beneficiaries	Annual value (£m)[1]	Present value (£m)[2]	Confidence
Supporting better decisions based on better evidence base	Biodiversity Net Gain market value[3]	Low	Wider society	-	1,419	Good
		High	Wider Society	-	17,751	Moderate
	Improvements in WFD status from water company actions	Central[4]	Wider society	-	3,921	Good
		Low	Wider society	-	4,779	Moderate
	Common Agricultural Policy payments		Wider society	485	12,728	Moderate
	Biodiversity targets[5]		Decision-makers	404	10,606	Poor
Species management	Avoided eradication costs from late detection of invasive species	High	Wider society	-	96	Good
		Low	Wider society	-	443	Moderate
Planning system	Avoided legal fees	Low	Planning system	5	127	Good
		High	Planning system	10	269	Moderate
	Avoided interest payment on development loans		Planning system	166	4,352	Moderate
	Avoided loss of ecosystem services delivered by maintaining priority habitats		Wider society	0.1	80	Good
Easier access	Time savings for existing data search		All users	8	216	Moderate
Cultural services	Cultural opportunity cost of not farming[6]		Wider society	Asset value	2	Moderate
	Domestic tourism expenditure for 'Watching wildlife, bird watching, other nature' activities[7]		Wider society	265	6,939	Poor
Improved health and well-being	Physical health benefits for volunteers[8]		Wider society	8	196	Good
	Opportunity cost of volunteer time		Volunteers	56	458	Moderate

Table Notes:

[1] Annual value refers to monetised value in year 0 (i.e., 2020). The benefit indicators without an annual value, are those where benefits are assumed to arise later in the time profile (i.e., they are lagged).

[2] Present value over 60 years with a declining discount rate starting at 3.5% as per HM Treasury Green Book (2020) guidance

[3] From Defra (2019) BNG Impact Assessment, where the low value is the low estimate of annual costs. Whilst the high value represents the best estimate of net benefit value of the BNG market.

[4] Analysis uses NWEBS 'central' value in the 'best' and 'maximum' results. Not labelled as 'high' to avoid confusion with 'high' NWEBs monetary unit values.

[54] Omitted from low and best estimates of the cost-benefit analysis results as there is a risk of partial double-counting, therefore only included in the high estimates.

[6] Omitted from final results as this is an example from a single site, it illustrates the value involved in principle, but does not reflect a national value.

[7] Omitted from final results as the link to the species data pathway is tenuous.

[8] Omitted from final results as these values are not additional to the species data pathway. It is assumed that active volunteers would continue to record and collected data in the absence of the pathway.

[9] See Section F.1 for the detailed description of calculations of each benefit.

6.1.1 Non-monetised benefits

The key non-monetised benefits not captured in the above results include two benefits that arise within the data pathway:

- Time savings for primary data gathering through avoided cost of (repeated) primary research by accessing data through the species data pathway.
- Social cohesion and integration benefits such as the value of partnerships and networks of stakeholders working in collaboration towards a common goal (e.g., the Partnership for Biodiversity in Planning), as well as the development of standards for species data collection.

And several benefits from subsequent uses of the data:

- The method applied to estimate domestic expenditure attributable to wildlife tourism lacks a robust link to the tole of species data.as a result in the values reported in Table 6.2 are omitted from the final results. However, the £265 million per year value helps put other benefits included in the CBA in context.
- The physical health benefits of volunteers, measured as avoided medical treatment costs is an important value to recognise, not only for the volunteers themselves, but for health services provided by the NHS. This value may not be additional to the species data pathway – much of the voluntary species recording activity might happen anyway, so it is excluded from the final results of the CBA, but it also provides context for other benefits.
- The value of scientific research publications that utilise species data. Breeze et al. (2020) conclude that there are significant research costs savings attributed to the use of pollinator monitoring schemes and through the provision of high-quality data they save at least £1.5 on data collection per £1 spent. In addition, the number of peer reviewed papers published using data from the NBN Atlas is approximately 785 (pers comm NBN Atlas).
- Additional policy outcomes and other public sector resource allocations, such as funding of Wildlife Crime Units which have received funding from Defra and the Home Office (approximately £1.2 billion from 2016-2020) (UK Parliament, 2016).
- There may be significant omitted benefits, such as the role of species data in several of the ecosystem services identified in the UK national ecosystem assessment (e.g., greenhouse gas sequestration and storage in vegetation and soils).
- Some benefits from species management in relation to preventing extinction, and from maintenance of species populations.
- The species data pathway also enables knowledge sharing and education through courses, training opportunities. For example, it provides access to past data to understand trends and context and enhances skills of volunteers through contact with experts in specific taxa (JNCC, 2018). The species data pathway also encourages the public to engage with nature through unstructured monitoring initiatives (e.g., RSPB's Big Garden Bird Watch).

6.2 Costs

Table 6.3 presents the estimated annual and present values of the costs of operating the species data pathway, based on the indicators identified in Section 5.3.3[16]. Overall, there is good to moderate confidence in the estimates, while the difficulty of attributing costs of species data to the pathway leads to poor confidence in the estimates. Indicators with good confidence reflect data from published reports (e.g.,

[16] Further details on methods can be found in Section F.2 .

organisational annual report, State of Nature 2019).

Minimum, best and maximum estimates are produced as part of the cost-benefit analysis reporting. Where an indicator has a low and high range, the low is included in the main results of the CBA. The minimum estimates reflect cost indicators with good confidence, and the 'best' estimates aggregate costs with good and moderate confidence. The exception to this is the estimated costs of volunteer effort, which are only included in the 'maximum' results. The maximum cost estimates represent the 'best' estimates and also includes indicators with a poor confidence rating.

The costs identified for the baseline include a small sum – approx. £100,000 per year – for capital replacement that are not currently funded. These represent additional spending that is considered necessary to maintain the existing system and the benefits from the species data pathway into the future (as assumed in the benefits analysis).

Investment/savings indicators are a source of double counting, since funding from some organisations (e.g., public sector) covers costs identified for other organisations (e.g., NBN Trust, LERCs, RSPB). For this reason, the investment/savings indicators are not carried through to the cost-benefit analysis reporting. Additionally, data requests from public sector bodies potentially double count with the estimated LERC chargeable data requests. Therefore, the public sector spending figure has a poor confidence rating and is only included in the 'maximum' results.

Table 6.3: Identified and monetised costs of the England species data pathway: no species data pathway compared to baseline

Cost category	Cost indicator		Annual Value (£m)	Present Value (£m)[1]	Confidence
Data pathway capital costs	Unmet capital replacement costs	Low	0.1	6	Good
		High	0.1	6	Moderate
Data pathway revenue costs	Species data pathway operating costs		9	232	Good
	Species data activity		19	498	Moderate
	Public sector spend on data generation		0.04	1	Poor
	Volunteer effort (data stewardship)		63	1,640	Good
Investments / Savings	Funding from UK public sector bodies attributable to species data[2,3]		0.1	4	Moderate
	Funding from grants[2]		0.1	0.2	Moderate
Cost of system users	Chargeable LERC data requests	Low	1	31	Good
		High	3	66	Moderate
	Data requests public sector bodies[2]		0.5	12	Poor
	Cost of EIA content relating to species data		2	62	Good

Table Notes:

[1] Present value over 60 years with a declining discount rate starting at 3.5% as per HM Treasury Green Book (2020).

[2] Omitted from the cost-benefit analysis results as they double-count with the LERC chargeable data requests and the species data pathway operation costs.

[3] Current sources of grant funding received by the NBN Atlas are due to expire in 2021.

[4] See Section F.2 for the detailed descriptions of calculations of each cost.

6.2.1 Non-monetised costs

The key non-monetised costs not captured in the above draft results include:

- The full operating costs of national recording and monitoring schemes – although this is expected to be partly captured by the inclusion of NGO expenditures under 'Species Data Activity'. For example, Breeze et al. (2020) estimated that pollinator monitoring scheme costs ranged between £6k to £2.7 million per year depending on the number of sites and skill level of the network (i.e., volunteer-based vs professional). Similarly, a recent proposal for an EU pollinator monitoring scheme estimates that implementing the scheme would cost the EU €13.3 million per year, with the UK contributing €1.5 million per year (Potts et al. 2021). However, these costs have been assessed against the benefits to society provided by pollinators (€15 billion per year).
- Private sector costs are partly captured under chargeable data requests reported by LERCs and EIA costs. But, understanding of the additional costs through increased reliance on primary data collection in the absence of the species data pathway is unclear.

In a wider sense, it would also be the case that, without species data provided through the pathway, decision-making systems would be different. For example, the current laws that protect species would probably have to be written differently, and the alternative would almost certainly be less effective at protecting biodiversity or cost more to implement. This supported by work for HM Treasury by the Natural History Museum and Vivid Economics (2021) conclude that "it is twice as expensive to delay action to stabilise biodiversity intactness globally as it is to act immediately" and that in delaying actions stabilising biodiversity intactness become infeasible (p.2). Specifying an alternative decision-making system scenario is beyond the scope of this study's quantified CBA analysis. However, these should be noted in considering the overall costs and benefits of the system.

6.3 Benefit-cost ratio

To compare the costs and benefits of the species data pathway, the annual values for the appraisal period must first be converted to the same price year (2020 prices) using the latest HM Treasury GDP deflators (HM Treasury, 2021) and then adjusted for social time preference (HM Treasury, 2020). To do so, in line with UK government policy, the Green Book standard discount rate (3.5% and declining) is applied to future values. Further details on specific assumptions can be found in Section 5.3.3 and Appendix F.

The cost-benefit analysis main results are presented in Table 6.4. As can be seen, in the 'best' present value of benefits (PVB) is approximately £23 billion, over the 60-year appraisal period. The PVB far outweighs the 'best' present value of costs (PVC) of £0.8 billion. The net present value (NPV) is estimated as the difference between the 'best' PVB and 'best' PVC, which is approximately £22 billion. The 'best' PVB is primarily driven by the CAP payments estimates (55%), followed by avoided interest payment on development loans (19%) and improvements in WFD status (17%). Whilst the 'best' PVC is driven by the cost of species data activity from NGOs (60%) and followed by estimates of the species data pathway and operation (28%).

The relative importance of cost and benefit estimates does vary across the minimum, best and maximum results in Table 6.4. Under the 'minimum' results, the key cost is the operation of the species data pathway (70% of minimum PVC) whilst the key benefit is the willingness to pay for improvement in WFD status (69% of minimum PVB). For the 'maximum' results, the dominant cost is associated with the inclusion of

volunteer effort (66% of total maximum PVC). However, on the benefits side the drivers of the PVB are more diverse due to the inclusion of the directed spend for biodiversity targets (31% of total maximum PVB) and the benefits of targeting agri-environment payments (37% of the total maximum PVB).

The benefit-cost ratio (BCR) is also given, which divides the 'best' PVB by the 'best' PVC, resulting in a BCR of 28:1. This BCR further indicates that the benefits of the existing species data flow pathway far outweigh the costs of supporting it. It should be noted that these estimates only include the monetised costs and benefits of the species data flow pathway in England, as defined in Section 2 and Section 5.3.1.

Table 6.4: Cost-benefit analysis main results: no species data pathway compared to baseline

Category	Cost/Benefit	Minimum (PV60 £m)	Best (PV60 £m)	Maximum (PV60 £m)
COSTS				
Data pathway capital costs	Capital replacement costs	6	6	6
Data pathway operating costs	Species data pathway operating costs	232	232	232
	Species data activity	-	498	498
	Public sector spend on data generation[1]	-	-	1
	Volunteer effort (data stewardship)	-	-	1,640
Costs of system users	Chargeable LERC data requests	31	31	31
	Data requests public sector bodies	-	-	12
	Cost of environmental impact assessment (EIA) content relating to species data	62	62	62
	Total present value costs (PVC)	**331**	**829**	**2,483**
BENEFITS				
Supporting better decisions based on better evidence base	Biodiversity net gain market value	1,419	1,419	1,419
	Improvements in WFD status from water company actions	3,921	3,921	3,921
	Common Agricultural Policy payments	-	12,728	12,728
	Biodiversity targets	-	-	10,606
Species management	Avoided eradication costs from late detection of invasive species	96	96	96
Planning system	Avoided legal fees	127	127	127
	Avoided interest payment on development loans	-	4,352	4,352
	Avoided loss of ecosystem service delivered by maintaining priority habitats	80	80	80
Existing data	Time savings for existing data search	-	216	216
Improved health and wellbeing	Opportunity cost of volunteer time	-	-	458
	Total present value of benefits (PVB)	**5,643**	**22,939**	**34,003**
RESULTS				
	Net present value (PVB – PVC)	**5,311**	**22,110**	**31,520**
	Benefit-cost ratio (PVB / PVC)	**17.0**	**27.7**	**13.7**

Table notes:

[1]Omitted from low and best estimates of the cost-benefit analysis results as there is a risk of partial double-counting, therefore only included in the maximum estimates.

Using the minimum, best and maximum figures in Table 6.4, additional BCRs can be estimated to reflect the range in impacts (see Table 6.5) where the BCRs in bold are the same as in Table 6.4. For each calculation, the benefits far outweigh the costs. Even under the 'worst case' (i.e., minimum benefits and maximum costs), the benefits outweigh the costs 2:1. Table 6.5 concludes that the cost-benefit analysis estimates a range of impacts between 2:1 to 103:1, although this represents extreme assumptions. The study approach has moderate confidence in estimated impacts between 14:1 to 28:1.

Table 6.5: Summary of benefit-cost ratios of the main results

Ratio calculation (PVB / PVC)	Benefit-cost ratio (PVB / PVC)
Minimum / Minimum	**17.0**
Minimum/ Best	6.8
Minimum / Maximum	2.3
Best / Minimum	69.2
Best / Best	**27.7**
Best / Maximum.	9.2
Maximum / Minimum	102.6
Maximum / Best	41.0
Maximum / Maximum	**13.7**

To put these figures into context:

- Bateman and Mace (2020) provide a summary of benefit-to-cost ratios for large-scale investments in built and natural assets in the UK. The general trend is that most public sector investments produce BCRs of less than 10:1.
- Griffin et al. (2019) found that investment in marine spatial data infrastructure delivered benefit-cost ratios between 2:1 and 18:1 (average of 7:1).
- Alluvium (2016) concluded that the direct benefits of Atlas of Living Australia outweigh the input costs and investment from Government programmes by a ratio between 1:1 and 8:1.

6.4 Sensitivity analysis

This section presents the results of the sensitivity analysis suggested in Section 5.3.5.

6.4.1 Sensitivity 1 – Lower discount rates

The Green Book (HM Treasury, 2020) recommends the use of reduced standard discount rate[17] to be applied "where policies have intergenerational impacts" (p.18) which could include irreversible changes to the natural environment. This rate reflects the equal weight that society puts on the welfare of future generations. In this analysis, the reduced standard discount rates (starting at 3.0% and declining over time) has been applied to all main results 'best' impacts. The results in Table 6.6 illustrate the importance of the choice of discount rate, as the guidance does not suggest that the reduced rates should be applied to the estimated costs of the analysis. Ultimately, the benefit-cost ratio has stayed the same.

[17] Reduced standard discount rate assume pure time preference is 0. Further details in Annex 6 of the Greenbook (HM Treasury, 2020).

Table 6.6: Lower discount rates sensitivity results

	Minimum (PV60 £m)	Best (PV60 £m)	Maximum (PV60 £m)
Total present value costs (PVC)	366	915	2,741
Total present value benefits (PVB)	6,479	25,570	37,783
Net present value (PVB – PVC)	6,113	24,655	35,042
Benefit-cost ratio (PVB / PVC)	**17.7**	**27.9**	**13.8**

6.4.2 *Sensitivity 2 – Time horizon*

The CBA follows HM Treasury Green Book guidance (2020) on time horizons, which is 60 years in the main results. However, the SBIF Review (Wilson et al., 2018) used a 30-year appraisal period, as recommended by previous versions of the Green Book for non-infrastructure projects. Therefore, to facilitate some comparison with the SBIF Review BCRs, a set of results estimating present values for 30 years is adopted.

Table 6.7 shows the 'minimum, 'best' and 'maximum' estimates of the PVC, PVB, NPV and BCR when using a shorter appraisal time horizon. Although overall the total PVC and PVB is lower with a shorter time horizon, the BCR is largely unchanged for the 'best' and 'maximum'. Relative to the SBIF Review investment appraisal (Wilson et al., 2018), where the biological recording infrastructure BCR ranges from 9.9 to 23.5, the estimates in this study are lower as the framing of the CBAs are different. But the overall conclusion remains the same that benefits from the species data pathway significantly outweigh costs.

Table 6.7: Shorter time horizon sensitivity results

	Minimum (PV30 £m)	Best (PV30 £m)	Maximum (PV30 £m)
Total present value costs (PVC)	241	602	1,802
Total present value benefits (PVB)	3,007	15,560	23,590
Net present value (PVB – PVC)	2,766	14,958	21,778
Benefit-cost ratio (PVB / PVC)	**12.5**	**25.9**	**13.1**

6.4.3 *Sensitivity 3 – Treatment of volunteer costs and benefits*

In the 'maximum' main results, volunteer effort is treated as a cost to reflect the contribution of volunteers to operation of the species data pathway and the opportunity cost of this volunteering time is a benefit. A set of results are produced showing the impact of excluding one or the other from the 'maximum' main results. Table 6.8 presents the results of this analysis. The inclusion or exclusion of these costs and benefits have significant effects on the estimated BCRs, producing a range of 14:1 and 40:1.

Table 6.8: Treatment of volunteer costs and benefits results

	Main results 'maximum' (PV60 £m)	Without volunteer costs or benefits	Only including volunteer costs	Only including volunteer benefits
Total present value costs (PVC)	2,483	843	2,483	843
Total present value benefits (PVB)	34,003	33,544	33,544	34,003
Net present value (PVB – PVC)	31,520	32,702	31,061	33,160
Benefit-cost ratio (PVB / PVC)	**13.7**	**39.8**	**13.5**	**40.3**

6.4.4 Sensitivity 4 – Inclusion of high indicator values

The main results of the CBA make use of the low indicator values of costs and benefits presented in Table 6.2 and Table 6.3. As an additional sensitivity test, the high values of these indicators have been included following the same aggregation rules for minimum, best and maximum estimates. The resulting analysis has therefore made use of the high values of benefit and cost indicators for:

- Biodiversity Net Gain market value (£17,751m, PV60);
- Improvements in WFD status from water company actions (£4,779m, PV60);
- Avoided eradication costs from late detection of INNS (£443m, PV60);
- Avoided legal fees (£269m, PV60);
- Capital replacement costs (£6m, PV60); and
- Chargeable LERC data requests (£66m, PV60).

The resulting minimum, best and maximum results are shown in Table 6.9.

Table 6.9: Inclusion of high indicator values sensitivity results

	Minimum (PV60 £m)	Best (PV60 £m)	Maximum (PV60 £m)
Total present value costs (PVC)	367	865	2,519
Total present value benefits (PVB)	23,321	40,617	51,681
Net present value (PVB – PVC)	22,954	39,752	49,162
Benefit-cost ratio (PVB / PVC)	**63.5**	**47.0**	**20.5**

As part of the sensitivity analysis, the indicator values were included in the main results one by one to identify which value would drive a significant change in the results[18]. In doing so it was found that the BCR remained unchanged across all cost indicators and the majority of the benefit indicators. The high value of the BNG market value increased the overall 'best' BCR to 47.7, as compared to 27.7 in Table 6.4. This indicates the results presented in Table 6.9 are primarily driven by the high estimate of the BNG market value.

Using the minimum, best and maximum figures in Table 6.9, additional BCRs can be estimated to reflect the range in impacts when including the high value of certain indicators (see Table 6.10). These BCRs are also compared to the range of impacts presented in Table 6.5. BCRs in bold refer to results in Table 6.4 and Table 6.9, for the main and sensitivity 4 results respectively. As expected, the inclusion of the high values of

[18] This means that PVCs and PVBs were adjusted to include the high value of a given variable, whilst all other values remained the same.

benefit and cost indicators results in the BCRs to increase. The range of impacts is between 9:1 to 141:1, although again this represents extreme assumptions. There is moderate confidence in estimated impacts between 21:1 and 64:1, under the sensitivity 4 results.

Table 6.10: Comparison of benefit-cost ratios between main results and sensitivity 4 results

Ratio calculation (PVB / PVC)	Main results	Sensitivity 4 results
Minimum / Minimum	**17.0**	**63.5**
Minimum/ Best	6.8	27.0
Minimum / Maximum	2.3	9.3
Best / Minimum	69.2	110.6
Best / Best	**27.7**	**47.0**
Best / Maximum.	9.2	16.1
Maximum / Minimum	102.6	140.8
Maximum / Best	41.0	59.7
Maximum / Maximum	**13.7**	**20.5**

6.5 Limitations

The results presented in this Section can be interpreted with a reasonable degree of confidence; however, there are limitations to the approach. The cost-benefit results are contingent on the structure of the CBA (as described in Section 5), in particular the framing of the 'no species data pathway' scenario and the boundaries of the costs and benefits assessed (further interpretation in Section 8.2).

The cost-benefit analysis assesses the value of the existing species data flow pathway but does not address efficiencies or evaluate the overall success of the existing system. It is beyond the current scope of this study to explore enhancements to the data flow pathway and the additional benefits that could be provided and the costs of providing those enhancements (see Section 8 for further details).

There are also methodological limitations as values have been estimated by applying significant assumptions to interpret existing data sources, as well as information from Stakeholder Engagement. This makes it difficult to assess the full scale of the costs and benefits of the species data flow pathway in England, therefore making estimates reliant on extrapolation. As a result, the analysis uses a variety of assumptions, described in Section 5.3.3 and Appendix F, and produces a range of values for the costs and benefits, as shown in Table 6.4.

7. Review of the SBIF recommendations and their applicability to England

This Section reviews the applicability to England of the SBIF recommendations on the biological recording system in Scotland. It is based on the analysis in Sections 2 and 3, and the findings from the cost-benefit analysis in Section 6 and informs this Study's conclusions and recommendations in Section 8.

The Scottish Biodiversity Information Forum (SBIF) published a comprehensive review of wildlife recording in Scotland in 2018 (the "SBIF Review"). The Review was overseen by an Advisory Group of 16 governmental bodies and NGOs, and its findings were based on interviews with 42 organisations, 290 public questionnaire responses and four workshops attended by 39 organisations. It took two years to complete. While thinking has evolved within SBIF for some of the recommendations and new approaches are being put forward, our analysis here is based on the recommendations as published in the 2018 Review. Other issues that lie outside the SBIF recommendations are discussed in Section 7.2.

7.1 Issues and challenges in Scotland

The main issues and challenges identified in the SBIF review were:

- No one can easily provide, access or use all of the biodiversity data collected in Scotland;
- Those involved with the collection, management and sharing of biodiversity data are under resourced and struggling just to maintain the status quo;
- As such they are unable to improve the infrastructure required to support the species data pathway;
- Issues with the infrastructure are sapping volunteer energy and risking the ongoing provision of data;
- There are geographic and taxonomic gaps in Scotland making provision of services and information required to undertake commercial assessments difficult, as well as conforming with the biodiversity duty;
- There is little consistency in biodiversity data services provided by different organisations within the network, and
- The review also provided some estimates of the tangible and intangible economic benefits of a transformed biodiversity data framework, which have been considered in scoping the cost benefit analysis undertaken in this study.

The review carried out for this study concludes that the above issues identified by SBIF also apply in England to some extent. Review points specific to each issue are presented below.

7.1.1 Access to resources and services

The issue of incomplete geographic coverage of Scotland by LERCs was identified as an issue by SBIF. Many parts of Scotland do not have a LERC. Although some of these have voluntary recording groups acting as LERCs, they provide a more limited level of service.

This is not the case in England, where there is complete LERC coverage[19], even if there are inconsistencies in the level, type and cost of LERC service-provision across the country. There are geographic gaps in species record collection and coverage in both countries.

7.1.2 Technology

There is an uneven distribution of IT and data management skills across the organisations involved in biodiversity data collection and management in Scotland and this is replicated in England. This is a particular problem in volunteer-led recording groups and schemes but is also true for some of the smaller LERCs. The SBIF Review also found that there was a lack of affordable and supported data management software with all the functionality necessary for effective handling of taxonomic information. Many data management systems are locally developed and do not integrate with regional or national systems. Systems that had been centrally developed and adopted by a number of organisations, such as Recorder, MapMate and Marine Recorder are no longer supported, either financially or with sufficient developer capacity to provide updates and technical support for users, or to develop enhancements that would enable integration with other systems in a more federated way.

The situation is similar in England. Where technology (such as recording apps) is available, it is inconsistently adopted, and 'operationally' unsupported, partly due to a lack of training in its use, but also due to some reluctance amongst recorders to change their behaviour and use new or different methods.

7.1.3 Funding

The SBIF Review concluded that there was a lack of adequate funding to support the current approaches to data collection, curation, sharing and use at the local, regional and national levels. This is also an issue in England. There is a strong reliance on volunteer effort in both countries and access to funding may be an underlying factor. Any funding that is available is often short-term and uncertain, making long-term strategic planning and improvements to the current infrastructure extremely difficult, if not impossible.

It should be noted that several SBIF recommendations were made in conjunction with a case for enhanced funding in Scotland to transition to an improved biological recording system. Such a transition for England is not yet defined, and this limits the direct applicability of some of the SBIF recommendations to the current situation in England. However, the issues identified in Scotland are relevant to the potential future development of the species data system in England – these are developed further as part of future management in Section 8.2.

7.1.4 Access to data

While the NBN Trust provides a digital data sharing infrastructure with the capability to provide a stable inclusive, central species data repository, not all recorders or local or regional data aggregators submit data to the NBN Atlas. This was exacerbated during the general movement towards more open data in Government (over approximately the last decade) by the NBN Atlas not having functionality that enabled high resolution data to be shared between approved users via the NBN Atlas without being publicly available (as was possible on the NBN Gateway the predecessor to the NBN Atlas). This resulted in

[19] The county of Essex has two LERCs currently with overlapping services.

approximately 50% of data providers not giving permission for their data to be transferred to the NBN Atlas and many of these have not yet recommenced sharing data.

As there is no central way to access data from numerous different sources seamlessly, it is difficult for data users to determine if they have accessed all relevant available data. Furthermore, the sheer number of potential data holders and the variation in procedures to seek permission to access and use data, often makes accessing data a time consuming and costly process. This often results in potential data users simply not accessing the data needed for good decision making.

7.1.5 Data flow and submission

A key finding of the SBIF Review was that poor data flow hampers the access or use of biodiversity data collected in Scotland. Evidence collected for this report shows that the same is true for some data in England. Ongoing diversification of routes for data submission has rapidly increased the volume of species data collected but has also increased the complexity of flow pathways and further there may be proportionately less data provided nationally for wider re-use, with increasing data held in silos at a localised or organisational level more effective integration of data pathways would improve the potential of species data for a range of uses.

Although the current diversity of systems reflects the rich history of biological recording in the British Isles, attempting to share wildlife observations to support conservation and research can be bewildering for data holders. In both Scotland and England there is a confusing plethora of data submission routes, which can lead to duplication of effort, the same data being sent to multiple organisations or, conversely, data not being shared at all. The correct route for data submission is not always sufficiently clear. Lack of transparency over data flows means that recorders do not always know where their records will end up or how they will be made accessible for use, and data users may face difficulties in finding all available data.

7.1.6 Data standards

While there are data standards for biodiversity data in the UK, (including the internationally recognised Darwin Core standard) they are not commonly adopted except at the national level. This may be due in part to a lack of awareness of the standards and/or a lack of skills and resources to make the transition to routinely using DarwinCore. Nevertheless, some organisations have started initiatives to progress towards adopting DarwinCore. Data are often submitted to data aggregators (county recorders, LERCs, schemes and societies) in non-standardised bespoke spreadsheets, via email or text. However, there is a perception that enforcing standards for data collection and submission might alienate volunteer recorders and reduce the amount of data submitted.

7.1.7 Data verification

There are a number of issues affecting the verification of species identification. The majority of verifiers are volunteers and there is a perception that the specialist taxonomic skills are mostly concentrated in an ageing population, with succession planning difficult to achieve. There has been a huge increase in the number of records submitted over the last 20 years due to increases in citizen science participation and technological advances. However, **technological advances that could assist with verification for some species, such as machine learning, have not been widely adopted as yet**, mainly due to known

limitations in their accuracy, or to a lack of confidence in systems whose accuracy has not yet been fully tested.

The lack of standardised submission formats increases the time taken to verify records as verifiers have to spend time reformatting data and/or requesting extra information from recorders. Many species often require physical specimens for accurate identification. There is also a resistance to the sharing of appropriately flagged unverified data, all of which leads to blockages in data flow and delays in records being made available for use.

7.2 Applicability of SBIF Recommendations to England

The SBIF Review provided 24 recommendations for the Scottish Government to consider in order to transform biological recording infrastructure in Scotland, grouped into four main themes:

- Changes to the way data moves around from recorders to organisations and to end users (data flows);
- Changes to how the 'services' provided by different parts of the infrastructure are delivered (services include many things - provision of ID training and support for recorders, managing and curating data and specimens, delivering data to end users, administration and financial support etc.);
- Changes to how the organisations within the biological recording sector are managed and organised (governance); and
- Changes to funding.

Many of the recommendations focus on improving flows of data and making the system more efficient and straightforward, particularly in terms of data management, verification and sharing. This would require changes to the structure and governance of the organisations currently involved in species data to ensure complete geographic coverage across Scotland. One central data repository was also recommended by SBIF, to act as a "one stop shop" for data verification. Recorders would be able to submit records to any affiliated organisations or schemes and they would become available for use by all.

The recommendations need to be considered in light of the preferred options for Service Provision, In-Country Governance and Funding (Box 7.1), which is where some of the practicalities of implementation make the recommendations unlikely to be feasible in England. The SBIF review was based on an assumption of open data, which effectively means free data in the sense that the costs to those recording and managing it cannot be recovered. This leads to its approach to seek 100% Government funding for species data pathway and infrastructure enhancements. Under such a public funding model it suggested options for service provision and governance that would require changing the legal entities and/or governance of independent organisations.

Full funding from government in perpetuity is currently seen as a highly unlikely proposition in England, with risks involved in such a major reliance on one funding source (and not the preferred option for many participants in this study). This difference in the context of the SBIF work and recommendations and the current situation in England mean that the SBIF recommendations, reviewed in Section 7.2.1 are, as worded, often not applicable to England. However, there are similarities in the issues considered a priority

to address (e.g., investing in standards and staff), and these are developed further in Section 8.

Even if the infrastructure were fully funded there are other issues around data ownership that would affect the potential to have fully open data. It may also require a change in culture and behaviour for staff and volunteers to adapt to the new infrastructure, so would need support for any transition. Such changes carry the risk that expertise could be lost and data sharing reduced if the benefits of any change are not communicated to and accepted by those involved. Full consultation of the organisations and stakeholders involved is recommended to help design and implement any changes.

Box 7.1: SBIF preferred options for implementation

Service Provision

Services provided by the Infrastructure are either delivered centrally (where it makes sense to avoid duplicating services that are needed in all countries of the UK), nationally (where it makes sense to devolve delivery of a service to be within each country due to jurisdictional differences) or regionally (where it makes sense to have a local presence). All services would be accessible online, and users would choose the relevant country or region to obtain a service from. There would potentially be four national Service Providers (for devolved services), one central Service Provider (for services common to all four countries of the UK) plus a network of NBN Regional Hubs within each country to provide regional and local services. Hosting arrangements for Regional Hubs would maximise the reach into, and mutual support between, all sectors through building partnerships and sharing back-office facilities. Existing LERCs in Scotland would be replaced by a network of NBN Regional Hubs acting as the regional level of the NBN to provide regional and local services with oversight from a single National Hub.

In-Country Governance

In-country governance would be delivered through the formation of NBN Scotland to oversee a network of affiliated Regional Hub Partners. Governance can be facilitated through the use of the terms and conditions of funding for each Regional Hub and through the presence and leadership of NBN Scotland. NBN Trust would be the Lead Governance Body at the UK level and at the national level in Scotland.

Funding

Sufficient public funding will be provided in perpetuity to enable UK, National and Regional Hubs and Super Partners to operate the preferred models to data flow, service provision and governance. A Single Framework Agreement or equivalent between Scottish Government and the Lead Governance Body would define the level of revenue from a biodiversity levy or other tariffs (e.g., a minor increase to business rates) necessary to cover relevant operating costs of the Infrastructure and its Super Partners. The Lead Governance Body would disburse this funding between National and Regional Hubs and Super Partners and Community Fund applicants to maximise the strategic impact and contribution of each one to the National Outcomes for Scotland and the Scottish Biodiversity Strategy. These disbursements would replace current funding arrangements and commercial charges, releasing funds currently committed by Public Bodies to LERCs and the NBN, enabling biodiversity data to be open and Infrastructure services to be free to all at the point of use.

The applicability of each recommendation to England, based on both the text of each recommendation by

SBIF (*in italics*) and the preferred option for implementation, is, is summarised in Table 7.1, together with an indication of whether the recommendation is applicable as it stands or might require some modification (e.g., in terms of method or timing of implementation). Further discussion is provided in subsequent text.

Table 7.1: Summary of the 24 recommendations made by the SBIF Review and their applicability to England

RECOMMENDATION		APPLICABLE TO ENGLAND
1	Primacy of the NBN & NBN Atlas	Partially
2	Affiliation of data submission routes	Yes
3	Single, central route for casual records	No
4	Primacy of affiliated data submission routes	Possibly
5	Provision of records collected under licence or for consent/status	Yes
6	Recognition & resourcing of a central data management portal	No
7	Recognition & resourcing for super partners	Possibly
8	System simplification	Yes
9	Establishment of an NBN national hub	Possibly
10	Establishment of a network of NBN regional hubs	No
11	Automated use feedback & showcasing	Yes
12	NBN regional hub service focus & branding	Possibly
13	Consistent service provision across Scotland	Possibly
14	NBN regional hub hosting arrangements	No*
15	National & regional hub service strategy	Not directly equivalent/applicable*
16	Recognition & resourcing of a central hub for the UK	Partially
17	Governance of NBN Scotland	Unknown
18	Unification of BRC & NBN trust data management services	Partially
19	Team building & professional development	Partially
20	A single framework agreement	Yes
21	Funding drawn from those who gain value or cause harms	Yes
22	A single approved body to disburse funds	Yes
23	Community funds to support verifiers, recorders & outreach	Yes
24	An implementation plan to achieve recommendations by 2025	No

7.2.1 Assessment of applicability of SBIF recommendations to England

This assessment is based on the SBIF recommendations in their 2018 report, which were the basis of the consultation meetings for this study. Due to differences in context, and between the scope of this study and the SBIF Review, many of the recommendations as worded are not fully and directly relevant to England. It is recognised that thinking has since evolved in Scotland, and this, along with some of the principles behind the SBIF recommendations, have been a useful input to the discussion in Section 8.

1. PRIMACY OF THE NBN & NBN ATLAS: *The National Biodiversity Network (NBN) and NBN Atlas platform remain the primary place for the submission, dissemination and discovery of biological records and added-value datasets and services.* ***Applicable to England: Partially***

With sufficient funding the NBN Atlas can continue as the primary place for the dissemination and discovery of biological records. The NBN Atlas user interface and functionality require improvement, but it remains suitable as a central data discovery and dissemination portal and a 'first port of call' for users, providing

some data on open licences, and signposting to where and how users can access the rest of the data they might need. Implementation of the two-tier system and improvements to other aspects of data flow could deliver a significant improvement in accessibility of species data via the NBN Atlas. However, unless there is full government funding and fully open data, the NBN Atlas will not be fit for purpose as a 'one-stop shop' for high resolution species data.

The NBN Atlas is not currently set up as a direct data submission route for recorders. Improvements to and perhaps rationalisation of current data submission routes, such as iRecord, could prove more cost-effective than adding yet another submission route though the NBN Atlas, as well as avoiding duplication of existing effort.

2. AFFILIATION OF DATA SUBMISSION ROUTES: *All biological records should be submitted online and channelled to the NBN Atlas via standard, affiliated routes.* ***Applicable to England: Yes***

Encouragement of more online recording is likely to improve the speed of data flows and reduce costs. However, it has to be noted that this would require a significant extension of a culture shift that is already underway within the recording community, and investment in online recording systems in order to be successful. The results of the SBIF Review questionnaire stated that 85% of records collected in the field are still recorded on paper and that records are submitted in many different formats, nearly two thirds of verifiers receive records by email or text, 40% receive excel spreadsheets and only 16% receive records online. Requiring all data to be submitted online would need to be phased in over a number of years, handled sensitively and properly supported. Affiliated routes for submission of data, which could include accreditation for compliance with data standards and transparency over data flows, could be applied even without all submission being online.

3. SINGLE, CENTRAL ROUTE FOR CASUAL RECORDS: i*Record should be the single, central affiliated channel through which to submit 'ad hoc' records for verification, inclusion in relevant National Recording Schemes and dissemination via the NBN.* ***Applicable to England: No***

Different systems have different pros and cons for individual recorders. The greatest volumes of records are entered via offline software such as MapMate and Recorder. There are also established online systems to match the requirements of recorders, e.g., BirdTrack, iNaturalist and local systems (Living Record, RECORD). Only accepting ad hoc records submitted via iRecord would risk significantly reducing the amount of data available.

There are detailed issues to breakdown in considering the submission routes for casual records. In terms of the recording software, having a single route for ad hoc records is likely to stifle innovation. Reliance on one or a few software systems would bring a risk unless they can secure long-term funding for hosting, software maintenance, operational support and future development. To enable integration of casual records into other data sets, consistent meta-data and transparent data ownership may be more important.

4. PRIMACY OF AFFILIATED DATA SUBMISSION ROUTES: *Biological records for a specific National Recording Scheme, recording group, project or organisation should be submitted via their affiliated route.* ***Applicable to***

England: Possibly

Affiliated data submission routes could be promoted and incentivised. However, it is unclear how affiliated submission routes would be identified and confirmed. A significant change in recorder behaviour and clarification of verification responsibilities would be required.

5. PROVISION OF RECORDS COLLECTED UNDER LICENCE OR FOR CONSENT/STATUS: *Biological records collected with public funding, under licence, for Environmental Impact Assessment or planning consent, or for an academic or professional qualification, should be provided to the NBN Atlas as a matter of good practice.* ***Applicable to England: Yes***

There was significant support for this recommendation to be applied to England expressed by consultees in this study. However, several important changes would be needed to enable this to happen, including legislation to require the sharing of data by private companies (e.g., terms of consenting), provision of best practice guidance on data collection, and development of standards for sharing data and accessible support for enablement. Resources would also be needed to enable follow up on where records have been collected to ensure they are being shared.

Doing so could significantly increase the volume of species records shared in the UK.

6. RECOGNITION & RESOURCING OF A CENTRAL DATA MANAGEMENT PORTAL: *Recorder 6 and Marine Recorder should evolve to become the common, central data management portal for data custodians to collate, view and manage their own biological records and datasets (unless a suitable internal business system is used).* ***Applicable to England: No***

While this seems like a laudable recommendation in principle, there are risks associated with 'putting all the eggs in one basket'. Maintaining a single system in perpetuity (or its successors as it would have to be updated regularly to keep up with advances in technology and changes in data collection) would be risky given current levels of funding and support. There are a number of practical considerations that would need to be worked through such as: how can independent organisations such as schemes and societies or LERCs be compelled to use one particular piece of software? Who would decide whether an internal business system was 'suitable' and how would this be policed? Would enforcing this actually reduce rather than increase the sharing of data? Further, for marine data the NBN Atlas is reliant on provision from DASSH and the affiliations which feed into the marine biodiversity data flow via DASSH.

Recorder 6 and Marine Recorder should be regarded as standards-compliant databases rather than portals. Investment in geospatial successors that facilitate data flow pathway tasks, and the sharing of FAIR data principles compliant datasets would be widely welcomed.

7. RECOGNITION & RESOURCING FOR SUPER PARTNERS*: *Super Partners should be fully recognised and sustained to a level that has the capacity to support verification on a major scale.* ***Applicable to England: Possibly***

The term 'Super Partner' refers to the organisations or systems that provide major components of the

Infrastructure or public services upon which the Infrastructure depends (other than LERCs and the NBN Trust).

If there were full government funding and fully open data, this could be appropriate. It is recognised that there are a number of organisations who provide essential services in support of the biodiversity data framework such as networks of recorders and verifiers, curated reference collections and species taxonomies and there is support for the adequate funding of these services. However, the level of support for Super Partners recommended could only be implemented in England with full public funding in perpetuity. There are some concerns over how Super Partner status is assigned or achieved, and what impact it may have on organisations or groups who are not awarded such status.

8. SYSTEM SIMPLIFICATION: *The systems and tools available for collecting, curating, aggregating and disseminating biological records across all environments (terrestrial, freshwater and marine) and sectors should be rationalised.* ***Applicable to England: Yes***

A system simplification exercise could be undertaken in England. Comprehensive mapping of the current data flows would need to be undertaken and recommendations for rationalisation made. However, an assessment would need to be made about the consequences of rationalisation and the potential for loss of data and volunteer recorders/verifiers.

9. ESTABLISHMENT OF AN NBN NATIONAL HUB: *An NBN National Hub for Scotland should be established to support a network of NBN Regional Hubs and to facilitate the flow of biological records into the NBN Atlas to create a definitive evidence base for Scotland.* ***Applicable to England: Possibly***

If there were full government funding and fully open data, it may be possible to implement this recommendation in England with the agreement of the NBN Trust, if it were only as described in the text above. However, the SBIF Review document recommends that the NBN National Hub is the lead governance body of the NBN Regional Hubs and the whole infrastructure for Scotland. This is potentially problematic, see response to recommendations 16 & 17 below. Furthermore, the concept of NBN Regional Hubs as described in the SBIF Review is unlikely to be feasible in England (see recommendation 10 below).

10. ESTABLISHMENT OF A NETWORK OF NBN REGIONAL HUBS: *A network of NBN Regional Hubs operating in partnership with the NBN Trust covering the whole of Scotland should be created.* ***Applicable to England: No***

The SBIF Review states that their preferred option for this recommendation is for existing LERCs in Scotland to be replaced by a network of NBN Regional Hubs. The hubs would either be hosted by a partner organisation or be part of the NBN Trust, with staff becoming employees of those partner organisations or the NBN Trust itself. This is unlikely to be a feasible option for England as there are 44 LERCs, all operating as separate businesses with their own governance structures, who would need to fundamentally change their operating models and/or merge with other LERCs or partner organisations to make this happen. There is little appetite or perceived need for this in the organisations involved.

11. AUTOMATED USE FEEDBACK & SHOWCASING: *Use feedback for Recorders and Data Providers should be built into all automated processes facilitated by the NBN Atlas.* ***Applicable to England: Yes***

The NBN Atlas can be developed to provide a suite of reporting tools to show when records and datasets have been viewed, reported or downloaded (indeed some of this is already available).

12. NBN REGIONAL HUB SERVICE FOCUS & BRANDING: *NBN Regional Hub Partners should provide services that i) support the flow of biological records to the NBN Atlas for Open use, ii) raise the awareness of, engagement in, and support for biological recording, and iii) support the effective interpretation and use of biological records in local and regional decision-making.* ***Applicable to England: Possibly***

This recommendation could only be implemented in its entirety if there was full government funding, fully open data and the Regional Hubs were implemented as described in the SBIF Review. There is some difficulty of interpretation as the word "Regional" has different contexts in the two nations. Recommendations in this report include the recognition of Local Data Centres with key functions similar to those described in the SBIF recommendation.

13. CONSISTENT SERVICE PROVISION ACROSS SCOTLAND: *NBN Regional Hub Partners should offer a set of core services in a consistent way so that service users from across Scotland can access the same core service from any location in Scotland.* ***Applicable to England: Possibly***

If there was fully government funding, fully open data and the Regional Hubs were implemented then this recommendation could be implemented as stated. However, in the absence of full transformation, improvements to consistency of service across England should be possible through working with ALERC and supporting the accreditation of LERCs. Recommendation 6 in this report advocates certification of standards-compliant processes in the BDF rather than accreditation of organisations.

14. NBN REGIONAL HUB HOSTING ARRANGEMENTS: *NBN Regional Hub Partners should be hosted by an organisation that can provide access to professional back office support (including finance, human resources and IT), line management and office facilities.* ***Applicable to England: No***

This could only be implemented if there was fully government funding, fully open data and the Regional Hubs were implemented as described in the SBIF Review. However, as NBN Regional Hubs are not thought to be feasible for England this recommendation is not applicable at that scale. It is relevant to ensure that all LERCs have appropriate back-office support, and it may be possible for an organisation or organisation(s) to provide that support to LERCs in order to achieve economies of scale. However, consultation with LERCs would be required to be able to state with any confidence that this could be achieved.

15. NATIONAL & REGIONAL HUB SERVICE STRATEGY: *A national service strategy for the biological recording infrastructure in Scotland should seek to perennially grow the contribution of the Infrastructure in support of the National Outcomes for Scotland.* ***Applicable to England: Not directly equivalent/applicable***

This could only be implemented if there was full government funding, fully open data and the Regional Hubs were implemented as described in the SBIF Review. As NBN Regional Hubs are not thought to be feasible for England this recommendation is not applicable. However, the principle of biological recording infrastructure increasing its contribution to national objectives is relevant to England.

16. RECOGNITION & RESOURCING OF A CENTRAL HUB FOR THE UK: *The NBN Trust should be given special status as the Lead Governance Body for the biological recording infrastructure in Scotland.* ***Applicable to England: Partially***

If the NBN Trust were fully resourced it could deliver the functions set out in the fuller description of the recommendation in the SBIF report: User & Partner Services (covering NBN Atlas product ownership, partner account management and data and list curation); Central Services (covering governance, public relations and human resources), and Technical Services (covering system support, system development and digital content management). During any implementation phase, a small Programme Office (covering programme and project management, business analysis and communications) should also be maintained. The National Hub for Scotland should be resourced to cover partner liaison, GIS and data analytics and education. However, the role of "the Lead Governance Body for the biological infrastructure in Scotland" requires clarification and, therefore, it is difficult to assess if this recommendation would be applicable to England.

It should be noted that a Central Hub for marine biodiversity data via DASSH already exists with functions delivered through a collaborative approach with contributors, national and international links. Additional centralised resource for this Marine Hub would facilitate greater envisaged and known implementation.

17. GOVERNANCE OF NBN SCOTLAND: *The NBN National Hub for Scotland should be established as a division of NBN Trust and should be known as NBN Scotland.* ***Applicable to England: Unknown***

This recommendation suggests setting up a Scottish Division of the NBN Trust which would oversee the regional and national infrastructure in Scotland. It is dependent on the establishment of an NBN National Hub and the NBN Regional Hubs. However, the role of "the Lead Governance Body for the biological infrastructure in Scotland" requires clarification and, therefore, it is difficult to assess if this recommendation would be applicable to England.

18. UNIFICATION OF BRC & NBN TRUST DATA MANAGEMENT SERVICES: *The data management services of the Biological Records Centre (within the UK Centre for Ecology & Hydrology) and the NBN Trust should be brought together either through amalgamation or through a formal partnership arrangement for maximum synergy.* ***Applicable to England: Partially***

The partnership between the NBN Trust and the Biological Records Centre could be strengthened through a more formal arrangement. However, amalgamation of the two entities would need to be fully supported by both the NBN Trust and BRC/UKCEH and provide significant benefits over partnership for both organisations.

19. TEAM BUILDING & PROFESSIONAL DEVELOPMENT: *The NBN Trust should invest in a National and Regional Hub professional development programme to build rapport, to encourage common ways of working and to grow collective capacity through developing the skills and capabilities of everyone involved.* ***Applicable to England: Partially***

If there were full government funding and National and Regional hubs in England, this recommendation could be implemented. However, as NBN Regional Hubs are not thought to be feasible for England this recommendation is not applicable. Nevertheless, the issue of professional development within the species data pathway is a relevant one to consider in developing an alternative proposition for England.

SBIF RECOMMENDATIONS 20 – 23

Recommendations 20 – 23 all relate to how the transformed infrastructure should be funded and theoretically could all be applicable to England if full government funding were achieved.

20. A SINGLE FRAMEWORK AGREEMENT: *Sufficient public funding should be provided to cover the core operating costs of the NBN Trust and its network of National and Regional Hubs, Super Partners and community groups in perpetuity where these are providing public services as a public good in support of the National Outcomes for Scotland.* ***Applicable to England: Yes***

MEDIN and DASSH for marine data handling similarly require sufficient public funding.

21. FUNDING DRAWN FROM THOSE WHO GAIN VALUE OR CAUSE HARMS: *The source of public funding should be designed to i) share the core operating costs of the Infrastructure between the sectors who need to access biodiversity data and realise value from doing so and to ii) base the greatest burden of funding upon those whose activities are key drivers of biodiversity loss.* ***Applicable to England: Yes***

22. A SINGLE APPROVED BODY TO DISBURSE FUNDS: *The NBN Trust should be the Approved Body for the disbursement of funding provided through any Framework Agreement.* ***Applicable to England: Yes***

This disbursement should take into account the needs across the biodiversity data framework, for example, with MEDIN involvement to ensure support to marine data management.

23. COMMUNITY FUNDS TO SUPPORT VERIFIERS, RECORDERS & OUTREACH: *A Community Fund should be established to facilitate the scaling up of public participation in biological recording to ease current pressure points and to encourage participation and equal access for all.* ***Applicable to England: Yes***

24. AN IMPLEMENTATION PLAN TO ACHIEVE RECOMMENDATIONS BY 2025*: The SBIF Review Working Group should develop a detailed Implementation Plan for the period from 2020 to 2025 that sets out how the transition from the current situation to the future situation in Scotland will be achieved and monitored.* ***Applicable to England: No***

This recommendation is not relevant in England at present since a 'future situation' is yet to be defined.

7.2.2 Relevance in context of the SBIF recommendations review and findings of the economic analysis

As noted above SBIF recommendations 20 – 23 all relate to how the transformed infrastructure should be funded in Scotland. They theoretically could all be applicable to England if full funding, from Government and/or a mixture of sources, were achieved. However, the case for that funding, and the future system it would support (e.g., potentially involving a transformed infrastructure, different data stewardship

processes, and expanded data functionality and services), have not been defined in England.

Section 6 identifies large benefits compared to the costs of the current species data pathway in England. The economic case for investment in the species data pathway should be investigated. What this should fund can draw on the SBIF recommendations, but a specific investment scenario would need to be devised for England. For example, a future national system that overcame coordination challenges (see Recommendation 6) and targeting future funding to 'super-partners' (Recommendation 7) are relevant options from SBIF that could enhance the system in England. These options are expanded on in Section 8.2.

8. Conclusions and discussion

This section provides conclusions from the analysis in this study and makes recommendations for management of the species data pathway in England. Principles to underpin future development and management of the species data pathway are suggested, but these would need to be taken forward in partnership with existing stakeholders, and further discussion of future options is essential.

8.1 Conclusions

8.1.1 The species data pathway in England

Species data flows in England (and the whole of the UK) can be complex, even for a single species group. Biological recorders, many of whom are volunteers, will typically choose a data flow route that reflects their personal experience and peer networks. As a result, there are inconsistencies in data pathways for similar records from different locations or different records from the same location. Despite this, the extent of volunteer input is a strength of the system, both in terms of volume of records and expertise.

The current English system has evolved over time and is not necessarily as efficient as it could be. This is not the fault of any one organisation, but a result of its evolution over time and the resources available. A key strength of the system is the large volume of species data that are recorded and shared, but there is also a widespread lack of clarity regarding roles, responsibilities, processes and even data accessibility. Historical under-funding is a contributory factor, which has also limited capacity to capitalise on new infrastructure (such as software or e-DNA techniques).

Novel data capture methods, together with rapid developments in computational and communication technology, are bringing great opportunities to accelerate rates of recording and data sharing, but also carry risks (see Section 3.3). The BDF will need to develop new systems to realise this potential while building capacity to verify and interpret data generated through technological advance.

8.1.2 Review of the SBIF 2018 Recommendations

This study has assessed the relevance of the 2018 SBIF recommendations to England. The assessment is based on the SBIF recommendations in their 2018 report, which were the basis of the consultation meetings for this study.

Due to differences in context, and between the scope of this study and the SBIF Review, many of the recommendations, as worded, are not fully or directly relevant to England. It is recognised that thinking has since evolved in Scotland since 2018 and this, along with some of the principles behind the SBIF recommendations, has been a useful input to the forward-looking interpretation presented in Section 8.2.

8.1.3 Economic analysis of the current species data pathway in England

This report has produced a cost benefit analysis of the current species data pathway in England. Doing so is a very challenging task because the main use of species data is to inform decisions, and:

- The species data pathway involves a variety of public, private and third sector organisations, with overlapping activities, partnerships and funding;

- A 'no species data pathway scenario' of how decisions would be taken with different information is not possible to observe and difficult to define and estimate; and
- Species data form one aspect of environmental data, and so species information is used to a greater or lesser extent to make different environmental decisions. Two important assumptions are made in this respect. Firstly, that society would try to pursue the same environmental outcomes through the same systems but do them less well without access to robust and systematic species data. Secondly, without species data, decisions and management would be reliant on habitat data that, without complimentary data on species, would be unable to design, target or assess habitat condition, quality or changes (e.g., restoration) in sufficient detail to deliver many policies. As a result, the analysis uses a variety of assumptions, described in Section 5 and Appendix F, and produces a range of values for the costs and benefits.

As a result of these challenges confidence levels in the data are noted and a range of benefit and cost estimates are calculated and compared, some of which reflect very conservative assumptions. The conclusions of the work are that the benefits from uses of species data almost certainly outweigh the costs of the species data pathway.

- The minimum estimate of benefits (those relating to benefits estimated with good confidence) is £5.6 billion, so outweigh the minimum estimate of costs (£0.3 billion) by 17:1.
- The best estimate of benefits is £23 billion, so outweighs the best estimate of costs (£0.8 billion) by 28:1 and outweighs the maximum identified costs (£2 billion) by 9:1.
- The maximum estimate of benefits is £34 billion, so outweighs the best estimate of costs (£0.8 billion) by 41:1 and outweighs the maximum identified costs (£2 billion) by 14:1.
- The minimum estimate of benefits is £5.6 billion, so outweighs the maximum estimate of costs (£3 billion) by 2:1, representing the lowest possible ratio.
- The maximum estimate of benefits is £34 billion, so outweighs the minimum estimate of costs (£0.8 billion) by 103:1, representing the highest possible ratio.

This indicates that the cost-benefit analysis estimates a range of impacts between 2:1 and 103:1, reflecting the most extreme sets of assumptions. Therefore, this study estimates, with moderate confidence, that the benefits of the species data pathway outweigh costs by between 14:1 and 28:1, with a best estimate of 28:1. These data demonstrate the high value of species data's role in a variety of decision-making processes in the UK, benefiting the private and public sectors, and wider society.

Non-monetised costs and benefits have been considered and identified in Section 5. There may be some omitted costs, although the main costs of the pathway are believed to be captured. There may be significant omitted benefits, such as the role of species data in the UK national ecosystem assessment. Therefore, the best estimates of the benefit cost ratio are likely to be an underestimate.

The wide range of results reflects moderate or poor confidence in some assumptions and data. Sensitivity analysis has been applied to the choice of discount rates (sensitivity 1), appraisal time horizon (sensitivity 2), treatment of volunteering effort (sensitivity 3) and inclusion of high value indicator estimates (sensitivity 4).

The results of the CBA are influenced by the framing of the analysis, in particular:

- The 'no species data pathway' scenario, which assumes the species data flow pathway does not exist, and therefore that:
  - The volume and range of species data recorded is the same, but the subsequent movement of the data through the species data flow pathway does not happen.
  - The decision-making and knowledge systems that use species data are the same, but they operate without organised data inputs from the species data pathway. Instead, secondary data is used, obtained through bespoke search activities and/or primary research.
  - Benefits identified relate to both of these aspects of species data: the benefits of having better access to data as a result of the species data pathway, and the benefits of having access to better data as a result of its analysis or other value-added activities in the species data pathway.
- The boundaries of the costs and benefits assessed. Benefits are assessed wherever the role of species data in decision-making can be clearly identified. Costs are assessed for the species data pathway, including use and analysis of species data where it is the primary information being considered. There are subsequent use and analysis activities which involve species data, but only as a minority of the information sets used. The costs of these activities are not included in the analysis. This is because it is assumed that these activities would happen in broadly the same way in the absence of species data – in practice the costs involved might be slightly lower (as species data would not be used) or higher (as more effort would be needed to source alternative information).

The cost-benefit analysis results have implications for the review of the applicability of the SBIF recommendations to England. Compared to the SBIF investment appraisal, this Report uses a different framing of key aspects of economic analysis, focussing on the additional benefits from use of species data. This results in economic values that are more directly attributable to the data and the pathway but lower values. Nevertheless, it supports the conclusion that benefits from the species data pathway significantly outweigh costs.

The SBIF recommendations relate to how a transformed biological recording infrastructure could be funded in Scotland. All these recommendations could, in principle, be applicable to England if full funding, from Government and/or a mixture of sources was provided. The economic analysis in this report is the start of developing the case for that funding.

Given the high level of benefits relative to the costs of the current species data pathway identified in Section 6, a case should be investigated for increased funding for the species data pathway in England. What this should fund can draw on the SBIF recommendations and other findings from Sections 2 and 3. These options are discussed further below.

8.2 Future management

The SBIF review and findings of this study identified a lack of adequate funding to support the current approaches to data collection, curation, sharing and use at the local, regional and national levels, in both Scotland and England. The analysis in this study is of the current species data pathway and its benefits but supports consideration of the future management of the species data pathway and the infrastructure that supports it.

8.2.1 Future management of the species data pathway

Future management of the species data pathway will continue to be a collaborative activity between a range of stakeholders from the public, private and third sectors. It has not been within the scope of this study to consult these stakeholders regarding future management approaches and requirements in sufficient depth for firm future plans to be laid out, making further discussion and partnership working essential.

Nevertheless, there are underlying principles, many of which are also reflected in the SBIF Review, that provide a good starting point for this process:

- The current system is underfunded and not achieving its full capability.
- A mixed future funding model is most realistic: in return for public funding, a purpose of the system should be to support public policy outcomes (including those obtaining high current value from the system), other public benefits and private benefits (e.g., planners and developers).
- Shared data principles should be adopted, allowing recovery of a reasonable proportion of costs from users in order to sustain the species data pathway.
- Channels for 'unaffiliated' records to be submitted can be improved to facilitate verification and standards-compliance.
- Incentives for good data access can be strengthened.
- Building on existing systems there are potential economies of scale in software development across the pathway, and also opportunities to invest in professional development (e.g., of verifiers, of using standards).
- Any future technical changes in the pathway (e.g., software, standards) will need to be supported by a well-resourced programme of cultural change, and through a partnership between a lead national body and local expertise nodes (i.e., LERCs).

The purpose of any future change for wider environmental outcomes needs to be made clear, highlighting potential benefits an enhanced system could produce, including but not limited to enabling:

- Multi-taxa integrated analyses, such as the UK "State of Nature" reports and future multi-taxa climate indicators.
- Use of all available data in planning system.
- Monitoring of public policy outcomes, such as for biodiversity net gain and agri-environment schemes.

8.2.2 Potential benefits from enhancing the species data pathway

The conclusion of the economic analysis is that benefits from the species data pathway significantly outweigh its costs. This has significant implications. Firstly, the range of benefits identified suggests that funding for the species data pathway should be possible from a variety of sources: the public sector (justified through the public goods benefits supported), the private sector (for the efficiencies the species data supports), and third sector/philanthropic sources.

Secondly, the benefit-cost ratios identified (of between 14:1 – 28:1) and the challenges in the data recording system suggest that increased funding for it could be justified in several ways. Extra funding to maintain and update elements of the system could be justified simply by avoiding the risks of parts of the pathway

breaking, due to factors such as loss of key individuals (e.g., verifiers with expertise on a specific taxa) or obsolescence of software. In the economics analysis in Section 6 the baseline costs of the current species data pathway in England include some increases against current funding, which are considered necessary to sustain the current data pathway, and also make an allowance for optimism bias in relation to future costs. As a result, they are consistent with sustaining the identified benefits over time, and therefore the present value of benefits calculation.

Therefore, a case can be made for increase in funding for the species data pathway in England, in line with the additional costs required to maintain the data pathway. The costs identified for the baseline include a small sum – approx. £100,000 per year – that are not currently funded. These represent additional spending that is considered necessary to maintain the existing system and the benefits from the species data pathway into the future (as assumed in the benefits analysis). Further to that, additional benefits could be obtained through enhancing the benefits identified in Section 6. Examples include better provision of data for the planning system, or for prioritising or monitoring policy outcomes (such as biodiversity net gain, or agri-environment schemes).

An additional investment of 50% of the current costs would only need to result in a small (5%) increment in the best-case benefits to have benefit-cost ratio of over 3:1 - greater than many other public sector spending options. Investment in developing the BDF and supporting the data pathway should be able to generate additional benefits to a wide range of beneficiaries. Enhanced revenues could be obtained through better provision and promotion of existing services, enhancement of those services, and through compulsion to access species data in line with existing good practice. For example, the land use planning system could require developers to prove access[20] to available data in their proposals, and a standard minimum fee could be agreed for data searches within the species data pathway, generating increased revenue for the system. This can be investigated through further economic appraisal, once an investment/ enhancement case for the species data pathway has been defined and agreed by relevant stakeholders.

To gain a more detailed understanding of the benefits of enhancing the species data pathway, more specific information could be gathered about data users who gain private benefit and therefore currently, or could be expected to, pay for data. A survey of these data users could investigate the type and size of the benefits they realise, the most valuable aspects of the services they currently receive or could receive in future, and views on how they would like those services to develop.

8.2.3 Future design of the species data pathway

Exactly how to design and fund future BDF and manage species data pathways is complex. Developing detailed recommendations will require further consultation with the variety of stakeholders involved, and full analysis of investment options. However, it is widely recognised that funding could be targeted to overcome coordination challenges at a mix of national and local scales. The best mix of spatial scales to target funds will depend on the function being funded and will require consultation with stakeholders. Some stakeholders identify that the current system involves significant duplication of efforts, inconsistent and inefficient approaches. With a better 'baseline' national infrastructure, flows and governance, a more federated, collaborative and shared approach could be adopted at England level – which may be able to

[20] At a minimum accessing data and stating that it is not thought useful to the planning process. This is easier to regulate that requirements to interpret data in a certain way.

realise significant efficiency savings.

Coordinating a national approach will face issues around data ownership that would affect the potential to have fully open data, and volunteers would have to change (with support) their culture and behaviour to adapt to new infrastructure. Such changes carry the risk that expertise could be lost and data sharing reduced.

Changes to the system would have to marry these cultural challenges with exploitation of new IT capabilities. Different data storage and processing approaches allow separation of the ownership, processing, storage and access-management for species data, which historically have been carried out by distinct organisations/ locations. Amongst many future options, there could be a system with:

- A single data storage, possible on the cloud (with relevant security and local information back-ups), allowing application of best available technologies and achievement of economies of scale,
- More standardised data recording and meta-data (e.g., with resource to support training and adoption of Darwin-Core and FAIR), but still allowing bespoke input processes for different organisations/ systems);
- Devolved management responsibilities for data, maintaining the current role of specialist bodies and local knowledge to manage data verification, collation and signposting, and
- Although the background infrastructure could change, the current organisational data input, access and use routes could be maintained. Enhanced services could also be provided, for example with more comprehensive access to searches for data existence (enhancing this current NBN atlas function).

Data access and searches are clearly commercially sensitive issues across the species data pathway and the organisations involved. A balance between open data principles and funding needs is required. This could involve fulfilling open data aims through the existence of all data being shared at a medium to coarse resolution (e.g., with no distinction between nil returns and absence of data). This would protect the ability for organisations to operate paid-for services linked to the data (e.g., at higher resolution, with trend or other analysed data, and advice on nil returns). Therefore, current organisations could retain their identity and purpose, their data stewardship role, and the value they add within the system, such as from local knowledge of LERCs, or specialist knowledge of taxa, (e.g., by National Recording Schemes). However, the data system on which they operate could be integrated.

Such proposals would involve significant culture change amongst the volunteers and organisations in the species data pathway. Funding would be needed for skills development across the species data pathway to enable changes of this type. For example, the adoption of standards takes resource and engagement and a greater understanding on the value of adoption.

A better ‘baseline’ BDF and governance could align with a more federated, collaborative and shared approach at a national level. This could increase the efficiency of species data flows, for example simplifying data flows by removing duplication of efforts or resolving inconsistent and inefficient approaches, in the current system. The shift to a new system including the case for funding it is a strategic move, which needs careful planning and coordination of stakeholder involvement, design/functionality, and the investment case.

Options for simplifying data flow pathways include:

- **Option 1**: Parts of the existing network could stop gathering records into their own systems, and instead encourage records to be added to one of the 'recognised' systems. This would reduce the array of routes for recorders to submit records, and potentially streamline the flow of data. In order for this to happen, however, those recognised systems must be able to guarantee the reliable and prompt supply of data for the purposes of the users and the organisations that supported such an approach would need to be recognised (by funders among others) as playing an important role, even if they don't take control of the resulting data.
- **Option 2**: Effort can be invested in making the current array of systems interoperable. Good progress is being made in this area, but as with other possible changes, it will be challenging to achieve across the board. The use of agreed standards would have a positive effect, but it could be challenging to convince all stakeholders on what those standards should be, and then to apply them (for example many systems currently used lack the ability to track unique record IDs, which would probably have to be a feature of a fully integrated system).
- **Option 3**. In the medium to long term (10+ years) a move towards a national BDF that supports FAIR principles, with publicly funded infrastructure and data stewardship and privately funded data use and services could be explored.

It should be noted that options 1 and 2, and the other steps, are not mutually exclusive, and it would be possible to make progress towards options 1 and/ or 2 while recognising that not all stakeholders will engage. It is likely that there will always be some complexity in this area (for all the reasons given above). Therefore, there will always be a need to put resources into collating, checking and cleaning data from multiple sources, so that it can be used for a variety of purposes. This is particularly the case in the short to medium term, even if action was being taken to implement options 1 and/ or 2.

The costs of these options are not fully understood and need to be established as part of their development. This is partly a reflection of the fact that so much biological recording relies on volunteers, whose input needs to be carefully managed to remain focussed on the activities for which they are motivated to volunteer. This is different to an entirely professional industry, where resources can be directed to areas of greatest need and standards can be more easily enforced.

8.3 Recommendations

This section sets out recommendations arising from this study, based on the interviews and workshops held, the analysis in Sections 2 and 3 and the results of the costs benefit analysis (Section 8), and on the expert opinions of the consortium members who carried out this study. It firstly lays out some principles for managing the species data pathway, and then provides recommendations in sections relating to:

- Defining the biodiversity data framework (BDF)
- Principles and standards
- Investment
- Data use and re-use

Steps in planning the way forward are also suggested. The overarching aim of these principles,

recommendations and future steps is to use the BDF as the basis for streamlining operations through joined-up approaches that deliver high quality, accessible species data for multiple purposes, based on FAIR data principles.

Principles for Managing the Species Data Pathway

The following principles are suggested as a basis from which to manage the species data pathway in future.

Firstly, the specific characteristics of the pathway need to be recognised, in particular the role played by volunteers. Unlike many other volunteering roles in society, which respond to social needs, many of the volunteers contributing to the species data pathway are self-motivated to record and/or steward species data. Species data collection and recording (and possibly stewardship along subsequent parts of the pathway) often do form part of funded activities, meaning useful data are recorded but budgets are not allocated to stewardship of that data. As a result, the definition of open data, whereby it is supplied "at cost of reproduction", is unsustainable for the species data pathway. Note that this is different from funded data collection by professionals, where the funder will usually cover some of the stewardship costs for the data, if only to ensure it fulfils the purpose for which they are funding it.

Secondly, to maximise the value of species data to society, the system should be organised to support FAIR (Findable, Accessible, Interoperable, and Reusable) principles. The relationships between organisations in the data pathway, organisational structures and governance, staff resources and skills, funding and the species data pathway should all aim to support the FAIR principles.

Thirdly, those who support the species data pathway through stewardship of data (in line with FAIR principles) need to be able to cover the costs of efficiently doing so and be financially viable. This means being able to retain rights over data use in order to recover reasonable costs such that system is sustained. What is 'reasonable' will depend on the mix of funding the system receives.

Fourthly, these principles mean that 'accessible' in FAIR means 'shared data' (rather than 'open data') is the appropriate default approach for species data. In return for sufficient funding of the core of the species data pathway and/or infrastructure, the public sector could obtain species data on an open data basis (i.e., available at the resolution it was collected at (other than for sensitive species data). What is 'sufficient' funding would need to be determined based on the characteristics of the system funded and its ability to provide value-added services with species data that generate revenue.

The challenge for managing the species data pathway is to balance the FAIR data principles with the need for financial viability. This needs to be done in a way that supports stewardship of higher quality, shared species data in the long term.

8.3.1 The biodiversity data framework

Define and recognise the biodiversity infrastructure

Consultation carried out for this study reinforces the need to recognise the strengths and advantages of current approaches, but also reveals a widespread lack of clarity regarding roles, responsibilities, processes and even data accessibility.

The "National Biodiversity Network" includes all the varied organisations and individuals involved in collecting and managing species data in the UK, providing an umbrella for the diversity of actors involved in the species data flow pathway from collection to use. Within this network, certain organisations have a key role, undertaking data stewardship on behalf of the NBN. The key physical nodes for species data under current arrangements are the NBN Atlas, LERC databases, NSS databases (including those hosted by BRC on behalf of NSS). The organisations managing these nodes are not part of government. They are not-for-profit organisations, generally with a core professional staff, drawing on the efforts of thousands of volunteers. Collectively, they also engage with hundreds of partner organisations at national and local levels across the whole country. One recommendation arising from this study is that the key role of these organisations (referred to as the "biodiversity data framework" in this report) needs to be clearly defined, identified and recognised to improve species data flows.

Stakeholders commented that marine species data are challenging to generate and to fund. Recognition of a marine dedicated node or partner within the BDF may be advisable due to the specialist expertise and knowledge required to interpret data. While marine data infrastructure is likely to remain substantially separate from terrestrial/freshwater systems (see Section 2), a level of interoperability around the coastal zone and species that use both environments is desirable. The key physical nodes for marine species data under current arrangements is the Archive for Marine Species and Habitats Data (DASSH), the biodiversity data archive that forms part of Marine Environmental Data and Information Network (MEDIN).

The key roles of the four parts of the BDF are identified as:

UK Data Portal

- The only UK wide multi-taxa aggregation of biodiversity data.
- The first port of call to discover biodiversity data, providing signposting to where more or higher resolution data can be accessed.
- Enable the visualisation and analyses of data from multiple data providers in one location.

Local Data Centres

- Provide a local data stewardship service, which facilitates the use of data in decision making, increases and improves recording and links recorders to the processes and policies that use their data.
- Provide a range of data services that support and improve local planning decisions.
- Provide data services to a range of other clients including decision makers and nature conservation bodies

National Schemes and Societies:

- Collate, verify and disseminate records of species within their area of taxonomic expertise. Often in close collaboration with the UKCEH Biological Records Centre.
- Support recorders through development of resources (such as published national atlases and identification guides), provision of feedback and mentoring/tutoring.
- Provide data and information for research and conservation at the national level, often in partnership with others such as the UKCEH Biological Records Centre and the statutory agencies.

Marine Data Portal

- Data Archive Centre for long term curation of marine biodiversity data
- A web portal to make it easy to find UK marine biodiversity data
- Integration of marine biodiversity data with wider environmental data

Clearer definition of the BDF and clarification and endorsement of roles of key elements within it would assist in improving data flows (recorders understanding of how to contribute data, how the contributed data fit into the overall picture, and users understanding how to access data quickly and easily).

Recommendation 1

Recognise the key components of the Biodiversity Data Framework as local data centres, National Schemes & Societies, UK-wide and marine data portals. Each should have a clearly defined role, enabling them to work together as a collaborative, connected community.

Communicate and promote the BDF to recorders, data managers and data users

Consultation carried out for this study has shown that there is a diverse array of actors with an interest in species data. These include central and local government, conservation agencies, conservation NGOs, academia, the private sector, research bodies, landowners and professional institutes. Many of them are supportive of existing BDF operations to an extent, but also recommend changes and improvements to deliver better quality, more joined up approaches, better data accessibility or faster response times. Collective commitment and resource investments will be essential to secure the improvements identified and carry out more effective promotion of the benefits of sharing species data through the BDF.

Recommendation 2

Promote to key sectors the public good, efficiency and conservation benefits of collecting, sharing and using data through a clearly defined and recognised BDF.

Maintain local and taxon-specific databases, but improve transparency of data flows

The general consensus between participants in this study and expert opinion is that local and taxon-specific databases were required to enable local and taxon-specific knowledge to be embedded into the data entry, data management and quality assurance processes. It was acknowledged that in order to improve this system there needs to be greater transparency of data flows and an adoption of the FAIR Data Principles, which could be facilitated via an accreditation system detailed in recommendation 5. The use of a proven, standards-based system should be promoted for data entry and data management, providing flexibility for a national, local or taxonomic focus and allowing the collected data to be shared for verification at different geographic scales and flow to the UK data portal and end use.

While there was some support for one centralised database for all species records, there are inherent risks in having a single system, such as over-reliance on a single IT framework and its managers and of loss of engagement of voluntary recorders through perceptions of remoteness from data stewardship and its use.

Recommendation 3

Maintain the local and taxon-specific biodiversity databases, with a greater emphasis on transparent data flows and data sharing via a UK data portal.

8.3.2 Principles and standards

Although the BDF comprises separate organisations, they should operate to shared principles and standards.

Adopt FAIR data principles

The FAIR principles (Wilkinson et al. 2016), were created to ensure scientific data are accessible for analysis in downstream studies, ensuring the maximum value can be gained from existing data products. Although originally intended to focus on data generated in research (to support peer-reviewed publication), the principles have attracted wider attention and support. They recognise the rapid acceleration of online data access and the need to apply principles that support machine discovery and use of data as well as human operation. The key FAIR data principles are:

- To be Findable, data should be assigned a unique identifier and indexed in a searchable resource.
- To be Accessible, data are retrievable using a standardised communications protocol.
- To be Interoperable, data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- To be Reusable, data are released with a clear and accessible data usage license.

Recommendation 4

Base the generation, management, collation and sharing of species data on FAIR Data Principles – to make species data in England Findable, Accessible, Interoperable and Reusable throughout the species data network. These principles should be adopted in the BDF and should underpin all parts of the system.

In the consultation carried out for this study with key stakeholders, this recommendation received almost unanimous support.

Communicate and promote data standards

While key data standards are in place and applied in parts of the system (see Section 2.3), many NBN participants are unclear regarding their existence, application and importance. Funding pressures sometimes mitigate against compliance. Ensuring compliance with standards might imply the need to invest more time and resources compared with "casual approaches" but can also save time and resource in the longer term. The need to communicate standards is most acute before the data collection stage, in both professional and volunteer recording.

Recommendation 5

Develop and promote data standards throughout the BDF and the wider network of species data collectors and contributors. Particular emphasis should be on the data collection stage to ensure data meets its potential for use.

Certification of species data processes

Consultation for this study has identified that while current data nodes do curate, verify, mobilise and make accessible much high-quality species data, there are weaknesses in the system that cause bottlenecks and constrain data quality and accessibility. Wider adoption of standards, data-sharing protocols and consistent use of clear data flow pathways would assist in providing wider access to high quality data. Training and demonstration of compliance with agreed processes would give assurance to recorders, data contributors, funders and users that improvements were being made as part of a level playing field. Certification of processes, with a focus on advancement and improvement, is preferred to accreditation of organisations.

Recommendation 6

Develop a system of certification of BDF processes to drive high common standards across the network and the adoption of the FAIR principles. This will aid clear data flow pathways and efficient interoperability between BDF nodes.

8.3.3 Investment

Shared costs and benefits approach

The commercial users of species data consulted in this study recognised the fairness of contributing to the costs of data stewardship at the point of use, provided that they would be able to access high quality and comprehensive data without delay. For the private sector it is often preferable to pay for data that is known to be comprehensive (and minimise uncertainty and delay) than it is to receive data for free but with uncertain comprehensiveness, delays in accessing data, contacting multiple providers and/or aligning incompatible datasets.

Across almost all stakeholder groups there was support for a mixed-funding model in which the BDF would not rely exclusively on financial support from any single source, within government or the private sector. A part of that approach could be that in return for sufficient funding of the core of the species data pathway and/or infrastructure, the public sector could obtain species data on an open data basis (i.e., available at the resolution it was collected at (see start of Section 8.3).

While most stakeholders supported the principle of Open Data where practicable, there was recognition that Open Data is not synonymous with free data for all; it is more concerned with data accessibility (https://www.theodi.org). Indeed, the evidence from the early years of the NBN Atlas shows that a drive towards Open Data with no charges levied at the point-of-use to any user, has reduced data accessibility.

This recommendation aligns with the UK Geospatial Strategy approach to Open Data, including the maximisation of economic, social and environmental value and the efficient and fair use of public money. The strategy highlights that the power of location data can increase when it is re-used by more people and businesses beyond its original purpose. Mission 2, to improve access to better location data, identifies high quality and FAIR (Findable, Accessible, Interoperable and Reusable) as the most important aspects of data.

Recommendation 7

Adopt a Data Sharing approach across the pathway, in which many end-users contribute to reasonable costs of data stewardship within the BDF and, where practicable, species data is accessible online.

Government investment in the BDF as a public good

The need for access to high quality species data to support planning, agricultural and environmental legislation and policy is urgent and is clearly recognised in the government's 25-year Environment Plan. The BDF is a public good because use of species data has benefits to a very wide range of functions and all of society. Some of these uses are non-rival (their use of the data does not preclude other uses) and non-excludable (benefits cannot be restricted to certain individuals or groups). Examples of such uses include designation of protected areas to conserve species that all in society gain value from knowing are being protected, and land use planning decisions that conserve species diversity in the wider landscape. Importantly, such decisions also have intergenerational benefits.

A mixed public-private sector funding model is preferred by the majority of stakeholders consulted in this study as it offers a diverse, sustainable framework that promotes innovation outside of government and levers in private sector funding to supplement core investment.

This recommendation aligns with the UK Geospatial Strategy's Mission 2 to improve access to better location data, specifically to support environmental outcomes. It is also supported by the range of public and private benefits, identified in Section 6, that the species data pathway provides.

Recommendation 8

Invest in the BDF, from central government budgets, as part of a mixed public/ private sector funding model, recognising the essential role of accessibility to high quality species data as a public good to deliver environmental legislation and policy.

Invest in verification

The SBIF review found that verification systems were under considerable strain and may be unsustainable in the face of rapidly rising demand (i.e., numbers of records requiring verification). This finding is confirmed for England by consultation in this study. Rising expectations and need for the most current data in decision-making processes make reductions in time delays (due to verification or other reasons) between recording and accessibility a high priority. While some of the demand increase arises from novel recording methods, technology may offer part of the solution through the recent growth in AI-based machine learning.

Much more research is required to improve reliability and design the optimum scope of verification by machine or human experts. Verification processes can also be improved by development and use of a protocol. Its development should involve those undertaking verification, and users of data who rely on verification.

Verification is an important part of the data pathway but suffers from a lack of clarity over what the process entails as a whole and who is responsible for each part of it. Currently there is a lot of variation in how verification is applied, dependent on taxonomy, geography and where in the BDF a record enters the system. A verification protocol could specify:

- The criteria by which record status is assigned; the circumstances in which machine verification could be applied;

- The procedures by which human verifiers are appointed;
- What their remit is; and
- How their roles relate to other parts of the BDF.

This would go a long way to ensuring everyone within the system understands their roles and that everyone submitting records understands exactly what is going to happen to their data. It would detail where there are gaps in the taxonomic expertise available to assist with verification and this would inform strategies for addressing these gaps.

These recommendations align with the UK Geospatial Strategy Mission 3: Enhance capabilities, skills and awareness and Mission 4: Enable innovation.

Recommendation 9
Developed a verification protocol with key stakeholders in the verification process, which aligns to current and future verification requirements and technologies. The resources required to support implementation of the protocol should also be identified.

Recommendation 10
Invest in building capacity for verification, through expert training and the use of new approaches such as automated assessment to support verification decisions.

Invest in data re-use
Consultation in this study revealed that high volumes of species data collected for specialised purposes are not reaching the BDF and do not become available for subsequent re-use. These data could potentially contribute to research and decision-making, development of collective knowledge and understanding of species distributions. Examples include data collected for research and used in peer-reviewed publications and data collected through automated bio-acoustic methods for assessment of a single private-sector development application. In future this issue may be partially addressed through implementation of other recommendations in this report (as in recommendation 5), but a targeted programme of capture and mobilisation of such datasets would improve access to high-quality species data.

Recommendation 11
Invest in processes to capture and mobilise species data generated by research and high volume, novel recording methods that can be used to supply the BDF.

8.3.4 Improving data use and re-use

Make transparent the deposit of data collected in processes that support regulatory compliance
Consultation in this study identified that species data collected by consultants to support assessments of projects that require regulatory approval, such as developments in the planning process, do not normally reach the BDF and therefore are unavailable for re-use. This represents a wasted resource that can add delays and costs to subsequent research, biodiversity assessment and wider decision-making. A new regulation would drive the development of efficient data flows into the BDF from this source.

In the marine environment the Marine Data Exchange provides an example of how this can work successfully for all parties. Established by The Crown Estate in 2013, to store, manage and share offshore survey data throughout project life cycles, this resource helps to make valuable data freely accessible, promote collaboration within the sector, reduce survey costs and ultimately de-risk investment offshore (Crown Estate, 2021).

This recommendation is supported by the Chartered Institute for Ecology and Environmental Management (CIEEM), where it is seen as a step towards the environmental consultancy profession contributing positively to the protection of biodiversity and the efficient re-use of collected data.

New regulation is believed to be essential to achieve widespread implementation of this measure. For over a decade it has been a requirement on CIEEM consultant members to share data, wherever possible, as part of its Code of Professional Conduct. There is scant evidence that this voluntary approach is followed within or outside of CIEEM, so it is regarded as ineffective. A system has been in place since 2013 (Vogel, 2016, Smith et al 2016) to facilitate re-use of consultants' data through a portal managed jointly by CIEEM, BRC, NBNT and ALERC. From 2013 to 2020 a total of 38,062 species records were captured through this portal. A further 23,638 Great crested newt records captured through the Natural England licensing system between 2013 and 2018 are shown on the NBN Atlas. It is believed that these represent a very small proportion of the new species data collected by consultants in assessment of potential developments.

Feedback to those who generate and submit data is required so that they can demonstrate compliance with this recommendation. However, this area is complex due to data protection (of recorder details - subject to GDPR). Emerging tools in academia give researchers a unique global IDs (e.g., ORCID), but this is not yet taken up by consultants. The potential to develop such a system as a measure to support FAIR data and recommendation 12 require further investigation.

This recommendation aligns with the UK Geospatial Strategy's Mission 2 to improve access to better location data, specifically to support environmental outcomes.

Recommendation 12

Mandate the re-use of species data collected by consultants in transparent processes that support regulatory compliance, potentially through new regulation. This will reduce survey costs, improve professional standards and support environmental outcomes.

Require use of data shared through the BDF in applications that may impact biodiversity

Many stakeholders expressed concern that, in a high proportion of planning applications affecting biodiversity, proponents and their specialist advisers were not accessing existing species data and that omission adversely affected the quality of decision making. This recommendation aligns with the UK Geospatial Strategy's Mission 2 to improve access to better location data, specifically to support environmental outcomes.

New regulation is believed to be essential to achieve widespread implementation of this measure. Current voluntary approaches do not achieve the level of take-up of access and use of existing biodiversity data required to ensure that avoidable biodiversity loss is prevented (GLA, 2016).

Recommendation 13

Require proponents of development to certify that best available species data through the BDF have been accessed in the preparation of applications where there is risk of impact on biodiversity, potentially through new regulation. This will help ensure that existing and newly collected species data is equally available to project proponents, regulators and evaluators and will support environmental outcomes.

Improved standards in surveys with public funds

Although there is a requirement for data collected using certain public funds, such as government or National Lottery, to be shared, there is little or no follow-up to ensure data are shared. Furthermore, data collection and sharing are often afterthoughts in project proposals or not given later consideration through project lifecycle nor close. By requiring a data plan with project proposals, with requirement to review and maintain this during a project, it would ensure this vital step is fully considered and therefore more likely to be actioned to facilitate appropriate data archival, coverage of storage costs and onward sharing.

This recommendation aligns with the UK Geospatial Strategy's objective to improve access to public sector data.

Recommendation 14

Require organisations collecting data funded by public money to provide a plan for data collection and sharing, in accordance with FAIR data principles, before funds are received.

8.3.5 Quick wins in the species data pathway

There are a number of actions to manage the species data pathway that can be implemented, and provide benefits, relatively quickly. This study has identified the following such 'quick wins':

- Enhance the contribution of the most knowledgeable contributors, by providing support to activities like skill sharing across the network, to ensure their contribution is retained, such as in verification activities.
- Shore up the system to stop collapse through investment in centralised IT support, provided through ALERC and to NBN and BRC, and additional staff to enhancement resilience and capacity for development.
- Initiate a culture change initiative to support sharing of records, highlighting the added value, and building on shared data principles, demonstrating benefits, including financial returns, that benefit the species pathway as a whole. Incentives could be provided, such as help with costs, recorder ratings/ recognition and feedback. Support could be sought from leading natural scientists and communicators to champion the importance of strong species data stewardship to generate science to underpin UK's role in addressing the climate and ecological emergencies.
- Develop data processes, in particular a verification protocol that is transparent and helps avoid duplication of effort (Recommendation 9).
- Research, develop and promote the implementation of FAIR data principles across the species data pathway.
- Engage funders of the species data pathway to build awareness of good practice, such as on data submission, and the important of a shared data, FAIR principles, and data standards.

Note that while these actions can provide quick wins, they can also generate returns in the longer term, for example a process of culture change will realise its full benefits over 5-10 years and maintain them beyond that time. In addition, several of these actions will initiate positive engagement which would help to overcome past mistrust between different organisations in the species data pathway.

8.3.6 Planning investment in the species data pathway

Investments to improve and simplify the species data pathway in England require consideration of what an enhanced species data pathway would look like and is beyond the scope of this study. The principles for managing the pathway (see start of Section 8.3) are a basis for future development, and any plans will require further economic analysis of the proposed investment. Therefore, this section only considers how such a future investment could be planned.

The process can start with comprehensive analysis and mapping of the current data flows to identify if there are viable options for rationalisation. The SBIF review, undertaken over a significantly longer timescale than this study, generated a detailed analysis and map of data flows in Scotland (see Figure 3.8 in Wilson et al., 2018). An equivalent detailed and quantified analysis for England would provide an informed basis for exploration of options for how to manage the pathway in the short and long term. This could also consider whether organisations in the pathway have appropriate capabilities, for example of staff resources, IT and back-office support. It may be possible for an organisation or organisations to achieve economies of scale in providing that support.

In developing investment options, it is recommended that:

- Full consultation of the organisations and stakeholders involved is carried out to help design and implement changes and
- Provision is made for training to develop the skills needed to realise the benefits and minimise the risks from any changes.

Potential benefits include more efficient integration of data, including unaffiliated records. Development of consistent meta-data and transparent data ownership should be carried out with this in mind.

The funding of the species data pathway from public funds is justified based on the benefits it supports for wider society and taxpayers. Further revenue can be obtained from the sectors who need to access biodiversity data and realise value from doing so. The focus should be on sustaining revenue for data stewardship from those who get value from uses of higher quality, shared species data.

As well as coordinating some UK scale change with the Scottish thinking stemming from SBIF, it would be useful to be able to draw on similar reviews in Wales and Northern Ireland to complete the understanding of the position in all UK nations, and which recommendations are UK wide or country specific. A collaborative approach has been developed in Wales in the last decade, led by the four Welsh LERCs working together and in full co-operation with Natural Resources Wales to deliver online data accessibility to all sectors. There may be good practice approaches in both nations that could be helpful pointers to making improvements to the BDF in England.

Planning of investment in England, and at UK scale, would also benefit from research into the needs of biodiversity data users. These include the short-term needs for data in private-sector decision-making, and the medium to long term biodiversity data needs for supporting policy and legislation. A survey of private sector users could help to quantify the potential benefits from improved data services, also informing the economic appraisal of investment in the system.

A key part of future management of the species data pathway will be technological change: new IT capabilities make it easier distinguish between the management of parts of the species data pathway (e.g., IT hardware and software), and the stewardship of data through that infrastructure. Associated with this it is increasingly possible for functions along the pathway to be divided into specialist functions. Examples of relevant functions include:

- Human skills, such as:
  - Place-based knowledge (e.g., local interpretation and expertise), and
  - Specialist work on data (e.g., data verification, analysis of species trends, adoption of standards such as DarwinCore);
- IT functions, such as:
  - Different data entry interfaces/ submission routes (but adopting data standards to help make them interoperable);
  - Tracking data with different commercial status within the system; and
  - Value-added service interfaces.

If there are potentially viable changes identified for the species data pathway, a theory of change should be developed, allowing consideration of how to support and fund these changes. In should be noted that changes in infrastructure and/or process would also require a change in culture and behaviour for staff and volunteers to adapt, and hence funding would be needed to support any such transition. If the benefits of any change are not communicated to and accepted by those involved through lack of transition planning and funding, this increases the risk that expertise could be lost and data sharing reduced. Funding this culture change should be seen as an investment in retaining and growing the total contribution of all volunteers who support the species data pathway (valued at over £55m per year).

References

Abcfinance.co.uk. (2021). Commercial mortgage rates & fees. [online]. Available at: https://abcfinance.co.uk/commercial-mortgages/calculator/rates-and-fees/#:~:text=If%20you're%20planning%20on,between%202.75%25%20and%207%25.

Agnesi, S., Mo, G., Annunziatellis, A., Chaniotis, P., Korpinen, S., Snoj, L., Globevnik, L., Tunesi, L., Reker, J., 2017, Assessing Europe's Marine Protected Area networks – Proposed methodologies and scenarios, ed. Künitzer, A. ETC/ICM Technical Report 2/2017, Magdeburg: European Topic Centre on inland, coastal and marine waters, 72 pp.

ALERC. (2015). Association of Local Environmental Record Centres - Strategic Plan 2015 – 2020. [online]. Available at: http://www.alerc.org.uk/publications-and-positions.html

Alluvium. (2016). Assessment of the Atlas of Living Australia's Impact and Value. [online]. Available at: https://www.csiro.au/en/About/Our-impact/Our-impact-in-action/Natural-environment/Atlas-of-Living-Australia

August, T., Harvey, M., Lightfoot, P., Kilbey, D., Papadopoulos, T., and Jepson, P. (2015). Emerging technologies for biological recording. *Biological Journal of the Linnean Society* 115 (3): 731–749. http://onlinelibrary.wiley.com/doi/10.1111/bij.12534/abstract

Bateman, I.J. and Mace, G.M. (2020). The natural capital framework for sustainably efficient and equitable decision making. Nat Sustan, 3, pp.776-783. [online]. Available at: https://www.nature.com/articles/s41893-020-0552-3

Beale, S., Bending, M., Trueman, P. (2007). An Economic Analysis of Environmental Interventions That Promote Physical Activity. University of York: York Health Economics Consortium.

Barratt Developments Plc. (2020). Annual Report and Accounts 2020. [online]. Available at: https://www.barrattdevelopments.co.uk/investors/results-reports-and-presentations/rp-2020

Berkeley Group. (2020). Annual Report. [online]. Available at: https://www.berkeleygroup.co.uk/about-us/investor-information/annual-report

Berry, R. J. (1988). Biological Survey: Need & Network. Report of a Working Party set up by the Linnean Society of London. [online]. Available at:

Bonnet, P., Goëau, H., Hang, S.T., Lasseck, M., Šulc, M., Malécot, V., Jauzein, P., Melet, J.-C., You, C., and Joly, A. (2018). Plant Identification: Experts vs. Machines in the Era of Deep Learning. In Multimedia Tools and Applications for Environmental & Biodiversity Informatics Multimedia Systems and Applications., A. Joly, S. Vrochidis, K. Karatzas, A. Karppinen, and P. Bonnet, eds. (Springer International Publishing), pp. 131–149.

Bradfer-Lawrence, T., Gardner, N., Bunnefeld, L., Bunnefeld, N., Willis, S.G. and Dent, D.H. (2019).

Guidelines for the use of acoustic indices in environmental research. Methods in Ecology and Evolution, 10(10), pp.1796-1807.

Breeze, T., Bailey, A., Balcbome, K., Brereton, T., Comont, R., Edwards, M., Garratt, M., Harvey, M., Hawes, C., Isaac, N., Jitlal, M., Jones, C., Kunin, W., Lee, P., Morris, R., Musgrove, A., O'Connor, R., Peyton, J., Potts, S., Roberts, S., Roy, D., Roy, H., Tang, C., Vangergen, A., Carvell. (2020). Pollinator monitoring more than pays for itself. Journal of Applied Ecology. [online]. Available at: https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/1365-2664.13755

Buglife (2018). Annual Review 2018. [online]. Available at: https://cdn.buglife.org.uk/2020/04/AnnualReview2018-FINAL.pdf

Burnett, J., Copp, C. and Harding, P. (1995). Biological recording in the United Kingdom - present practice and future development. Summary report. London, Department of the Environment, 27pp. [online]. Available at: http://nora.nerc.ac.uk/id/eprint/7867/

Bush, A., Compson, Z.G., Monk, W., Porter, T.M., Steeves, R., Emilson, E.J., Gagne, N., Hajibabaei, M., Roy, M. & Baird, D. (2019). Studying ecosystems with DNA metabarcoding: Lessons from biomonitoring of aquatic macroinvertebrates. Frontiers in Ecology and Evolution, 7, p.434.

Butterfly Conservation (2020). Annual Report and Accounts 2019/20. Available at: https://butterfly-conservation.org/sites/default/files/2020-09/BC%20Annual%20Report%20&%20Accounts%202020.pdf

Chatham House. (2021). Chatham House Rules. [online]. Available at: https://www.chathamhouse.org/about-us/chatham-house-rule

Christie and Rayment. (2012). An Economic Assessment of the Ecosystem Service Benefits Derived from the SSSI Biodiversity Conservation Policy in England and Wales, Ecosystem Services. [online]. Available at: https://www.sciencedirect.com/science/article/pii/S2212041612000095

Claxton K, Martin S, Soares M, Rice N, Spackman E, Hinde S, et al. (2015). Methods for the Estimation of the NICE Cost Effectiveness Threshold. Health Technology Assess. [online]. Available at: https://www.journalslibrary.nihr.ac.uk/hta/hta19140/#/full-report

Crown Estate (2021). Marine Data Exchange. https://www.thecrownestate.co.uk/en-gb/what-we-do/on-the-seabed/marine-planning/

Darras, K., Furnas, B., Fitriawan, I., Mulyani, Y. and Tscharntke, T. (2018). Estimating bird detection distances in sound recordings for standardizing detection ranges and distance sampling. Methods in Ecology and Evolution, 9(9), pp.1928-1938.

Dasgupta, P. (2021), The Economics of Biodiversity: The Dasgupta Review. (London: HM Treasury)
DWC. DarwinCore standard. [online]. Available at: https://dwc.tdwg.org/

Department for Environment, Food and Rural Agriculture. (2013a). UK CAP allocations announced. [online]. Available at: https://www.gov.uk/government/news/uk-cap-allocations-announced

Department for Environment, Food and Rural Agriculture. (2013b). Implementation of CAP reform in England Evidence Paper. [online]. Available at: https://consult.defra.gov.uk/agricultural-policy/cap-consultation/supporting_documents/131022%20CAP%20Evidence%20Paper%20%20Final.pdf

Department for Environment, Food and Rural Agriculture. (2015). Evaluation of the catchment based approach - Economic assessment of the catchment based approach.

Department for Environment, Food and Rural Agriculture. (2019). Biodiversity net gain and local nature recovery strategies. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/839610/net-gain-ia.pdf

Department for Environment, Food and Rural Agriculture. (2020a). Guidance: Enabling a Natural Capital Approach. [online]. Available at: https://www.gov.uk/guidance/enabling-a-natural-capital-approach-enca#enca-featured-tools

Department for Environment, Food and Rural Affairs. (2020b). UK Biodiversity Indicators 2020: A2. Taking action for nature: volunteer time spent in conservation. [online]. Available at: https://jncc.gov.uk/our-work/ukbi-a2-conservation-volunteering/

Defra and Environment Agency. (2018). Press release: £5 billion investment by water companies to benefit the natural environment. [online]. Available at: https://www.gov.uk/government/news/5-billion-investment-by-water-companies-to-benefit-the-natural-environment

Department for Transport. (2020). Guidance: TAG Data Book. [online]. Available at: https://www.gov.uk/government/publications/tag-data-book

Doohan, B., Fuller, S., Parsons, S. & Peterson, E.E. (2019). The sound of management: Acoustic monitoring for agricultural industries. Ecological Indicators, 96, pp.739-746.

Eble et al. (2020) Chapter Two - Marine environmental DNA: Approaches, applications, and opportunities. Advances in Marine Biology, 86, 141-169. https://doi.org/10.1016/bs.amb.2020.01.001

eftec and ABPmer (2019). Marine Environmental Data and Information Network (MEDIN) Cost Benefit Analysis. Final report for MEDIN. [online]. Available at: https://medin.org.uk/about/key-documents

eftec, SQW, CEH, ABPmer (2019). The ecosystem contribution to tourism and outdoor leisure. Report to Defra, September 2019. [online]. Available at: http://randd.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&ProjectID=20245

eftec, WSP and ABPmer. (forthcoming). Biodiversity Net Gain Market Analysis. Report for Defra

Environment Agency. (n.d.). Working paper: Estimates for the benefits and costs for the control of floating pennywort (Hydrocotyle ranunculoides). Unpublished.

Environment Agency. (2020). Water Industry National Environment Programme (WiNEP v3, National Dataset, Public version). [online]. Available at: https://data.gov.uk/dataset/a1b25bcb-9d42-4227-9b3a-34782763f0c0/water-industry-national-environment-programme

Environment Agency. (2021). WFD River, Canal and Surface Water Transfer Water Bodies Cycle 2 Classification 2019. [online]. Available at: https://data.gov.uk/dataset/a2a10b3a-2049-48ba-9ab5-fbc3ae26c9f9/wfd-river-canal-and-surface-water-transfer-water-bodies-cycle-2-classification-2019

Eldridge, A., Guyot, P., Moscoso, P., Johnston, A., Eyre-Walker, Y. & Peck, M. (2018). Sounding out ecoacoustic metrics: Avian species richness is predicted by acoustic indices in temperate but not tropical habitats. Ecological Indicators, 95, pp.939-952.

Fezzi, C., Bateman, I. and Ferrini, S. (2013). Using revealed preferences to estimate the value of travel time to recreation sites, Journal of Environmental Economics and Management, 67(1), p.58-70. [online]. Available at: https://www.sciencedirect.com/science/article/pii/S0095069613000880?via%3Dihub

Fingas M. (2019). Chapter 5 - Remote Sensing for Marine Management. In, World Seas: An Environmental Evaluation, Volume III: Ecological Issues and Environmental Impacts Book • Second Edition • 2019. https://doi.org/10.1016/B978-0-12-805052-1.00005-X

Fox, R., Bourn, N.A.D., Dennis, E.B. et al. Opinions of citizen scientists on open access to UK butterfly and moth occurrence data. Biodivers Conserv 28, 3321–3341 (2019). https://doi.org/10.1007/s10531-019-01824-6

Gibb, R, Browning, E, Glover-Kapfer, P, Jones, KE. Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. *Methods Ecol Evol*. 2019; 10: 169– 185. https://doi.org/10.1111/2041-210X.13101

Goëau, H., Bonnet, P. & Joly, A. (2018) Overview of ExpertLifeCLEF 2018: how far automated identification systems are from the best experts? In: CLEF working notes 2018

GHK. (2010). Update of the UKBAP Costs and Funding Data. [online]. Available at: http://sciencesearch.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&Completed=0&ProjectID=17218

Greater London Authority. (2016). Planning for Biodiversity?. [online]. Available at: https://www.london.gov.uk/sites/default/files/biodiversity_and_planning_research_report_0.pdf

Greenhalgh, J.A., Genner, M.J., Jones, G. and Desjonquères, C., 2020. The role of freshwater bioacoustics in ecological research. Wiley Interdisciplinary Reviews: Water, 7(3), p.e1416.

Groom, Q., Weatherdon, L. & Geijzendorffer, I.R. (2017). Is citizen science an open science in the case of biodiversity observations? https://doi.org/10.1111/1365-2664.12767

Griffin, E., Crompvoets, J. and Coote, A. (2017). A marine spatial data infrastructure in New Zealand: a systematic review on the Cost-benefits. Journal of Spatial Science, 64:1, 33-47, DOI: 10.1080/14498596.2017.1372227

Hassall, I., Cheffings, C., Robinson, A. & Robinson, P. (2020). Review of biodiversity data use in the Country Nature Conservation Bodies. *JNCC Report No.670*, JNCC, Peterborough, ISSN 0963-9091.

Hayhow DB, Eaton MA, Stanbury AJ, Burns F, Kirby WB, Bailey N, Beckmann B, Bedford J, Boersch-Supan PH, Coomber F, Dennis EB, Dolman SJ, Dunn E, Hall J, Harrower C, Hatfield JH, Hawley J, Haysom K, Hughes J, Johns DG, Mathews F, McQuatters-Gollop A, Noble DG, Outhwaite CL, Pearce-Higgins JW, Pescott OL, Powney GD and Symes N. (2019). The State of Nature 2019. The State of Nature partnership. [online]. Available at: https://nbn.org.uk/wp-content/uploads/2019/09/State-of-Nature-2019-UK-full-report.pdf

Heritage Fund. (2020). Application guidance. [online]. Available at: https://www.heritagefund.org.uk/funding/application-guidance-national-lottery-grants-heritage-ps3000-ps10000-and-ps10000-ps100000

HMRC. (2015). Annual UK Property Transactions Statistics 2015. [online]. Available at: https://www.gov.uk/government/statistics/annual-uk-property-transactions-statistics-2013

HM Treasury. (2020). The Green Book: appraisal and evaluation in central government. [online]. Available from: https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-governent

HM Treasury. (2021). GDP deflators at market prices, and money GDP, December 2020 (Quarterly National Accounts). [online]. Available at: https://www.gov.uk/government/collections/gdp-deflators-at-market-prices-and-money-gdp

JNCC (2018). Terrestrial evidence review part 1: The value of JNCC's terrestrial evidence programme. [online]. Available at: https://data.jncc.gov.uk/data/819af873-5b9c-424a-9c0f-c73aec2e181f/JNCC18-01-Terrestrial-evidence-review-part1.pdf

JNCC. (2020). MPA network assessments. [online]. Available at: https://jncc.gov.uk/our-work/mpa-network-assessments/

Machado, R.B., Aguiar, L. and Jones, G., 2017. Do acoustic indices reflect the characteristics of bird communities in the savannas of Central Brazil?. Landscape and Urban Planning, 162, pp.36-43.

Ministry of Housing, Communities and Local Government. (2020). Live tables on planning application statistics. [online]. Available at: https://www.gov.uk/government/statistical-data-sets/live-tables-on-

planning-application-statistics

MEDIN. (2020). MEDIN annual report for 2019/2020. Available at: https://www.medin.org.uk/sites/medin/files/documents/MEDIN_Annual_Report_2020_final.pdf

Metcalfe, P. (2012). Update of CRP WFD Benefit Value - Economic Component, report for the Environment Agency.

Mora-Fallas, A., Goëau, H., Mazer, S., Love, N., Mata-Montero, E., Bonnet, P. & Joly, A. (2019) Accelerating the Automated Detection, Counting and Measurements of Reproductive Organs in Herbarium Collections in the Era of Deep Learning. Biodiversity Information Science and Standards 3: e37341. https://doi.org/10.3897/biss.3.37341

National History Museum and Vivid Economics. (2021). The Urgency of Biodiversity Action. Report prepared for HM Treasury. [online]. Available at: https://www.nhm.ac.uk/events/the-urgency-of-biodiversity-action.html

Natural England (2019). The Biodiversity Metric 2.0 (JPO20). [online]. Available at: http://publications.naturalengland.org.uk/publication/5850908674228224

National Trust. (2021). Selborne Common. Available at: https://www.nationaltrust.org.uk/selborne-common

NBN Trust. (2020). National Biodiversity Network (NBN) Tools, resources and publications: NBN Record Cleaner. Available at: https://nbn.org.uk/tools-and-resources/nbn-toolbox/nbn-record-cleaner/

NERA Economic Consulting. (2007). The benefits of Water Framework Directive Programmes of Measures in England and Wales. Final report to Defra.

ONS. (2019). UK natural capital: urban accounts. Available at: https://www.ons.gov.uk/economy/environmentalaccounts/bulletins/uknaturalcapital/urbanaccounts#cultural-services

ONS. (2020). Dataset: Annual Survey of Hours and Earnings time series of selected estimates. [online]. Available at: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/ashe1997to2015selectedestimates

Piper, A., Batovska, J., Cogan, N., Weiss, J., Cunningham, J., Rodoni, B. & Blacket, M. (2019). Prospects and challenges of implementing DNA metabarcoding for high-throughput insect surveillance GigaScience 8(8). https://dx.doi.org/10.1093/gigascience/giz092

Pocock, M.J., Newson, S.E., Henderson, I.G., Peyton, J., Sutherland, W.J., Noble, D.G., Ball, S.G., Beckmann, B.C., Biggs, J., Brereton, T. and Bullock, D.J., 2015. Developing and enhancing biodiversity monitoring

programmes: a collaborative assessment of priorities. Journal of Applied Ecology, 52(3), pp.686-695.

Potts, S.G., Dauber, J., Hochkirch, A., Oteman, B., Roy, D.B., Ahrné, K., Biesmeijer, K., Breeze, T.D., Carvell, C., Ferreira, C., FitzPatrick, Ú., Isaac, N.J.B., Kuussaari, M., Ljubomirov, T., Maes, J., Ngo, H., Pardo, A., Polce, C., Quaranta, M., Settele, J., Sorg, M., Stefanescu, C., Vujić, A., Proposal for an EU Pollinator Monitoring Scheme, EUR 30416 EN, Publications Office of the European Union, Ispra, 2021, ISBN 978-92-76-23859-1, doi:10.2760/881843, JRC122225. [online]. Available at: https://wikis.ec.europa.eu/display/EUPKH/EU+Pollinator+Monitoring+Scheme

Rhinehart, TA, Chronister, LM, Devlin, T, Kitzes, J. Acoustic localization of terrestrial wildlife: Current practices and future opportunities. Ecol Evol. 2020; 10: 6794– 6818. https://doi.org/10.1002/ece3.6216

RSPB (2020). Annual report 2020. Available at: https://www.rspb.org.uk/globalassets/downloads/annual-report-2020/rspb-annual-report-08-10-2020-signedoff-interacrtive-pdf.pdf?sourcecode=GENWEB0026

Sommerville, A. (1975). A Guide to Biological Recording in Scotland. [online]. Available at: https://nbn.org.uk/about-us/where-we-are/in-scotland/the-sbif-review/sbif-review-archives/handbooks-guidelines/

Sugai, L.S.M., Desjonqueres, C., Silva, T.S.F. & Llusia, D. (2019). A roadmap for survey designs in terrestrial acoustic monitoring. Remote Sensing in Ecology and Conservation.

Teixeira, D., Maron, M. & van Rensburg, B.J. (2019). Bioacoustic monitoring of animal vocal behavior for conservation. Conservation Science and Practice, 1(8), p.e72.

Terry, J.C.D., Roy, H.E. & August, T.A. (2020). Thinking like a naturalist: Enhancing computer vision of citizen science images by harnessing contextual data. Methods in Ecology and Evolution, 11(2), pp.303-315.

UK Parliament. (2016). Future funding of the national wildlife crime unit. [online]. Available at: https://questions-statements.parliament.uk/written-statements/detail/2016-03-01/HCWS561

van Horn, G., Mac Aodha, O., Song, Y., Cui, Y., Sun, C., Shepard, A., Adam, H., Perona, P., and Belongie, S. (2018). The iNaturalist Species Classification and Detection Dataset. In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (IEEE), pp. 8769–8778.

Villanueva-Rivera, L.J., Pijanowski, B.C., Doucette, J. & Pekin, B. (2011). A primer of acoustic analysis for landscape ecologists. Landscape ecology, 26(9), p.1233.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., Bonino da Silva Santos, L., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.R., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J.,

Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. & Mons, B. 2016. Comment: The FAIR guiding principles for scientific data management and stewardship. Scientific Data, 3, 160018, https://doi.org/10.1038/sdata.2016.18.

Williams, F., Eschen, R., Harris, A., Djeddour,D., Pratt, C., Shaw, R.S., Varia, S., Lamontagne-Godwin, J., Thomas, S.E. and Murphy S.T. (2010). The Economic Cost of Invasive Non-Native Species on Great Britain. Report no. CABI/001/09. [online]. Available at: https://www.researchgate.net/publication/298559361_The_Economic_Cost_of_Invasive_Non-Native_Species_on_Great_Britain

Wilson, E., Edwards, L., Judge, J., Johnston, C., Stroud, R., McLeod, C. and Bamforth, L. (2018). A Review of the Biological Recording Infrastructure in Scotland by the Scottish Biodiversity Information Forum: Enabling Scotland to be a global leader for biodiversity. Scottish Biodiversity Information Forum Commissioned Report No. 1. Available at: https://nbn.org.uk/about-us/where-we-are/in-scotland/the-sbif-review/

Wäldchen, J. and Mäder, P., 2018. Machine learning for image-based species identification. Methods in Ecology and Evolution, 9(11), pp.2216-2225.

Xie, J., Hu, K., Zhu, M. & Guo, Y. (2020). Data-driven analysis of global research trends in bioacoustics and ecoacoustics from 1991 to 2018. Ecological Informatics, 57, p.101068.

Appendix A - Figure 2.3

5a) Data flow models: Model 0 – current situation

Overall, in the current situation resources are in play at the local level and within schemes; little is centralised or nationalised so efficiency is minimised but local value is maximised.

Currently, users need to go to multiple data providers (including recorders, LERCs, schemes, organisations, groups, social media, etc) to ascertain and access all possible sources of data

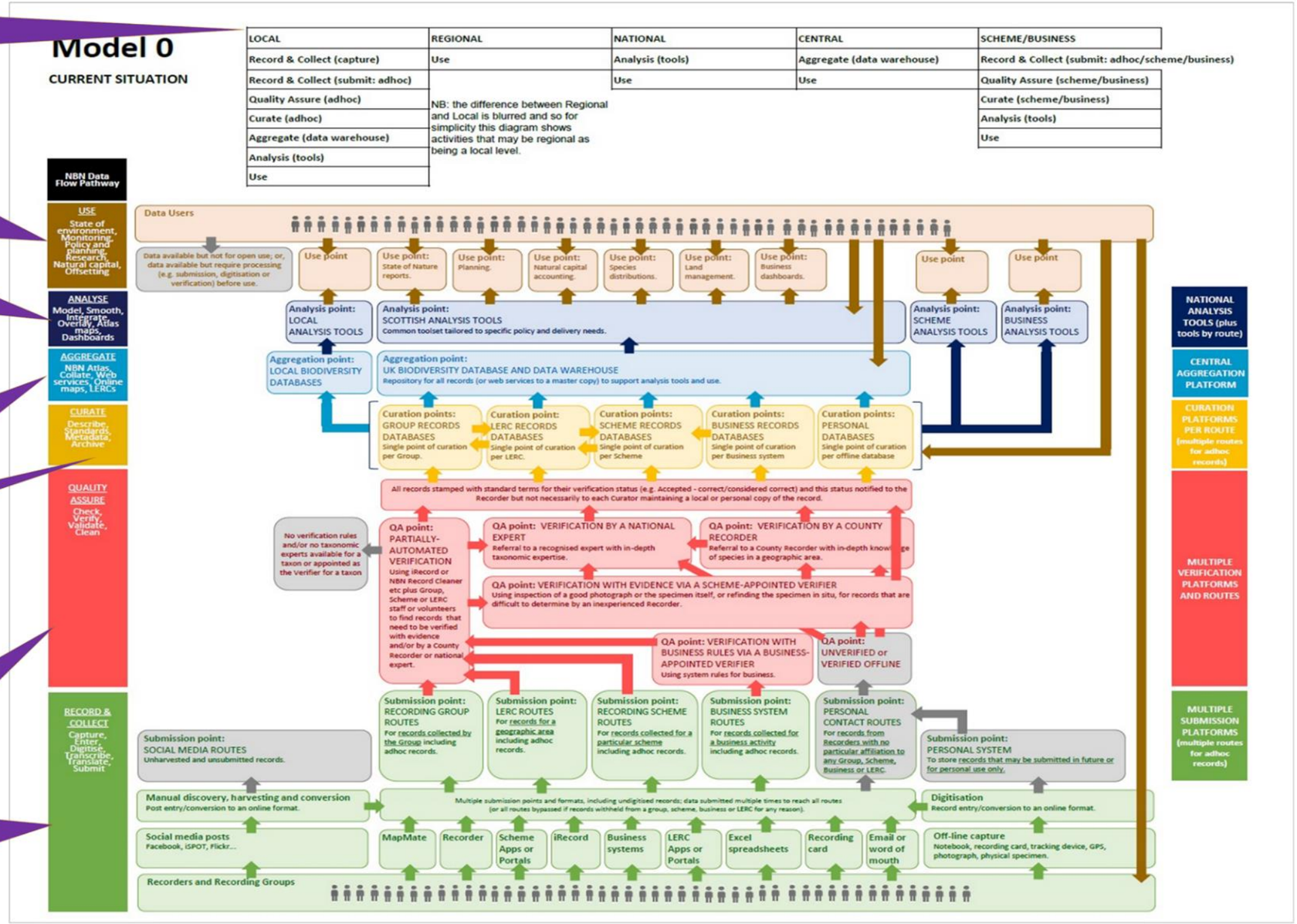
Currently, analysis tools are shaped by national government or LERC service user needs; schemes and organisations also provide analytical tools.

Currently, the NBN Atlas is the primary aggregation point for the UK but there are others which also aggregate (potentially a different version of) records for local and regional use.

Currently, records may be held by more than one curator and/or in databases that aren't openly available. There is effort involved in exchanging records between curators and aggregation points.

This is a generalised view of verification; there are bottlenecks and verification rules may not be comprehensive or current. It can be hard to contact recorders as there is no register of contact details nor contact made.

Currently, lots of different formats are welcomed whether on or offline; social media posts require manual 'harvesting'. Recorders may submit a record multiple times to ensure it reaches all curators or users.



Source: Outputs from the SBIF Review Workshop on Data Flows

Appendix B - Stakeholders participating in the consultation

Anglia Ruskin University
Association of Local Government Ecologists (ALGE)
Austin Foot Ecology
Bat Conservation Trust
Bees, Wasps and Ants Recording Society
Botanical Society of Britain and Ireland
Bristol Museums, Galleries & Archives
Bristol Regional Environmental Record Centre
British Dragonfly Society
British Lichen Society
Bumblebee Conservation Trust
Butterfly Conservation
Camden Council
Centre for Environment, Fisheries and Aquaculture Science (Cefas)
Chartered Institute of Ecology and Environmental Management (CIEEM)
Cheshire WildLife Trust
Conchological Society of Great Britain and Ireland
Department for Environment, Food and Rural Affairs
Derbyshire Wildlife Trust
Devon Biodiversity Records Centre (DBRC)
Environment Agency
Essex Field Club
Exmoor National Park Authority
Golder
Greenspace Information for Greater London (GiGL)
Ground Beetle Recording Scheme
Hoverfly Recording Scheme
Joint Nature Conservation Committee (JNCC)
Marine Biological Association
Natural England
Nature Scot
NatureSpot
New Forest National Park Authority
North and East Yorkshire Ecological Data Centre (NEYEDC)
Northumberland County Council
Riverfly Recording Scheme
Royal Society for the Protection of Birds (RSPB)

SeaSearch
Soldier Beetles and Allies Recording Scheme
Somerset Wildlife Trust
Sunderland City Council
Surry Biological Information Centre (SBIC)
Swift Ecology Ltd.
Telford and Wrekin Council
Thames Valley Environmental Records Centre (TVERC)
The Earthworm Society
The Mammal Society
The University of Sheffield
Woodmeadow Trust
Worcestershire Biological Records Centre (WBRC)
Yorkshire Naturalists' Union
Yorkshire Wildlife Trust

Appendix C - Data flow pathway process

Data pathway group	Process	Process meaning in the context of species data
Record and Collect	Capture	Human observation or automated detection of species record in the wild
	Enter	Enter the record in a format that can be communicated
	Digitise	Transfer data from one format (such as paper) to another (such as digital), to facilitate its accessibility and integration with other data
	Translate	Transfer data held in one schema (e.g., Common species names) into another (e.g., Scientific names)
Quality Assure	Check	Ensure that the data has been communicated from the originator correctly
	Verify	Confirm the truth of the recorded data (e.g., that a species has been correctly identified)
	Validate	Confirm the accuracy of the recorded data (e.g., its geoprecision)
	Clean	Identify or remove confirmed errors in the recorded data, replace errors with corrected data, maintain audit trail
Curate	Describe	Make a description of the dataset to assist data users in assessment of its relevance and utility to their purpose
	Standards	Compare the supplied dataset with recognised standards, communicate with data provider, ensure compliance with standards
	Metadata	Assist the data provider with completion of standard metadata to assist data users in assessment of its relevance and utility to their purpose
	Archive	Ensure that the data is secure and accessible for all current and future uses
Aggregate	Collate	Combine data from multiple sources into a coherent database to improve its accessibility and utility
	Web services	Provide and support a digital interface to enable users to find and use data
	Online maps	Combine multiple spatial datasets and provide online access in map view format
Analyse	Model	Extrapolate or interpolate from recorded data to predict likely presence of species at locations where they are currently not recorded
	Integrate	Analyse species data in combination with other environmental datasets to improve understanding
Use	State of Environment	
	Monitoring	
	Policy and planning	
	Research	
	Natural Capital	
	Offsetting	

Appendix D - Stakeholder interview questions

About Your Organisation

Sector - In which of the indicated stakeholder sectors (Table 1) do you place your organisation?

1. Geographical Focus / Coverage -
    a. Do you have a local or national remit or work at multiple levels?
    b. What is the main geographic focus of your work/ data?

2. Environmental Scope - Which of the following are you mainly involved/interested in?
    a. Terrestrial
    b. Freshwater
    c. Coastal
    d. Marine
    e. All of the above

3. Current Biodiversity Information Involvement – Is your organisation involved in:
    a. Collecting/capturing species data?
    b. Managing species data?
    c. Using information?
    d. Supporting or otherwise facilitating others in these activities?
    e. Seeking to encourage/enable more effective ways of operating?

4. Future Biodiversity Information Involvement - In relation to the points in Q2 and Q3.
    a. Is this likely to change in future? Yes/No?
    b. Whether Yes/No - Why?
    c. If Yes, how might the status quo change?
    d. If Yes, what particular measures/needs do you envisage?

COBP Interview Questionnaire

Most of the questions are presented as statements. You are invited to comment as follows: [Ranked responses: 0 – Don't know/Not in a position to comment; 1 – Strongly disagree; 2 – Disagree; 3 – Neither agree/disagree; 4 – Agree; 5 – Strongly agree].

Views on species data & wider biodiversity information provision, use and value

Q1 Biodiversity information needs

a) Biodiversity information needs for different purposes at national and local level in England have been adequately identified.

b) Data needs have been identified in relation to:

i. All species
ii. All marine species
iii. All terrestrial species
iv. Species assemblages/communities
v. Key Indicator species

vi. Protected species
vii. Other priority species
viii. Invasive non-native species

c) If you disagree with Q1(a), what are the major sectoral problems or wider issues?

d) Please suggest examples of where species data needs have been identified.

Q2: Biological recording in the UK is presently well set up to meet species data needs at different geographical scales (from local to national):

a) Overall

b) for your sector

c) Please give examples where species data needs are being met [text answer]

d) Existing infrastructure provides a sound basis for meeting future data needs

e) If you disagree with 2d) why is this? [text answer]

	(a)Species data needs	
Geographical scale	**1. Overall**	**2. My sector**
i. UK/England	[0 1 2 3 4 5]	[0 1 2 3 4 5]
ii. Local/Regional	[0 1 2 3 4 5]	[0 1 2 3 4 5]
iii. To integrate/scale effectively between levels	[0 1 2 3 4 5]	[0 1 2 3 4 5]

Data collection and capture

Q3 Recording effort

a) Current recording activity successfully targets the gathering of all the critical or important species data that are required.

b) A large quantity of species data is collected overall but for many species groups the resulting picture is incomplete.

c) Where it is possible to target and coordinate recording effort, the resulting data are likely to be more useful.

d) It is important to provide opportunities to help engage and develop new recorders.

Q4 What do you identify as the main obstacles or bottlenecks to records being made?

a) To what extent do these issues undermine the value of the collected datasets? [text answer]

b) What measures can you suggest in order to address these issues (Q3, Q4) – and what resources or other change would this require? [text answer]

c) How necessary will the resulting changes be in relation to the additional data needs likely to result from legislative or policy drivers? (e.g., Environment Bill, Planning White Paper) [text answer]

Q5 What are the opportunities to improve the species data landscape in relation to recording, with regard to:

a) Changes of approach?
b) Adopting existing good practice?
c) Further development of handheld recording apps?
d) Recruitment of additional recorders?
e) Recruitment of volunteers in administrative/coordinating roles?

f) External support for recorders/recording organisations?
g) Other possibilities?

Q6 What do you see as the main potential benefits of more novel sampling methods? (e.g., eDNA sampling, acoustic monitoring, camera trapping, drones, etc)

Data Submission

Q7 To what extent do you consider that the lack of transfer between the data originator and interested parties (e.g., NSS, LERCs, iRecord) to be an issue that needs to be addressed?

a) Overall?
b) In relation to data not being verified or validated?

Q8 Do you have any suggestions to facilitate and encourage the submission of data to interested parties by the originators?

a) Are there any sectors where this particularly needs to be addressed?

Quality Assurance

Q9 Effective quality assurance is vital to the process of checking records prior to their incorporation within a dataset of records of known quality.

Q10 There is value in sharing unverified data, e.g., in relation to the presence of invasive species but consideration will need to be given to where and in what form.

Q11 If unchecked data are to be incorporated in a database or published dataset their status need to be adequately flagged to prevent inappropriate usage.

Q12

a) To what extent, if any, do current quality assurance processes cause significant delay to data flow?
b) Are there enough resources (people/time) available to carry out quality assurance, and if not, how can we improve things?
c) To what extent could the QA process be assisted by

a) sharing the load (e.g., with sub-verifiers to handle common/distinct species)
b) increasing the extent to which spatial/species rules 'intelligently' process records?

Data management: aggregation, curation, analysis [This section, Q13,14) only used for data managing organisations]

Q13 If you use a bespoke database system (e.g., Marine Recorder, Recorder 6, MapMate) for managing species records, to what extent:

a) does this meet your existing needs, or could it be improved?
b) Is it likely to meet your future needs and aspirations?
c) Are you concerned that there is a risk of it not being supported in future? (e.g., under new operating systems or corporate IT arrangements)?

Q14 If a new 'universal', standards-driven biological recording database or modular system were to be developed, what would be the key features you would want it to incorporate in order to address issues that you have highlighted, including future needs?

a) What other systems would it need to work with?
b) What benefits would result from such investment?

Data Sharing

Q15 What measures do you think could be employed to ensure that species data collected in connection with the planning system can be made more widely available?

Q16 A considerable quantity and much wider range of species data are collected by academic and research bodies. Few data from these sources are provided to NSS, LERCs or the NBN Atlas.

a) What measures to you suggest are needed to address this
b) What benefits do you envisage if this can be achieved?

Support and Facilitation

Q17 What support or facilitation measures:

a) Do you rely on?
b) Do you provide to others?
c) Do you think are most important or currently lacking?

Funding

Q18 Current funding for recording infrastructure in the UK restricts and undermines what could be achieved

Q19 Potential funding mechanisms - What funding support/ mechanisms ought to be in place to meet meeting critical/priority/other information needs - with particular reference to species data?

To what extent do you agree with the outlined options below?
What do you suggest as an alternative to or variant on these?
What do you see as key advantages or drawbacks in relation to each of these suggested options?

i. Government departments/agencies should fully fund the entire biodiversity information infrastructure to meet identified cross-sector information needs and to unleash the social and economic benefits that biodiversity data can provide. Since data collection and management and the information systems would be wholly funded from public money the expectation is that species records, other data and datasets would be all be made available under an Open Gov or Creative Commons (CC 0) open licence with no restrictions on use (subject to sensitivity issues). This is the model that was proposed for Scotland as a result of the Scottish Biodiversity Forum (SBIF) review.
ii. Government departments/agencies should fund the core activities of the NBN Trust and Local Environmental Records Centres and continue to fund key NSS in support of national biodiversity indicator provision, and information system provision, with the indicated bodies providing support to other parts of the NBN. Agencies and other official users would have the use of the subsidised data for their purposes, but commercial clients would be

required to pay for their data use. This is the general basis of the current approach in Wales.

iii. Government departments/agencies provide little or no subsidy with official and commercial users required to pay for access to data shared via a Creative Commons non-commercial (CC-BY-NC) licence. This is closer to the current situation in England but will require some more effective mechanism to ensure that commercial, statutory, research and other funded users will, collectively be covering the costs of data curation and dissemination undertaken by voluntary and non-profit bodies

iv. Some other approach – what do you suggest in order to meet the needs of your organisation/sector, recorders, data-providers and information users whilst ensuring the viability, and ongoing development of the NBN and the biodiversity information infrastructure

Routes to Data Access

Q20 Looking forwards, which of these statements do you (most) support?

- A single UK/England database into which all (species) data flow and from which they can be accessed for whatever purpose with Web APIs provide access to the data and services from the websites of different users according to their purpose.

- A focus of data flowing into local databases with a national data portal able to poll these for data on request
- A focus on individual species group databases with national and local portals each able to poll these.
- A mixed approach with NSS and LERC able to exchange and provide access to their data and able to share data to the NBN Atlas and benefit from Web APIs or provide access to their data solely via the Atlas.
- Some variants of these or something else entirely?
- What are the reasons for your answer?

Q21 All data providers should embed the FAIR Data Principles in their operations so that species data or datasets and other shared data are adequately supported by meta-data and fully findable, accessible, interoperable and reusable for many different purposes.

- If you disagree with Q21, why?

Q22 Data providers should seek to make datasets as open as practicable, securing the support of and any required authorisations from the original recorders or compilers of contributory data, for reuse under the terms of a standard, machine-readable licence.

Q23 What measures or additional information needs do you envisage a need for in relation to the Environment Bill and Planning White Paper Proposals?

Appendix E - Comparison of benefit and cost typologies

Table A.1: Comparison of benefit typologies

This report		SBIF Review (Wilson et al., 2018)
Benefit category	**Benefit description**	**Benefit area**
Input benefits from easier access to species data		
Easier access	Data search – time savings: Avoided time that would have been spent accessing data under the no species data pathway scenario	Opening up of biodiversity data; Providing support and insights
Data generation	Avoided cost of (repeated) primary research	Maximising confidence and pace;
Output benefits for decision-making		
Supporting better decisions based on better evidence	Biodiversity management itself, management of associated activities (e.g. wildlife tourism) and activities for which biodiversity acts as an indicator (e.g. drought resilience of soils, water quality risks); Cost of achieving biodiversity targets; Local, national and international; Biodiversity net gain; 25 Year Environment Plan; WFD status; WiNEP driven by WFD; Agri-environment payments directed; Value for understanding impact of pressures and conservation action; Value for informing management; Value for reporting biodiversity status and meeting legislative requirements; Other strategic objectives	Increasing national and global impact; Transforming decision-making; Optimising efficiency and sustainability; Achieving a sustainable economy
Planning system	Planning fees and applications; Planning for biodiversity; Avoided time-delays	
Facilitating innovation	Use of machine learning & AI; Development of tools (e.g., NEVO)	Facilitating digital innovation
Outcome benefits from data and decisions		
Species management	Monitoring change in biodiversity; Use of biodiversity data; Controlling invasive species; Preventing extinction; Conservation efforts; Resilience of ecosystem services	Maximising confidence and pace; Opening up of biodiversity data; Avoiding costs/nugatory spend; Mainstreaming biodiversity; Enhancing ecosystem resilience; Preventing species extinctions; Transforming decision-making
Supporting research	Research publications and projects; Understanding of existence value	
Avoided loss of opportunities related to future use and re-use of data	See technology developments discussed in Section 3, e.g., DNA based technology	
Education	Courses; Training/Skills enhancement; Public engagement	Developing capacity; Engendering participation
Improving public health and well-being	Health & well-being benefits of volunteers	Engendering participation; Improving public health
Cultural services	Opportunity cost of land use (i.e., maintaining priority habitats); Tourism and recreation activities driven by species	Encouraging wildlife & landscape tourism
Social cohesion and integration	Networking; Development of partnerships; Facilitating the conversation;	Engendering participation; Optimising efficiency and

	Development of standards	sustainability

Table A.2: Comparison of cost typologies

This report		SBIF Review (Wilson et al., 2018)
Cost category	**Cost description**	**Cost area**
Costs to data providers		
Data pathway capital costs	Data management infrastructure (e.g., communication, processing and storing); hardware; software	Computer hardware; Computer software; Legal and professional fees
Data pathway revenue costs	Staff time; Volunteer time; In-kind contributions; Advice and services; Other resource contributions	Development staff; Operational staff
Investments/Savings	Statutory funding; Other funding sources	Statutory agency funding for LERCs/Recording groups; Statutory agency funding for NBN trust; Statutory agency funding for Super Partners; Statutory agency funding for Communities; Grants and donations; Community grants; Super Partner grants
Costs to data users		
Costs for system users	Data searches/requests, data access, and use; Public sector (e.g., policy) and private sector (e.g., EIA)	

Unallocated benefit areas from SBIF review: Optimising efficiency and sustainability; Upholding standards; Reducing land use conflicts; Empowering communities

The benefits assessed in the SBIF review focused primarily on direct benefits of the benefits of the biodiversity data framework, therefore it is expected that some benefit areas do not align with this report.

Unallocated cost areas from SBIF review: Regional partnerships; Office overheads; Event and project overheads; Education overheads; Promotional materials; Legal and professional fees; Other overheads and contingency.

This is expected as the scope of the analysis of this report focuses on the costs required to support the existing species data pathway to provide its services, rather than the establishment of a new or improved species data pathway.

Appendix F - Quantification of costs and benefits

This appendix summarises the methods used to quantify the costs and benefits identified in Section 5.3.3 (see Table A.3 and Table A.4) as well as providing more detail on these methods in Sections F.1 and F.2 All evidenced cited in this section are reported in the main report reference list. As explained in Section 5.3.1, the approach to the cost-benefit analysis is based on two underlying assumptions:

- The scenario looks at the absence of the species data pathway, but all else remains the same (Ceteris Paribus). This assumes society would try to pursue the same environmental outcomes through the same systems but do them less well without systematic species data. This ignores potential alternative approaches, which are assumed to be less beneficial (otherwise they'd be done now) and probably substantially different.
- The importance of species data in overall natural environment management. Without species data, decisions and management would rely heavily on habitat data. This is rather crude (e.g., identifying improved/unimproved grassland). Identification of specific habitat types (e.g., chalk grassland) and/or assessment of the condition or quality of habitats, requires data on species (i.e., on specific species or representative species groups).

Table A.3: Summary of benefit methodologies

Benefit	Indicator	Link to species data	Valuation metric and key variables	Key assumptions
Supporting better decisions based on better evidence base	Biodiversity net gain market value	Species data is a key component in assessing habitat condition.	Metric: Proxy for market value -Expected average annual value of BNG market (low) (£/yr.) -Estimated net benefit of BNG market (high) (£/yr.)	Full attribution to the species data pathway Benefits start in 2023 to reflect time required for market to be set up.
	Improvements in Water Framework Directive status from water company actions	Species data pathway allows for targeted interventions to be implemented on rivers to improve species status (biological quality).	Metric: Willingness to pay to for improved WFD status - Length of rivers improved by all WiNEP interventions and species related WiNEP interventions (km) - WFD status of species related WiNEP interventions (km/status) - England annual average low and central NWEBS values for rivers (£/km)	Partial attribution to the species data pathway Benefits start in 2025 and are cumulative for the first 25 years and remain constant thereafter.
	Common agriculture policy payments	Species data allows funding to be directed to Pillar 2 of the Common Agricultural Policy (CAP).	Metric: Proxy for market value - Total spend on Pillar 2 (£ 2014-2020) - Annual benefit of transferring funding from Pillar 1 to Pillar 2 (£ 2014-2020)	Full attribution to the species data pathway

Benefit	Indicator	Link to species data	Valuation metric and key variables	Key assumptions
	Biodiversity targets	Species data is a key component in assessing habitat condition	Metric: Directed expenditure/resources Costs of delivering UK Biodiversity Action Plan (BAP) in England by Habitat Action Plans widespread species costs and individual species costs (£/yr.)	Action plan costs attributable to species data pathway
Species management	Avoided eradication costs from late detection of invasive species	Access to species data facilitates monitoring of INNS and enables early eradication.	Metric: Avoided eradication costs - Total annual control costs of INNS to the English economy - Total annual costs of INNS to the English economy	- Current control costs of INNS relative to total annual costs of INNS (%) - Average difference in control costs and all other costs of INNS (%)
Planning system	Avoided legal fees	Developers can face legal action if they do not do due diligence on species data.	Metric: Avoided legal fees - Number of planning permissions applications received by district LPAs in England - 1% of planning applications received by London Boroughs used a search of GiGL's data records - 79% of data requests received by English LERCs were for supporting planning applications - Number of enforcement notices issued by district LPAs in England - Legal fees for a straightforward (1 day) high court judicial review hearing (£/day)	Proportion of planning applications that used GiGL's services is representative of England
	Avoided interest rate payments	Developers that face legal challenges are also subject to delays in their construction process.	Metric: Avoided additional interest rate payments - Value of UK property development - Proportion of developments that could be delayed without access to species data pathway	- Scale factor based on number of enterprises in England vs UK - Average time of development - Average loan repayment schedule Commercial interest rate
	Avoided loss of ecosystem services delivered by maintaining priority habitats	Species data is a key component in assessing habitat condition and identifying priority habitat types.	Metric: Avoided loss of benefits - Area of England that is developed each year and area of priority habitat area this is developed on each year (hectares) - Average rate of annual development - Consumer surplus values of ecosystem services delivered by SSSI conservation activities (£/ha)	Ceteris paribus and the importance of species data in overall natural environmental management; Land area developed occurs at the average rate of development; SSSIs are the best representative of sites of priority habitats

Benefit	Indicator	Link to species data	Valuation metric and key variables	Key assumptions
Easier access	Time savings for existing data search	Reduced time searching for existing species data due to the availability and accessibility of data through data repositories (e.g., data portals).	Metric: Time savings - Number of MEDIN users - Number of download events via the NBN Atlas (excl. for personal use) - Average annual time saving per user per year - Average wage for full-time employment across all occupations	Non-wage labour cost uplift
Cultural services	Cultural opportunity cost of farming	Landholdings that maintain a systematic record of species is motivated by species data.	Metric: Opportunity cost - Area of Selbourne Common (acres) - Average market price of pastureland for English Estates and Farmland (£/acre)	Estimated value represents an asset value. Therefore, is not profiled overtime.
	Wildlife tourism expenditure	Wildlife tourism is an activity partly driven by the visitors understanding of species existence.	Metric: Attributable expenditure - Total expenditure on wildlife tourism in England by domestic day visits and overnight stays (£) - Total attributable activity expenditure by day visits and overnight stays in England (£) - Proportion of UK visits where wildlife tourism was the sole reason for a day visit or an overnight stay	UK proportion of visits is representative for England visits
Improved health and wellbeing	Physical health benefits of volunteer recorders	In addition to improving the general welfare of volunteers, if people are active during their volunteer time, volunteering can also have measurable physical health benefits.	Metric: Avoided medical treatment costs - Number of volunteer hours collecting biodiversity monitoring data - Breakdown of volunteer time by activity - Avoided cost to the health service of delivering 1 QALY (£/QALY)	- Attribution of volunteer hours by sub-category group that are deemed to be 'active' - Active volunteer visit is at least 2 hours
	Opportunity cost of volunteer time	Volunteers' personal enjoyment of partaking in recording and collecting species data.	Metric: Opportunity cost of leisure time - Number of volunteer hours collecting biodiversity monitoring data - Breakdown of volunteer time by activity - Value of time for non-work purposes (£/hour)	- Volunteer time spent on administrative tasks and support are attributable to the species data pathway - 75% of monetary unit value is (£/hour) is equal to volunteer opportunity cost of leisure time

Table A.4: Summary of cost methodologies

Cost category	Indicator	Link to species data	Key variable	Key assumptions
Data pathway capital costs	Unmet capital replacement costs	Capital and replacement costs for LERCs, NBN Atlas and MEDIN portals, i.e., cost of the biodiversity data framework	- Number of NBN Atlas developers and other staff (FTE) - Average capital replacement costs for the NBN Atlas (£/per FTE) - Lower bound optimism bias adjustment factors for capital expenditure for equipment/development	NBN Atlas capital replacement costs per FTE are representative of LERCs and MEDIN Staff requirements remain constant
Data pathway operating costs	Species data pathway operating costs	Operating costs of species data pathway for LERCs, NBN Atlas and MEDIN portals	- Total costs to ALERC members and non-ALERC members to enable delivery of their services - NBN Atlas operation costs (cloud hosting and staff salaries) - MEDIN operation costs	Current operating costs are representative of future costs.
	Species data activity	Expenditure by select non-government organisations on species data activities.	- Butterfly Conservation total expenditure on continuing operations - Buglife's expenditure on research, survey and advice - RSPB total spend on managing nature reserves, research policy and advice, education, and inspiring support and supporter care	Attribution of spend to species data pathway; Current spend is representative of future years
	Public sectors spend on data generation	Public sector spend on generating new data records.	- Natural England spend on species data generation by external contracts	Current spend is representative of future years
	Volunteer effort (data stewardship)	Species data recording and collecting is primarily undertaken by volunteers.	- Number of volunteer hours collecting biodiversity monitoring data - Breakdown of volunteer time by activities - Value of skilled volunteer time (£/hour)	- Assume volunteers classed as skilled volunteers - Volunteer time spent on administrative tasks and support are attributable to the species data pathway
Investments /Savings	Funding for UK public sector bodies	Funding received by NBN Atlas and MEDIN is used to support their operations	- Total funding received by NBN Atlas from English public sector bodies - Total funding received by MEDIN from English public sector bodies	Current funding levels are representative of future funding levels.
	Funding from grants	Funding received by NBN Atlas and MEDIN is used to support their operations	Total grants received by NBN Atlas	Costs start in Year 0; however, funding sources are due to end in 2021 and therefore do not continue into the future.
Costs of system users	LERC chargeable data requests	Access to species data via LERCs	- Number of chargeable data requests received by LERCs - Average (mean) LERC charge rate, including VAT (£/request)	Split between non-chargeable and chargeable requests from 2013 are representative of current and future years

Cost category	Indicator	Link to species data	Key variable	Key assumptions
	Data requests from public sector bodies	Public sector bodies such as Natural England and the Environment Agency purchase data through the species data pathway. However, it is difficult to ascertain the total spend on data requests by these organisations.	Environment Agency approximate spend on importing data from LERCs and NBN Atlas	
	Environmental Impact Assessment costs	Environmental impact assessments are a statutory requirement in some planning decisions. Part of the content generated for an EIA requires species data.	- Number of planning decisions involving a Statutory Environmental Statement - Cost of environmental impact assessment content that relates to species data (£/EIA)	Number of EIA's undertaken by developers remains constant

F.1 Benefits

Benefit – Biodiversity net gain market value

Species data is a key component in assessing habitat condition within the Natural England Biodiversity Metric 2.0 (Natural England, 2019)[21]. The metric is designed to be used as the basis of a forthcoming biodiversity offsets market in England. Without species data for context and to benchmark habitat condition, we assume the market would not be able to function and some or all of its benefits would be lost.

Therefore, a range of values are identified from the impact assessment of Biodiversity net gain (BNG) and other policy proposals for the Environment Bill (Defra, 2019). One (underestimate) of value is based on the size of this market was assessed in the impact assessment concludes that the expected average annual value of the BNG market is between £55-£682 million per year, in 2017 prices. The low estimate of expected average annual value (i.e., costs of activities in) the BNG market is approximately £211 million per year. This is comprised of net cost to developers (£199 million) and ongoing costs to government (£11 million). This is equivalent to £236 million per year, in 2020 prices, and is used as a lower bound in the CBA analysis. The best average annual net benefit value is estimated as the difference between best estimate of average annual market cost (£211 million) and the best estimate of average annual benefits from the BNG market (£1,395 million, in 2017 prices). This produces the upper bound value for the CBA of approximately £1,184 million per year, in 2020 prices. The monetary unit value in the market is assumed to remain constant over time, however benefits are assumed to arise starting in Year 3 to reflect the time required to set up the BNG market.

The Environment Bill was placed before parliament following this impact assessment, so this is interpreted to indicate acceptance of these costs (by Government, on behalf of wider society) of enhancing biodiversity. We therefore assume that the gross benefits to society of the policy to enhance biodiversity exceed £236m per year. The policy is only possible to implement with systematic and consistent current and past species data (i.e., species data is necessary to distinguish between specific habitat types) for England. Without the species data pathway, we assume that the future planning system would be the same or worse than the current system in achieving biodiversity policy outcomes.

An alternative approach to valuing this benefit could be to estimate the additional costs of data collection in the absence of the species data pathway. This approach would assume that other alternative metrics or direct measurements of data would be used to fulfil the objectives of the policy and thus not be reliant on the species data pathway. Understanding of the additional costs of data collection are currently unknown, and therefore the next best approach (i.e., the one described above) has been applied in this study.

Benefit - Improvements in Water Framework Directive status from water company actions

The benefit of maintaining freshwater quality is estimated by the welfare gain from maintaining or improving the Water Framework Directive (WFD) quality status of river waterbodies that are included in the Water Industry National Environment Programme (WiNEP) intervention list with a quantified environmental outcome (Environment Agency, 2020). The benefit of the species data pathway in this context is that it

[21] Note that Natural England are due to launch the Biodiversity Metric 3.0 in Spring 2021.

allows for targeted interventions to be implemented on rivers to improve species status (biological quality). In the absence of the species data pathway, improvements would still occur, but they would be at random (i.e., untargeted due to a lack of location specific information). For context, the total length of WFD rivers in England is approximately 53,000 kilometres (Environment Agency, 2020).

Total investment in WiNEP is roughly £5 billion between 2020-2025 (Defra and EA, 2018). This represents the total resources directed across 11,084 interventions, of which 1,789 have a quantified environmental outcome on 16,090 km the river length improved or protected (Environment Agency, 2020). Although this is 30% of the total length of all rivers in England, the 31,300 km downstream of improved stretches, represents 59% of all rivers in England[22], also benefit. WiNEP interventions are categorised by a three-part code consisting of i) core obligation, ii) action and WLB priority; and iii) additional driver codes. After determining which drivers require species data, 1,278 WiNEP interventions were identified as being primary linked to species data[23], with just under 900 interventions having a quantified environmental outcome of river lengths (km) improved.

The total length of rivers that would have an improvement under a species related WiNEP intervention is just over 6,064 km, representing 11% of all WFD rivers in England. The total length of these waterbodies (i.e., accounting for downstream impacts) is approximately 11,300 kilometres, which is 21% of all rivers under the WFD in England. Under the no species data pathway scenario, improvements to biological quality would arise randomly based on other actions. This is estimated as the difference between the quantified river improvements from all WiNEP interventions and the quantified river improvements from targeted WiNEP interventions (i.e., 16,090-6,064 = 10,026 km). The attributable improvement in river length as a result of species data is estimated as 10,026 km relative to river improvements from all WiNEP interventions, which is approximately 62%. This proportion is applied to the total length of rivers that would have an improvement under a targeted WiNEP intervention (6,064 km), resulting in 3,779 kilometres of river length being improved due to the species data pathway. As each WiNEP intervention with an environmental outcome is linked to a specific waterbody, the current Cycle 2 WFD status (Environment Agency, 2021) can be determined and thus the improvement in status as well. The majority of species related WiNEP interventions occur on river lengths with moderate status (48%), followed by poor status (17%), bad status (2%) and good status (1%). These proportions are used to disaggregate the 6,062 kilometres of targeted improvement attributed to the species data pathway into each WFD status category.

The improvement is estimated by a given status (i.e., change in the WFD status from Poor to Moderate) or maintenance of Good status. The economic value of the WiNEP interventions is based on the National Water Environment Benefits Survey (NWEBS) values (NERA Economic Consulting 2007; Metcalfe, 2012). The NWEBS values provide low, central and high estimates of values for river water bodies, in 2012 prices.

[22] Based on Cycle 2 WFD (2016) data, total length of WFD rivers in England is just under 53,000 kilometres.

[23] This refers to interventions with the following primary driver codes: BW_IMP1, BW_IMP2, BW_IMP3, BW_INV1, BW_INV2, BW_INV3, BW_INV4, BW_MON, BW_ND, BW_NDINV, DrWPA_IMP, DrWPA_INV, DrWPA_ND, EE_IMP, EE_INV, EPR_MON1, HD_IMP, HD_INV, HD_ND, INNS_INV, INNS_MON, INNS_ND, L_IMP, LWFD_IMPg, LWFD_IMPm, MCZ_INV, NERC_IMP1, NERC_IMP2, NERC_INV1, NERC_INV2, SSSI_IMP, SSSI_INV, SSSI_ND, SW_IMP, SW_INV1, SW_INV2, SW_MON, SW_ND, U_IMP1, U_IMP2, U_IMP3, U_IMP4, U_IMP5, U_IMP6, U_INV, U_INV2, U_MON1, U_MON2, U_MON3, U_MON4, U_MON5, WFD_IMP_CHEM, WFD_IMP_FISH, WFD_IMP_WRFlow, WFD_IMP_WRHMWB, WFD_IMPg, WFD_IMPm, WFD_IMPp WFD_INV, WFD_INV_CHEM1 WFD_INV_CHEM10, WFD_INV_CHEM11, WFD_INV_CHEM12, WFD_INV_CHEM13, WFD_INV_CHEM14, WFD_INV_CHEM2, WFD_INV_CHEM3, WFD_INV_CHEM4, WFD_INV_CHEM5, WFD_INV_CHEM6, WFD_INV_CHEM7, WFD_INV_CHEM8, WFD_INV_CHEM9, WFD_INV_FISH, WFD_INV_WRFlow, WFD_INV_WRHMWB, WFD_MON_CHEM, WFD_ND, WFD_ND_WRFlow, WFD_ND_WRHMWB, WFD_NDINV_WRFlow, WFD_NDINV_WRFlow, WFD_NDINV_WRHMWB, WFD_NDLS_Chem1, WFD_NDLS_CHEM2, WFDGW_IMP_GWQ, WFDGW_IMP_GWR, WFDGW_INV_GWQ, WFDGW_INV_GWR, WFDGW_ND_GWQ, WFDGW_ND_GWR, WFDGW_NDINV_GWQ, WFDGW_NDINV_GWR, WFDGW_NDINV_GWR&GWQ

NWEBs values have been inflated to 2020 prices using the HM Treasury (2021) GDP deflator. The NWEBS values represent survey respondents' willingness to pay (WTP) for six equally weighted ecosystem components (Defra, 2015, p.69):

- **Fish**;
- **Other animals such as invertebrates**;
- **Plant communities**;
- The clarity of water;
- The condition of the river channel and flow of water; and
- The safety of water for recreational contact.

This assessment uses the England annual average low and central value estimates for improving the status of river water bodies (£/km) and is applied to the length of river in each WFD status category. As only three of the six NWEBS ecosystem components are related to species (those in **bold**), and as they are equally weighted, half of the unit values are applied. The total willingness to pay for improvement/maintenance of WFD status is between £48-£59 million per year. The average annual unit values are assumed to remain constant over the assessment period. Benefits from the WiNEP interventions are expected to take effect in 2025 and are cumulative at five-year intervals until 2050 (i.e., 2030 = £48m + £48m). This 25-year period reflects water companies water resource management plan time horizon. After 2050, the benefit value remains constant over the remaining appraisal period as it is assumed that society does not cause future damage that would result in a deterioration in status.

Benefit – Agri-environment policy payments

Systems of agricultural payments in England are changing after the UK's exit from the EU. The future system in England (the 'Environmental Land Management Scheme (ELM)), remains under development. It is expected to have similar environmental objectives to the EU's Common Agricultural Policy (CAP). Therefore, evidence on CAP spending is used here as a proxy for the impacts of future ELMS payments.

Species data has allowed funding to be directed to Pillar 2 of the CAP. The way Pillar 2 funding is designed, allocated and its outcomes judged, relies heavily on species data. Without systematic national current and past species data, the Pillar 2 funding would not be able to be managed in this way. The total spend on Pillar 2 was £1,520 million between 2014 and 2020 (Defra, 2013a), assuming a constant annual rate, the annual Pillar 2 funding is £253 million per year.

Defra (2013b) identifies that Pillar 2 achieves higher benefits per £ compared to Pillar 1. Without species data, government would not be able to allocate Pillar 2 to locations where it realises higher benefits, and so it is assumed that it would achieve the same results as a subsidy distributed based on agricultural data (i.e., Pillar 1 of CAP). Estimates the benefit of transferring 15% (£1,889 million) of funding from Pillar 1 to Pillar 2 is, on average, £3,041 million over a 5-year time period, giving an annual benefit per pound of £1.61. This benefit per pound is applied to the annual Pillar 2 funding (£253 million). Therefore, the estimated additional annual benefit using species data for Pillar 2 funding is £486 million per year, which is assumed to remain constant over the appraisal period.

Benefit - Biodiversity targets

GHK Consulting (2010) provides estimates of the costs of delivering UK Biodiversity Action Plan (BAP) in England, broken down by Habitat Action Plans (£354 million per year) widespread species costs (£197 million per year) and individual species costs (£31 million per year). Assuming 50% of the Habitat Action Plan cost can be attributed to species data and 100% of the widespread and individual species costs, the total cost of delivering priority habitats, as a result of species data, in England is estimated at £404 million per year. However, this is a target, and does not reflect what is spent on priority habitat delivery, the actual spend is unknown. Therefore, it is assumed that half of the delivery costs are spent. This results in an estimated spend on delivering priority habits of £202 million per year. The annual value is assumed to remain constant over the appraisal period.

The estimated annual value represents an estimate of resources directed, rather than the additional value of species data. The estimate is omitted from the minimum and best results of the cost-benefit analysis as there is a risk of partial double-counting with benefit estimates for BNG market value and CAP payments. Therefore, it is only included in the maximum results. Note that the biodiversity targets policy landscape has changed significantly since 2010, and therefore there is poor confidence in the resulting estimates on resource directed to deliver the UK BAP in England.

Benefit – Invasive species management

The benefits of species data to species management are estimated in terms of the avoided costs of eradicating later-detected invasive non-native species (INNS). The known costs of controlling[24] floating pennywort is £1.4 million and the estimated avoided damage cost of controlling floating pennywort is £5.6 million (Environment Agency, n.d., Williams et al., 2010). The difference in these costs is 293%. Williams et al. (2010) provides estimates for earlier stage and later stage eradication costs for five different species. The highest estimate, for Water primrose, has a difference of 3,330:1 between early and late-stage eradication. The lower end of the range of data given by Williams et al. (2010) is the difference between mid-stage eradication costs (£4.7 million) and late-stage eradication costs (£18.8 million) for Coypu is 3:1. The annual cost of INNS to the English economy estimated by Williams et al. (2010) is approximately £1,623 million per year, in 2020 prices across 12 sectors. This includes control costs, damage costs and wider society costs, of which the difference in cost of control and these other costs is approximately 57% of total annual costs.

To estimate the attributable total annual cost of INNS to early-stage eradication, an assumption of the proportion of costs must be made. As a low estimate, 1% of the current annual costs of INNS is assumed to be related to early-stage eradication, this reflects the contribution of the total early-stage eradication costs for the five case study species in Williams et al. (2010) relative to the estimated £1.6 billion in total annual costs. As a high estimate, the proportion applied is estimated using the average early-eradication costs of the five case study species (£1.9 million in 2020 prices) multiplied by the total number of species or species groups (i.e., 23[25]) assessed in Williams et al. (2010) which provides a total early-eradication costs for all species considered by Williams et al. (2010) equal to £44 million. This represents 3% of the total annual cost of INNS to the English economy and represents the higher range of costs related to early-stage

[24] This includes management (e.g. removal, spraying or other eradication measures) and, monitoring and evaluation costs, as well as spend on protecting assets and preventing spend.

[25] Count of species or species groups in Table 17.3 in Williams et al. (2010).

eradication. The 57% difference in costs based on Williams et al. (2010) is used to provide a conservative estimate for avoided costs of eradicating later-detected INNS. Applying the estimated cost difference (57%) and the assumed proportion of costs attributable to early-stage detection (1-3%) to the total annual cost of INNS, provides an estimate of the avoided eradication costs from late detection of invasive species between £5 million and £25 million per year. The lower end of the range is used in the main results of the CBA and the higher value is used in estimates for sensitivity 4 analysis. Benefits do not arise until year 10 and are assumed to remain constant over time.

Benefit – Avoided legal fees

Research published by the Greater London Authority (2016) concluded that out of 88,000 planning permission applications received by London Boroughs, only 1% had used a search of GiGL's[26] data records, whilst a further 18% of applications should have used GiGL's data record search service. Developers can face legal action if they do not do due diligence on species data. Therefore, a benefit of the species data flow pathway can be estimated as the avoided legal fees incurred by developers that follow due diligence procedures.

Extrapolating from London, under the existing species data pathway we assume that the proportion of planning applications that used GiGL's services is representative of England. The number of planning permissions applications received by district LPAs in England in the last year[27] is 399,000 (MHCLG, 2020). Assuming that 1% of planning permission applications received by district Local Planning Authorities (LPAs) in England have used a data record search, implies that roughly 5,100 planning applications received by district LPAs used data record searches. In 2013, ALERC estimated that 79% of data requests received by English LERCs were for the purposes of supporting the planning system, this amounts to roughly 10,900 requests – which are assumed to be for planning applications (pers comm, ALERC). These figures provide a range for the minimum (5,100) and maximum (10,900) number of planning applications that use data record searches.

However, only a fraction of these applications would be subject to a legal challenge. The number of enforcement notices issued by district LPAs in the last year[28] is approximately 3,100 (MHCLG, 2020). This represents 1% of the total planning applications received over the same period. This indicates that between 40-85 planning applications that used data record searches would be issued an enforcement notice. The number of applications received is assumed to remain constant over time, as is the proportion that uses data record searches and the proportion of enforcement notices issued.

The cost of a legal challenge is highly variable and will depend on the complexity of the case, the number of grounds of the challenge, and other things. The estimated legal fees for a straightforward (1 day) high court judicial review hearing are £120,000 per day (pers comm, confidentially from law firm). This results in the avoided legal fees by planners following due diligence processes on species data to be estimated as between £4.8 and £10.2 million per year. The monetary unit value is assumed to remain constant over the appraisal period.

[26] GiGL is the Greenspace Information for Greater London.
[27] Planning permission applications received between Q4 2019 and Q3 2020 (MCLG, 2020).
[28] Planning permission applications received between Q4 2019 and Q3 2020 (MCLG, 2020).

Benefit – Avoided interest rate payments

In addition to legal fees, developers that face legal challenges are also subject to delays in their construction process. This can cause delays in loan repayments resulting in added interest payments. Therefore, a benefit of the species data flow pathway is that some delays are avoided. The value of this can be estimated as the avoided opportunity costs of delaying investment, calculated based on the additional interest rate payment on a longer loan period.

The value of UK property development is estimated as £300 billion per year (HMRC, 2015.). This has been scaled to England, using an assumption from the Defra (2019) Biodiversity Net Gain Impact Assessment, where UK values were scaled down to England based on the number of enterprises[29]. Therefore, using a scale factor of 90%, the value of England property development is estimated as £271 billion per year. The proportion of developments that could be delayed without access to the species data pathway is estimated as 1%, based on GLA (2016). Therefore, the total value of development supported by species data is £2.6 billion per year.

Developers are assumed to take out a loan equal to the total property development value, and therefore pay interest on this loan. Average development time is assumed to be 1.5 years (pers comm)[30], and the average loan repayment schedule is assumed to be 3.5 years[31] at a commercial interest rate of 4.8% (abcfinance.co.uk)[32]. Therefore, total interest paid during a loan repayment, without any delays to the assumed development time period, is equal to £581 million in total over the repayment period. If these developments were subject to a legal challenge that would result in a delay of 1 additional year, the total interest paid on the development loan increases to £747 million per year. The avoided additional interest payment on a development loan is derived as the difference between the interest paid with and without a delay, which is £166 million per year. This remains constant over the appraisal period.

Benefit – Avoided loss of ecosystem services delivered by maintaining priority habitats

The land area of England is 13.9 million hectares of which 1.6 million hectares (11%) is already developed land (eftec et al. forthcoming). Whilst priority habitats (PH) cover 1.3 million hectares (9%) of total land area in England. The area of England that has potential for development is estimated as the difference between the total land area and the already developed land area (12.3 million hectares).

eftec et al. (forthcoming), as part of a project commissioned by Defra, have estimated the area of England that is developed each year (6,330 hectares) and the area of PH that is developed on each year (487 ha). These hectare values can be used to estimate the proportion of land in England that is developed annually as well as the proportion of priority habitat land area in England that is being developed each year (i.e., lost). The former is estimated by dividing 6,330 hectares developed each year by the 12.3 million hectares of land that is potentially developable (0.05%) – this is assumed to be equal to the average rate of development each year. The estimated proportion of priority habitat area that is developed each year

[29] As stated in Defra (2019): Data missing for England on turnover, GVA and purchase of inputs, but available for the UK. UK values scaled down to 90.27% to estimate England values, as this is the ratio of England to UK number of enterprises https://www.ons.gov.uk/businessindustryandtrade/business/businessservices/bulletins/uknonfinancialbusinesseconomy/previousReleases.

[30] Parameters used are based on confidential advice from a property finance expert.

[31] Average development loan repayment period is assumed to be 3.5 years, based on review of Berkeley Group (2020) and Barratt Homes (2020) annual reports, which indicates that loan terms of 2-5 years is the most common.

[32] Evidence suggests that commercial investment mortgages will have a rate between 3.5% and 6% (abcfinance.co.uk). An average is used for this analysis.

(0.04%) is derived by dividing the annual PHI developed on (487 ha) by the total area of PH in England (1.3 million ha). Assuming development occurs at an average rate, the area of PH lost per year due to development is estimated by multiplying average rate of annual development (0.05%) by the total area of PH in England (1.3 million hectares), resulting in approximately 649 hectares each year. Therefore, the avoided area of PH development is the difference between the 649 hectares lost each year (i.e., no species data pathway) and the estimated 487 hectares of PH actually developed each year (i.e., baseline). This results in 162 hectares per year of PH development that is avoided and is assumed to remain constant over the appraisal period.

Christie and Rayment (2012) produced an economic assessment of the ecosystem service benefits associated with current biodiversity conservation policy for Sites of Special Scientific Interest (SSSIs). As part of this research, consumers surplus values of ecosystem services delivered by SSSI conservation activities (£/ha) were estimated for two scenarios: maintain funding and increased funding. The average per hectare value of ecosystem services delivered by SSSI conservation activities under the maintain scenario is £882/hectare, in 2020 prices. SSSIs are assumed to be the best representative sites of PH. Therefore, the average £ per hectare value is multiplied by the area of PH development that is avoided each year (162 hectares). This produces an avoided loss of ecosystem services delivered by maintaining priority habitats of £143,000 per year. This benefit is cumulative over time, meaning that the undiscounted annual value in year 2 is year 1 plus £143,000 (i.e., £143,000 + £143,000 = £285,000).

Benefit – Time savings for existing data search

Reduced time searching for existing species data due to the availability and accessibility of data through data repositories (e.g., data portals). This search time saving relates to uses of species data such as desk-based studies, planning applications, research or Environmental Impact Assessments. As such, references to downloads for 'personal use' are not included in these figures.

The approach follows the method described in the CBA of MEDIN by eftec and ABPmer (2019) and is applied to user statistics for the MEDIN portal and NBN Atlas. The estimate of active users of the MEDIN portal is 1,400 (eftec and ABPmer, 2019) and the number of download events registered on the NBN Atlas, excluding for personal use, is 13,428. The registered download events through the NBN Atlas are assumed to be representative of users. The number of users is assumed to remain constant over the appraisal period.

The average annual time saving per user was calculated using data generated by the MEDIN survey questions[33] related to how many hours in an average week[34] the use of MEDIN saves the respondent, versus a scenario of doing the same work with MEDIN not existing (eftec and ABPmer, 2019)[35]. The average time savings across all respondents (including non-respondents) was estimated as 30.2 hours per user per year[36] and is assumed to remain constant over the appraisal period. Multiplying the average hours saved

[33] Benefits from existing data searches is derived from survey question number 10, full details can be found in the eftec and ABPmer (2019) final report for MEDIN.

[34] The survey clarified that "if there is no such thing as an average week, please approximate across the year. For example, if you have shorter periods of high use and longer periods of moderate or no use, please consider how this would average out over the course of a year".

[35] Users selected their time saving from a list of time brackets (0 hours, 0-4 hours, 5-9 hours etc) with the mid-point value of each range being taken as the time saving. The options were: 0 hours per week, 0 – 4, 5 – 9, 10 – 14, 15 – 19, 20 – 24, 25 – 29, 30 – 34, 35 +, I'm not sure.

[36] Average weekly time savings across respondents was multiplied by the number of work weeks in a year, calculated as 52 weeks minus 7 weeks for holiday and sickness.

per year by the total number of users (NBN Atlas and MEDIN), results in 448,000 hours saved on existing data searches.

The value of time is estimated using ONS ASHE (2020) data on wages for full-time employees across all occupations. In 2020 the average wage is estimated as £15 per hour, which has been uplifted for non-wage labour costs of 21.8% to cover additional employee benefits such as pensions, National Insurance contributions, sickness pay as well as maternity and paternity pay[37]. The uplift results in an average wage of approximately £18 per hour, which is applied to the total hours saved to produce an annual value of time savings for existing data search of £8.2 million per year. The monetary unit value is assumed to remain constant over time.

Benefit – Cultural opportunity cost of farming

The cultural value of species data is inherently difficult to measure and value. However, landholdings that are managed in order to maintain a systematic record of species are clearly motivated by species data. An example of this is Selbourne Common, a National Trust property of 267 acres (National Trust, 2021).

The monetary value of maintaining Selbourne Common can be estimated as the opportunity cost of not exploiting it commercially (e.g., not farming the land). The average market price of pastureland for English Estates and Farmland is £7,200 per acre (Strutt & Parker, 2020). Applying this to the Selbourne Common area yields a cultural opportunity cost of farming of £1.9 million, which is an asset value.

This value provides an example of the societal benefit of the history of species data recording, and therefore the species data pathway, in England. It could potentially be extrapolated to other sites, but this would require detailed site by site research.

Benefit – Wildlife tourism expenditure

Wildlife tourism is an activity partly driven by the visitors understanding of species existence, which provides a link to the use of species data. eftec et al. (2019) uses tourism data from the Great British Day Visits Survey and Great British Tourism Survey to attribute domestic expenditure from tourism activities to the natural environment. Results are estimated for Great Britain, devolved nations and nine English regions. For the purposes of this work, the interest lies in the expenditure attributed to visits and trips motivated by the activity 'watching wildlife, bird watching and other nature' (henceforth referred to as 'wildlife tourism').

Total England expenditure on wildlife tourism is £2.5 billion, in 2020 prices, where day visits accounts for 82% of total spend. Note that this is an unadjusted figure and does not reflect motivation of trips nor the number of other activities undertaken on a single trip. In England, total attributable activity expenditure by day visits and overnight stays for wildlife tourism is estimated as £745 million per year[38], in 2020 prices. Where day visits represent approximately 91% and overnight trips represents 9% of the England total. This represents visits and trips where the main motivation was wildlife tourism, this is assumed to represent

[37] Follows the approach used for the MEDIN CBA (eftec and ABPmer, 2019). The uplift factor is consistent with the approach applied in Impact Assessments by BEIS (Pers. Comm. BEIS, July 2019). Source: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lc_lci_lev&lang=en

[38] eftec et al. (2019) produced values in 2017 price year. HM Treasury (2021) GDP deflators have been used to inflate values to 2020 prices. For further details on the attribution process please see the final report produced for Defra.

the baseline.

The proportion of UK visits where wildlife tourism was the sole reason for a day visits is 13% and for an overnight trip is 1% (eftec et al., 2019). Multiplying these proportions for each visit type (day or overnight) by their respective total (unadjusted) wildlife tourism expenditures, provides an estimate of total expenditure on 'wildlife tourism' that is solely motivated by the activity. This is roughly £265 million per year and is assumed to represent the baseline activity that is dependent on the species data pathway. This value is assumed to remain constant over the appraisal period. The exact role of the data pathway in supporting this spending is unclear – historic and current data about species status clearly plays a role in wildlife tourism (e.g., in knowing which scarcer species occur in which areas of the country). However, wildlife tourism would still occur under the 'no species data pathway' scenario, but only based on anecdotal evidence.

Note that this approach represents a re-distribution of expenditure, and therefore is not a marginal benefit calculation as would be recommended by the Green Book (HM Treasury, 2020). To fully assess the additionality of this benefit place-based impacts would need to be factored in. Therefore, overall confidence in the approach is 'poor'. This value is omitted from the CBA results; however, it does serve as important contextual information and is therefore included as a key non-monetised benefit.

Benefit – Physical health benefits of volunteer recorders

In addition to improving the general welfare of volunteers, if people are active during their volunteer time, volunteering can also have measurable physical health benefits. This benefit is estimated from the proportion of the visits that are active. The health benefits of active recreation are measured in Quality Adjusted Life years (QALYs[39]), and the economic value calculated based on the average avoided health cost due to improvement in a QALY. This approach has been used in ONS (2019) Urban natural capital accounts and with the subsequent data sources cited in Defra's 'Enabling a Natural Capital Approach' (2020a).

The State of Nature Report (Hayhow et al., 2019) estimates that 7.5 million volunteer hours go into collecting biodiversity monitoring data per year. As part of the UK biodiversity indicator data (Defra, 2020b), volunteer time is divided into four sub-categories: survey, data input and analysis (65% of 2018 total hours), countryside management and advisory support (19%), administrative or office support (2%) and other conservation work (14%). The proportions of volunteer time in each category are applied to the estimated volunteer hours reported in the State of Nature Report to estimate volunteer hours for each sub-category. An assumption on the proportion of hours in each category that are 'active' is applied, where countryside management is deemed 100% active, survey, data input and analysis and other conservation work is 50% active, and time spent on administrative, or office support is 0% active. This results in an estimate of active volunteer time of 4.4 million hours.

To estimate the number of visits from the 4.4 million hours of active volunteer time, assumptions are made around the length of volunteer time that is active. It is assumed that each active volunteering session is at least two hours, resulting in an estimate of active volunteer visits of 2.2 million per year. The estimated number of volunteers, and thus active volunteers is assumed to remain constant over the appraisal period.

[39] QALY is a health measurement used widely in health and health economics research. QALY of zero denotes death, and 1 denotes full health.

The benefit of active volunteering is measured as the improvement in QALYs. Beale et al. (2007) analysed Health Survey for England data, estimating that 30 minutes a week of moderate-intense physical exercise, if undertaken 52 weeks a year, would be associated with 0.0106768 QALYs per individual per year. It is assumed two hours of collecting biodiversity data is equivalent to 30 minutes of moderate-intense physical exercise. Beale et al. (2007) assume this relationship between physical activity and QALYs is both cumulative and linear. Claxton et al. (2015) estimate a cost-effectiveness threshold of a QALY to be roughly £12,900/QALY in 2008 prices. This figure represents the additional cost to the health system to avoid one lost QALY of health. It is used as a proxy for health costs, reflecting the avoided health costs when QALY is improved by one unit. Based on this information, the avoided health cost per active visit is estimated as £3.29 in 2020 prices. The monetary unit value is assumed to remain constant over time.

The total avoided medical treatment costs are £7.4 million per year. However, this value is not additional to the species data pathway as this study assumes that this type of volunteering would continue under the 'no species data pathway' scenario in order to ensure that the same volume of records is collected as in the baseline.

Benefit – Opportunity cost of volunteer's leisure time

The State of Nature Report (Hayhow et al., 2019) estimates that 7.5 million volunteer hours go into collecting biodiversity monitoring data per year and is based on data collected as part of the UK Biodiversity Indicator 'A2. Taking action for nature: volunteer time spent in conservation' (Defra, 2020b)[40]. The time spent on volunteer activities is treated as an opportunity cost of volunteers' leisure time (i.e., volunteers use their free time to contribute to the data collection). This is interpreted as volunteers' private benefit from volunteering.

Of the 7.5 million volunteer hours, it is estimated that 58% are active volunteer hours[41], indicating that 42% are inactive hours (i.e., primarily administrative support). This equates to 3.1 million volunteer hours, which are attributable to the species data pathway as under the 'no species data pathway' this administrative support and desk-based volunteer effort would not occur.

The Department for Transport TAG Data Book (DfT, 2020) provides values of travel time savings (VTTS) for non-work purposes by income band and as an average (£/hour), in 2010 prices. These figures represent a market value. Values of time are presented for commuting, other and all non-work time. In this study, the average value of all non-work time (£7.5/hour in 2020 prices) is used as the starting point in this analysis. This value represents the opportunity cost of travel time relative to the value of work, rather than leisure time.

To represent the opportunity cost of volunteer's leisure time, the DfT (2020) average unit value for travel time is adjusted to relate to leisure time. Fezzi et al. (2013) conclude that value of travel time is approximately 70-80% of the average wage rate. Therefore, the midpoint (75%) is applied to the DfT TAG average value of all non-work time, producing an adjusted monetary unit value of £5.6/hour in 2020 prices.

[40] This is an indicator of the number of hours worked by volunteers for 14 UK conservation charities and public bodies. Note that total hours spent in the underlying dataset has been updated since the publication of the State of Nature Report.

[41] Based on estimated number of active volunteer hours relative to total volunteer hours. See calculation for 'Benefit – Physical health benefits of volunteer recorders'.

This produces an annual value of £17 million when applied to the relevant recorded volunteer hours (3.1 million hours). The monetary unit value and number of volunteer hours are assumed to remain constant over the appraisal period. This value is fully attributable to the species data pathway as the personal benefit from volunteering is associated with the engagement of the volunteers with the species data pathway (i.e., gain satisfaction from knowingly contributing to society).

F.2 Costs

Costs – Unmet capital replacement costs

Capital costs and replacement of LERCs, NBN Atlas and MEDIN are aggregated and adjusted for optimism bias to represent the total capital costs of the species data pathway, assuming straight line depreciation. The resulting present values over 60 years are adjusted for lower bound optimism bias in line with HM Treasury Green Book guidance (2020).

Capital replacement costs for the NBN Atlas have been estimated based on the budget required to ensure that developers have the quality of equipment needed to maintain the NBN Atlas; this is roughly £2,500-£3,000 per developer (pers comm, NBN Trust). An average cost is applied in the analysis (£2,750). The associated budget for all other staff is approximately £1,000 per person (pers comm), NBN Trust. The current full-time equivalent (FTE) staff requirements of the NBN Atlas include a lead developer, developer, data manager and project manager at 1 FTE each, and part-time project officer and administrative support (0.8 and 0.6 FTE respectively). This results in total capital costs of between £5,500 for the NBN Atlas. These unit values are assumed to remain constant over the appraisal period.

To estimate the capital costs of MEDIN and LERCs, the estimated NBN Atlas capital costs are used as the basis. As operating costs for these organisations (see below) have been estimated with good confidence, the ratio of the NBN Atlas to MEDIN operating costs (1.8) and NBN Atlas to LERCs (27.2) operating costs has been used as a scale factor. For MEDIN and LERCs this results in total capital costs of approximately £9,700 and £149,900.

The total one-off capital costs of species data pathway are between £322,300. Equipment replacement is between 3-5 years which forms the basis of the low and high values estimated (pers comm, NBN Trust). Assuming straight line deprecation, depreciation costs range between £64,000 and £107,000 each year. It is assumed that new equipment is purchased in year 1 (i.e., 2021).

The resulting 60-year present values (both the main results, sensitivity 1 and sensitivity 2) are adjusted for optimism bias. The Green Book provides an upper (200%) and lower (10%) adjustment factor that should be applied to capital expenditure for equipment/development projects (HM Treasury, 2020). The Green Book (ibid) recommends starting with the upper bound adjustment and then move towards the lower bound adjustment after having mitigated for appropriate risks. The lower bound adjustment figures for the low depreciation estimates, are reported in the main results, sensitivity 1, sensitivity 2 and sensitivity 3 as the capital costs used in this assessment are considered known costs, therefore the majority of the risks associated with optimism bias are deemed to be appropriately mitigated. The high depreciation estimates are included as part of sensitivity 4.

Costs – Species data pathway operating costs

Operating costs of LERCs, the NBN Atlas and MEDIN are aggregated to represent the costs of the species data pathway. The annual species data pathway operating costs are estimated as £9 million per year and are assumed to remain constant over time. A breakdown of this figure is described below.

For LERCs, this is estimated as the costs to ALERC members and non-ALERC members to enable delivery of their services to partners, stakeholders, clients and the public (pers comms, ALERC). This estimate includes non-salary and salaried costs and is inclusive of in-kind contributions by host organisations where applicable, and a value in £ provided by office-based volunteers assisting with core outputs. It does not include OS map licencing costs which are mainly provided in-kind through partnership agreements with hosts or strategic partners, as this figure is unavailable (pers comm, ALERC). Values were estimated for the 2016/17 financial year and are assumed to be representative of current and future operating costs. Therefore, the total operating cost for LERCs is £7.4 million per year in 2020 prices.

For the NBN Atlas, operation costs reflect cloud hosting and staff salaries which equates to £272,000 per year in 2020 prices (pers comms). The MEDIN annual report for (2019/20) provides a total expenditure value of £480,000, per year. This includes employment costs of the core team, travel and subsistence as well as external contracts.

Costs – Species data activity

Species data activity is estimated based on relevant expenditures identified for the Royal Society for the Protection of Birds (RSPB), Butterfly Conservation and BugLife. Total species data activity expenditure is estimated as £19 million per year and is assumed to remain constant over time. A breakdown of this figure is described below.

Butterfly Conservation's annual report details that 19% of total expenditure is on charitable conservation activities (Butterfly Conservation, 2020). Their total expenditure on continuing operations is reported as £4.8 million, therefore 19% of this value is attributed to the species data pathway (£0.9 million per year).

Buglife's expenditure on research, survey and advice is assumed to be attributable to species data activities (£0.2 million per year) (Buglife, 2018).

RSPB's 2019/20 annual report concludes that the organisations spend £93 million on managing nature reserves, research policy and advice, education and inspiring support and supporter care (RSPB, 2020). It is assumed that the Butterfly Conservations 19% is representative, therefore 19% of £93 million is attributed to the species data pathway (£18 million) per year. It is possible that this is an overestimate, but there are many other species conservation charities whose spend on the species data pathway is not known. Therefore, the potentially overestimated RSPB figure is included in the total to compensate for underestimates in other parts of the third sector.

Costs – Public sectors spend on data generation

In a follow-up interview, Natural England noted that they generate species data both internally and through external contracts. This differs to the spend on data requests, as this indicates that Natural England are generating new records rather than paying to access existing ones. It was estimated that the spend on

species data generation by externals was £40,000 per year (Natural England, pers comm). This value is assumed to remain constant over the appraisal period. Due to the inherent difficulty of estimating spend on data generation, this value is given a poor confidence rating and only included in the maximum CBA results.

Cost – Volunteer effort

The State of Nature 2019 report estimates that 7.5 million volunteer hours go into collecting biodiversity monitoring data per year (Hayhow et al., 2019). Of which 3.1 million volunteer hours are attributable to the species data pathway as under the 'no species data pathway' this administrative support and desk-based volunteer effort would not occur.

The Heritage Fund provides estimates of the value of volunteer time as part of their application guidance (2020). The value of volunteer time, depending on skill level and activities, ranges between £10-£50 per hour. The lower end of the range reflects the value of volunteer time for administrative work and the higher end reflects professional volunteer time. Assuming these are skilled volunteers, their inputs can be valued at £20 per hour (Heritage Fund, 2020). The total value of these volunteer hours is £63 million per year, which is assumed to remain constant over time.

Costs – Funding from UK public sector bodies

Funding received by the NBN Atlas and MEDIN, primarily comes from UK public sector bodies, this study only accounts for the English public sector bodies funding, as it is assumed that this funding is provided to support English species data. Total English public sector funding received by NBN Atlas is £55,000 (NBN Trust, 2020). Whilst MEDIN receives total funds of approximately £393,000 (MEDIN, 2020). However, this does not solely support species data on the MEDIN portal. Therefore, based on the distribution of data types on the portal, 20% of which is related to marine biodiversity, 20% of total funds are included in this study (£78,600).

This results in a total value of £134,000, which is assumed to occur annually and remain constant over the study time period. There is a risk that this figure double-counts the operating costs of the NBN Atlas and MEDIN, so is excluded from the main results of the CBA.

Costs – Funding from grants

Grants are also a source of funding within the species data pathway. The NBN Atlas reports a total grant received equal to £118,000, however these funding sources are due to end in 2021 and therefore are not factored into the analysis after 2021.

Costs – LERC chargeable data requests

Based on a 2013 assessment by the Association of Local Environmental Record Centres (ALERC), LERCs in England received a total of 14,000 data requests, of which 21% were non-chargeable[42]. From this it is estimated that 79% of the data requests (10,880) received by LERCs were chargeable, mainly for the purposes of supporting the planning system, and paid for accordingly. Though this is based on data from 2013, the proportions and number of requests are not expected to have changed over time, and therefore

[42] Non-chargeable requests were for a range of purposes (e.g., student projects), and the expectation is that most the requests would be for the benefit of volunteers and volunteer organisations (pers comm, ALERC).

remain constant over the appraisal period (pers comms ALERC). Following the approach set out under the 'Avoided legal fees' calculation, 1% of planning permission applications received by district LPAs in England have used a data record search which equates to approximately 5,100 applications in the last year[43]. These figures provide a range for the minimum (5,100) and maximum (10,880) number of planning applications that use data record searches.

The charge rates of data requests for LERCs vary significantly, with the average (mean) cost44 (including VAT) for the same product estimated as £233, ranging between £120 and £450 (pers comm ALERC). Applying this average cost to the high and low estimated number of chargeable data requests received, produces a total value of between £1.2-2.5 million per year, which is assumed to remain constant over time. The low estimate is used in the CBA main results and the high value is included as part of sensitivity 4.

Costs – Data requests from public sector bodies

Public sector bodies such as Natural England and the Environment Agency purchase licenses for data through the species data pathway. However, it is difficult to ascertain the total spend on data requests by these organisations. Furthermore, due to the nature of the chargeable requests provided by ALERC (pers comm), this indicator is likely to double-count.

In a follow-up interview with the Environment Agency, an approximation of the spend associated with importing data from LERCs and NBN Atlas amounts to a value of £400,000 per year.

In a follow-up interview, Natural England stated that their local teams spent between £150-£200 per data cut per project on data request from LERCs (pers comm, Natural England). Each data cut is licensed to a specific project and therefore cannot be re-used in another project. In general, statutory bodies will request and pay the rate set by LERCs or recording schemes for data of an appropriate resolution for project delivery. To provide an estimate of Natural England spend on species data a proportion of the Environment Agency value (£400,000/year) is used. The ratio of Natural England staff (2,000) to Environment Agency staff (10,600) is estimated as 19%. This is applied to the value provided by Environment Agency (pers comm), resulting an annual expenditure figure of £75,500 per year.

These values are only included in the maximum results estimates of the cost-benefit analysis as they potentially double-count with the LERC chargeable data requests.

Costs – Environmental Impact Assessment costs

The number of planning decisions involving a Statutory Environmental Statement in 2019 is recorded as 238 (MHCLG, 2020). This reflects the number of environmental impact assessments (EIA) undertaken by the developers, which will have content relating to species data in them. This is assumed to remain constant over the appraisal period.

The cost of EIA content relating to species data is estimated to range between £10,000 and £20,000 per assessment (pers comm eCountability). The lower end of this range is used to estimate costs relating to species data from the data pathway (other data in EIA may be from primary survey). Applying this to the

[43] See calculation 'Benefit – Avoided legal fees' for more detail.
[44] Based on a sample of 12 of the 44 English LERCs.

number of statutory environmental statements involved in planning decisions provides an estimate of the cost of EIA content relating to the species data pathway of roughly £2.4 million per year. The monetary unit value is assumed to remain constant over time.

Appendix G - Net present value calculations

In developing the CBA model, the annual benefits and costs over the assessment period are aggregated in present value (discounted) terms. The formal calculations are presented below.

$$\mathbf{PVB} = \sum_{\mathbf{i=0}}^{\mathbf{5}} (\frac{\mathbf{1}}{\mathbf{1+r}})^{\mathbf{i}} * \boldsymbol{Benefits}_{\boldsymbol{i}}$$

Where PVB is present value benefit, *i* is an index for the year and *r* is the discount rate.

$$\mathbf{PVC} = \sum_{\mathbf{i=0}}^{\mathbf{5}} \left(\frac{\mathbf{1}}{\mathbf{1+r}}\right)^{\mathbf{i}} * \boldsymbol{Costs}_{\boldsymbol{i}}$$

Where PVC is present value cost, *i* is an index for the year and *r* is the discount rate.

Costs and benefits are compared in present value terms. The net present value (NPV) of an intervention is calculated as the difference between Present Value of the Benefits (PVB) and the Present Value of the Costs (PVC):

$$\mathbf{NPV} = \mathbf{PVB} - \mathbf{PVC} = \left[\sum_{\mathbf{i=0}}^{\mathbf{5}} \left(\frac{\mathbf{1}}{\mathbf{1+r}}\right)^{\mathbf{i}} * \boldsymbol{Benefits}_{\boldsymbol{i}}\right] - \left[\sum_{\mathbf{i=0}}^{\mathbf{5}} \left(\frac{\mathbf{1}}{\mathbf{1+r}}\right)^{\mathbf{i}} * \boldsymbol{Costs}_{\boldsymbol{i}}\right]$$

eftec
economics for the environment

4 City Road, London EC1Y 2AA
+44 (0) 20 7580 5383
eftec@eftec.co.uk
eftec.co.uk
@eftecUK

Final report | May 2021