



GBIF-ICLEI Best Practice Guide for Biodiversity Data Publishing by Local Governments

Version 1.0



May 2012

Suggested citation:

GBIF (2012). GBIF-ICLEI Best Practice Guide for Publishing Biodiversity Data by Local Governments, (contributed by Cadman, M.J.; Chavan, V.; Patrickson, S.; Galt, R.; Mader, A.; Sood, R.; Hirsch, T.) Copenhagen: Global Biodiversity Information Facility, Pp. 62, ISBN: 87-92020-37-2, Accessible at http://links.gbif.org/gbif_best_practice_guide_data_publishing_by_local_governments_en_v1

ISBN: 87-92020-37-2

Persistent URL:

http://links.gbif.org/gbif_best_practice_guide_data_publishing_by_local_governments_en_v1

A stand-alone Concise Summary of this document is available at :

http://links.gbif.org/concise_summary_gbif_best_practice_guide_data_publishing_by_local_governments_en_v1

Language: English

Copyright © Global Biodiversity Information Facility, 2011

License:

This document is licensed under a [Creative Commons Attribution 3.0 Unported License](http://creativecommons.org/licenses/by/3.0/)

Project Partners: The Global Biodiversity Information Facility (GBIF), ICLEI- Local Governments for Sustainability (ICLEI), and the Secretariat of the Convention on Biological Diversity (CBD)

Document Control:

Version	Description	Date of release	Author(s)
0.9	Content development	26 November 2011	Mandy Cadman
1.0	Review, edits, revision	23 May 2012	Mandy Cadman, Vishwas Chavan, S. Patrickson, A. Mader, R. Sood, T. Hirsch

Cover photo credits: Arpad Benedek. Photo used with permission.



Local governments have a critical contribution to make towards the implementation of the Convention on Biological Diversity and the achievement of its Aichi targets. It is well known that they are responsible for much of the action required to fulfill these aims but their role is also essential when it comes to the assessment.

Local governments have access to a wealth of fine-scale data which, if shared in a consistent format, enables effective policy implementation at local level as well as contributing to improved data resolution across the landscape for the benefit of biodiversity practitioners at all levels. GBIF plays a central role in driving this process and has now facilitated it further with this set of guidelines.

Braulio Ferreira de Souza Dias
Executive Secretary
Secretariat of the Convention on Biological Diversity

About GBIF and the project partners

GBIF: The Global Biodiversity Information Facility

GBIF was established by countries as a global mega-science initiative to address one of the great challenges of the 21st century – harnessing knowledge of the Earth’s biological diversity. GBIF envisions a world in which biodiversity information is freely and universally available for science, society, and a sustainable future. GBIF’s mission is to be the foremost global resource for biodiversity information, and engender smart solutions for environmental and human well-being (GBIF 2011a). To achieve this mission, GBIF encourages a wide variety of biodiversity data holders, generators and users across the globe to discover and publish (make discoverable) data to global standards through the GBIF network. Website: <http://www.gbif.org>

ICLEI: Local Governments for Sustainability

ICLEI is an international association of local governments and national and regional government organizations committed to sustainable development. ICLEI’s mission is to build and serve a worldwide movement of local governments committed to achieving tangible improvements in global sustainability through cumulative local actions. ICLEI provides technical consulting, training, and information services to build capacity, share knowledge, connect leaders, and support local government in the implementation of sustainable development initiatives at local government level. Website: <http://www.iclei.org>

The Local Action for Biodiversity (LAB) Pioneer Project, which is run by ICLEI’s Cities Biodiversity Centre in partnership with the IUCN, is a global urban biodiversity project that was initiated in 2006. LAB’s approach is action-oriented and customized for local and regional governments and their partners around the world, with the goal of improved and effective biodiversity management at the local level. It is a key component of, and contributor to, the Global Partnership on Cities and Biodiversity, chaired by the Convention on Biological Diversity.

CBD: The Convention on Biological Diversity

The Convention on Biological Diversity (CBD), which was opened for signature in June 1992 at the United Nations Conference on Environment and Sustainability (the ‘Rio Earth Summit’), is an international agreement between 193 Parties (192 national governments and the European Union) to ensure the conservation of biological diversity, the sustainable use of its components and the fair and equitable sharing of benefits arising from use of its genetic resources. The implementation of the CBD is supported by a Secretariat which has its institutional home within the United Nations Environmental Programme (UNEP).

The CBD’s comprehensive and integrated approach to biodiversity conservation and management acts as a framework within which Parties can guide national policies that are subsequently implemented at the national, sub-national and local level. Because local authorities represent the level of government closest to people, and are important implementers of plans and policies, their support is crucial in achieving the objectives of the Convention. With this in mind, the CBD established the Global Partnership on Local and Sub-national Action for Biodiversity. At the 10th meeting of the Conference of the Parties to the Convention, a CBD decision further endorsed the Plan of Action on Sub-National Governments, Cities and other Local Authorities for Biodiversity, providing the Parties with guidelines on how they can support local governments and, thereby, increase their contribution to implementation of the Convention. Website: <http://www.cbd.int>

Table of Contents

Contents

Executive Summary	1
Section 1: Why we need this guide	4
Section 2: Local Government Processes and Publishing Biodiversity Data	9
Section 3: Publishing primary biodiversity data - principles of best practice	21
Section 4: Operational principles for publishing primary biodiversity data	26
4.1. How to capture biodiversity data	27
4.2. Tools for preparing data for publishing	29
4.3. Publishing metadata	34
4.4. Criteria for selecting a data publishing option	36
4.5. Registering your data	38
Section 5: Quick Guide to data publishing by local governments	40
5.1. Summary of recommendations for local governments	40
5.2. Where to find assistance	41
Section 6: Concluding remarks	48
References	49
Appendix 1: Glossary of Terms	52
Appendix 2: Acronyms used in this document	54
Appendix 3: Other useful resources	55

List of Figures

- Figure 1: Schematic representation of an ecosystem
- Figure 2: Municipal functions that can incorporate biodiversity data
- Figure 3: Enhancing EIA practice through the biodiversity data publishing framework, showing anticipated benefits and different steps of EIA process
- Figure 4: The data publishing workflow (adapted from GBIF, 2011b)
- Figure 5: GBIF tools for preparing biodiversity data for publishing
- Figure 6: Generating a Darwin Core Archiving using the GBIF Spreadsheet Processor
- Figure 7: Options for preparing and publishing primary biodiversity data
- Figure 8: Data publishing options in relation to IT and data management capacity
- Figure 9: Document map: The metadata publication workflow together with links to relevant guides
- Figure 10: Document map: The occurrence data publishing workflow together with links to relevant guides
- Figure 11: Document map: The checklist publication workflow together with links to relevant guides

List of Boxes

- Box 1. Definition of biodiversity
- Box 2. Ecosystems and Ecosystem Services
- Box 3. Challenges faced by local governments in dealing with biodiversity
- Box 4. GBIF Participants
- Box 5. Additional guidance on issues of data quality
- Box 6. How EIA practice can be enhanced through the increased availability of published, verifiable biodiversity data.
- Box 7. Local biodiversity strategies and action plans and data publishing
- Box 8. Data Papers - a mechanism for incentivizing data publishing
- Box 9. Categories of biodiversity data.
- Box 10. Terminology used to describe datasets.
- Box 11. What is data publishing?
- Box 12. Further reading on this topic (data publishing).
- Box 13. How to use this tool (GBIF Excel Templates)
- Box 14: GBIF Tools for publishing biodiversity data.
- Box 15. A note on data standards and formats.

Box 16. How to use this tool (GBIF Spreadsheet Processor)

Box 17. Data Hosting Centres

Box 18. GBIF Participant Nodes

List of Tables

Table 1. Selected examples of ecosystem services.

Table 2. Quick Guide to publishing occurrence data using GBIF-supported tools.

Table 3. Summary of GBIF User Guides and other documents.

Executive Summary

Introduction: local governments as users and generators of biodiversity data

Primary biodiversity data are the digital text or multimedia records detailing facts about the occurrence of organisms. Knowledge about the identity and occurrence of organisms forms the backbone of our understanding of the biological world, and is essential for monitoring the state of natural ecosystems, developing sound environmental management policies, and making ecologically sound, sustainable development decisions.

Local governments are becoming increasingly important as managers and users of biodiversity assets and ecological infrastructure. They are responsible for environmental management and planning, regulation of land-use through planning and decision-making, and supporting the implementation of global, national and sub-national policies and strategies relating to biodiversity and sustainable development. Biodiversity management options arise in almost all traditional fields of activity for which local governments are responsible, such as integrated development planning, service provision and management of urban green spaces. They are, therefore, both important users and generators of biodiversity data.

For a variety of reasons, local governments face many challenges when it comes to dealing with biodiversity, and access to readily usable and verifiable biodiversity data is often problematical for them. Furthermore, much of the biodiversity data collected as part of local government planning processes is either lost after completion of the report, or is collected in inconsistent formats that cannot be easily archived or shared. This has largely been due to the lack of awareness of the tools and protocols suitable for capturing, sharing, archiving and accessing primary biodiversity data, or lack of knowledge about how to use these tools.

The purpose of this best practice guide

This best practice guide describes the tools, standards and infrastructures that are available to practitioners in local government for publishing biodiversity data, and explains when and how to use them. It also explains important principles that underlie the data publishing process. It also sets out to:

- Make local governments aware of the benefits to them of being able to access biodiversity data via the GBIF network, and highlights the important role they can play as contributors of biodiversity data; and,
- Explain how data publishing can be incorporated into planning, policy development and decision-making processes in local government.

Publishing biodiversity data

Through the Global Biodiversity Information Facility (GBIF), digital biodiversity data are being made freely and openly available via the Internet for scientists, researchers, national and local authorities and the general public. GBIF provides a suite of standards and tools that can be employed to discover and publish primary biodiversity data. 'Publishing' is the process through which biodiversity datasets are made publicly accessible in a standardized format, via an online access point (typically a web address, or URL). This access point is recorded in a registry managed by GBIF. Published datasets can then be discovered and accessed via the GBIF Data Portal.

The data publishing workflow

Data publishing through the GBIF network involves a number of clear steps including:

- Capturing the data in a consistent, exchangeable format;
- Preparing the data for publication (i.e. converting it into a standardized format known as a Darwin Core Archive that can be accepted on the GBIF network);

- Publishing the dataset (i.e. making it publicly accessible via a web address using the GBIF tools); and
- Registering the data in the web-based data access point in the GBIF registry.

Once the data have been published and registered with GBIF, then they are freely and openly accessible through the GBIF network and the GBIF data portal (<http://data.gbif.org>).

GBIF Tools and protocols for data publishing

At each step in the data publishing workflow there are simple tools and protocols available for data publishers to use. In this best practice guide, we describe these tools, explain how and when to use them, and refer the user to other GBIF guides that can provide additional assistance.

There are a variety of tools available for capturing data and preparing it for publishing, and the option selected should be matched with the technological and data management capacity of the user. These tools include:

- pre-configured Excel spread sheets for capturing data in a standardized way;
- the **Spreadsheet Processor** or **Darwin Core Archive Assistant** for generating the Darwin Core Archive file, and
- the **Integrated Publishing Toolkit** or other manual tools for publishing the Darwin Core Archive file through the GBIF network.

The simplest route for publishing biodiversity data would be to use the pre-configured GBIF Excel Spreadsheet templates, prepare them for publishing using the GBIF Spreadsheet Processor and then publish the datasets using the GBIF Integrated Publishing Toolkit (IPT) or through a Data Hosting Centre, if one is available. These steps are explained in greater detail in Section 4 of the best practice guide.

Biodiversity data publishing by local government

It is difficult to make a 'one-size-fits-all' set of recommendations to local governments as to which data publishing option they should follow, as these institutions vary widely in respect of capacity and resources. However, some general rules of thumb that should be applied include:

- Use the GBIF Excel spreadsheet templates to capture all biodiversity data that are gathered as part of local government planning processes - this will mean that data will automatically be suitable for publishing via the GBIF network.
- Build the use of these tools, and this best practice guide into local biodiversity management activities and into the Terms of Reference for all consultants that are contracted to do assessments or prepare biodiversity reports or plans for any relevant local government process. You should also recommend to consultants that they make use of the GBIF network to source appropriate data when embarking on a new study.
- Make data discovery and publishing a standard practice in Local Biodiversity Strategies and Action Plans (LBSAPs) as well as any broader sustainable development plans.
- Become part of the GBIF network and benefit from the knowledge exchange and support it offers.

Benefits to local governments of publishing biodiversity data

There are a number of compelling reasons for local governments to publish biodiversity data using the GBIF tools and network. Chief amongst these are that data publishing will:

- enable free and open access to biodiversity data, which is essential for biodiversity-inclusive planning and development at local government level;
- facilitate the ongoing expansion and improvement of the local, national and global biodiversity databases on which environmental planning, EIAs, land-use management, policy development and areas of scientific work frequently rely, improving baseline knowledge of the ecosystems of a particular site, region or country;
- help practitioners who conduct specialist work for local governments to gain recognition for their work by enabling them to be cited in future uses of their data;
- enhance the quality, predictive value, verifiability and transparency of local government planning processes, thus improving the land-use decisions that they inform and the confidence civil society can place in these decisions.

Where to find further assistance:

The principles, tools and processes described in this Executive Summary are explained in greater detail in the GBIF/ICLEI/CBD best practice guide to data publishing by local governments. There are also numerous other GBIF User Guides that are available online to assist with the publication of primary biodiversity data and their associated metadata, using the GBIF tools. These guides provide detailed, step-by-step instructions in the use of all the key tools used at different steps in the data publishing process. The table below summarises the key documents that can provide assistance at each step of the data publishing pathway.

Should further assistance be required, you can contact the wide network of GBIF country and organization Participant Nodes, which can be looked up on the GBIF website and are described in the full local government best practice guide. The purpose of these Nodes is to encourage, coordinate and assist in biodiversity data publishing activities within their respective jurisdictions and domains.

Section 1: Why we need this guide

Publishing biodiversity data is defined as making biodiversity datasets publicly accessible in a standardized format via an online access point (typically a web address). This access point is recorded in a registry managed by the Global Biodiversity Information Facility (GBIF) and can be accessed (or 'discovered') via the GBIF Data Portal. This best practice guide serves to enable local governments, their consultants and other interested and affected parties to discover, capture, manage and publish biodiversity data via the GBIF network.

In this section we provide the rationale for the development of this guide, by explaining the importance of biodiversity data and its relevance to local government, and by placing the best practices that are presented in the guide into the broader GBIF data publishing context.

The overall **purpose** of this guide is to:

- make local governments aware of the **benefits** to them of being able to access biodiversity data via the GBIF network, and the **important role** they can play as contributors of primary biodiversity data to the GBIF network through regional, national and global databases and information systems;
- describe how local governments, their consultants and other partners can **use the tools, practices and infrastructures** that are available through the GBIF network to discover, capture and publish primary biodiversity data to common standards;
- explain how **use of these tools can be incorporated into planning, policy development and decision-making processes** within local government, to improve biodiversity management and strengthen environmental sustainability across all core functions.
- show how data publishing can **strengthen Local Biodiversity Strategies and Action Plans** and explain the importance of developing action plans for data publishing as part of the LBSAP process.

Important note: This best practice guide provides a **summary** of information about biodiversity data publishing that is relevant to users in local government. It should be used **in conjunction with** other guides that GBIF has published to provide more detailed information and guidance on the use of specific tools or methods related to different steps in the data publishing pathway. Cross-reference to these guides is made in each section, and detailed references are provided in the Appendices.

The importance of biodiversity

Biodiversity (see Box 1) is a critical foundation of human well-being and contributes significantly to shaping the development path of a region or country. In addition to its own intrinsic value, biodiversity is an invaluable resource for building sustainable livelihoods, creating jobs and alleviating poverty; enhancing rural development, food security and land-use; delivering water and other critical resources; and assisting communities with adaptation to the effects of climate change. Sound biodiversity management, therefore, is not only good for the environment, but contributes to local development, underpins economic prosperity and is essential for maintaining the quality of life of all people.

Biodiversity, natural resources and healthy ecosystems are a valuable source of capital - called natural capital - that is needed for building the economy and achieving environmentally sustainable development. Natural capital underpins other forms of capital (i.e. financial, manufactured and human) that are well-understood components of a healthy economic system. Sustainable local development, and the economic activities that make urban settlements prosper, are based primarily on goods and services delivered by healthy ecosystems (See Box 2).

Rural communities usually depend directly on biodiversity for their daily needs (food, shelter, fuel, medicines) and livelihoods. Similarly, urban communities and cities rely on forests, wetlands and other natural areas to provide clean water and protect against the effects of natural disasters such as floods and droughts. Biodiversity and healthy ecosystems also provide the basis of many production activities that contribute to food security, provide employment, and shape the development path of a country or region. This means that wise management of biodiversity and ecosystems is in itself a development objective.

BOX 1: Definition of biodiversity

Biodiversity is defined in the Convention on Biological Diversity as: “the variability amongst living organisms from all sources including, *inter alia*, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems.”

United Nations (1992): Text of the Convention on Biological Diversity, accessible at <http://www.cbd.int/doc/legal/cbd-en.pdf>

Biodiversity is under pressure

Globally, biodiversity resources are being eroded at an unprecedented pace by unsustainable land-use practices, over-extraction of natural resources, inappropriately-located development, loss of habitat, invasive alien species, pollution and other environmental changes, including those that are attributable to climate change (MA 2005). Biodiversity loss disrupts ecosystem functioning, holds negative impacts for human health and well-being and compromises opportunities for achieving sustainable socio-economic development. Disrupted ecosystems are more vulnerable to shocks and disturbances, less resilient and less able to supply humans with essential ecosystem services. Because the consequences of biodiversity loss and ecosystem disruption are often harshest for the poor, inappropriate biodiversity management poses a significant barrier to achievement of the Millenium Development Goals (MDG). Ecosystems and biodiversity need to be managed and used in ways that cater for the development needs of the world's population, but that maintain ecological functioning across landscapes.

Global efforts to improve environmental decision-making in order to stem biodiversity loss and maintain functioning ecosystems are impeded by, amongst other things:

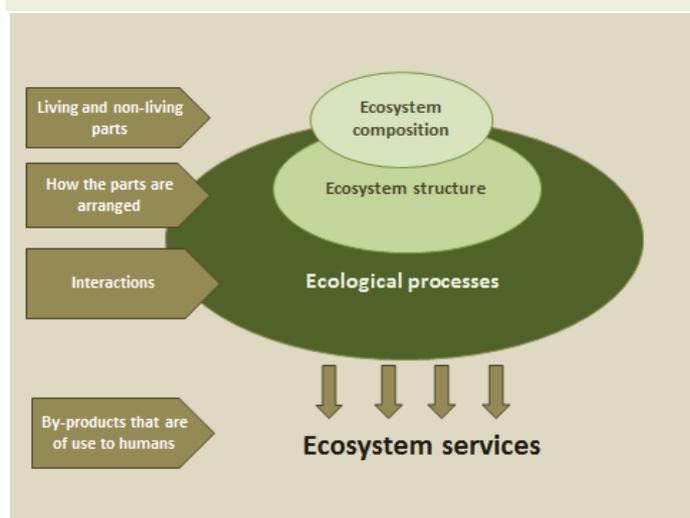
- Incomplete knowledge of biodiversity (and a lack of species inventories) for many parts of the world;
- Limited information on the occurrence of important species; and,
- Major gaps in local and global monitoring systems due to a lack of well-documented, comparable data for most major ecosystems (MA, 2005).

Knowledge about the identity and occurrence of organisms forms the backbone of our understanding of the biological world, and is essential for monitoring the state of natural ecosystems, for developing sound environmental management policies and making ecologically sustainable development decisions. Local governments have much of the regulatory mandate to control and optimize human impact on ecosystems, and to do this need ready access to verifiable biodiversity data that is in exchangeable formats.

Box 2: Ecosystems and Ecosystem Services

Ecosystems are collections of living organisms, the interactions between them, and between them and their physical environment. Each ecosystem is characterized by its *composition*, *structure* and the *ecological processes* (interactions and functions such as nutrient cycling, water flows and dispersal) that maintain the ecosystem and keep it functioning as a unit (See Figure 1, below).

Healthy, functioning ecosystems deliver a range of *ecosystem services*, which are the by-products of ecological functioning that are of benefit to people. Ecosystem services can be categorized into a number of types, including:



- **Regulating services**, that maintain natural processes;
- **Provisioning services**, which are harvestable or usable goods;
- **Cultural services**, that are the non-material benefits people derive from natural systems; and
- **Supporting services**, that are necessary for the production of all the others.

None of these ecosystem services can be replaced by man-made substitutes.

Figure 1: Schematic representation of an ecosystem (adapted from Cadman *et al.*, 2010)

Table 1: Selected examples of ecosystem services:

Provisioning	Regulating	Supporting	Cultural
Clean water	Climate regulation	Soil formation	Recreation
Food for people, grazing for livestock	Disease control	Primary production	Education
Fibre	Protection from natural disasters	Nutrient cycling	Spiritual and religious values
Fuel	Pollination	Production of oxygen	Aesthetic value

GBIF and the development of best practices for publishing biodiversity data

Through the *Global Biodiversity Information Facility* (GBIF), digital biodiversity data are being made freely and openly available via the Internet for scientists, researchers, decision-makers and the general public. GBIF promotes a suite of standards and data publishing tools that can be used to capture, publish and discover primary biodiversity data in standardized formats. To date, these tools have been used for publishing biodiversity data gathered from observational records (accounting for about 60% of some 360 million records accessible through GBIF as of May 2012) and biological specimens housed in natural history collections (about 40%).

This best practice guide describes these tools, standards and infrastructure, and explains when and how they should be used, with specific reference to local government. It also indicates sources of additional help, should this be required.

What this guide is, and what it is not

Primary biodiversity data are defined as 'digital text or multimedia data records detailing facts about the instance of occurrence of an organism', i.e. the 'what', 'where', 'when', 'how' and 'by whom' of the occurrence and the recording (GBIF, 2009).' This best practice guide is one of a series of publications developed by GBIF relating to publishing primary biodiversity data (see the resources list in Appendix 3 for details). If widely adopted, the best practices described here should ensure that:

- **Local governments benefit** from the use of GBIF tools and infrastructures for discovery and publishing of primary biodiversity data in their existing fields of activity, such as environmental and development planning, municipal service provision, procurement, green space management, awareness raising and public participation;
- Primary biodiversity data gathered through local government processes **contribute to building national or regional datasets**, and are **freely available in the public domain** for further access and use;
- Biodiversity-inclusive environmental assessment, management and planning at local government level are promoted and strengthened, **enhancing decision-making processes** and contributing positively to **stemming biodiversity loss**.

The document is divided into six parts:

Section 1 provides the rationale for developing the best-practice guide;

Section 2 summarizes the opportunities and challenges that working with biodiversity data present for local government;

Section 3 describes key concepts and principles that local government practitioners should understand in order to make effective use of the biodiversity publishing tools and infrastructures that are available to them;

Section 4 (the bulk of the guide), is more practically-orientated, providing 'how-to' guidelines on what practitioners should do to publish or access primary biodiversity data via the GBIF network.

Section 5 provides a Quick Guide to biodiversity data publishing for local governments, and includes a summary of where additional help can be found if needed.

Section 6 is a brief conclusion.

The guide also includes supporting information, such as a glossary of terms, and a list of references and useful websites, that are included in **Appendices** at the end of the document.

Please Note: This guide is not a comprehensive reference work on local government planning and decision-making processes; the Global Biodiversity Information Facility; biodiversity informatics standards, tools, processes and infrastructure; or biodiversity. For readers who may require more detailed information on these topics, references are given, where possible, for other documents that provide in-depth guidance.

Section 2: Local Government Processes and the Publishing of Biodiversity Data

This section of the guide provides a brief explanation of the types of local government planning processes that could be enhanced by greater access to verifiable primary biodiversity data. It also provides an overview of the challenges facing data publishing in this context, and outlines the benefits to local governments of publishing primary biodiversity data to global standards using the GBIF network.

Local governments as users and generators of biodiversity data

Local governments or authorities are administrative units of an area smaller than a state or province. They are becoming increasingly important as users and managers of biodiversity assets and natural resources, because they set local environmental and development policy, are responsible for land-use planning and decision-making, and develop and manage infrastructure that depends and/or impacts on biodiversity. They are also responsible for local action that supports the implementation of several global,

national or sub-national strategies and policies relating to biodiversity and sustainable development (such as the CBD; the CBD Plan of Action; Local Agenda 21; and national or local biodiversity strategies and action plans - NBSAPs and LBSAPs). Furthermore, environmental sustainability underpins the success of almost all of the core areas of activity for which local authorities are responsible, such as integrated development planning, municipal service provision, economic development and job creation (See Figure 2).

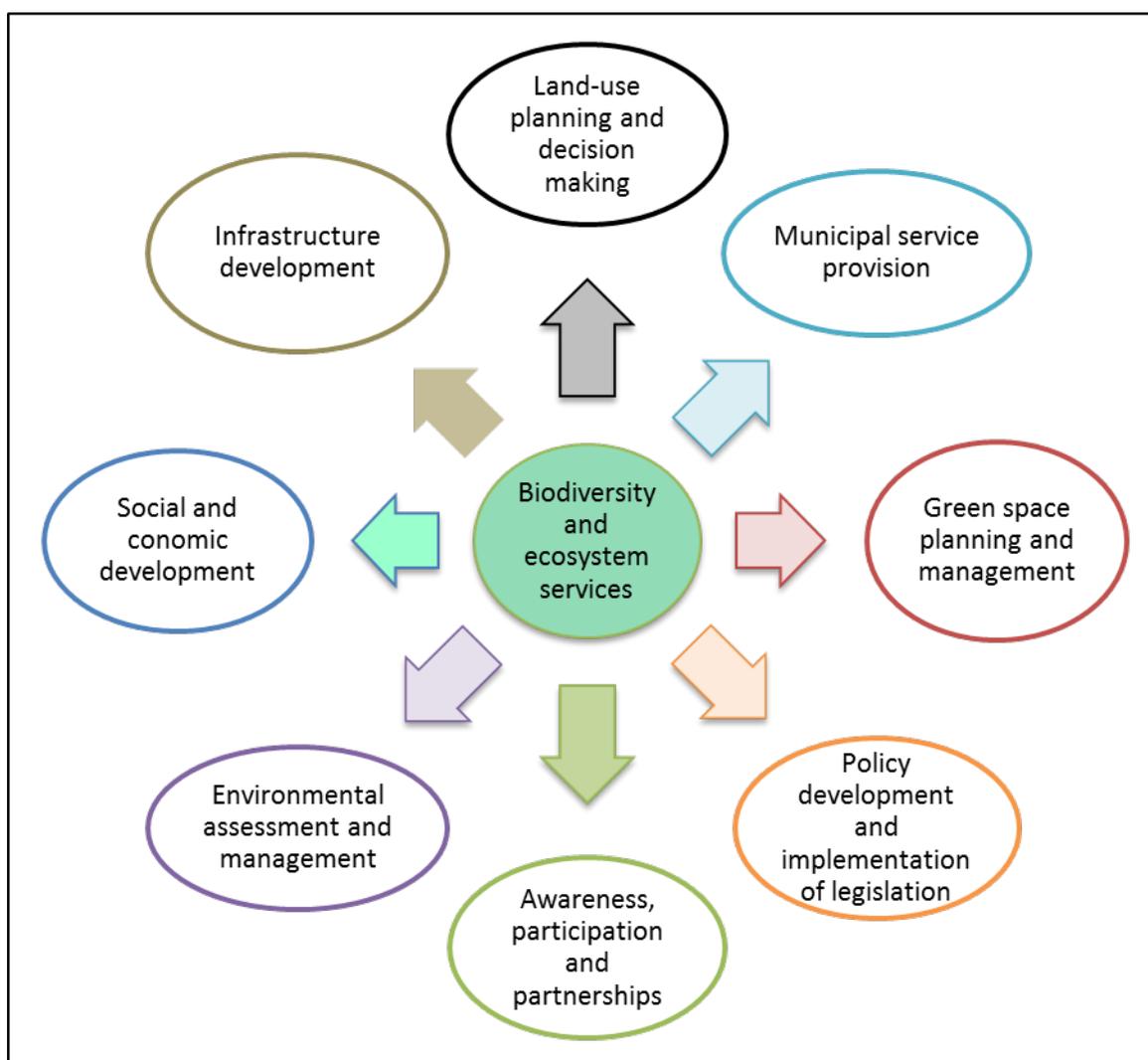


Figure 2: Municipal functions that can incorporate biodiversity data (adapted from ICLEI, 2008)

In the course of doing their work local governments can, therefore, both (i) use biodiversity data that are needed for planning, and (ii) generate biodiversity data about sites. Data generated by local governments could contribute positively to the state of knowledge of a region and to the growth of local, national and even global biodiversity

databases. However, these data currently vary widely in format and purpose and are seldom captured in forms that make them accessible for later use. Furthermore, it is widely recognized that much of the biodiversity data generated through various planning processes and other activities carried out by local governments, is lost after completion of the project. There are many reasons for this, but key amongst them has been a lack of awareness of the tools and protocols suitable for capturing, publishing, archiving and accessing biodiversity data, or lack of knowledge about how to use these tools. In addition, local governments face many other challenges in dealing with biodiversity-related work (See Box 3).

Box 3: Challenges faced by local governments in dealing with biodiversity

There are many impediments to the sound management of biodiversity by local governments, including that:

- Of all the burgeoning and often conflicting agendas of local government, conventional *economic development* and *service provision* normally receive highest priority;
- *Awareness* amongst political decision makers of the positive links between biodiversity, human well-being and sustainable development is often low, and so they may not support investments in biodiversity-related activities;
- Levels of *capacity* for meaningful incorporation of biodiversity priorities into integrated development planning, land-use management and other key activities are often low; few local authorities employ their own specialist biodiversity staff, and duties relating to biodiversity are often added onto the already-full workloads of staff with other core portfolios; local authorities are often reliant on consultants whose knowledge of biodiversity may vary widely, and officials may not be able to assess whether or not biodiversity has been given adequate attention in reports submitted by these consultants;
- *Legal mandates* in respect of the environment generally, and biodiversity in particular, are sometimes unclear to local governments, and where there is no legal requirement to do something, other incentives may be necessary to encourage publishing of biodiversity data;
- *A shortage of funding* may limit local government action, or raise concerns that taking on biodiversity-related activities will place undue strain on already-stretched budgets.

There are a number of important initiatives around the globe that have been instituted to address some these issues (such as the ICLEI-IUCN Local Action for Biodiversity Project). Ensuring that biodiversity receives proper attention, however, remains a challenge. Assigning responsibility for ensuring that biodiversity data are gathered and ultimately published appropriately might be a challenge for under-capacitated local authorities, but this guide includes recommendations, tools and protocols that should assist local governments with this task.

Sources of primary biodiversity data used by local governments

In general, the collection of biodiversity data for local government reports is *ad hoc* and opportunistic, making use of a variety of sources that happen to be available to the practitioner doing the work. These sources may include: literature (scientific or popular); floral and faunal inventories; personal communication and consultation with experts (data based on expert knowledge of a site or species); and information available at universities and other academic institutions, including specimens in natural history collections. Observational data are also gathered through field studies, but these tend to be 'once-off' and of limited duration, due to the time constraints under which most planning processes are expected to be completed at local government level. Local, regional, national and global data resources can also be used, where practitioners are aware of them, but these provide patchy geographic and taxonomic coverage and do not serve their data using consistent formats.

Some of these traditional sources of biodiversity data are not easy for local governments to access, and officials may not know if the data are reliable or up to date. The formats in which the data are collected also vary widely, which means that they are not easily shared or exchanged. This often results in duplication of effort and expense, as the collection of biodiversity data is commissioned independently for the different reports and planning processes in which they are needed.

Why should local governments access data and publish them via the GBIF network?

GBIF and its Participants (see Box 4) provide a rich source of verifiable biodiversity data, served in consistent formats that are freely and openly accessible. GBIF has already

'proven the concept' for global biodiversity data publishing, access and discovery in support of policy-making. This means it should be a relatively short step to secure uptake by local government of the tools and protocols that are available via the GBIF network for publishing biodiversity data. Primary biodiversity data that are published and accessible can be used in other reports and processes, and can contribute significantly to the state of knowledge of biodiversity. Using the common or global standards that have already been established and accepted holds the real benefit that no one has to invent new standards, or invest in developing them, and everyone can benefit from having access to more and more data - this is especially important for local governments, which are often relatively small geographic entities and totally dependent on being able to access and exchange information with neighbouring local and national governments, or other institutions.

Box 4: GBIF Participants

GBIF works through its Participants to mobilize biodiversity data and put them to use, to improve search mechanisms, data and metadata standards, web services and other components of an Internet-based information infrastructure for biodiversity. A GBIF Participant can be any country or organization that has signed the GBIF Memorandum of Understanding (MoU) and that has expressed its intention to abide by its provisions. Voting Participants are countries that contribute financially to the GBIF budget, with voting rights on the GBIF Governing Board, and Associate Participants are countries or organizations that do not contribute to the GBIF budget, but that are willing to observe the terms of the MoU. Associate Participants may take part in the deliberations of the Governing Board, but do not have voting rights. As of mid-2011, GBIF had 33 voting countries, and 24 associate countries as Participants. In addition to this, 47 international and regional organizations were GBIF Associate Participants. The lists of these Participants are available at: <http://www.gbif.org/governance/governing-board/voting-participants/>, <http://www.gbif.org/governance/governing-board/associate-country-participants/>, and <http://www.gbif.org/governance/governing-board/other-associate-participants/>

Challenges facing data publishing and how they can be overcome

Local governments may have concerns about data publishing relating to: the nature of the data themselves (how to describe and manage it); practical challenges relating to capacity, technology and costs; and issues of principle and mindset (psychological or behavioural barriers).

Challenges relating to describing data:

Key challenges relating to the nature of the primary biodiversity data include:

- *Data standards*: compatibility and integration;
- *Data types* - what data types are needed, and by whom?
- *Data quality* - fitness for use, and by whom?
- *Data volumes* - how many data are enough?

Primary biodiversity data collected during the compilation of inventories, EIA, SEA, EMPs, LBSAPs and other municipal planning processes, tend to be gathered 'once off', and vary greatly in accuracy, precision and type. Until recently, the methods used to present, store and archive such data have also been inconsistent. Primary biodiversity data often do not appear in reports, and are seldom captured in any forms that are accessible and re-usable. Specialist reports (which usually contain the biodiversity data) are often submitted in summarized form and may not include full datasets or any supporting field research. This means that the biodiversity findings of many of these planning and decision-making processes are often not verifiable.

The suite of tools presented in this guide provides ways of overcoming these challenges through:

- the use of standardized templates for data collection and management (see Section 4); and
- the inclusion of supporting information called metadata (see Section 4), that make it possible to authenticate the data, assess their quality and strengthen the confidence with which they can be used (see Box 5).

Box 5: Additional guidance on issues of data quality

Practitioners seeking more guidance on issues of data quality are referred to the GBIF Training Manual, or to the following GBIF publications:

1. Principles of Data Quality (Chapman, 2005a) - accessible at http://www.gbif.org/orc/?doc_id=1229&l=en
2. Data Cleaning (Chapman, 2005b) - accessible at http://www.gbif.org/orc/?doc_id=1262&l=en
3. Georeferencing - (Chapman and Wieczorek, 2006) - accessible at http://www.gbif.org/orc/?doc_id=1288&l=en
4. Generalizing Sensitive data (Chapman and Grafton, 2008) - accessible at http://www.gbif.org/orc/?doc_id=1233&l=en

There may be misconceptions regarding the *cost of the technology or professional time* required to publish biodiversity data. All of the GBIF publishing tools are freely and openly available via the Internet, and they do not require expensive software or hardware to operate, so technology costs should not be a barrier to the use of GBIF tools by most local governments.

The most significant investment of professional time lies in the initial learning process required for a designated staff member to become familiar with the tools and procedures - thereafter, the ongoing investment of time is relatively small. However, in many local governments there are few or no staff members with portfolios dedicated to this type of work, and a very real consideration will be who should be responsible for ensuring that biodiversity data are published. We recommend that a requirement to publish the data according to the GBIF/ICLEI Best Practice Guidelines should be built into the terms of reference of the consultants, academic institutions, NGOs or other professionals, who are hired to undertake biodiversity surveys, or any other work that generates biodiversity data. This best practice guide could be included as an information document in such terms of reference.

It is important to acknowledge that *capacity for data management and information technology* varies considerably between institutions and amongst individuals. A variety of tools and procedures have been developed specifically to accommodate these variable levels of capacity.

Challenges relating to mindset:

Potential barriers to publishing biodiversity data include concerns about intellectual property rights and data security, and other issues such as potential delays and obstacles that data publishing may present to plans for meeting development objectives.

- *Intellectual property rights (IPR):* Misconceptions around intellectual property rights and data 'ownership', are often put forward as a reason by consultants or developers for not sharing primary biodiversity data. However, the GBIF suite of tools for publishing biodiversity data allows for free and open access to data, with a mechanism for due attribution and credits to data publishers. This means that data ownership rights are respected, whilst promoting free and open sharing of data (King, *et al.*, *in press, Impact Assessment and Project Appraisal*).

- *Data security:* There are legitimate concerns around making the exact locality data of rare, threatened, endangered or otherwise sensitive species openly available. The practices outlined in this guide include mechanisms for restricting or managing access to sensitive data, without compromising the utility and transparency of the data publishing system.

Potential delays and other barriers: There may also be concerns that publishing biodiversity data will cause undue delays in planning and decision-making processes and, potentially, an increased cost in professional fees. Similarly, there may be concerns that making biodiversity data publicly available may present an obstacle to development, especially if biodiversity of high conservation value is found to be present.

We propose that the few additional steps required to publish the biodiversity data, are easily offset by the benefits to be gained. Practitioners are already collecting occurrence data for the reports they prepare - using the GBIF tools for data collection and publishing will ensure that the data are captured more efficiently, in formats that make it possible to exchange data, without significant investment of time or resources. This means that biodiversity data collected for one planning process (such as an EIA or SEA, for example), could potentially be used for the development of other plans or policies, thus cutting down the need to commission further studies.

Furthermore, we propose that instead of being an obstacle to development, greater transparency regarding the collection of biodiversity data should promote wise decision-making, increase the confidence that is placed in development decisions, and build a positive reputation for local authorities. Addressing this particular issue comprehensively is beyond the scope of this best practice guide, but, further insights can be gained from the recently published Conservation Commons white paper, which can be found at <http://www.conservationcommons.net/conservation-commons-prepares-draft-on-barriers-to-biodiversity-data-sharing>

Responding to these challenges: Each of the challenges mentioned in this section arises from legitimate and understandable concerns. However, most are based on widely-held misconceptions and can be overcome relatively easily, but this requires a change of mindset. Demonstrating the benefits that data publishing holds for improving the effectiveness of local government planning and policy development should help bring about that change.

Incentives for, and benefits of, biodiversity data publishing

The most compelling argument for publishing biodiversity data through the GBIF network is that it will **strengthen and streamline planning and policy development** within local government. It will do this by:

- Increasing the pool of reliable data that is available for informing planning and decision making in local government;
- Increasing the reliability, credibility and transparency of local government planning processes;
- Helping ensure that development planning and decision making are well-informed and carried out within the opportunities and constraints posed by the environment, thereby reducing possible delays further down the line and facilitating proactive compliance with environmental legislation;
- Making it possible to exchange data readily and cheaply with other regions and between projects.

As an example, Box 6 illustrates how EIA practice can be enhanced through the increased availability of published, verifiable biodiversity data.

Box 6: How EIA practice can be enhanced through the increased availability of published, verifiable biodiversity data

Environmental Impact Assessment is a pro-active planning and decision-support tool used by local governments and other stakeholders to assist them in assessing the viability of proposed development activities. It is the principal tool through which the economic, social and ecological impacts of a development proposal can be assessed, enabling a planning decision on whether the proposed development should be approved or not. If approved, the EIA can help identify measures for improving the economic, social and ecological viability and sustainability of the proposed development. Impact assessment holds great potential to ensure that biodiversity values are recognized and taken into account in decision making about land use and development, and almost every EIA uses and generates biodiversity records of some sort. Historically, however, global experience shows that the treatment of biodiversity in EIAs has generally been inadequate (King *et al.*, *in press.*, *Impact Assessment and Project Appraisal*).

To help implement the provisions of the CBD for biodiversity-inclusive EIA, the International Association for Impact Assessment (IAIA) has developed a set of guiding and operating principles for integrating biodiversity into EIA and a best practice manual for biodiversity-inclusive EIA has been published for practitioners and reviewers in Southeast Asia (Rajvanshi, *et al.*, 2007). GBIF, in association with IAIA, has published two best practice guides focussed specifically on publishing EIA-related primary biodiversity data, and it is recommended that local governments advise their EIA consultants to adopt these practices. They can be accessed at:

http://links.gbif.org/eia_biodiversity_data_publishing_guide_en_v1, and http://www.iaia.org/publicdocuments/special-publications/sp7_web.pdf. The ways in which publishing EIA-related biodiversity data can enhance the EIA process are summarized in Figure 3.

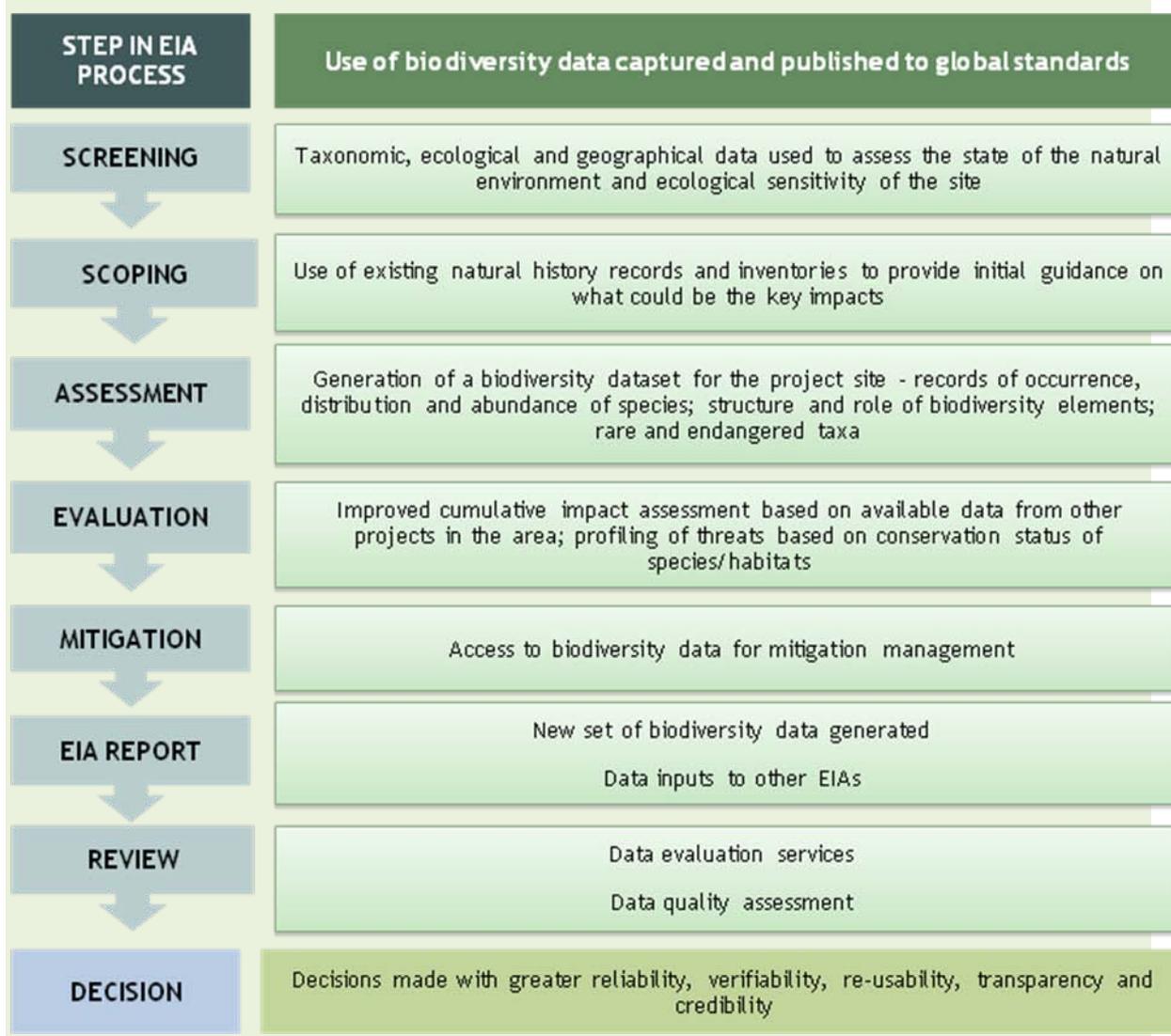


Figure 3: Enhancing EIA practice through the biodiversity data publishing framework, showing anticipated benefits and different steps of EIA process (King *et al.*, in press, *Impact Assessment and Project Appraisal*)

In addition, use of the publishing infrastructure provided through GBIF and its national and institutional nodes will ensure that local governments benefit from the support provided through these networks, whilst contributing to the continuous improvement of local, national and global biodiversity databases. If data discovery and publishing are incorporated into Local Biodiversity Strategies and Action Plans (see Box 7), this can strengthen their contribution to achieving national biodiversity goals.

Box 7: Data publishing and local biodiversity strategies and action plans (LBSAPs)

Halting the loss of biodiversity, and ensuring its sustainable use, cannot be achieved effectively without careful and informed planning that enables strategic action. It is for this reason that the Convention on Biological Diversity (CBD) places great importance on the development of National Biodiversity Strategies and Action Plans (NBSAPs) and their local-level counterparts, LBSAPs.

Local Biodiversity Strategies and Action Plans (LBSAPs) are documents that detail the broad strategies and specific programmes of action that a local government implements in order to conserve and manage its biodiversity assets effectively, and ensure their equitable and sustainable use. LBSAPs can play an important role in supporting, *inter alia*, biodiversity-inclusive land-use planning and decision making, ecosystem management, environmental monitoring and reporting, environmental education and public awareness programmes.

In order for LBSAPs to be effective in achieving their goals, they need to be underpinned by sound science and rely on the availability of accurate biodiversity data that are readily accessible; because LBSAPs usually include some form of biodiversity assessment, they may also generate new biodiversity information about a region, and so may contribute importantly to the body of knowledge of local biodiversity. However, in order for this to happen, local government practitioners (or their consultants) need easy-to-use mechanisms for discovering and accessing verifiable biodiversity data, and for publishing the data that they generate. Publishing primary biodiversity data via the GBIF network addresses these needs. Incorporating data publishing as a key area of activity in LBSAPs provides local governments with an effective way of doing this. GBIF has developed a guide that explains best practices for developing strategies and action plans for data discovery and publishing (see http://www.gbif.org/orc/?doc_id=2755), and these could be incorporated into LBSAPs.

ICLEI, in partnership with the United Nations University - Institute for Advanced Studies and the Secretariat of the CBD, have compiled a set of guidelines to aid local governments in the compilation of LBSAPs. These LBSAP Guidelines support data publishing through the GBIF network.

An additional incentive for publishing biodiversity data through the GBIF network is that it provides opportunities through the mechanism of Data Papers (See Box 8) for data

gatherers, holders or managers to be given due credit for their efforts in data collection, thus addressing concerns regarding intellectual property rights or control over data.

Box 8: Data Papers - a mechanism for incentivizing data publishing

Why do we need data papers?

Knowledge about the identity and occurrence of organisms is essential for informed planning and decision-making that integrate biodiversity conservation and sustainable development goals. However, many of the primary biodiversity data that underpin traditional, published scientific work are often lost, are not accessible to other users, or exist in such varied forms that they cannot readily be used in other applications. Although many scientists and other data gatherers are happy, in principle, to share their primary biodiversity data with other users, they have understandable concerns regarding intellectual property rights, a possible loss of 'control' over sensitive data and a lack of incentives for publishing the data (e.g. receiving due credit for the time and effort spent gathering the data). Furthermore, to date, there has been no mechanism available for publishing primary biodiversity data using the traditional format of a peer-reviewed, citable scholarly publication.

What is a data paper?

Data papers are a mechanism through which data publishers can author and publish metadata describing primary biodiversity datasets as citable academic papers. The primary purpose of a data paper is to expose and describe the data, rather than analyse and interpret the data. Although the concept of a data paper is not new, to date there has been no mainstream mechanism or associated software tools for generating data paper manuscripts (Chavan and Penev, 2011, accessible at: <http://www.biomedcentral.com/1471-2105/12/S15/S2>).

A biodiversity data paper comprises a searchable metadata document that describes a particular dataset (or group of datasets) that is available online; it is published in accordance with accepted academic practices and complies with a standardized format that has been pioneered by the Global Biodiversity Information Facility and Pensoft Publishers. An important feature of data papers is that they should always be linked to the published datasets they describe and that this link (ideally a URL) should be published within the data paper itself.

Biodiversity data papers hold great promise as a mechanism to incentivize data publishing, as they provide the opportunity to credit and cite a wide range of data publishers including scientists, data curators, data managers and other individuals and institutions that play a role as creators of data. This strengthens traditional scientific practice and publication and facilitates data sharing, re-use and preservation by a wide range of data users.

The first GBIF data paper has already been published (see <http://dx.doi.org/10.3897/zookeys.150.200>) . For further information on data papers, visit <http://www.biomedcentral.com/1471-2105/12/S15/S2>, or, for the Pensoft data publishing guidelines including data paper authoring, visit [http://www.pensoft.net/JFILES/Pensoft Data Publishing Policies and Guidelines.pdf](http://www.pensoft.net/JFILES/Pensoft%20Data%20Publishing%20Policies%20and%20Guidelines.pdf)

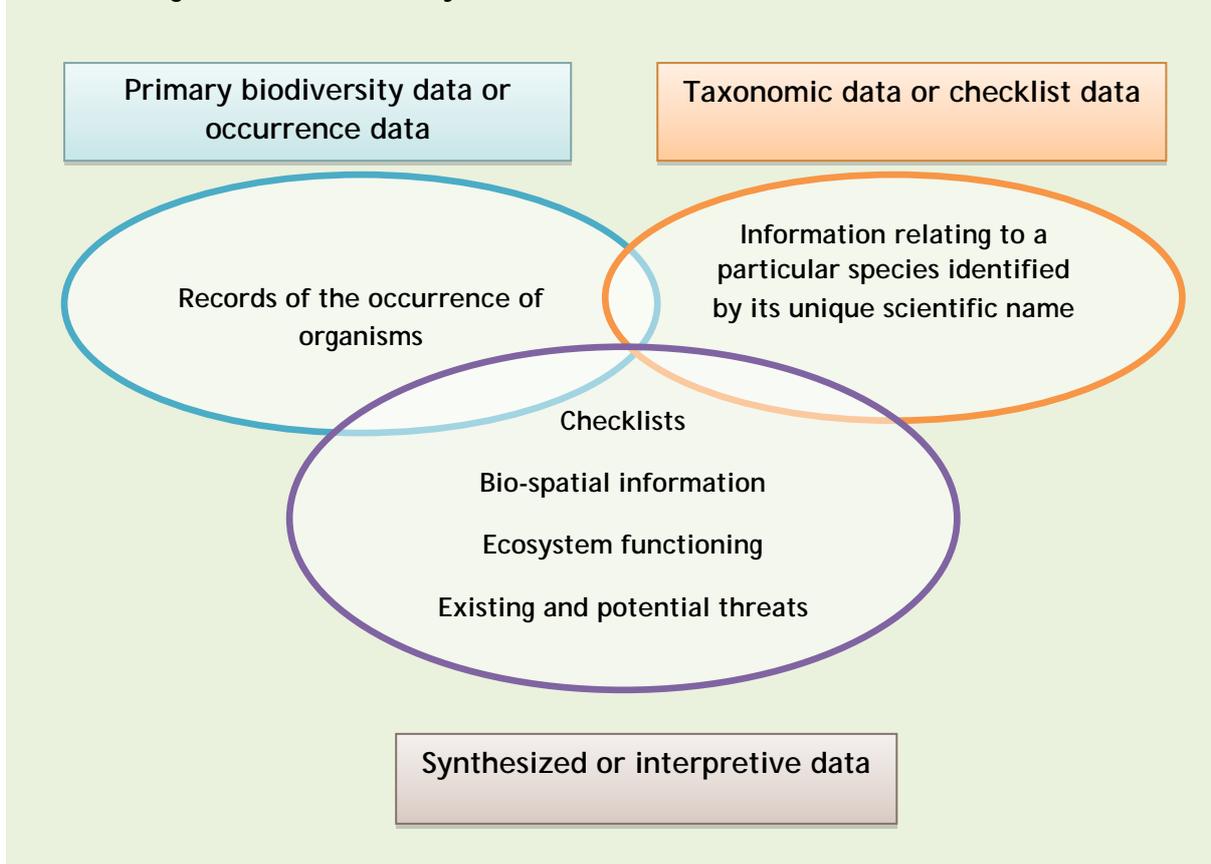
Section 3: Publishing primary biodiversity data - principles of best practice

This section of the guide describes key concepts and principles that local government practitioners should understand in order to make effective use of the available tools and infrastructures for publishing primary biodiversity data. It describes key types of biodiversity data that exist, defines important terms, and outlines the guiding and operational principles that underpin best practice for publishing biodiversity data.

Types of biodiversity data

To be meaningful and useful, biodiversity data need to be comparable. There are several categories of biodiversity data, or levels at which data can be gathered and used, and it is important to distinguish between these. The first distinction to be drawn is between *primary biodiversity data* (precise records of occurrence, distribution and abundance of species), *taxonomic data* (checklists and information about the correct scientific identity of organisms), and *synthesized or interpretive (secondary) data* (see Box 9).

Box 9: Categories of biodiversity data



Primary biodiversity data are the digital text or multimedia data records that detail the *occurrence* of organisms. These data provide information about *what* organisms are present, *where* they occur (specified by geographic coordinates), *when* they are found and *who collects* the data (GBIF, 2010; GBIF, 2011b).

Synthesized or interpretive data include a wide range of information regarding habitats and ecosystems, species checklists, the role of particular biodiversity elements, rarity and conservation status, indigenous knowledge of biological resources, existing and potential pressures on biodiversity, and other similar kinds of information. Much of the biodiversity information contained in the planning and decision-making reports handled by local governments is synthesized or interpretive biodiversity information, but all of it is based on large volumes of primary biodiversity data.

The best practices described in this guide apply to primary biodiversity data only.

Some important data terms

From a data publication perspective, GBIF makes the distinction between several terms relating to biodiversity data, and it is essential to use these precisely to avoid confusion. These terms, which are described in Box 10, as well as in related GBIF publications (GBIF, 2011b), include: data resources or datasets, data elements, data values and metadata.

Metadata refer to the descriptive information that accompanies a dataset - they are the data about data. Metadata are required for all datasets published through the GBIF network and can describe both digital and non-digital datasets (GBIF 2011b). Metadata improve dataset discovery, and provide potential users with details on the fitness for use of the data they describe.

Primary biodiversity data, taxonomic data and metadata are each supported by a different data publishing option within the GBIF network (See Section 4).

Box 10 : Terminology used to describe datasets.

What it is called	What it is	Example
Metadata	Information about the dataset	Who collected the data, when it was collected
Dataset or data resource	Collection of data records	List of species from a site
Data elements	Categories of information about the data records	Scientific name, latitude, longitude
Data values	Data	A data value for the element 'scientific name' could be <i>Acacia karoo</i>

Data publishing

GBIF provides a means of sharing biodiversity data, through a process known as 'publishing' (See Box 11), that makes it universally accessible through the use of standard procedures and protocols. This guide will help officials in local government, their consultants and other interested and affected parties to choose the most suitable option or tool for publishing the primary biodiversity data they have gathered, enabling it to be accessed through the GBIF network.

Box 11: What is data publishing?

In this context, the term '*publish*' refers to making biodiversity datasets publicly accessible in a standardized form, via an online (Internet) access point, which is typically a web address (a URL). This access point is recorded in a GBIF Registry, which then serves to make the virtual location of the dataset freely and openly available. The original data published through the GBIF network is never 'handed over' to GBIF, but instead remains under the authorship of the originator of the data. GBIF maintains a Data Portal (<http://data.gbif.org>) which facilitates discovery and access to data indexed from published datasets in the required formats.

Best practices for publishing primary biodiversity data

These best practice guidelines are intended to enhance biodiversity-inclusive planning and decision-making in local government, and to improve access to, and availability of, primary biodiversity data. In using the term 'best practice', it is acknowledged that the principles and procedures outlined here represent the best possible practice based on current knowledge and technology, and that future improvements may be possible as this field of work develops further.

There are six key principles that must be applied at all stages of the data publishing workflow to underpin best practice: accuracy, precision, fitness for use, effectiveness, efficiency and transparency (Chapman, 2005a):

Accuracy: refers to how correct the data are. For example, is the organism correctly identified? Or, does the locality information match with the known distribution of the organism? If the organism is incorrectly identified, then the accuracy of this information is low. If it is correctly identified, the accuracy of the data is high.

Precision or resolution: refers to the exactness or level of detail of the data. For example, if an organism is identified only to the level of family, the precision of this record is low, even if it is accurate (correct). For occurrence data, the precision is low if only the broad area of distribution is given, but the precision of the data would be high if exact geographic co-ordinates are supplied.

Quality, or 'fitness for use': In the context of this guide, data are described as 'fit for use' or 'potential use' (Chapman, 2005a), if they are suitable for the intended use in planning and decision making in local government processes. Data of low accuracy and low precision are poor quality data that will not be fit for use. High quality data are both accurate and precise, as well as being comprehensive, complete, up to date, easy to access and interpret and consistent with other sources. GBIF strives to ensure as far as possible the quality and fitness for use of the data published through its network.

Effectiveness: this is the likelihood that the data, or a method, might have of achieving the intended outcomes.

Efficiency: relates to the ratio of output (fit-for-use data) to input (data capture and publishing).

Transparency: relates to how accurate, precise and complete the information is that describes the primary dataset (i.e. this relates to the metadata). Transparency enhances accessibility, and also the fitness for use of the data.

Each of these principles can be applied to the primary biodiversity data themselves, and to the tools, protocols and practices that are employed at each step of the data publishing workflow.

The data publishing workflow

Data publishing through the GBIF network follows a series of clear steps, shown in Figure 4, below. Each of these steps is described further in the subsequent sections of this document, and detailed GBIF User Guides are available for each step in the process (these are summarized in Figures 8 - 10 and Table 3, in Section 5).

Suggestions for further reading on these topics are provided in Box 12.

Box 12: Further reading on this topic (data publishing)

GBIF website: www.gbif.org

GBIF (2010). Best practice guide for 'Data Discovery and Publishing Strategy and Action Plans' version 1.0. Authored by Chavan, V. S., Sood, R. K., and A. H. Arino. Copenhagen: Global Biodiversity Information Facility, 29 pp. ISBN: 87-92020-12-7. Accessible online at http://www.gbif.org/orc/?doc_id=2755

GBIF (2011b). Getting started: An overview of data publishing in the GBIF network, (contributed by Remsen, D., Ko, B., Chavan, V., Raymond, M.), Copenhagen: Global Biodiversity Information Facility, 16 pp. ISBN: 87-92020-28-3. Accessible at http://links.gbif.org/getting_started_publishing_en_v1

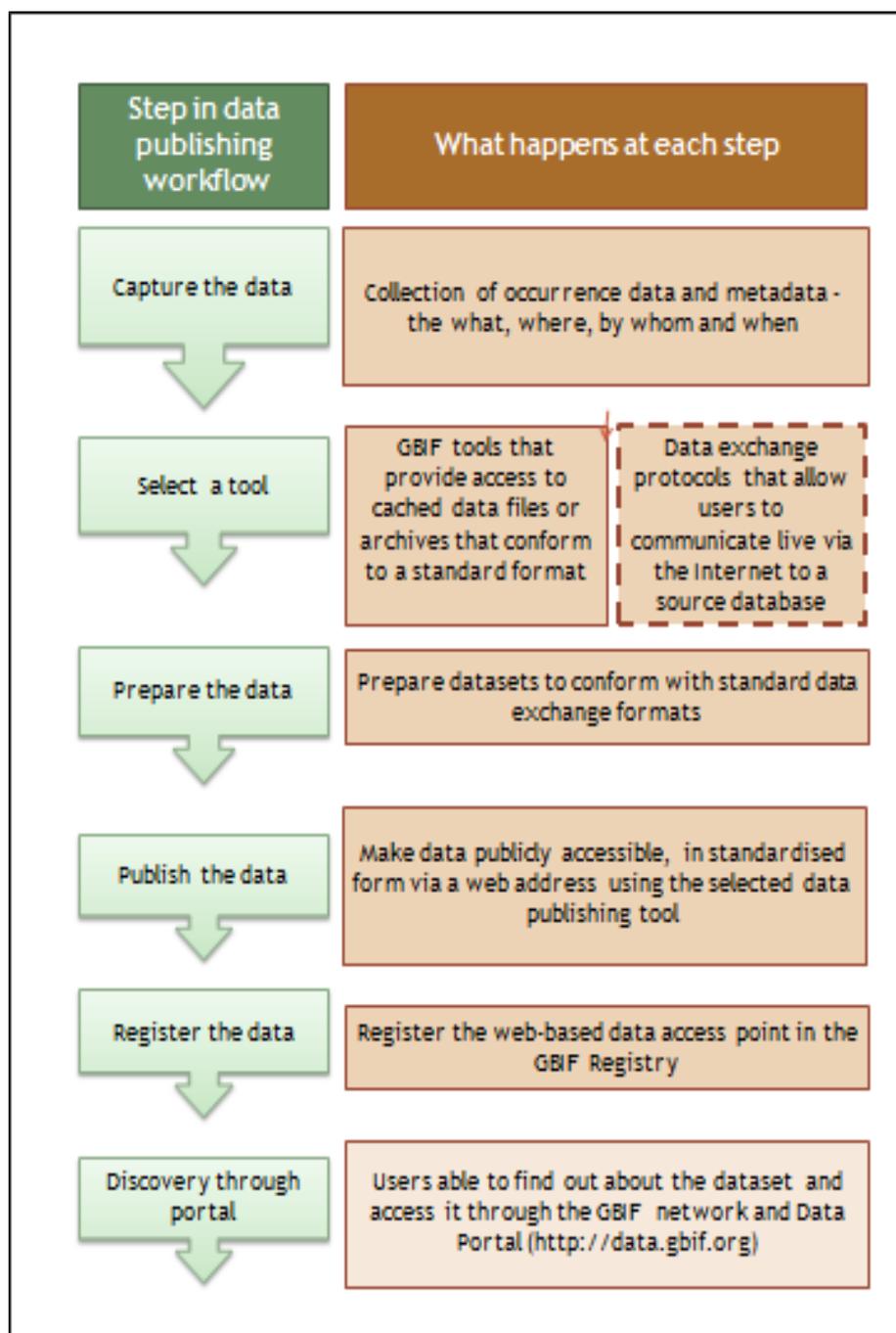


Figure 4: The data publishing workflow (adapted from GBIF, 2011b)

Section 4: Operational principles for publishing primary biodiversity data

This section of the guide explains in detail what happens at each step of the data publishing workflow . These steps are summarized in Figure 3, above, and Table 2 (see page 48)

4.1. How to capture primary biodiversity data

It is important to collect data in consistent formats so that they can be compared between sites or projects, and efficiently integrated with national and global biodiversity databases.

Several GBIF tools exist for capturing, publishing and discovering primary biodiversity data (digitized or otherwise) using consistent formats. Capturing data for inclusion in databases is work that is usually carried out by collections managers (in herbaria and museums) or their data management staff, or by specialist researchers who possess the specific technical or technological capacity for doing this work. Local government officials are rarely experts in biodiversity or well-versed in the technicalities of biodiversity informatics, and may not be familiar with the use of data management tools. However, such expertise is not necessary for the use of GBIF tools for data capture, which are simple to use and have the advantage of:

- (a) making it easier to collect and manage primary biodiversity data;
- (b) improving the consistency and utility of data collection; and,
- (c) collecting the data in a form that is suitable for publishing using GBIF infrastructure.

Using GBIF Excel Templates

GBIF provides a set of pre-configured Excel spreadsheet files that serve as templates for capturing occurrence data (primary biodiversity data), metadata and simple species checklists. These spreadsheets are simple tools that provide a common format and standard for collecting data, using consistent terminology. The spreadsheets include a large number of data fields (or data elements), described using standardized terms (called the **Darwin Core**), into which data (or data values) can be captured. Although it is recommended that as many fields as possible are used in order to maximize the quality of the data, there is a minimum set of six compulsory fields that must be filled in. Essential data elements include: taxon name, latitude/longitude, date/time, name of data collector, name of identifier, reference or link to specimen or photograph. Using these data elements means that the data are not only precise, but are easier to authenticate, thus increasing the confidence with which they can be used.

There are three GBIF Excel spreadsheet templates available:

- (i) *Metadata template*: suitable for composing a metadata document, i.e. including all the information needed to describe the dataset. This template is available for download and use at http://tools.gbif.org/spreadsheet-processor/templates/metadata/metadata-1_v1.xlsx
- (ii) *Occurrence template*: suitable for collating or capturing primary biodiversity (species occurrence) data. These data can be specimens housed in natural history collections or species occurrence observations recorded in the field. This template is available for download and use at http://tools.gbif.org/spreadsheet-processor/templates/occurrence/occurrence-1_v1.xlsx
- (iii) *Species (checklist) template*: suitable for recording and storing simple annotated species checklists. Currently, GBIF recommends three types of checklist (species) templates, depending on the taxonomic information that one is using.

These templates are available at: http://tools.gbif.org/spreadsheet-processor/templates/checklist/checklist-1_v1.xlsx (Template 1); http://tools.gbif.org/spreadsheet-processor/templates/checklist/checklist-2_v1.xlsx (Template 2); and http://tools.gbif.org/spreadsheet-processor/templates/checklist/checklist-3_v1.xlsx (Template 3). The three templates differ in terms of the data fields they include.

The spreadsheet templates that will be of most use to local authorities are the metadata and occurrence templates. These spreadsheets are easy to use and include inline help, which is accessed by hovering the cursor over spreadsheet cells with red upper-right corners. For a quick summary of how to use these tools, consult Box 13, below, or Table 2 in Section 5. There are also a number of GBIF User Guides (GBIF 2011c, GBIF 2011d) that provide detailed step-by-step assistance for use of the Excel spreadsheets (see the Reference List and Appendix 3 at the back of this document).

Once the spreadsheets have been filled in and uploaded, a web-based service called the GBIF Spreadsheet Processor is used to check the validity of the data and to convert them into a standardized format suitable for sharing. This is described in Section 4.2.

Box 13: GBIF Excel spreadsheets - how to use these tools

Step 1: Access the spreadsheets by logging onto the GBIF website and downloading the appropriate template (i.e. occurrence or metadata templates).

Step 2: Populate the spreadsheet with your data using at least the compulsory fields; make use of the inline help if you need to, by hovering the cursor over the cells marked with red upper corners.

Step 3: Upload the completed spreadsheet to the Spreadsheet Processor so that it can be turned into a format suitable for publishing.

4.2. Tools for preparing data for publishing

To 'publish' data means to make them publicly available on the Internet, followed by registering the access point (URL or web address) with the GBIF Registry.

GBIF provides a rich array of support and tools for customizing data formats and for publishing primary biodiversity data in compliance with global standards. The GBIF tools for publishing standardized data files include: the **GBIF Integrated Publishing Toolkit (IPT)**, the **GBIF Spreadsheet Processor** and the **Darwin Core Archive Assistant**, as shown in Figure 5, below. These tools have largely replaced the use of other community tools (data exchange protocols) that were traditionally used for publishing and making data accessible (See Box 14). We recommend that, wherever possible, local governments use the Integrated Publishing Toolkit for publishing their biodiversity data, for reasons explained below. In future, data publishers will be able to access the IPT *via* Data Hosting Centres (See Box 17).

Before data can be published they must first be 'prepared' for publishing - that is, they must be converted into a standardized format known as a *Darwin Core Archive file* (DwC-A) that is supported by the GBIF network (See Box 15 for a note on Darwin Core Archives). Data publishers do not have to generate Darwin Core Archive files themselves, unless they choose to do so, in which case they can make use of a tool called the Darwin Core Archive Assistant (See Appendix 1 or visit <http://tools.gbif.org/dwca-assistant/>).

Box 14: A note on 'community tools' or data exchange protocols:

Data exchange protocols enable the user to interact with a database, and then have their data returned to them in a standardized response format. There are three data exchange protocols that GBIF can accept (TAPIR, DiGIR and BioCASE), but these are no longer the preferred method for publishing biodiversity data, as they either use older data standards that are no longer supported (TAPIR), or they are no longer under active development (DiGIR), or they focus on data from a restricted geographic area (BioCASE), and are mostly suited to taxonomic data only. They also require the use of specialized software called 'wrappers'.

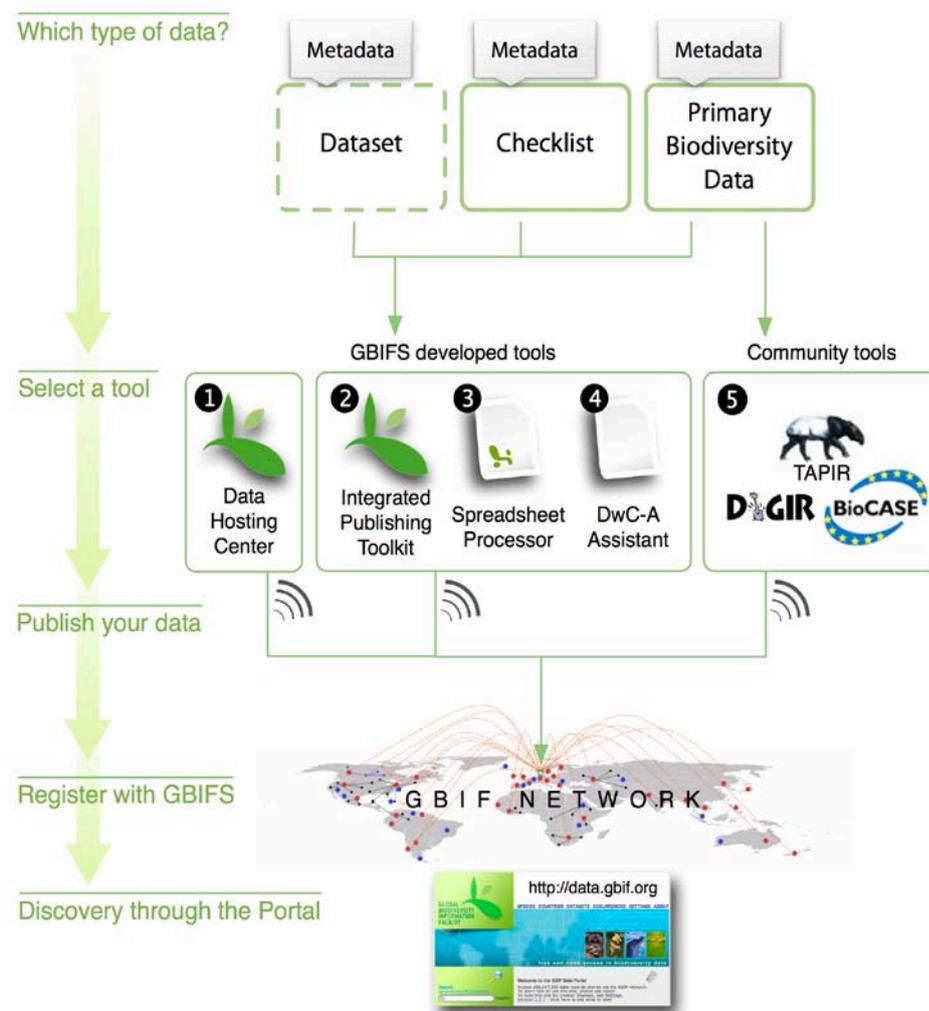


Figure 5: GBIF Tools available for preparing biodiversity data for publishing

Box 15: A note on data standards and formats

Adherence to standards is essential to facilitate data sharing, but publishers of biodiversity data do not have to become bogged down in the technical detail underpinning these standards. Use of good application tools can ensure compliance with standards whilst providing an interface that allows the data publisher to focus on data content, rather than technical details of the system.

There are, however, two standards of which publishers of biodiversity data should be aware, as they underpin the publishing process. The first of these is the **Darwin Core (DwC)** body of standards, which is a standardized glossary of terms for describing and documenting the occurrence of species. GBIF has used the Darwin Core guidelines to develop a standardized format that can be used to publish both species occurrence data and checklist data. This format is called the **Darwin Core Archive**, and is the preferred means for publishing primary biodiversity data.

An expanded explanation of these standards is provided in Appendix 1 at the end of this guide, and interested readers can download the relevant user guides from the GBIF website.

The GBIF Spreadsheet Processor

The Spreadsheet Processor is a web-based application that transforms pre-configured Excel spreadsheet files for occurrence data or metadata into Darwin-Core Archive Files that are suitable for publishing via the GBIF network (GBIF 2011c). The Spreadsheet Processor accepts the completed Excel spreadsheet templates as a web form or as an email attachment. It then performs a series of data checking (validation) and transformation steps, and returns a validated Darwin Core Archive file to the user, suitable for publishing via GBIF or other biodiversity networks that support this format. This process is illustrated in Figure 6. The spreadsheet processor is hosted at <http://tools.gbif.org/spreadsheet-processor/>.

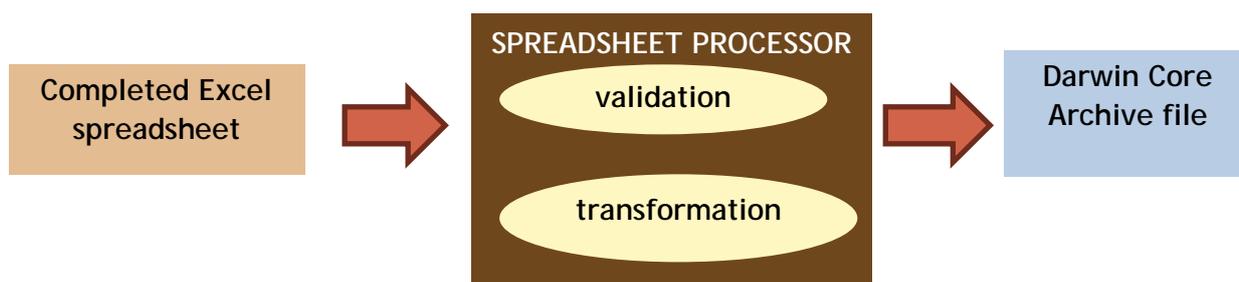


Figure 6: Generating a Darwin Core Archive using the GBIF Spreadsheet Processor

Box 16: The GBIF Spreadsheet Processor - how to use this tool:

Step 1: Access the Spreadsheet Processor at <http://tools.gbif.org/spreadsheet-processor/> and upload your completed Excel spreadsheet by following the instructions provided.

Step 2: Once Spreadsheet Processor has checked and transformed the data, a Darwin Core Archive file will be returned to you, saved in the same folder as your original spreadsheet.

Step 3: Publish the DwC-A file in one of the following ways:

- (a) do it yourself by posting it on a web server and registering the URL with GBIF through a Participant Node (note: registering of datasets is explained in further detail in section 4.5. of this best practice guide);
- (b) send it by FTP or email to a Data Hosting Centre for publication via the GBIF IPT; or
- (c) use the IPT yourself to publish the file.

If the data are already digitized or in the Darwin Core Archive format, then there are two publishing options:

- using the GBIF Integrated Publishing Toolkit
- using the Darwin Core Archive Assistant

The GBIF Integrated Publishing Toolkit (IPT)

The Integrated Publishing Toolkit (IPT) is a software platform developed by GBIF to facilitate easy and efficient publishing of biodiversity data on the Internet.

Why using the IPT is recommended for local governments:

- It can be used to manage and publish primary occurrence data, taxonomic checklists and resource metadata.
- It can be used for data that have already been converted to Darwin Core Archives (using the Excel spreadsheets and the Spreadsheet Processor) or it can accept any delimited text files (e.g. text files using comma or tab-separated values).
- The IPT supports automatic registration of the dataset (see Section 4.5).
- The IPT can be used to author the metadata files, and can be used to create Data Papers (Section 4.3)
- The same version of the IPT can be used by many different data publishers. For example, if ICLEI were to host a version of the IPT on its systems, then any number of local governments could use it to publish their data, while keeping clear attribution and a distinct identity for the datasets.

Currently, data publishers wishing to use the GBIF IPT need to instal and host a local version of the IPT at their home institution. Information on installing and operating the IPT can be found in the IPT user manual or on the IPT website, at <http://code.google.com/p/gbif-providertoolkit/>

In future, it will be possible to access the IPT via a GBIF-endorsed Data Hosting Centre (See Box 17), and this will be the easiest option for practitioners in local governments to use.

The Darwin Core Archive Assistant

This facility can be used when data are already digitized or in a relational database. It would be suited to those users who have access to high levels of data management and IT capacity. It is not recommended for use by local government, but, for those who are interested, more information can be found at <http://www.gbif.org/orc/?docid=2827&l=en>.

Box 17: Data Hosting Centres

These will, in future, provide a service to scientists and practitioners who generate large volumes of primary biodiversity data, but who are unable to share or publish it directly themselves due to a lack of suitable informatics infrastructure, or the skill-sets required to use publishing tools. The Data Hosting Centres will provide an ideal 'one stop shop' through which local government data publishers can capture, prepare, publish, register and archive their data. Once the practitioner has submitted data in the required format (i.e. Darwin Core Archives), the rest of the data publishing process can be channelled through the Data Hosting Centre, although the originator of the data always retains authorship and control of the dataset. There are currently GBIF-supported Data Hosting Centres under development in Denmark (DanBIF) and South Africa (Endangered Wildlife Trust). ICLEI could play an important role as a Data Hosting Centre for local government, especially as it is a GBIF Associate Participant and therefore able to endorse publishing institutions and their datasets for registration.

4.3. Publishing metadata

Documentation describing datasets (resource metadata), is an essential part of any data management system, and allows data users to assess the quality of the dataset and its fitness for use (GBIF, 2011f). It is not possible to publish primary biodiversity data on the GBIF network without the accompanying metadata.

Metadata standards and data elements

As with the publication of occurrence data, there are clearly defined standards, known as a ***metadata profile***, with which a metadata document must comply. These are described in detail in the GBIF Metadata Profile Reference Guide (GBIF, 2011f).

The types of data elements (or fields) described in a metadata document include, *inter alia*: who collected the data, where and when the data were collected (basic information); geographical and temporal coverage of the data; sampling methods; physical data; project data; associated parties; where voucher specimens are housed.

There are five compulsory data fields, designed to allow a prospective end-user of the data to discover the name and a brief description of the dataset, details of the key contact person and information regarding data management rights. Once the compulsory data fields have been filled in, the file can be saved, then modified and updated at any later time.

Options for authoring and publishing metadata

As with occurrence data, there are three ways of writing (authoring) a metadata file:

- Using GBIF Excel spreadsheet templates;
- Using the GBIF Integrated Publishing Toolkit;
- Modifying an existing sample document.

If occurrence data are being published using Excel spreadsheets or the GBIF IPT (either directly or via a Data Hosting Centre, when these are available), then there is a built-in metadata authoring function that can be used to write the accompanying metadata document, following the same general procedures as are outlined in Table 2, below. Using the GBIF IPT makes it easier to handle large numbers of metadata documents, but requires the data to be already in a digitized form, and involves loading the IPT software onto your computer or data management system (if no Data Hosting Centre is available). In the context of local governments, it is probably quicker and easier to use the GBIF Excel spreadsheet template for metadata (http://tools.gbif.org/spreadsheet-processor/templates/metadata/metadata-1_v1.xlsx), and the Spreadsheet Processor, and then to publish it through a Data Hosting Centre.

There are three GBIF guides that provide assistance with the steps and procedures to follow for authoring metadata files. These include:

- GBIF spreadsheet templates: User Guide (http://www.gbif.org/orc/?doc_id=2823&l=en)
- GBIF Metadata Profile: Reference Guide (http://www.gbif.org/orc/?doc_id=2820&l=en)
- GBIF Metadata Profile: How-to Guide (http://www.gbif.org/orc/?doc_id=2821&l=en)

4.4. Criteria for selecting a data publishing solution

As described above, there are different options for publishing biodiversity datasets once they are in the required format. For data publishers in local government, who are more likely to use Excel spreadsheets and the Spreadsheet Processor, the datasets need to be manually posted on a website, and there are different options for doing this. If the IPT is used, then publication is handled automatically.

Important criteria to apply in selecting the most appropriate publishing solution for your data include:

- (i) Whether the data have been digitized or not, and how many datasets you need to manage; and,
- (ii) The level of technological and data management capacity available to you.

GBIF Excel spreadsheets and the Spreadsheet Processor offer simple solutions for preparing data for publishing in situations where data have not yet been digitized (as is likely to be the case with many local government data-generating activities). An attractive feature of using these particular tools is that they do not require any specialized software, and they offer easy and quick options for managing smaller numbers of datasets. Once the DwC-Archive has been generated, it can be published using either the GBIF Integrated Publishing Toolkit (or, in future, via a Data Hosting Centre) (see Figure 7).

In cases where data have already been digitized or are in a relational database, it is possible to use the Darwin Core Archive Assistant to generate the DwC-A file, if adequate data management and IT capacity are available (See Figures 7 and 8).

The levels of IT and data management capacity available to a practitioner will also influence the choice of publishing pathway. When both the data management and IT capacity are low, the best option is to use the Excel spreadsheet templates and Spreadsheet Processor to generate the Darwin Core Archive, followed by publication via a Data Hosting Centre. This is the situation that is most likely to apply to data publishers in local government.

If the data management capacity is low, but the IT capacity is high, then the institution or practitioner may choose to install the Integrated Publishing Toolkit software itself, and publish the dataset directly using the IPT. If both the data management and IT capacity is high, then the publisher may choose to generate its own Darwin Core Archives and publish

them by installing the IPT to its own system. The relationship between IT and data management capacity in determining data publishing solutions is illustrated in Figure 8, below.

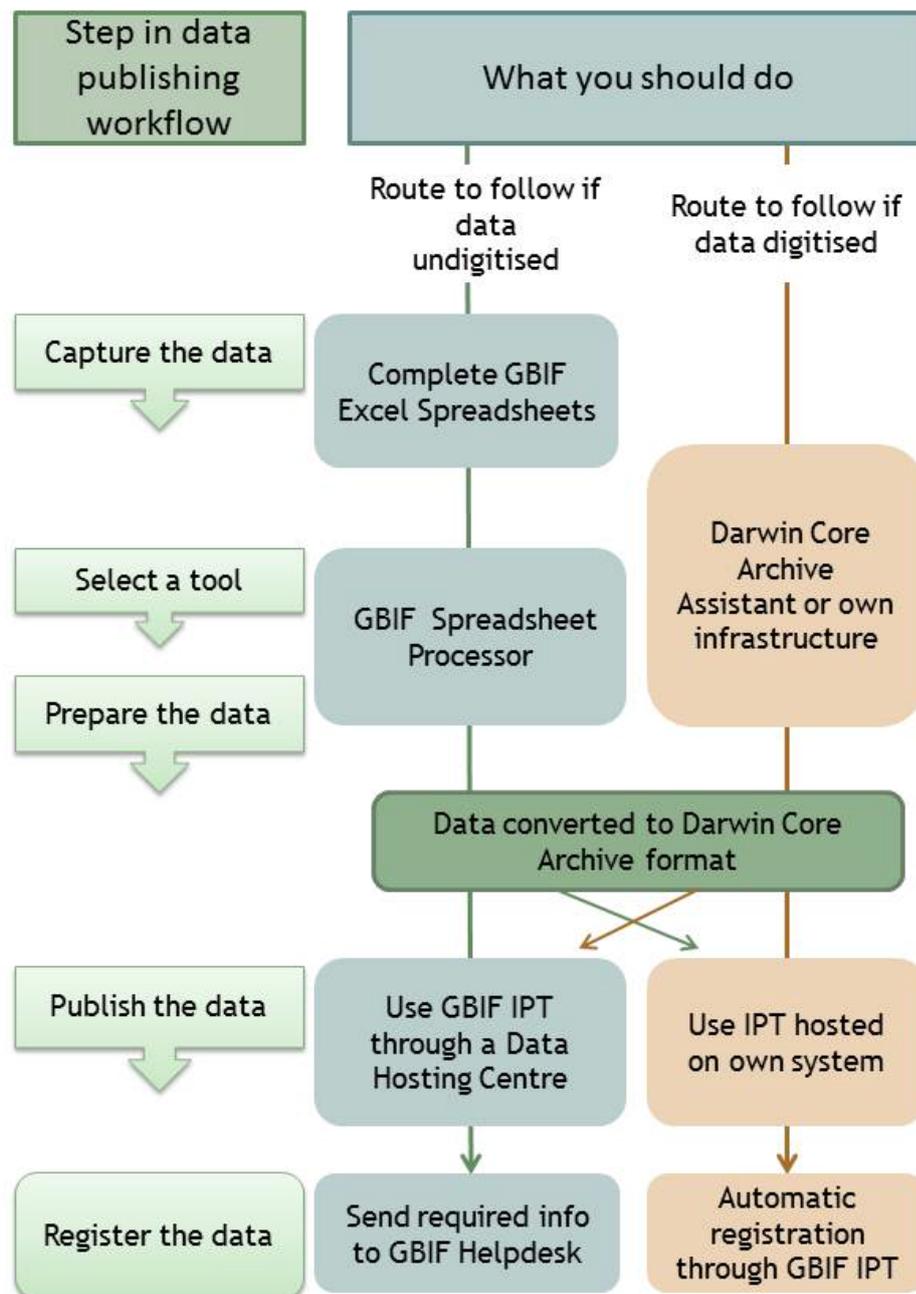


Figure 7: Options for preparing and publishing primary biodiversity data

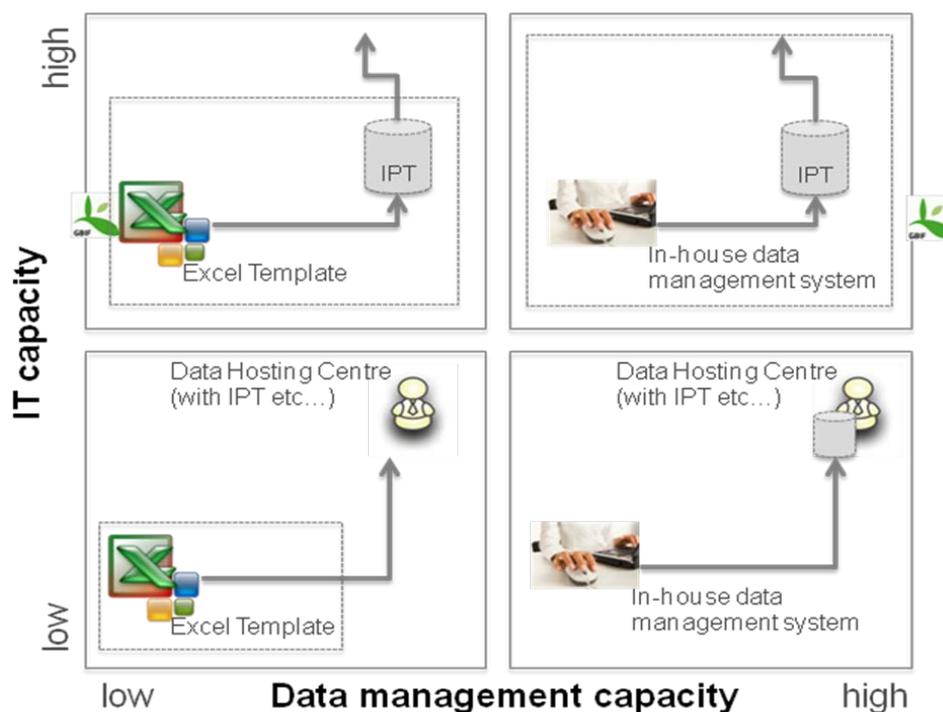


Figure 8: Data publishing options in relation to IT and data management capacity.

4.5: Registering your data

Registration is the final step in the data publication process. An entry for the dataset URL is made in the GBIF registry that enables the resource to be discoverable and accessible. There are three options for registration of datasets:

- (a) Using the GBIF Integrated Publishing Toolkit;
- (b) Using the Spreadsheet Processor; and
- (c) Using other tools.

How to register your dataset

The GBIF IPT supports automatic registration in the GBIF network (see the online manual for the IPT). Using the Spreadsheet Processor or other tools there is no automatic registration. An email must be sent to helpdesk@gbif.org with the following information:

1. Dataset title.
2. Dataset description.
3. Technical contact (the person to be contacted in matters regarding technical availability or resource configuration issues on the side of the dataset or data publisher).
4. Administrative contact (the person to be contacted in all matters regarding scientific data content and usage of a specific dataset or data publisher)
5. Institution name.
6. Your relation to this Institution.
7. The name of the GBIF Participant Node (see Box 18) that can endorse the publishing institution - the first choice would always be your own country's node, but, if not yet a Participant, alternative options are outlined in Box 18, below).
8. The dataset URL: either the access point URL (if you are publishing using one of the provider softwares), or the DwC-Archive URL (if you are publishing via a zipped DwC-Archive).
9. The metadata document URL.

The GBIF Helpdesk will attend to your registration request as quickly as possible. The Helpdesk will first contact the GBIF Node selected with the registration, and ask them whether they want to endorse the new data publisher installation in their domain. Each new registration needs formal endorsement from a GBIF Participant Node manager (who best knows the institutions and databases in their country/organization) before it is allowed into the GBIF Registry. This is a simple quality control step required by the GBIF Participant Node Managers Committee.

Once endorsement has been received and the registration is completed, the registered dataset can be found on the [GBIF Registry website](http://gbrds.gbif.org) (<http://gbrds.gbif.org>), through searching by institution name or dataset title.

Indexing of datasets by GBIF

Following registration, the GBIF Helpdesk will queue the newly registered dataset for indexing. Depending on the size of the dataset, indexing can take anywhere from minutes to weeks. If problems are encountered during indexing, the GBIF Helpdesk will work with you to resolve them as quickly as possible.

When indexing is successful, the new dataset will become publicly available (or discoverable) in the GBIF Data Portal (<http://data.gbif.org>). At present, GBIF attempts to update each registered dataset at least once every three months. During indexing, a set of core data elements is retrieved from your dataset and is stored in the GBIF index, so that the dataset will become accessible for searches.

GBIF has published a manual (GBIF, 2011g) describing the step-by-step process for publishing and registering data.

Section 5: Quick Guide to biodiversity data publishing by Local Governments

This section provides a summary of which publishing tools are available for use by local government practitioners, when they should be used and how to use them. It also provides lists of the documents that are available to provide more detailed assistance in the use of these tools. In this section we also make some general recommendations on how local governments might deal with the issue of data publishing and the benefits to them of doing so.

5.1. Summary of recommendations for local governments

It is difficult to make a 'one-size-fits-all' set of recommendations to local governments as to which data publishing option they should follow, as these institutions vary widely in respect of capacity and resources. However, some general rules of thumb that could be applied include:

- (i) Use the Excel Spreadsheet templates and Spreadsheet Processor to capture all biodiversity data that are gathered as part of local government planning processes. This means that your data will automatically be suitable for publishing via the GBIF network.
- (ii) If you have access to a Data Hosting Centre, make use of this facility to publish and archive your datasets, using the Integrated Publishing Toolkit.

- (iii) If you do not have access to a Data Hosting Centre, you can either publish your datasets manually, or host a version of the IPT on your own system to deal with publishing and registration.
- (iv) Build a requirement to publish biodiversity data according to the GBIF/ICLEI/CBD best practice guidelines into the Terms of Reference for all consultants that are contracted to do assessments or prepare biodiversity reports or plans for any relevant local government process. These best practice guidelines can be included as an information document in such terms of reference. You should also recommend to consultants that they make use of the GBIF network to source appropriate data when embarking on a new study.
- (v) Make data discovery and publishing a part of your Local Biodiversity Strategy and Action Plan.
- (vi) Become part of the GBIF network and benefit from the knowledge exchange and support it offers. You can become involved in the following ways:
- Local governments from countries which are GBIF Participants can get in touch with their national Biodiversity Information Facilities (BIF) or GBIF Node Managers. A list of the GBIF Participant NODES is accessible at <http://www.gbif.org/participation/participant-nodes/who-we-are/countries/>
 - Local governments from countries which are not yet GBIF Participants have two options. In the first place, they can contact ICLEI which is an Associate Participant in GBIF. Secondly, they can contact one of the Associate Organization Participant Nodes. A list of the Associate Organization Participants is accessible at <http://www.gbif.org/participation/participant-nodes/who-we-are/other-associate-participants/>
 - Contact one of the GBIF promoted Data Hosting Centres (DHCs) - there are currently two DHCs under development: DanBIF (Denmark) and the EWT DHC (South Africa).

5.2. Where to find assistance

Table 2 summarizes the tools that are available at each step in the data publishing pathway, and explains how and when they should be used. Figures 9 - 11 and Table 3

summarize the GBIF User Guides that are available online to assist with the publication of primary biodiversity data and their associated metadata, using the GBIF tools. These guides provide detailed, step-by-step instructions in the use of all the key tools used at different steps in the data publishing process.

Figures 9 - 11 are document maps, indicating the sequence in which the GBIF User Guides should be consulted in the publishing process. They are complemented by the information presented in Table 3, which lists the key documents that should be used to guide you at each step of the data publishing process.

Should further assistance be required, you can contact the wide network of GBIF country and organization Participant Nodes (see Box 18). These nodes encourage, coordinate and assist in biodiversity data publishing activities within their respective jurisdictions and domains.

Table 2: Quick Guide to publishing occurrence data using GBIF-supported tools

Publishing tool	When it should be used	How to use this tool
GBIF Excel Spreadsheet templates	<ul style="list-style-type: none"> ✓ When occurrence or species checklist data are not already digitized and you need to generate a DwC Archive OR when you need to generate a metadata document ✓ When publishers want a quick and easy publishing solution that does not require high data management or IT capacity ✓ When there are relatively few datasets involved 	<ul style="list-style-type: none"> ✓ Access spreadsheets by logging on to the GBIF website and downloading the appropriate template at http://tools.gbif.org/spreadsheet-processor/templates ✓ Populate the spreadsheet with your data, using at least the compulsory data fields; make use of the inline help by hovering the cursor over the cells with red upper corners, or use the Guides shown in Figs. 7 & 8 ✓ Upload the completed template to the Darwin Core Spreadsheet Processor at http://tools.gbif.org/spreadsheet-processor/
GBIF Spreadsheet Processor	<ul style="list-style-type: none"> ✓ When you need to convert a completed GBIF Excel Spreadsheet Template into a DwC-Archive file that is 	<ul style="list-style-type: none"> ✓ Access the Spreadsheet processor at http://tools.gbif.org/spreadsheet-processor/ and follow the instructions for uploading and processing the filled in Excel

	suitable for publishing	<p>Spreadsheet</p> <ul style="list-style-type: none"> ✓ A DwC archive file will be returned to you and saved in the same folder as your template ✓ Publish the DwC-A yourself by posting it on a web server and registering the URL with GBIF OR send it by FTP or email to a Data Hosting Centre for publication via the GBIF IPT
GBIF Integrated Publishing Toolkit	<ul style="list-style-type: none"> ✓ When you need to publish occurrence data, taxon data and associated metadata that are already digitized ✓ When you have an already-created Darwin Core Archive (e.g. that has been created using the Spreadsheet Processor), OR when you need to generate the Darwin Core Archive from pre-digitized data ✓ When you need to validate, publish and register DwC-A files ✓ when large numbers of datasets are being managed 	<ul style="list-style-type: none"> ✓ To work directly with the IPT you must install the IPT software on your computer; information on installing and operating the IPT can be found in the IPT User Manual or on the IPT project website (visit http://code.google.com/p/gbif-providertoolkit/) ✓ To generate a DwC-A file using the IPT, follow the instructions provided in the GBIF Darwin Core Archive How-to Guide (GBIF 2011e) ✓ Publishing (and registration) are built in automatically
EIA Primary Biodiversity Data Publishing Facility	<ul style="list-style-type: none"> ✓ A facility that allows you to generate DwC Archives and that provides access to biodiversity data, mapping services and other functions that are geared specifically for EIA practitioners ✓ Soon to be available in South Africa (SANBI) 	<ul style="list-style-type: none"> ✓ Register as a user (many functions on this system are only available to registered users) ✓ Create a user profile and create a project ✓ Within a project, create your dataset(s) using the GBIF Excel Spreadsheets. (The Facility will also accept plain and delimited text files) ✓ Once the DwC Archive is returned to you, either post the file on a website yourself,

		or do so using the GBIF IPT, OR send the dataset to a Data Hosting Centre for publishing and registration.
Data Hosting Centres	<ul style="list-style-type: none"> ✓ A one-stop shop at which you can download spreadsheet templates, publish DwC Archives and deposit data, if you do not have the time, resources or skills sets to install and work with the GBIF IPT yourself 	<ul style="list-style-type: none"> ✓ Access the IPT either at DanBiF Data Hosting Centre; the Endangered Wildlife Trust (EWT) Data hosting Centre, or SANBI (www.biodiversityadvisor.sanbi.org.za) ✓ Contact helpdesk@gbif.org for further information, or contact your Participant Node (see Box 18) <p>DanBiF: Currently under revision EWT: available at www.ewt.org.za/foryou/datasharing.aspx</p>

Table 3: Summary of GBIF User Guides and other documents that are available online to assist users with publishing biodiversity data through the GBIF network

Step in data publishing workflow	GBIF Guides available to assist you		
	For metadata	For occurrence data	For checklists
Familiarizing yourself with concepts, tools and procedures	Getting Started: an Overview of Data Publishing through the GBIF Network		
Selecting a tool for capturing data and describing the data	The Integrated Publishing Toolkit Guide		
	GBIF Spreadsheet Templates: User Guide		
	GBIF Metadata Profile: Reference Guide	Create Your Own Darwin Core Archive Step-by-Step Guide	
	GBIF Metadata Profile: How-To Guide	Darwin Core Archive Assistant User Guide	
Preparing data for publishing (also called 'mapping' data in some documents)	Darwin Core Archive How-to Guide		
		Darwin Core Quick Reference Guide	GBIF GNA Profile Reference Guide
		Darwin Core Archive: Reference guide to the XML Descriptor File	Publishing Species Checklists Best Practices
Publishing the data	Publishing and Registering Your Data with GBIF		

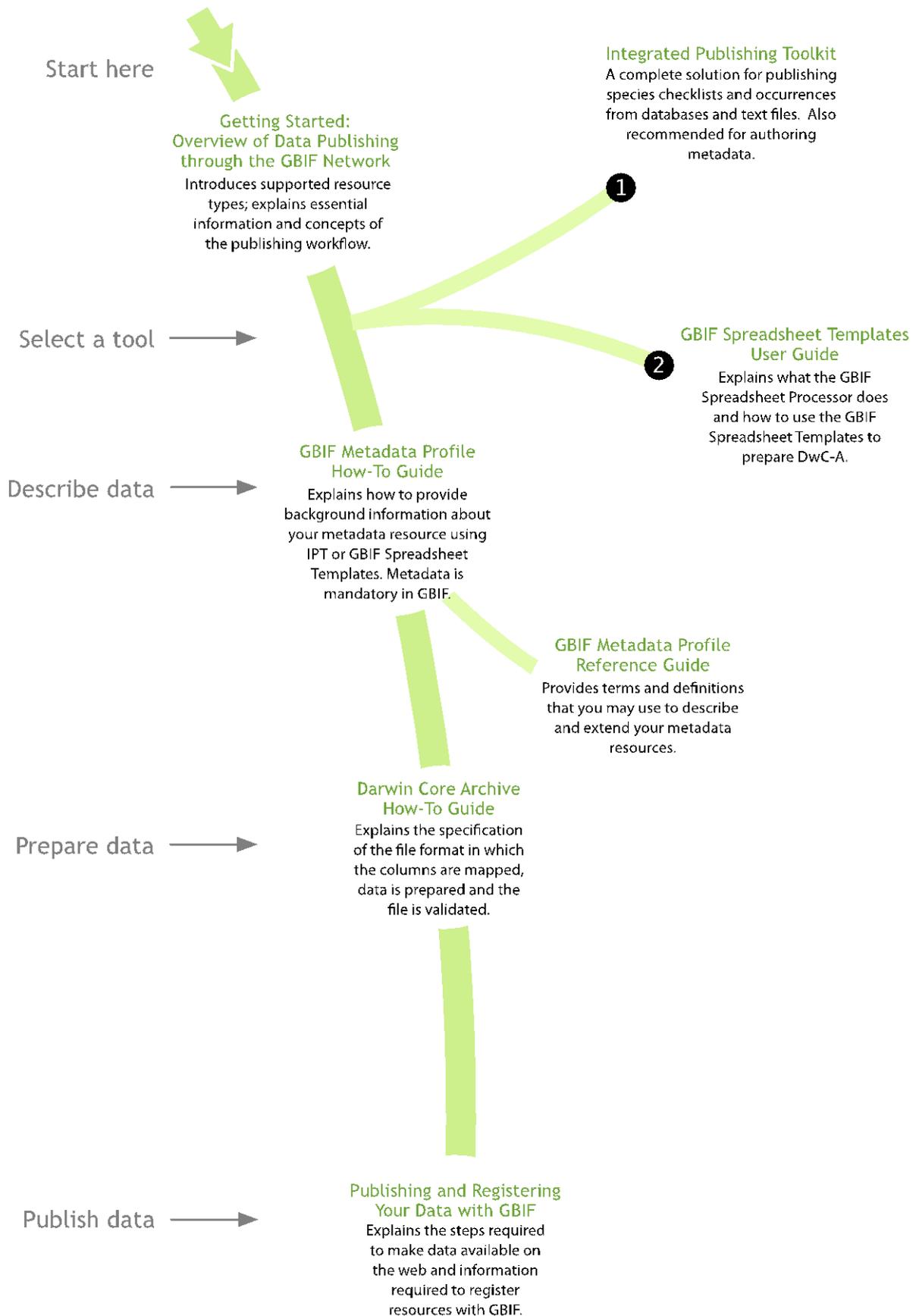


Figure 9: The documentation map for metadata publishing in GBIF.

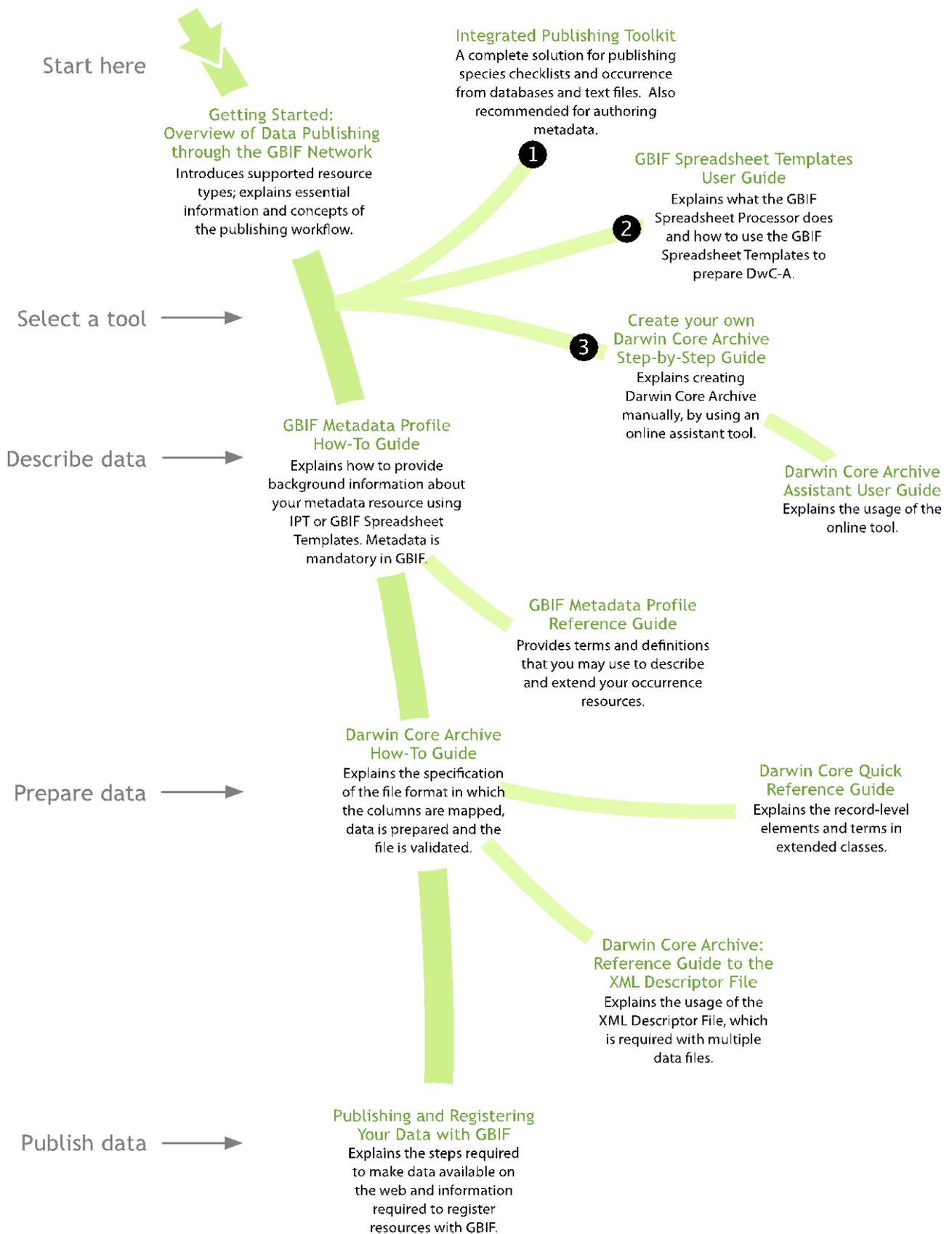


Figure 10: The documentation map for occurrence data publishing in GBIF

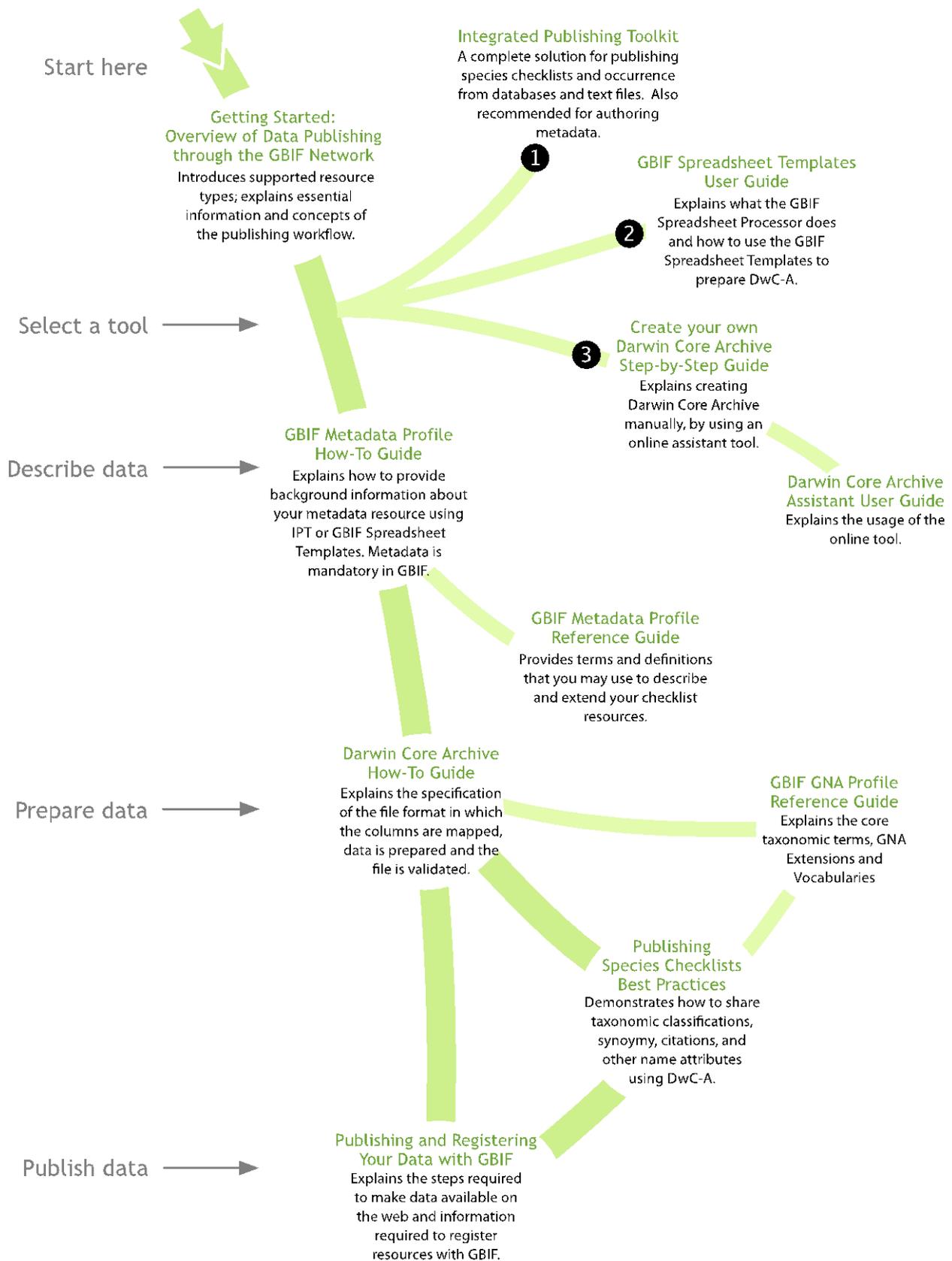


Figure 11: The documentation map for checklist publishing in GBIF.

Box 18: GBIF Participant Nodes

A GBIF Participant Node is the team that coordinates the network of data holders, users and other stakeholders for each GBIF Participant (see Box 4) and coordinates data-sharing activities within its domain. The Participant Node is a conduit by which the GBIF Participants meet their own biodiversity information needs within their respective countries or organizations, while benefitting from and contributing to the GBIF network's mission and goals on making biodiversity data globally accessible. To find out who your Participant Node is, check on the GBIF website at

<http://www.gbif.org/participation/participant-nodes/who-we-are/countries/>, for a list of country members, and <http://www.gbif.org/participation/participant-nodes/who-we-are/other-associate-participants> for the list of regional and thematic Participant Nodes.

Section 6: Concluding remarks

This best practice guide describes a suite of simple, inexpensive tools and procedures that can be used by local governments to capture, publish and discover primary biodiversity data. Publishing these data using consistent, internationally standardized formats is a relatively quick and easy procedure that can be easily adopted as an integral part or step of other planning or policy-development processes. Uptake of the tools and processes described in this best practice guide will:

- enable free and open access to biodiversity data, which is essential for biodiversity-inclusive planning and development at local government level;
- facilitate the ongoing expansion and improvement of the local, national and global biodiversity databases on which environmental planning, EIAs, land-use management, policy development and areas of scientific work frequently rely, improving baseline knowledge of the ecosystems of a particular site, region or country;
- help practitioners who do specialist work for local governments to gain recognition for their work by enabling them to be cited in future uses of their data;
- enhance the quality, predictive value, verifiability and transparency of local government planning processes, thus improving the land-use decisions that they inform and the confidence civil society can place in these decisions.

References:

Cadman, M., Petersen, C., Driver, A., Sekhran, N., Maze, K. and Munshedzi, S. 2010 *Biodiversity for Development – South Africa’s landscape approach to conserving biodiversity and promoting ecosystem resilience*. South African National Biodiversity Institute, Pretoria. 171 pp., ISBN: 978-1-919976-59-4. Accessible at http://cmsdata.iucn.org/downloads/primer_11_2_mb.pdf

Convention on Biological Diversity. (1992). United Nations Environment Programme, Publication no.: Na 92-7807

Chapman, A. (2005a). Principles of Data Quality, version 1.0. Copenhagen: Global Biodiversity Information Facility. 58 pp. ISBN: 87-92020-03-8. Accessible at <http://www2.gbif.org/orc/?doc id=1229&l=en>

Chapman, A. (2005b). Principles and methods of Data Cleaning, version 1.0. Copenhagen: Global Biodiversity Information Facility. 77 pp. ISBN: 87092020-04-6. Accessible at http://www2.gbif.org/Data_Cleaning.pdf

Chapman, A., Wieczorek, J. (2006). Guide to best practice for georeferencing. Copenhagen: Global Biodiversity Information Facility. 84 pp. ISBN: 87-92020-00-3. Accessible at <http://www2.gbif.org/Georeferencing.pdf>

Chapman, A., Grafton, O. (2008). Guide to best practices for generalising primary species occurrence data, version 1.0. Copenhagen: Global Biodiversity Information Facility. 27 pp. ISBN: 87-92020-06-2. Accessible at <http://www2.gbif.org/Generalising.pdf>

Chavan, V. and Penev, L. (2011). The data paper: a mechanism to incentivize data publishing in biodiversity science. *Bioinformatics* 12 (Suppl. 15): S2. Accessible at: <http://www.biomedcentral.com/1471-2105/12/S15/S2>

Conservation Commons white paper (2011): A review of barriers to the sharing of biodiversity data and information, with recommendations for eliminating them. Prepared by the Secretariat for the Friends of the Conservation Commons, available at <http://www.conservationcommons.net/conservation-commons-prepares-draft-on-barriers-to-biodiversity-data-sharing>

GBIF (2008). GBIF Training Manual 1: Digitisation of natural history collections data, version 1.0. Copenhagen: Global Biodiversity Information Facility. ISBN: 87-920-07-0. Accessible at <http://www.gbif.org>.

GBIF (2009). GBIF Work Programme 2009-2010. Copenhagen: Global Biodiversity Information Facility, 59pp. Accessible at <http://www2.gbif.org/WP2009-10.pdf>

GBIF (2010). Best practice guide for 'Data Discovery and Publishing Strategy and Action Plans' version 1.0. Authored by Chavan, V. S., Sood, R. K., and A. H. Arino. Copenhagen: Global Biodiversity Information Facility, 29 pp. ISBN: 87-92020-12-7. Accessible online at <http://www2.gbif.org/BestPracticeGuide-final.pdf>

GBIF (2011a). GBIF Strategic Plan 2012-2016: Seizing the Future. Copenhagen: Global Biodiversity Information Facility, 7pp. ISBN: 87-92020-18-6. Accessible at http://links.gbif.org/sp2012_2016.pdf

GBIF (2011b). Getting started: An overview of data publishing in the GBIF network, (contributed by Remsen, D., Ko, B., Chavan, V., Raymond, M.), Copenhagen: Global Biodiversity Information Facility, 16 pp. ISBN: 87-92020-28-3. Accessible at http://links.gbif.org/getting_started_publishing_en_v1

GBIF (2011c). GBIF Spreadsheet templates: User Guide, version 1.0. (contributed by Remsen, D., Doring, M., Robertson, T.), Copenhagen: Global Biodiversity Information Facility, 20 pp., ISBN: 87-920-27-5. Accessible at <http://links.gbif.org/dwca-spreadsheet-processor-guide>

GBIF (2011d). Publishing Species Checklists: Best Practices, version 1.0 (contributed by Remsen, D., Doring, M., Robertson, T.), Copenhagen: Global Biodiversity Information Facility, 20 pp., ISBN: 87-92020-26-7. Accessible at <http://www.gbif.org/orc/?docid=1229&l=en>

GBIF (2011e). Darwin Core Archives: How to Guide, version 1.0. (contributed by Remsen, D., Braak, K., Doring, M., Robertson, T.). Copenhagen: Global Biodiversity Information Facility, 21 pp. Accessible at http://links.gbif.org/gbif_dwca_how_to_guide_v1

GBIF (2011f). GBIF Metadata Profile: Reference Guide. (contributed by O Tuama, E., Braak, K.), Copenhagen: Global Biodiversity Information Facility, 19 pp. Accessible at http://links.gbif.org/gbif_metadata_profile_how_to_en_v1

GBIF (2011g). Publishing and registering data with GBIF, version 1.0. Copenhagen: Global Biodiversity Information Facility, 8 pp. ISBN: 87-92020-29-1. Accessible at

http://links.gbif.org/dwc-a_publishing_guide_en_v1

IAIA (2005). Biodiversity in impact assessment: IAIA Special Publication Series no. 3, July 2005, accessible at <http://www.iaia.org/publicdocuments/special-publications/SP3.pdf>

King et.al. (in press). Improving access to biodiversity for, and from, EIAs - a data publishing framework built to global standards. Accepted for publication *in Impact Assessment and Project Appraisal*.

MA (2005). *The Millenium Ecosystem Assessment, 2005: Ecosystems and Human Well-being - Synthesis*. Island press, Washington D.C.

Rajvanshi, A., Mathur, V., Iftikhar, U. A. (2007). Best-practice guidance for biodiversity inclusive impact assessment: a manual for practitioners and reviewers in South East Asia. CBBIA-IAIA Guidance Series, Capacity Building in Biodiversity and Impact Assessment (CBBIA) Project, International Association of Impact Assessment, North Dakota, USA.

United Nations (1992). Text of the Convention on Biological Diversity. Accessible at <http://www.cbd.int/doc/legal/cbd-en.pdf>.

Appendix 1: Glossary of Terms

(with special notes on Darwin Core and Darwin Core Archives)

Biodiversity: “the variability amongst living organisms from all sources including, *inter alia*, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems.” (CBD definition)

Citation (or data citation): a process in which a data publisher can be formally acknowledged and cited as the creator of the data.

Data publishing: a process through which biodiversity datasets are made freely and openly available in standardized formats, via an Internet access point that is indexed in the GBIF Registry.

Darwin Core: an internationally standardized set of terms for describing the identity and occurrence of organisms (also see Special Notes, below)

Darwin Core Archive: a standardized format in which data must be presented in order to publish it through the GBIF infrastructure (also see Special Notes, below)

Ecosystem: a collection of living organisms, the interactions between them and with their physical environment.

Ecosystem services: the benefits that people obtain from ecosystems.

Fitness for use (use when describing data): the suitability, effectiveness or usefulness of GBIF-mediated data in delivering accurate, authenticated, replicable and scientifically valid data for analysis and forecasting in conservation and management of natural resources.

Local government or local authority: an administrative unit of government responsible for an area that is smaller than a state or province

Metadata: information (data) about a dataset

Primary biodiversity data: digital text or multimedia data records of the occurrence of organisms

Special Notes on Darwin Core, Darwin Core Archives and the Darwin Core Archive Assistant

The Darwin Core body of standards

Darwin Core (DwC) is a body of standards that provides a stable set of terms for sharing information about biological diversity (Remsen, Braak, Döring and Robertson, 2011). It is used primarily to describe taxa and their occurrence in nature, as documented by specimens in natural history collections (physical or digital), observations and other related information (<http://rs.tdwg.org/dwc/>). The Darwin Core glossary includes a standardized set of terms, each accompanied by definitions and examples, and guidelines for use of the terms. The terms defined in the Darwin Core facilitate data sharing by providing a well-defined, standard vocabulary in a flexible framework that minimizes barriers to use and maximizes re-usability (Wieczorek, Doring, de Giovanni, Robertson and Vegliias, 2009). Although Darwin Core was conceived to facilitate the discovery, retrieval and integration of information about modern biological specimens and their occurrence in space and time, its application today is somewhat broader, and it can be used to build enriched data exchange formats (Remsen *et al.*, 2011). At this stage, it is considered that the Darwin Core glossary is suitable for describing primary biodiversity data from local government data-generating activities, but could be extended in future, if the need arises.

Darwin Core Archive

Darwin Core Archive (DwC-A) is an internationally recognized, biodiversity informatics standard that is the preferred format for publishing primary biodiversity data to the GBIF network. It makes use of Darwin Core terms to produce a single, self-contained dataset for sharing species-level (taxonomic) and species-occurrence data (Remsen *et al.*, 2011). The Darwin Core Archive consists of a series of one or more plain text files written in a particular format, including a compulsory **core data file** and a number of optional **extension files**. The core data file includes occurrence or taxonomic data about species (e.g. a list of species of mammals from a particular location), whereas the extension files include additional information about the core data file (such as the common names of the animals, habitat information, and so on). Each file is accompanied by a descriptor (or *metafile*) explaining how the files are organized. The core data file, extension files, descriptor (metafile) and metadata file (file describing the data) are all zipped together to make up the Darwin Core Archive for a particular dataset.

Appendix 2: Acronyms used in this publication

CBD: Convention on Biological Diversity

COP: Conference of Parties

DanBIF: Danish Biodiversity Information Facility

DwC: Darwin Core

DwC-A: Darwin Core Archive

DHC: Data Hosting Centre

GBIF: Global Biodiversity Information Facility

EIA: Environmental Impact Assessment

EMP: Environmental Management Plan

EWT: Endangered Wildlife Trust

IA: Impact Assessment

IAIA: International Association for Impact Assessment

ICLEI: Local Governments for Sustainability

IPT: Integrated Publishing Toolkit

LAB: Local Action for Biodiversity

LBSAP: Local Biodiversity Strategy and Action Plan

MDG: Millennium Development Goals

NBSAP: National Biodiversity Strategy and Action Plan

UNEP: United Nations Environment Programme

Appendix 3: Useful web addresses and References

Key References:

Getting Started:

1. Getting started: overview of data publishing in the GBIF network - http://links.gbif.org/getting_started_publishing_en_v1

Capturing data (Spreadsheets):

1. GBIF Spreadsheet templates: User Guide - <http://links.gbif.org/dwca-spreadsheet-processor-guide>
2. GBIF (2011c). GBIF Spreadsheet Templates: User Guide. <http://links.gbif.org/dwca-spreadsheet-processor-guide>
3. GBIF (2011d). Publishing species checklists: Best Practices. http://links.gbif.org/checklist_best_practices
4. Metadata Template: <http://tools.gbif.org/spreadsheet-processor/templates/metadata/metadata-1 v1.xlsx>
5. Species Occurrence Template: <http://tools.gbif.org/spreadsheet-processor/templates/occurrence/occurrence-1 V1.xlsx>
6. Checklist Template: <http://tools.gbif.org/spreadsheet-processor/templates/checklist/checklist-3 v1.xlsx>

Metadata:

1. GBIF Metadata Profile: Reference Guide - http://links.gbif.org/gbif_metadata_profile_how-to_en_v1
2. GBIF Metadata Profile: How-to-Guide - http://links.gbif.org/gbif_metadata_profile_how-to_en_v1

Checklists:

1. Publishing Species Checklist: Best Practices - http://links.gbif.org/checklist_best_practices
2. Publishing Species Checklists: Step-by-Step Guide - http://links.gbif.org/checklist_how_to

Darwin Core:

1. Create your own Darwin Core Archive: Step-by-Step Guide - http://links.gbif.org/dwca_manual_create_en_v1
2. Darwin Core Archive Assistant: User Guide - http://links.gbif.org/gbif_dwc-a_guide_en_v1.1
3. Darwin Core Archive Format: Reference Guide to the XML Descriptor File - http://links.gbif.org/gbif_dwc-a_metafile_en_v1/
4. Darwin Core Quick Reference Guide - http://links.gbif.org/gbif_dwc-a_guide_en_v1.1
5. Darwin Core Archive: How-to-Guide - http://links.gbif.org/gbif_dwc-a_how_to_guide_en_v1

Registering/Discovering/Publishing data:

1. Publishing and Registering data with GBIF - http://links.gbif.org/dwc-a_publishing_guide_en_v1

Developing strategies and action plans for data discovery and publishing:

1. http://www.gbif.org/orc/?doc_id=2755

Publishing Data Papers:

1. <http://www.biomedcentral.com/content/pdf/1471-2105-12-S15-S2.pdf>
2. [http://www.pensoft.net/JFILES/Pensoft Data Publishing Policies and Guidelinespdf](http://www.pensoft.net/JFILES/Pensoft_Data_Publishing_Policies_and_Guidelinespdf)

General websites:

CBD: www.cbd.int

GBIF: www.gbif.org

ICLEI: www.iclei.org/biodiversity and www.iclei.org/lab

UN-HABITAT: www.unhabitat.org