

Supplementary material 2 of  
“The key to bringing DNA collections to the next level”  
(Veltjen et al., 2024):

Interactive format of the Key,  
including guidance documentation

[Emily Veltjen](#)<sup>1\*</sup>, [Pieter Asselman](#)<sup>2</sup>, [Wim Baert](#)<sup>3</sup>, [Steve Baeyen](#)<sup>5</sup>, [Lise Beirinckx](#)<sup>4</sup>, [Liselot Breyne](#)<sup>5</sup>,  
[Dimitri Brosens](#)<sup>1,13</sup>, [Tim Claerhout](#)<sup>2</sup>, [Sari Cogneau](#)<sup>6</sup>, [Karen Cox](#)<sup>1</sup>, [Laura Cuypers](#)<sup>7,8</sup>, [Lynn Delgat](#)<sup>9</sup>,  
Philippe Desmeth<sup>10</sup>, [Jordi de Raad](#)<sup>11</sup>, [Lore Esselens](#)<sup>12</sup>, [Maria-Rose Eves Down](#)<sup>1</sup>, [Philippe](#)  
[Helsen](#)<sup>11</sup>, [Frederik Leliaert](#)<sup>3</sup>, [Kenny Meganck](#)<sup>12</sup>, [Zjef Pereboom](#)<sup>11</sup>, [Nathalie Smitz](#)<sup>12</sup>, [Gontran](#)  
[Sonet](#)<sup>7</sup>, [Maarten Trekels](#)<sup>3</sup>, [An Vanden Broeck](#)<sup>1</sup>, [Charlotte Van Driessche](#)<sup>1,2</sup>, [Aaike De Wever](#)<sup>1</sup>

<sup>1</sup> Research Institute for Nature and Forest, Brussels, Belgium

<sup>2</sup> Ghent University, Ghent, Belgium

<sup>3</sup> Meise Botanic Garden, Meise, Belgium

<sup>4</sup> Vrije Universiteit Brussel, Brussels, Belgium

<sup>5</sup> Flanders Research Institute for Agriculture, Fisheries and Food (ILVO), Ostend/Merelbeke, Belgium

<sup>6</sup> Belgian Coordinated Collections of Microorganisms / Institute of Tropical Medicine - Mycobacteriology Unit, Antwerp, Belgium

<sup>7</sup> Royal Belgian Institute of Natural Sciences, Brussels, Belgium

<sup>8</sup> University of Antwerp, Antwerp, Belgium

<sup>9</sup> Flanders Marine Institute, Ostend, Belgium

<sup>10</sup> Belgian Coordinated Collections of Microorganisms / Belgian Science Policy Office, Brussels, Belgium

<sup>11</sup> Royal Zoological Society of Antwerp, Antwerp, Belgium

<sup>12</sup> Royal Museum for Central Africa, Tervuren, Belgium

<sup>13</sup> Belgian Biodiversity Platform

\* Corresponding author: [emily.veltjen@inbo.be](mailto:emily.veltjen@inbo.be)

Start situation: There is a DNA collection in your organisation that you would like to mature. (Tip: If you are looking into starting a new DNA collection, read: [Harati et al. 2018](#) and/or [ISBER, 2023](#)).

1. Do you have an up-to-date overview of all direct, internal **stakeholders** of the institute's DNA collection and are you involving them in the (current) intent to "bring the DNA collection to the next level"?
  - Yes: continue to 2.
  - No or I need help: go to "[1. Involving internal stakeholders](#)".
2. Is preserving a DNA collection **within the scope of the institute** and is the DNA collection officially recognized within the institute?
  - Yes: continue to 3.
  - No or I need help: go to "[2. Having the DNA collection within the institute's vision](#)".
3. Do you have, on paper, a clear description of the **scope** of the DNA collection?
  - Yes: continue to 4.
  - No or I need help: go to "[3. Outlining a scope](#)".
4. Have you outlined the **current overarching workflow** of the DNA collection?
  - Yes: continue to 5.
  - No or I need help: go to "[4. Outlining the current workflow](#)".
5. Go to [Table 1](#), establish and log the **starting level**. Have you been able to establish the starting level and is the assessment properly logged?
  - Yes: continue to 6.
  - No or I need help: go to "[5. Establishing the starting level](#)".
6. **Level up**, one level at a time, and log the process. Have you reached all of the goals in level 3?
  - Yes: continue to 7.
  - No or I need help: go to "[6. Levelling up](#)".
7. Do you have a **re-evaluation strategy** for the DNA collection?
  - Yes: Perfect, all done ... for now!
  - No or I need help: go to "[7. Re-using the Key](#)".

**Table 1: DNA collection maturation chart.** A DNA collection is considered to start at level 0 in all categories (rows). If the DNA collection meets all the goals within a level (column), it achieves that level (e.g. a collection conforming to all goals within 'level 1' would be a level 1 collection). Making progress (e.g. reaching one specific goal, or reaching a complete new level) in this chart is considered 'maturing' the collection, with a fully matured DNA collection being 'level 3'.

Categories	Level 1	Level 2	Level 3
<a href="#"><u>Involvement of suppliers</u></a>	<a href="#"><u>Overarching workflow and expectations + communication strategy.</u></a>	<a href="#"><u>DMPs (+ versioning) as a tool for genetic research is standard practice.</u></a>	<a href="#"><u>A project conclusion protocol is in place.</u></a>
<a href="#"><u>Quality management</u></a>	<a href="#"><u>Quality meetings happen at fixed time intervals.</u></a>	<a href="#"><u>A written documentation policy is in place and executed.</u></a>	<a href="#"><u>Quality management is effective, up-to-date and FAIR.</u></a>
<a href="#"><u>Legal compliance</u></a>	<a href="#"><u>Informing on regulatory frameworks + central workspace for logging.</u></a>	<a href="#"><u>Five legal checkpoints are established.</u></a>	<a href="#"><u>The complete DNA collection is legally held.</u></a>
<a href="#"><u>Physical storage</u></a>	<a href="#"><u>A designated current space and a best-fit plan for the near future.</u></a>	<a href="#"><u>Best-fit storage method is functional for new DNA specimens.</u></a>	<a href="#"><u>The complete DNA collection is stored in the best-fit storage method.</u></a>
<a href="#"><u>Contingency planning</u></a>	<a href="#"><u>A basic contingency plan for the physical collection and data.</u></a>	<a href="#"><u>A final contingency plan for the physical collection and data.</u></a>	<a href="#"><u>The contingency plan is revised at agreed time intervals.</u></a>
<a href="#"><u>Identifiers</u></a>	<a href="#"><u>There is a unique ID-system for each physical sample and storage location.</u></a>	<a href="#"><u>There is a unique ID-system for each digital record + stable links.</u></a>	<a href="#"><u>Unique and persistent identifier-system within the global collection.</u></a>
<a href="#"><u>Digital management</u></a>	<a href="#"><u>Centralise and standardise data of the DNA collection + DMP.</u></a>	<a href="#"><u>Convert unstructured data to structured data.</u></a>	<a href="#"><u>Durable linkage to files and other data(bases).</u></a>
<a href="#"><u>Unlocking the collection</u></a>	<a href="#"><u>The DNA collection is on GrSciColl and the institute's website.</u></a>	<a href="#"><u>A test dataset in a repository + a publication strategy.</u></a>	<a href="#"><u>The complete DNA collection is in a public repository.</u></a>
<a href="#"><u>Loans</u></a>	<a href="#"><u>Responsible person appointed, public statement on how-to-loan and loan agreement template.</u></a>	<a href="#"><u>A FAIR loan policy + incoming and outgoing loan procedures.</u></a>	<a href="#"><u>Loaning policy and procedures are operational + a Loan Agreement breach protocol.</u></a>
<a href="#"><u>Stability</u></a>	<a href="#"><u>Financial responsibilities are outlined.</u></a>	<a href="#"><u>Optimised staffing.</u></a>	<a href="#"><u>The DNA collection has a stable budget and team.</u></a>
<a href="#"><u>Community</u></a>	<a href="#"><u>Become part of at least one community.</u></a>	<a href="#"><u>Active participation in co-creation.</u></a>	<a href="#"><u>Engagement in maintaining a community.</u></a>

Supplementary material 2 of "The key to bringing DNA collections to the next level (Veltjen et al., 2024)":  
Interactive format of the Key, including the guidance documentation

# 1. Involving internal stakeholders

**Question:** Do you have an up-to-date overview of all direct, internal stakeholders of the institute's DNA collection and are you involving them in the (current) intent of “bringing the DNA collection to the next level”?

**Explanation:** A stakeholder is a person or organisation with an interest or concern in the DNA collection. Internal stakeholders are those affiliated to the institute that is hosting the DNA collection, external stakeholders are all stakeholders not affiliated to the institute that is hosting the DNA collection. Stakeholders could be any of the following:

- Individuals that can help you mature your DNA collection (... or obstruct...). For example: the ICT team, technicians, researchers, the principal investigator or professor, the management team, ...
- Individuals that you have to convince of the collection's importance. For example: the management team, funding agencies, the general public, ...
- Individuals that are or will be affected by the changes to the collection and will have to comply. For example: students, researchers, external scientists that make a request to use specimens from the collection, ...
- Individuals that have the power to cancel your collection. For example: the management team, the government, the professor, the director, ...
- Organisations. For example: the institute the DNA collection is part of, the legal representative of the DNA collection, the DiSSCo research infrastructure, ...

**Importance:**

- The more support for the decisions being made, the higher the chance that your changes will have an effect and will be followed.
- DNA collections have a large “supplier-community” that can be diverse in scope, for example: different taxonomic or geographic scopes; different organisational units such as research groups etc. The collections are downstream of many actions and people, yet require all of these stakeholders to adhere to quality standards.
- DNA collections are typically linked to a genetic lab and its workflow.
- DNA collections are sometimes off-radar for either management or researchers, not considered as collections with long-term needs and stability.

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific Key-question: “Involving the stakeholders”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions by contacting the corresponding author.

### DiSSCo Flanders DNA working group experience with “Involving internal stakeholders” (2024)

Direct internal stakeholders of the DNA collections of the DiSSCo Flanders DNA working group vary depending on the collections’ context. Most of the participating DNA collections consist of one (or several) research unit(s) and one genetic lab, that create the majority of the DNA specimens.

The genetic lab is sometimes part of a larger infrastructure, for example for the DNA collection of the Research Group ‘Systematic and Evolutionary Botany’ at [Ghent University](#) the DNA collection is embedded in the CeMoFe (Centre for Molecular Phylogeny and Evolution) lab, making three other research units direct stakeholders as well. This mainly because decisions on physical management of the DNA collection affect the shared lab space and organisation.

The research unit that is directly involved in the management of the DNA collection usually has a mix of different profiles: lab technicians, field technicians, researchers on short term projects (and contracts) and researchers with a permanent position.

Next to a research unit, the DNA collections of the DiSSCo Flanders DNA working group often have direct stakeholders that are involved in centralised collection and/or data management. For example at the [Royal Belgian Institute of Natural Sciences](#) and the [Royal Museum for Central Africa](#), the collection services, who take care of many collections other than the DNA collection, are considered direct stakeholders. Another example is that of the DNA collection of the [Research Institute for Nature and Forest](#) where centralization of policy, procedures and digital infrastructure considering collection management is currently an ongoing effort of a supporting data team linked to the DiSSCo Flanders project.

As this is the first version of the tool, the DNA collections can now formally announce the planned usage of this tool to the internal stakeholders and discuss the involvement of the stakeholder in the intent of “bringing the DNA collection to the next level”.

**Return to the [Key](#).**

## 2. Having the DNA collection within the institute's vision

**Question:** Is preserving a DNA collection within the scope of the institute and is the DNA collection officially recognized within the institute?

**Explanation:** An alternative to be considered is to work together with an “external DNA collection” for storage (synonym: archiving) of DNA specimens generated by your institute.

**Importance:**

- If the management is on board with housing and curating a DNA collection and can follow the progress of maturation (e.g. by usage of this tool), it is expected that you'll be able to level up more easily. By actively involving them (e.g. communicating the needs and plans), resources (time, money, people) can be more easily allocated to collection management.
- DNA collections are sometimes off-radar within an institute.
- DNA collections sometimes organically develop as a by-product of research, without formal conceptualization.
- After about 20-30 years of genetic research, DNA collections that started small and were “in-check” can reach a point where they are in need of more structured organisation.

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific Key-question: “Having the DNA collection within the institute's vision”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions by contacting the corresponding author.

*DiSSCo Flanders DNA working group experience with “Having the DNA collection within the institute's vision” (2024)*

Within the working group, collections ranged from a carefully planned biodiversity biobank collection (e.g. the [EZA biobank](#)) to spontaneously created collections recently designated as official collections, to (to this day) ongoing debate on whether DNA collections are within the scope of the linked molecular lab; and lastly, even a DNA collection for which it was decided that the DNA collection does not fall in scope of the institute's vision and therefore can no longer be hosted. In the working group, an executed strategy to put (DNA) collection management more on the radar of the institute's management, was outlining a vision, mission and target audience statement for the (DNA) collection, sent for official approval by the management. Another simple, yet effective strategy that occurred in the working group, was dialogue of the genetic laboratory staff with the staff focused on collection management (if this type of staff is present).

**Return to the [Key](#).**

### 3. Outlining a scope

**Question:** Do you have, on paper, a clear description of the scope of the DNA collection?

**Explanation:** In an ideal situation, biobanks should establish the purpose of their collections in advance and organise them according to their expected use ([Corrales, Leliaert & Astrin, 2023](#)). However, having or managing a DNA collection is not always an actively or consciously made decision; it can also come about as a by-product of a genetic lab and/or research project, combined with the “gut feeling” of the people involved that the DNA specimens are too valuable to dispose of. In cases when a DNA collection was not carefully planned, this step has to be undertaken *ad hoc*.

Describing (the scope of) a collection can be as brief or detailed as necessary: in the format of a few sentences or bullet points which are only available internally, such as in guidelines or on an intranet page; on an institute’s web page; or even structured using Latimer Core terminology ([Woodburn et al. 2022](#), [Norton et al., continuously updated](#)) in a database or document. As long as the managers of the collection have a defined concept of the collection, this checkpoint is fulfilled.

Some inspiration about what can be stated in the scope: the collection purpose (why are you keeping the DNA collection?), the content (which biodiversity, from where, which type of specimens: DNA only?, which type(s) of DNA specimens: high molecular weight, eDNA (environmental DNA), aDNA (ancient DNA)?); the context (which organisational unit: institute, consortia, .. any policies or standards that are (to be) taken into account?); the future perspectives: how long are you planning to maintain the collection?

It is of significant added value if all of your stakeholders agree on the scope of the DNA collection, especially regarding the last question: ‘how long are you planning to maintain the collection?’. This is an excellent question to engage the management of your institute with.

**Importance:**

- To reflect on the purpose and importance of the DNA collection. This is relevant because: 1) to reuse a DNA specimen, you always need to execute destructive sampling and a specimen can become consumed. For example, in [Corrales, Leliaert & Astrin, 2023](#), there is differentiation between a core collection (no loans allowed) and a working collection (loans allowed); 2) DNA collections are derived from tissues, environmental samples or organisms (the “source material”): is the scope to keep all specimens where possible? Or do we decide to archive one representative specimen for the occurrence only?
- To identify similar collections and check for specialisation and/or collaboration with other collections.
- To evaluate if the specimen to be archived in your collection is within scope

- To evaluate if loan requests are within scope: you could decline a loan request, for example, if the scope of your DNA collection is: usage for non-commercial purposes only. Genetic material can be used in applied research or even commercial contexts and is regulated by Access and Benefit Sharing regulations (e.g. Nagoya protocol).

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific Key-question: “Outlining a scope”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions by contacting the corresponding author.

*DiSSCo Flanders DNA working group experience with “Outlining a scope” (2024)*

Within the working group, there was variation in scope and how the scope was outlined. There were collections that had a public statement on their scope (e.g. the [EZA biobank](#)), an internal document outlining the collection (e.g. the [Research Institute for Nature and Forest](#) compiled a draft “Strategic Collection Plan”, inspired by [Huxley et al. 2020](#) and [Luger et al. 2021](#), where the manager(s) of each subcollection, including the DNA collection, completed a standard questionnaire to position the collection in the wider collection landscape, to describe its importance, etc.), or, an, up until now undocumented scope. Most often the scope follows the history of the related research unit.

**Return to the [Key](#).**



## 4. Outlining the current workflow

**Question:** Have you outlined the current overarching workflow applicable for the DNA collection?

**Explanation:** Outlining a workflow can take a multitude of forms. It can be a simple one-pager of bullet points somewhere in the lab guidelines, a flowchart, a manually drawn sketch, a handbook, ... An outlined workflow allows all stakeholders to see the full process that influences the DNA collection.

**Importance:**

- DNA collections are often the end result of a complex workflow where many actions take place and multiple actors have responsibilities, all leading to a qualitative DNA specimen. There are many places where data can be logged or linked and hence possibly also (partly) get lost. DNA collection managers often have little control over upstream processes and events.
- By having a concept of how the collection is currently organised, you outline who has which responsibility at which stage of the workflow and how specimens and data are currently entering the collection.
- By having an overview of how the current organisation works, one can start analysing where it can be improved or even where it could be interesting to start adapting or changing the workflow. For example, if it is noted that there is too much variation on how, when and which information is communicated from the researcher that collected samples in the field to the DNA collection curators, it could be decided to invest in standardised workflows and writing out clear responsibilities for each actor. Furthermore, the process of describing these rules and the rationale behind them can be illuminating in itself (i.e. for the writers) and help to generate insight into possible future improvements to the workflow.
- An overarching workflow can also help you communicate more clearly why there is a need for certain rules when working together. This is useful both to engage new stakeholders (e.g. new students, new researchers), and also to explain to already existing stakeholders why rules are being initiated or changing as the DNA collection matures (e.g. by using this tool).

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific Key-question: “Outlining the current workflow”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions by contacting the corresponding author.

### DiSSCo Flanders DNA working group experience with “Outlining the current workflow” (2024)

In the working group, different genetic laboratories have written some form of a (draft) quality handbook or lab guidelines. Examples of lab workflows are:

- The Research Group ‘Systematic and Evolutionary Botany’ ([Ghent University](#)) has written a “General Guidelines and workflow v1.0 (2022)”, an internal document that gives an overview of what is expected when working with the genetic lab and DNA samples. It structures projects in different phases: the start of the project, sampling: before, sampling: during, sampling: after, molecular work and the end of your project. For each section, there is an introduction and a to-do list. After the written steps of the workflow, the outline is visualised in a workflow diagram, and lastly all Ghent University legal documentation and support is listed. At the end of the document, the reader confirms that he/she read the guidelines, including a statement that if the criteria are not met, the DNA sample cannot go into the collection. The writing of the document was inspired by the working group, and facilitated by the DiSSCo Flanders project. Some members of the working group also provided feedback on the document, while reading it for inspiration.
- The Laboratory for Molecular Systematics of the [Royal Belgian Institute of Natural Sciences](#) and the [Royal Museum for Central Africa](#) have developed an internal document: “Guidelines for sample and data management version 1.0 (May 2023)”, adapted from the document shared with the DNA working group by the Research Group Systematic and Evolutionary Botany. The document starts with a general overview of the workflow including a diagram, followed by a more detailed description of the workflow including the concept of “to-do’s” cfr. the guidelines of the UGent experience. The guidelines refer to a zipped file of empty folders and templates to work in, to collect data from different research projects in a standardised way.

Alternatively, there can also be project (type) workflows leading to DNA specimens in collection. For example:

- At the [Research Institute for Nature and Forest](#), recently an initiative started to streamline the eDNA data management, as different researchers have different practices. Hence for the eDNA projects (i.e. a project type), an overarching workflow is being outlined. Here it is being decided where to take snapshots of the data either for internal archivation or external data publication, and in the meanwhile of doing so, agreements are made on naming of files etc. The work is inspired by [Borisenko, Young & Hanner, 2024](#).
- The EMO BON project (European Marine Omics Biodiversity Observation Network) of which the [Flanders Marine Institute](#) is a partner has outlined its workflow in a comprehensive handbook ([Santi et al., 2021](#)).

Lastly, an interesting approach in the working group was that of the genomics core at the [Flanders Research Institute for Agriculture, Fisheries and Food](#). Here they made a template data model, to be filled in by each individual associated with genomic research for his/her

project. The overview has checkboxes and allows researchers to leave behind comments at each step. On the one hand, the exercise was aimed to map out the different responsibilities and actors, which will allow to identify in which parts of the workflow a more robust data framework should be developed. On the other hand the exercise was aimed to raise awareness on the importance of good documentation as every person is only part of a larger overall workflow.

**Return to the [Key](#).**

## 5. Establishing the starting level

**Question:** Go to [Table 1](#), establish and log the starting level. Have you been able to establish the starting level and is the assessment properly logged?

**Explanation:**

- “Establishing the starting level”: this means that you read the level-1-goal, level-2-goal and level-3-goal of each category. You note which level your collection is in per category. If none of the three goals apply to your current state of your DNA collection, the collection is at level 0 for that category. More information can be obtained on a goal by reading the aligning guidance documentation. A DNA collection therefore starts at level 0 and when all categories comply with the goal in the level 1 column, the DNA collection has reached level 1. The same applies for reaching levels 2 and 3.
- “Properly logging”: This means to actively log the date when the assessment was executed and who executed it. If relevant, it can be explained why certain goals are checked or unchecked in the case of the DNA collection at hand.

**Importance:**

- You first need to know where your collection is to see which categories you should invest in first (i.e., the lowest levels)
- A holistic maturation seems most durable and should be the goal; often, the less “fun” or less “obvious” aspects (such as legal aspects, disaster management, collection size projection to the future) are not completed, while more “sexy” aspects such as digitisation (e.g. with the current DiSSCo initiatives) move forward.

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific Key-question: “Establishing the starting level”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions by contacting the corresponding author.

We invite DNA collections to share their experiences. How was the process of mapping your DNA collection on [Table 1](#)? How many people were involved in the mapping? How long did it take? Was there discussion on any of the goals and why?

*DiSSCo Flanders DNA working group experience with “Establishing the starting level” (2024)*

For the next funding phase of DiSSCo Flanders (2025-2028) we plan to provide user-support-sessions (online) for using this tool as a service, as part of a first testing phase for the DiSSCo Flanders partners. This user-support will encompass the explanation of the tool to new users and support with mapping DNA collections on [Table 1](#).

**Return to the [Key](#).**

## 6. Levelling up

**Question:** Level up, one level at a time, and log the process. Have you reached all of the goals in level 3?

**Explanation:**

- The areas of improvement within each level can be tackled in parallel or one by one, as slow or as fast as need be. You are the architect of your collection maturation plan and you can decide what works best for your collection.
- Logging the process: which steps are reached when and how, who is responsible/who has contributed.

**Importance:**

- To communicate the process to your stakeholders.
- To break down the task of maturation into smaller, more easily achievable tasks.
- To set (and achieve) realistic goals.

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific Key-question: “Levelling up”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions by contacting the corresponding author. We invite DNA collections to share their experiences: Which goal was levelled up? How long did it take? How was the experience for the particular context of the DNA collection at hand?

*DiSSCo Flanders DNA working group experience with “Levelling up” (2024)*

For the next funding phase of DiSSCo Flanders (2025-2028) we plan to provide user-support-sessions (online) for using this tool as a service, in a first testing phase for the DiSSCo Flanders partners. This user-support will encompass: jointly plan goals, share experiences and exchange tips and tricks. Where relevant, experiences will be added to this tool after a user-support-session and included in future releases.

**Return to the [Key](#).**

## 7. Reusing the Key

**Question:** Do you have a re-evaluation strategy for the DNA collection?

**Explanation:** The re-evaluation strategy should state when (at which time interval) and how to re-evaluate the DNA collection, and who is responsible to start and carry out the re-evaluation procedure. Re-evaluation is bifold: you evaluate:

1. The DNA collection: does the collection still comply with all the level goals as it did previously?
2. This tool: is this tool still up-to-date with the current best practices of DNA collection management?

Once the collection has reached level 3, it is (theoretically) possible to drop back one or even two levels once the procedure is restarted.

**Importance:** On the one hand, it is to be expected that there can be changes in vision, techniques, facilities and context of a DNA collection that has reached level 3. On the other hand, it's important to reach a final stage at some point and decide that the collection is good enough to continue in its current state for a number of years, without feeling internal or external pressure to constantly revise workflows and storage. A compromise for both aspects is investment in maturation, followed by an agreed upon period of time of stasis. A good collection manager will do their best to keep (relatively) up to date with developments in the field. If there are any important changes in best practice which are relevant for your collection before the agreed-upon re-evaluation period, then of course a sooner re-evaluation process would be prudent.

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific Key-question: “Re-using the Key”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions by contacting the corresponding author. Level 3 DNA collections are invited to share their re-evaluation strategy and what happened at re-evaluation.

### *Re-using the Key: Hypothetical experience 1 (2024)*

Imagine that the institute's vision and, subsequently, the scope of the DNA collection change. For example, your genetic laboratory now wants to be accredited. This might set higher standards for storage methods (e.g. a colder storage method), to which you revise the best-fit plan in storage methods for the near future. This results in the collection dropping from level 3 to level 1 for the time being, as it no longer complies with the best-fit for your vision for the DNA collection.

*Re-using the Key: Hypothetical experience 2 (2024)*

Imagine that there is a new repository established to house all, in our case, Belgian DNA aliquots, as a national infrastructure service. It is possible that when restarting this tool, with this new information, you might decide not to manage the DNA collection in-house anymore and instead transfer the collection to the new facility.

**Return to the [Key](#).**

## 8. Involvement of suppliers

**Category:** Involvement of suppliers

**Level 1 goal:** Overarching workflow and expectations + communication strategy.

**Level 2 goal:** Data Management Plans (DMPs) (+ versioning) as a tool for genetic research is standard practice.

**Level 3 goal:** A collaboration termination protocol is in place.

**Explanation:**

Level 1: Have an easily understandable overview of the overall workflow to communicate to the suppliers of DNA specimens at the start of each new project that will generate DNA specimens. The document should be self-explanatory and complete. List the expectations clearly for collaboration with the genetic laboratory or DNA collection: what material, which templates, which responsibilities, etc. Also, clearly detail whose responsibility it is to communicate this overview, and when. This could also be less schematic but in the format of a policy, for example [Ståhls et al. 2021](#).

Level 2: Each new project that will generate DNA specimens (which are destined to enter the DNA collection), uses research DMPs (i.e. differing from operational DMPs, see [14. Digital Management](#)) with versioning throughout the project lifecycle, to monitor the material and data flow throughout. The research DMP has ideally been co-created and will be co-monitored with all stakeholders of the project throughout the DNA specimen lifecycle. This includes genetic laboratory technicians/managers, DNA collection curators/managers, (centralised) data managers or data stewards, ... . An interesting read on data management of biodiversity genomics is [Forsdick et al., 2023](#) and their provided Biodiversity Genomics Data Management Hub (<https://genomicsaotearoa.github.io/data-management-resources/>).

Level 3: Steps in level 1 and 2 should facilitate good, effective collaboration. However, for those cases where collaboration still proves difficult, there should be an additional protocol on “collaboration termination”. This protocol describes different scenarios, along with minimal actions to undertake to mitigate an unconcluded project and - as a last resort - the actions required to terminate the project if mitigation fails. This protocol can be versioned, using precedents to list relevant scenarios.

**Importance:** For the suppliers, there is no direct or immediate benefit in working towards qualitative archiving of their DNA specimen and neither are there immediate negative consequences for the lab technician or DNA collection curator for not managing the archiving process. The negative consequences only arise when someone requests access to DNA specimens to validate previous analyses or to re-use the specimens, causing the need for a specimen to be retrieved, or when there is insufficient space to archive new DNA specimens. Suppliers and managers need to co-invest in clearly stating their expectations to one another, as well as planning together to work towards the common goal of conducting good research



practice: qualitative archiving is a joint responsibility. All involved parties should be able to signal their expectations to each other. If a collaboration has gone awry and results in rejection of DNA samples to go into the collection (and potentially: disposal of DNA samples from the genetic lab), confirm this action and the rationale clearly in writing to the relevant parties.

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific Key-category: “Involvement of suppliers”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions.

#### *DiSSCo Flanders DNA working group experience with “Involvement of suppliers” (2024)*

Level 1: There are different guideline-documents present in the working group, targeted towards researchers and students. These documents go beyond the scope of the DNA collection and are generated to ensure a good collaboration with a genetic laboratory from the start, leading to qualitative DNA specimens in the linked DNA collection. See experiences in “[4. Outlining the current workflow](#)” for a more detailed description of institution-specific-experiences. Although not stated explicitly, the lab technician is responsible for sharing this information at the start of the collaboration with the genetic laboratory.

Level 2: In the next years, many of the institutes in the working group are expected to see an increase in usage of DMPs as the Flemish Open Science Board (FOSB) data stewards are focussing on raising awareness among researchers of the importance and added value of using DMPs, as well as supporting the process of their creation and implementation. Researchers still all too often regard a DMP as a purely administrative obligation, rather than as a beneficial tool to manage research data. Within the DiSSCo Flanders DNA collection working group, we follow up the progress of the FOSB and work closely together with data stewards. There are some (often larger) projects generating DNA specimens that are investing in having a DMP.

Level 3: Until now, we are not aware of any DNA collection in the DiSSCo Flanders DNA collection working group having this protocol explicitly written down.

#### *Involvement of suppliers: Hypothetical experience 1 (2024)*

Level 3 example scenario: A visiting scientist has generated DNA specimens in a genetic lab; in the DMP it is written that these are to be archived in the DNA collection. The visiting scientist was informed that he has to share data in the given folders and templates with a guidelines document that he signed. However, after the visiting scientist leaves, the agreed upon data is not shared with those that have to execute the archiving of the specimens. The visiting scientist does not answer emails asking for the data. A hypothetical “project conclusion protocol” for this example scenario could be as follows:

1. Communicate that there has been a breach of agreement and that the project conclusion protocol is being enforced, clearly stating the consequences and time frames, and add the relevant documentation. This communication is first sent to the visiting scientist.

2. If there is no response after one month: contact at least one listed colleague with the same information.
3. If there is no response after two months, specimens will be disposed of.
4. If there is a response, but no data after three months of the initiation of the “project conclusion protocol”: repeat communication.
5. If there is no data after four months of the initiation of the “project conclusion protocol”, dispose of DNA specimens.

**Return to [Table1](#).**

## 9. Quality management

**Category:** Quality management

**Level 1 goal:** Quality meetings happen at fixed time intervals.

**Level 2 goal:** A written documentation policy is in place and executed.

**Level 3 goal:** Quality management is effective, up-to-date and FAIR.

**Explanation:** “Quality management” in this Key refers to actively managing and overseeing the time invested towards qualitative operations of the DNA collection. Doing so encompasses a variety of important tasks: management of the creation or upkeep of any form of documentation such as policies, guidelines, Standard Operating Procedures (SOPs) and protocols; investing in making and documenting decisions sufficient planning; outlining strategy and managing dissemination of information to the relevant stakeholders.

Level 1: Have at least one yearly collection-quality meeting with one representative from each stakeholder group to plan which actions will be tackled in the next working year. Actively define and log who is responsible for the planned actions. Appoint (a) responsible person(s) = the quality manager to convene and structure this yearly meeting.

Level 2: Have a written (internal) documentation policy that clearly states how new documentation is generated, by whom, in which format (SOP, guidelines, policies, ...), where it is stored, who can access the documentation, when new documentation should be ready, how and when to improve the documentation (versioning), .... This documentation policy should be co-created with and approved by all stakeholders. The quality meetings (level 1) have a fixed checkpoint where actions between the current and the next quality meeting, regarding documentation, are listed.

Documentation that is anticipated to be created, based on this tool, and systematically revised:

- Documentation on the DNA collection ("[3. Outline a scope](#)")
- Documentation on decisions made ("[11. Physical storage](#)")
- Loan documentation ("[16. Loans](#)")
- Documentation that supports collaboration with depositors ("[8. Involvement of suppliers](#)")
- SOPs for checking legal compliance ("[10. Due diligence & legal compliance](#)")
- Contingency plans ("[12. Contingency planning](#)")
- SOPs on logging data in Digital Management System ("[14. Digital management](#)")
- Documentation on the usage of this tool (i.e. logs).

Level 3: Quality management is effective, up-to-date and FAIR.

- Effective quality management: there are enough checkpoints/meetings to manage the quality and agreed-upon quality standards are followed.

- Up-to-date: from the moment the documentation policy is introduced (level 2), all documents needed are created and kept up-to-date, thereafter following the current workflows and procedures.
- All documents are FAIR: Findable, Accessible, Interoperable and Reusable ([Wilkinson et al., 2016](#)). This means that the documentation itself can be easily found, accessed and used by the people that need the information. Furthermore, the documents are published FAIRly or openly, so that other DNA collections can (request to) reuse ideas and good practices to aid their own collection management. Documentation could be published on [Zenodo](#), for example.

### Importance:

- DNA collections can be managed in either an integrated organisational structure or a specialised one ([Huxley et al. 2020](#)). An integrated organisational structure means that the DNA collection is managed next to laboratory management and/or research activity. A potential pitfall of which can be that other responsibilities are considered to be more urgent or more rewarding in the short term and consequently quality management of the DNA collection is procrastinated. This can lead to little time being invested in DNA collection management and, more frequently, what time is invested is done so in an inconsistent manner. An advantage of an integrated organisational structure is that there is more involvement of the suppliers in the management of the collection: suppliers can see the consequences of their activities on the laboratory activity and/or DNA collection activity. When in a specialised organisational structure, meaning that the DNA collection (or all Natural Science collections) are a separate (internal) entity to the lab and research activity; quality management is expected to be an important category to invest in, and less prone to the pitfall of inconsistent and/or little devoted time. For DNA collections that are in a specialised organisational structure, the greater challenge is often that of communication and collaboration with suppliers ("[8. Involvement of suppliers](#)").
- Quality management is crucial for day-to-day operations: the independent history of each DNA specimen should be properly logged. In addition, documenting the process, for example by writing SOPs, promotes data comparability between different methods and aspects of the biobank workflow ([Corrales, Luciano, Astrin, 2023](#)).

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific Key-category: “Quality Management”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions.

### *DiSSCo Flanders DNA working group experience (2024) with “Quality Management”:*

During the working group meetings, most institutes that shared experiences did execute quality management of the DNA collections; however it mostly happens spontaneously (not systematically) by collection managers and direct stakeholders (e.g. researchers, laboratory technicians) if the inefficiency of a process became large enough and if enough resources allowed for revision of that specific inefficiency. Larger genetic labs typically need more

organisation and division of tasks compared to smaller genetic labs, hence, organically these labs have usually established a division of responsibilities and systems for more standardisation and hence quality control.

**Return to [Table1](#).**

## 10. Due diligence & legal compliance

**Category:** Due diligence & legal compliance

**Level 1 goal:** Informing on regulatory frameworks + central workspace for logging.

**Level 2 goal:** There are 5 legal checkpoints instated.

**Level 3 goal:** The complete DNA collection is legally obtained.

**Explanation:** For more in-depth information on the consequences of the Nagoya protocol on DNA collection management, [Corrales, Luciano and Astrin, 2023](#) (Discussion: Policies) is an interesting read.

Level 1: This goal has two actions: the responsibility to inform researchers and to provide a central (digital) workspace where relevant documentation can be stored. Both actions can be undertaken at the level of the institute or the DNA collection itself.

Inform the researchers of every new project associated with the institute's genetic laboratory / DNA collection (i.e. the suppliers of new DNA specimens) of the relevant regulatory frameworks to comply with (Nagoya, CITES, ...) when working with DNA specimens. "Informing" could be in the format of running through an "intake-checklist" at project start-up, either in the general project workflow of the institute or in collaboration with the genetic laboratory. During the checklist process, the researcher is made aware that it is their responsibility to check for compliance with the regulatory frameworks and that illegal specimens will be neither processed by the genetic lab nor stored in the DNA collection. To keep tabs on the different regulatory frameworks, it could help to have a designated contact person in your organisation who follows the developments of a certain type of regulatory framework. These contact people can be actively mentioned on the proposed intake-checklist. Multiple regulatory frameworks can be followed by one person, however, this workload should, of course, be kept in balance with the other responsibilities of said person. Alongside actively informing the researchers at the start of projects which plan to generate DNA specimens, there is also passive informing. This can take many different forms, such as listing responsibilities and relevant contact persons on an internal web page of the institute ("who does what") or informing via internal newsletters. Additionally, researchers can be redirected to contact information and regulatory frameworks in documentation that support researchers at the project writing stage, so that as early as possible in the research process steps are undertaken to start any needed administration processes.

For (DNA) specimens, a shared folder should be provided in a central institutional workspace. In this shared folder, researchers log all relevant documentation that puts restrictions on the (usually a great number of) specimens that are planned to become part of the (DNA) collection. This could include:

- Scans of any types of permits.
- Emails/letters to local authorities, private persons, ...

- Contracts that specify any restrictions or agreements on the forthcoming specimens. This could be expectations from project funders, expectations from collaborators.

Level 2: There are 5 checkpoints instated: pre-fieldwork, pre-labwork, pre-deposit, post-project and collection-action. The checkpoints help ensure legal compliance. The process or activity of legal compliance is called: “due diligence”. Actively appoint (a) responsible person(s) for each checkpoint, preferably supporting this responsibility with a Standard Operating Procedure (SOP) explaining how to carry out the role. Each of the five checkpoints can have a different responsible person. It is recommended that the responsible person is objective and that the checking responsibility does not lie with the researcher (especially important for the checkpoint ‘post-project’). However, if the checkpoint is deemed to be evaluated as negative, the researcher is the responsible party to undertake action to resolve the compliance issue. At the end of the checkpoint, it has to be actively documented whether the checkpoint was evaluated as positive. The appointed responsible person has the authority to stop the research in case of non-compliance and to contact focal points outside the DNA collections organisation (e.g. Nagoya specialist of University, VLIR, ..) for assistance. In [de Mestier et al. 2023](#): “Managing legal information, rights and restrictions related to specimens and samples”, there is a procedure visualised in Figure 3.1 ‘Due diligence for Collections Items In or Entering the Organization’, that can be adapted for establishing a procedure / checkpoint.

The five different checkpoints have been adopted from the CETAF Code of Conduct on ABS Annex 5 ([CETAF Legislations and Regulations Group, 2019](#)). The institute should also individually analyse at which points in their workflow a checkpoint should best be created:

- 1) Pre-fieldwork: To have a fixed checkpoint prior to executing fieldwork. This would require the establishment of a standardised procedure that reports planned fieldwork to a responsible person to execute due diligence.
- 2) Pre-labwork: To have a fixed checkpoint prior to commencing genetic lab work that screens incoming material (e.g. tissue, organisms, environmental samples) for DNA to be extracted from. Only applicable if a genetic lab is linked to the DNA collection. This is especially important for DNA specimens not collected in-house (i.e. those that already passed through checkpoint 1).
- 3) Pre-deposit: To have a fixed checkpoint prior to accepting a deposit into the DNA collection. This is especially important for DNA specimens not created in-house (i.e. those that already passed through checkpoint 1 and 2).
- 4) Post-project: To have a fixed checkpoint at the end of a research project. This could be linked to publishing an article or dataset. Here, it should be checked if the agreed terms of the contracts were all met and that all that was promised, is delivered (i.e. all contractual obligations are fulfilled).
- 5) Collection-action: This checkpoint refers to actions such as specimen disposal, transfer or loan. Loans can include the decision whether to consume (i.e. completely use it, until there is no specimen left) a DNA specimen or not. It is crucial to carefully check if a proposed action complies with the restrictions linked to each individual specimen.

Level 3: The complete DNA collection is legally held. This means that all DNA specimens in the DNA collection have been thoroughly checked for legal compliance. It is actively logged when the DNA specimens were checked and by whom. This goal comes with the following disclaimer: “the collection is legally held, to the best of the knowledge of the DNA collection managers”. It is, however, always possible that information is missing, and so it is impossible to state this with 100% certainty.

**Importance:** Compliance to both hard and soft law is important for the institute hosting the DNA collection, as it is the institute which is legally accountable for non-compliance. It is the shared responsibility of all people involved in the DNA collection to ensure that compliance is met. Rules and regulations are established to protect biodiversity, ecosystems, people, property, integrity, ... and make expectations transparent. Although not always perceived as such, they should facilitate good research practice, not hamper scientific research.

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific Key-category: “Due diligence & legal compliance”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions.

*DiSSCo Flanders DNA working group experience (2024) with “Due diligence & legal compliance”.*

Most institutes in the working group have invested in informing their employees about relevant rules/regulations, for example about Nagoya, in the form of info sessions. There is also passive information sharing, such as web pages with information and recommended literature and the listing of experts who can be contacted. For example, in Flanders, <https://nagoya.vlir.be/contacts-nl/> lists who the Nagoya contact person is for each Flemish University. In addition, some of the working group members included a section on legal information in the (recently developed) “lab guidelines” that are actively communicated at the start of projects that will generate new DNA specimens. None of the institutes / DNA collections in the working group have, at this point, fixed legal checkpoints, largely due to the question of who has to take this (extra) responsibility on? At this point, the responsibility of legal compliance is often left to the researchers themselves when they are collecting specimens or exchanging material. Most DNA collection managers in the working group do not know if there are illegally obtained DNA specimens in their collection, and if so, how many DNA specimens still need to be legalised. The difficulty of doing a due diligence check on DNA specimens already in the collection lies in factors such as the common issue of limited resources or having someone to do take the responsibility, combined with the specificity of each DNA specimen (different biodiversity, time, place), the sometimes scattered information to puzzle together, as well as the potential repercussions that could arise when “outing” DNA specimens which still need to be made legally compliant.

**Return to [Table1](#).**



## 11. Physical storage

**Category:** Physical storage

**Level 1 goal:** A designated current space and a best-fit plan for the near future.

**Level 2 goal:** Best-fit storage method functional for new DNA specimens.

**Level 3 goal:** The complete DNA collection is stored in the best-fit storage method.

**Explanation:**

Level 1: Achieving level one has two conditions: having a designated space for the DNA collection and a best-fit plan for physical storage for the near future.

The first action to undertake is to decide on the designated drawer, cabinet(s) or room to store the DNA collection. This designated storage space should preferably be physically separated from the research collections, isolated as much as possible from other activities. Actions should be undertaken to control accidental access to the collection, by clear labelling and/or by working with locks or a badge system.

In terms of the volume of the storage space: calculate or estimate the following information with information at disposal:

- The number of DNA specimens in the DNA collection
- The average yearly growth in the number of DNA specimens in the DNA collection
- The number of DNA specimens that the currently designated storage space (options) can hold.

Based on this information, estimate how many years the designated storage space (options) will likely be able to be of service to store the DNA collection.

Regarding isolation: the more truly physically separated from the genetic laboratory work or other similar collections, the better. This to control access to the collection: when close to the genetic laboratory or in a shared freezer (with another collection for example), there is a higher possibility that: 1) DNA specimens are moved away from their logged position (e.g. while searching for another DNA specimen), 2) other material finds its way into the DNA collection storage units (e.g. when thawing another freezer for maintenance), 3) accidental freeze-thaw cycles can occur when opening the DNA collection freezers for purposes other than the reuse of a DNA specimen. If moving of DNA specimens to a more isolated storage option (e.g. a new designated room or freezer) is required, this should be integrated in the best-fit plan for the near future (see next paragraph, level 2 and level 3 actions).

With this information, the internal stakeholders ([1. Involving internal stakeholders](#)) should be included in deciding on the storage space for the DNA collection, making sure to consider the context of the DNA collection ([scope](#) and resources): should the DNA collection be moved to another storage space? Should the current storage space be expanded? Does the current storage space suffice for the time being?

The second action to undertake is to actively evaluate the storage containers and materials within the designated storage space. Similarly, it is recommended to carry this evaluation out together with the [internal stakeholders](#), taking the context of the collection ([scope](#) and resources) into account. Different aspects regarding physical storage should be considered:

Storage parameters	Options
Type of preservation/ main container	Room temperature (see <a href="#">Corrales et al., 2023a</a> : Chapter 8, subtitle: Storage for concrete examples) / -20°C / -70°C / -80°C / -196°C. An interesting initiative to look into: <a href="https://www.freezerchallenge.org/">https://www.freezerchallenge.org/</a>
Strategy of preservation	It is not necessary to work exclusively with one preservation choice. Combinations are possible: you can strategize. Example 1: you can store a more valuable part of the DNA collection in a colder main container. Example 2: you can decide to continue to keep the old DNA collection stored at -20°C, while investing in -80°C storage for newly generated DNA specimens.
Main container organisation	Option 1: Continuous. Every incoming DNA specimen is added to the next available storage space. Option 2: Division. This can be in taxonomic concepts, geographic concepts or organisational concepts. For example: you can decide to have one freezer for research group A and a second for research group B.
Storage nesting & types	Tertiary: racks, crates, drawers Secondary: boxes, bags Primary: vials, plates  Overall: compatible with the chosen preservation temperature. Cryogenic and water-resistant for cold storage.
Drawers organisation	Continuous usage / organised per year, per project, according to usage (working collection versus archive collection), ...
Box-vials organisation	Grouped per project/ incremental on ID/ a box per user (group) /...
Label material	Written/ thermal-printed/ self-laminating/ laser-etched barcodes. There is a decision tree made by the Canadian Museum of Nature on label material ( <a href="#">link</a> ).
Where to label?	Vial and cap / bottom of the vial / side of the vial.

When to label?	If not all generated DNA specimens of the associated genetic laboratory go into the DNA collection, it might be worthwhile to work with a cheaper (yet less persistent) labelling method first, and a more sustainable labelling method once the DNA specimen goes into the DNA collection.
----------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

It is recommended that all made decisions, both that of the storage space and that of the best-fit strategy for storage containers and materials, are outlined in a date-stamped, written report that is archived so that (new) colleagues can easily follow the past decision making and rationale, for example when re-using the Key (["7. Reusing the Key"](#)).

Level 2: Best-fit storage method is functional for new DNA specimens. If, after evaluation, the outcome is that changes in material or strategy are the best-fit for the DNA collection at hand, it should be planned to actively allocate means (time, people, money) that allow newly generated or deposited DNA specimens to be stored in the DNA collection accordingly. This means that the existing collection is kept in the former configuration, for now. Once the newly generated or deposited DNA specimens are successfully being stored according to the best-fit method (materials and strategy) level 2 is achieved.

Level 3: The full DNA collection is stored in the best-fit storage method. Means have been allocated (time, people, money) to transfer the specimens that were still stored in the old configuration to the new, best-fit storage method (if it was decided that this was necessary at the planning stage (level 1)).

**Importance:** Overall, DNA (specimen) quality depends on the preservation method used and the duration of storage ([Walters & Haner, 2006](#)). If kept in cold storage, repeated freeze-thaw cycles should be avoided because they will degrade the DNA ([Corrales & Astrin \(2023\): chapter 8, subtitle: Storage](#)).

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific Key-category: “Physical storage”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions.

#### *DiSSCo Flanders DNA working group experience (2024) with “Physical storage”:*

An interesting variation in storage methods was present in the working group. There was also great interest in learning more about room temperature storage compared to cold storage, as well as the impact of initiatives such as the [freezer challenge](#) on the quality of DNA specimens. For the larger part of the working group, cold storage is used, with (part of) the DNA collections stored at -20°C, -70°C and -80°C, in both plates or vials in boxes. Often, multiple preservation methods are used (or are planned to be used) simultaneously, based on the value of the DNA specimens. None of the institutes have DNA collections stored in liquid nitrogen. For some collections, the investment was made to use 2D laser-etched barcodes at the bottom of vials,

while others used freezer resistant printed labels and others decided that for the DNA collection handwritten labels suffice, if placed in a predictable order (e.g. incremental).

The laboratories for Molecular Systematics of the [Royal Belgian Institute of Natural Sciences](#) and the [Royal Museum for Central Africa](#), rely on GenTegra™ for long-term room temperature DNA storage. This technology enables DNA samples to be stored in separate tubes at room temperature in a water-free environment. The mineral matrix at the bottom of the tube forms a protective coat around the DNA, shielding it from hydrolysis, oxidation and microbial growth. The responsibility for drying and storing DNA samples rests with individual researchers or research groups. However, there is a centralised DNA bank where GenTegraDNA tubes can be safely stored in a dry and accessible space for users of the molecular laboratories. Moreover, the GenTegra™ tubes offer practical advantages beyond preservation. They are designed to be easily transportable and suitable for shipping, enabling convenient sharing of DNA samples between collaborating institutes and researchers. Importantly, the successful long-term storage of DNA using this technology has been demonstrated by the retrieval and use of samples dried several years prior. For instance, in 2022, whole genome sequencing was successfully performed on DNA samples that had been dried back in 2011, showcasing the efficacy and durability of GenTegra™ in preserving DNA integrity over extended periods.

**Return to [Table1](#).**

## 12. Contingency planning

**Category:** Contingency planning

**Level 1 goal:** A basic contingency plan for the physical collection and data.

**Level 2 goal:** A final contingency plan for the physical collection and data.

**Level 3 goal:** The contingency plan is revised at agreed time intervals.

**Explanation:** Contingency planning means risk mitigation and disaster planning. It is an exercise of thinking about what could go wrong (potential risks), and agreeing beforehand what actions to undertake if and when they do. Potential risks for a DNA collection can include failure of physical or digital infrastructure (e.g. freezer breakdown, server breakdown), unexpected changes in the DNA collection organisation (e.g. the DNA collection manager pursues a new career), or (natural) disasters (e.g. flooding, fire, quarantine due to disease outbreak, ..). The more likely a potential risk is, the stronger the need to have a plan in place. In [Corrales, Leliaert & Astrin, 2023 \(Chapter 1\)](#) two worthwhile sections to read are 'Collection backups' and 'Risk Management', as well as references mentioned therein.

Level 1: Have a written down (internal) contingency plan for the current physical storage and the current data storage. The contingency plan clearly states who has which responsibility. For the data storage the basic plan addresses data backups and versioning, which, if the data is centralised ("[14. Digital management, level 1](#)"), often aligns within the institutional IT and data management policies. For the physical collection the contingency plan provides solutions for emergencies (e.g. natural disasters, pandemic, war, ..) as well as operational problems (e.g. freezer malfunction, shortage of personnel, ..) for the current (separated, ("[11. Physical storage, level 1](#)") DNA collection. The management has approved the basic contingency plan and ensured that sufficient means (time, money, people) are allocated to ensure successful execution of the contingency plan if need be.

Level 2: Have a written down (internal) contingency plan for the decided best-fit physical storage ("[11. Physical storage, level 2](#)") and the chosen structured digital organisation ("[14. Digital management, level 2](#)") of the collection. It is clearly stated in the contingency plan who has which responsibility. If the chosen structured data solutions are commercial or hosted solutions, worst-case scenarios such as the end of the product or service must also be considered. The management has approved the basic contingency plan and ensured that sufficient means (time, money, people) are allocated to ensure successful execution of the contingency plan if need be.

Level 3: The contingency plan is revised at agreed time intervals - there is a responsible person who oversees this revision and structures necessary meetings and actions. The contingency plans are FAIR (Findable, Accessible, Interoperable and Reusable). Here we advise that the documentation itself can be easily found by providing good metadata, however due to potential sensitivity of perhaps some of the data (i.e. misuse), we recommend only sharing the

contingency plans with researchers that wish to deposit specimens or other collection managers upon request.

**Importance:** Contingency planning is one of the key components of long-term sustainability for all repositories ([ISBER, 2023](#)). While hoping that the contingency plans will never need to be used, it is by anticipating potential problems that personnel linked to the DNA collection are empowered to correctly act and the collection itself can be kept safe.

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific Key-category: “Contingency planning”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions.

*DiSSCo Flanders DNA working group experience (2024) with “Contingency planning”:*

Classically, with cold storage, the DNA collections in the working group work with alarm systems and back-up freezers in case of mechanical breakdown. During the first brainstorming session, it was also noted that for most contingency planning, this is a task larger than solely the DNA collection, and that it should be tackled on the institute’s level.

During the visit to the [EAZA biobank hub at Zoo Antwerpen](#), it was demonstrated that for the specific context of this biodiversity biobank (not only DNA specimens) the four EAZA biobanking hubs worked in pairs, holding duplicates of each other’s collections as a contingency strategy in a geographically separate location. Duplicates of the samples collected by one of the hubs are exchanged with the partner hub annually.

Return to [Table1](#).

## 13. Identifiers

**Category:** Identifiers

**Level 1 goal:** There is a unique and persistent ID for each physical specimen.

**Level 2 goal:** There is a unique and persistent ID for the digital record.

**Level 3 goal:** There is a unique and persistent ID for the digital record, within the global collection.

**Explanation:** There are two types of identifier: the physical identifier, present on the physical DNA specimen itself, and the digital identifier, a unique code for the digital record of that DNA specimen which is present in the digital inventory (e.g. a Google sheet, a LIMS, a CMS etc.). This digital record, with all its linked information (e.g. sequence information), could serve as a “digital twin” of the DNA specimen. The same identifier can be used for both the physical and the digital specimen, or they can be different identifiers that are linked to one another.

The general recommendation for identifiers is that they are persistent, unique and human-readable ([ISBER, 2023](#)). In [Corrales et al. 2023b](#) (Chapter 2: Metadata and data management), it is recommended to work with Unique Universal Identifiers (UUID) and avoid using semantic information in the *collecting number*, as this data is subject to change and it may cause problems in databases, transcription errors and lead to misleading results.

In all cases, identifiers should be constructed according to a certain method or rationale. It is recommended to document how identifiers are generated for a collection and link to this information for each specimen that has had its identifier constructed following that rationale. This is especially important if there is specific meaning within the identifiers, such as a coding system for a certain taxon group, code pertaining to a certain collector, project or year numbers; or if a certain algorithm is used (e.g. <https://www.uuidgenerator.net/>). The latter should also be noted in case of future detection of a problem with the algorithm.

It is possible that one (DNA) specimen has multiple identifiers linked to it throughout its lifecycle. It is recommended to keep all old or non-preferred identifiers in the data of the (DNA) specimen (including its construction information), to ensure linkage to information in case that such a previously used but currently inactive identifier was used at the time that it was still the preferred identifier.

Level 1: Have a physical identifier on each container. This means a written or printed code on the vial or plate containing the DNA specimen(s). Depending on the choices in physical organisation of the DNA collection, other containers such as boxes, racks, drawers, freezers, rooms, ... are also coded with a persistent, unique and human-readable identifier so that the collection inventory allows retrieving the DNA specimen from within the hierarchy of containers. ‘Persistence’ in the context of a physical identifier, means that the label with the identifier on it cannot get separated from the container or that the written identifier cannot become unreadable.



Here, fail-safes should preferably also be initiated, such as predictable numbering (e.g. incremental). ‘Uniqueness’ in this context means that the identifier should be (minimally) unique within its direct physical organisation: the DNA collection it is part of (i.e. the group of DNA specimens in the same lab/building/institute). However, when a DNA specimen labelled with this type of identifier (only unique in its current collection-concept) is moved from this context, a new identifier has to be added to the physical specimen to make it unique in its new context (for example with the addition of a prefix or a suffix). Every physical identifier has to be listed in the digital inventory of the DNA collection, recorded in the centralised workspace (["14. Digital management, level 1"](#)). The physical identifier links the physical specimen to the digital record.

Level 2: Have an identifier for each digital record of the DNA specimens and storage locations. This identifier should be unique within the database. Depending on how many collections or labs share the same database, this means that the identifier will often be unique in a larger context than solely the DNA collection. A database in the context of DNA specimens is often a LIMS (Laboratory Information Management System) or a CMS (Collection Management System). If the digital identifier is different from the physical identifier, linkage between the two identifiers must be ensured. Linkage can be put in place by listing the physical identifier in the digital record's data or by explaining how the two identifiers are related. For example, the physical identifier might contain a part of the digital identifier, or vice versa, the digital identifier adds more information to the physical identifier, making it unique in the institute or global context.

Level 3: Institutes can ‘unlock’ their own collection in a stable https-URL format if they unlock the inventory of their collection via the institute’s or consortiums catalogue. ‘Unlocking’ in this context means that the specimen records and (part of) their data are open for the public to query, to browse through, as well as to use. This can be done using the concept of [CETAF Stable Identifiers](#) ([Güntsch et al., 2021](#)) or Darwin Core triplets (however: see [Guralnick et al. 2014](#)) - which can work well. However, the most persistent option will be the DiSSCo-provided “Natural Science Identifier” (NSId) ([Hardisty et al., 2021](#)). It is expected that, once operational, if DNA collection data is unlocked via the DiSSCo Research Infrastructure (RI), the DiSSCo RI will generate DOI’s - a “Natural Science Identifier” (NSId) - for each unlocked digital twin of the DNA specimens. For more information on the topic of persistent identifiers in the global/universal context, [this DiSSCo Tech post](#) is a recommended read.

**Importance:** Identifiers allow information to be linked to an object, in this case a DNA specimen. This can be crucial information for DNA collection management, such as occurrence information and storage location. In comparison to other Natural Science specimens that have their data physically stored with the specimen, for example: written on the label, DNA specimens often consist of relatively small vials or plates to which only the identifier is added and all other information has to be retrieved via the identifier.

Every DNA specimen needs a separate, unique identifier that differs from the identifier of the source specimen, voucher specimen or subsample. This is because each physically separated



DNA specimen has (or can have), different data linked to it, such as the exact storage location, management history, reuse history (including number of freeze-thaw cycles), volume and linked data, even though they can represent the exact same occurrence (point in time and space).

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific Key-category: “Identifiers”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions.

#### *DiSSCo Flanders DNA working group experience (2024) with “Identifiers”.*

In practice, for physical identifiers of the DNA collections in the working group, (i.e. those present on the specimen), it is noted that, depending on the resources, different ID-coding practices are applied. All allow identification, as long as the DNA specimens stay in their current context / DNA collection:

- Example 1: in terms of the limited space on a DNA specimen vial - making it necessary to shorten the full identifier of the DNA specimen when printed in human-readable format on the physical identifier.
- Example 2: there is merit to having a meaningful coding system (i.e. non UUID), to make it easier to spot a DNA specimen that is out of place without having to scan or search the database. In this case, if the identifier has a taxon name code or year code that doesn't match the other samples in the same storage container: in the set of consecutive specimens it stands out. Especially in smaller DNA collections, this practice has good feedback from users.
- Example 3: using a permanent 2D-barcode that is provided by the manufacturer on their vials saves a lot of time to physically label the samples. Furthermore, the permanent 2D barcodes are compatible with automated barcode readers, allowing very efficient identification and inventorying of physical samples. The 2D codes are not UUIDs, the barcode is a 10 digit tube-ID with an additional 3 character location-code. The IDs are unique to the vial supplier and are guaranteed to be unique for all the tubes in all their racks, but being only 12-14 digits in length, they cannot be considered UUIDs.

In practice, for the digital identifiers in the working group, this depends on the current workflow of the genetic lab and linked DNA collection, or/and the chosen digital solution ([“14. Digital management”](#)). Similarly to the physical identifiers, the digital identifiers all allow identification, as long as the DNA specimens stay in their current context / DNA collection.

Return to [Table1](#).

## 14. Digital management

**Category:** Digital management

**Level 1 goal:** Centralise and standardise data of the DNA collections + DMP.

**Level 2 goal:** Convert unstructured data to structured data.

**Level 3 goal:** Durable linkage to files and other data(bases).

**Explanation:** This category is about in-house digital management of the DNA specimens. This is a necessary step prior to unlocking the relevant data of users, because:

- 1) There is operational information that is not relevant for potential requesters to see
- 2) There can be sensitive information such as personal data (e.g. name and address of requestors, colleagues), logs of relevant mailing conversations, specimen location data in the collection (to avoid manipulation by unauthorised persons), sensitive occurrence data (e.g. endangered species, species with (lucrative) pharmaceutical applications, ..)

Level 1: Centralise and standardise data of the DNA collection + use an operational Data Management Plan (DMP) to structure the work. Regardless of already having a digital management system implemented or not, the steps below are a useful exercise either to optimise or to choose a structured data management system (see [level 2](#)). The results of this exercise are translated into an operational DMP (i.e. differing from research DMPs, [see “1. Involving internal stakeholders”, level 2](#)) with versioning that is updated when focussing on [level 2](#) and [level 3](#) of this category.

Firstly, it is agreed (e.g. with the relevant stakeholders ([“1. Involving internal stakeholders”](#)) and in accordance with the workflow ([“4. Outlining the current workflow”](#)):

- which data fields and data files are recorded with the DNA specimens of the DNA collection,
- in which format data is recorded,
- which of these fields or files are optional to log and which are mandatory.

This includes both relevant sampling data, lab data and relevant collection management data.

For each data field (e.g. a column in an Excel table) there should be a clear description, including the format in which data should be entered. Optionally, an example entry could be provided. Data fields can either be self-defined fields or fields as provided by data standards. Here, it is far more important to find something that works for the stakeholders who supply and process the data, rather than having a complicated, yet, state-of-the-art data sheet. If data fields already exist for the current digital management of the DNA collection (at the time of revising this goal), they can always be mapped to data standards when this becomes necessary (e.g. when publishing the DNA collection specimens to a repository like DiSSCo). Therefore, it is not necessary to change the current procedures, forms or data system to data fields as defined by data standards when they are already working well, rather, it is recommended to clearly describe the fields that are currently present and to make small changes where it is deemed necessary (e.g. mistakes, misuse of data fields, ..). Data standards can be applied, and by

---

Supplementary material 2 of “The key to bringing DNA collections to the next level (Veltjen et al., 2024)”: Interactive format of the Key, including the guidance documentation

design should fulfil the need of clearly defining data fields. However, if it is more practical to define the data fields otherwise or if the data standard does not suffice for the type of information you want to log, custom data fields are equally useful - if, of course, they are clearly defined. If data standards are used, data fields can simply be linked to the persistent identifier of that specific data concept and the linked definition and recommendations can be applied when logging information. For example, if it is decided that for the DNA specimens of your collection, the decimal latitude is logged in the data, you could refer to <http://rs.tdwg.org/dwc/terms/decimalLatitude> when giving the header for the column in the Google sheet in which you collect data from the suppliers.

Data standards that may be especially useful (or to map to when publishing, "[15. Unlocking the collection](#)") for DNA collections are the following:

- ISO data standards for notation (e.g. [ISO 8601 on date notation](#) (International Organization for Standardization [ISO], 2016));
- [Darwin Core](#) ([Wieczorek et al., 2012](#)) for Biodiversity data + [DNA derived data extension](#);
- [ABCD](#) (Access to Biological Collection Data task group, 2007);
- [ABCDEFG](#) ([Petersen et al., 2018](#)) for geoscience;
- [the GGBN data standard](#) for DNA specimens ([Droege et al., 2016](#));
- [The ENA Metadata model](#) for DNA specimens and sequences;
- The [GSC Minimum Information about any Sequence \(MIxS\)](#) for sequences;

With or without data standards; the following data concepts are recommended to be used for both biological and environmental DNA specimens (adapted from [Corrales et al. 2023](#), Chapter 2):

- The ID code that is physically present on the DNA specimen ("[13. Identifiers](#)", [level 1](#))
- The ID code of the directly linked Natural Science specimen(s): there is always a source specimen that the DNA was extracted from.
- The type of the source specimen, such as muscle tissue, buccal swab, water filter, ..
- The ID code of the voucher specimen. There should be a voucher specimen that represents the occurrence, population or taxonomic unit associated with the DNA specimen.
- The linked occurrence data that the DNA specimen represents:
  - Sampling date (and time) of the source specimen
  - Sampling location
    - i. Sampling X- and Y-axis: geographic location: descriptive and preferably also GPS coordinates (including GPS method, accuracy, ..)
    - ii. Sampling Z-axis if relevant (e.g. water samples, soil samples)
  - Associated taxon name(s): it is possible to have multiple taxon names associated with one DNA specimen. For example, older and newer identifications of the voucher specimen that change the taxon name associated with the DNA specimen; morphological versus sequencing identifications; sequencing identification at point in time 1 (e.g. reference database incomplete) versus

sequencing information at point in time 2 (e.g. reference database complete);  
metabarcoding of eDNA specimens, ...

- Sampling and transport methods (field to lab) of the source specimen
- DNA extraction method of the DNA specimen.
- Quality data of the DNA.
- Quantity data of the DNA.
- Preservation method of the DNA specimen.
- Storage location (and, by proxy, condition, such as temperature) of the DNA specimen.
- Linked information e.g. sequence identifiers, project identifiers, permit identifiers, ...

Secondly, the clearly described and defined data that was logged with the DNA specimens is captured in an unstructured (e.g. Google sheet, Excel) or structured format (relational databases), that is stored in a central digital space. This could be a shared online folder (e.g. Google folder or a OneDrive) containing different documents, one larger shared document, or a secured server in terms of a database. [Corrales et al. 2023b](#) (Chapter 2: Metadata and Data Management) recommend the use of a central digital space on internal organisational servers with strict access management to avoid misuse of sensitive information (e.g. personal data, vulnerable wildlife populations). The use of external servers may, however, be considered, as long as the necessary digital security measures are in place.

When working with unstructured data, the data fields are most likely translated into two types of practical documents:

- 1) A form that the suppliers (e.g. researchers, students) have to fill in for the specific project (i.e. batch of DNA specimens), when adding new DNA specimens to the collection. Even with structured data and applications (see [level 2](#)), such a form often still remains in place to harvest data from the suppliers. The form is then imported, batch by batch, into the database.
- 2) An inventory of the full collection for the DNA collection managers. This is usually a Google sheet or Excel inventory.

Commonly, both the form and inventory documents start as unstructured data (e.g. Excel, Gsheet) and can continue to be managed as such, as long as the number of people working with the files and the number of DNA specimens remains within the limits of efficient working. Such a system can be workable without having to search for other solutions. However, when it becomes hard to manage the files efficiently in this way, as well as when looking for a more durable manner to store this information, switching to a structured data system ([level 2](#)) needs to be planned.

When working with unstructured data formats, ideally the files have an accompanying 'readme' file or clear instructions provided to help both with data entry and to maintain linkage with the definitions of the data fields. Optionally, restrictions can be built in the control entry format or there can be automated or human-operated control of the data entries at one point in the workflow. For example, dropdown lists in Excel are a way to automatically control data entry. A

situation where project-based submission forms are checked by the DNA collection manager prior to adding the batch of DNA specimens to the central inventory, is an example of human-operated control of data entry. Alternatively, having one exemplary record at the top of each blank supplier form is also an easy extra guide to provide which facilitates correct data provision. It is recommended to make the supplier form and inventory list(s) as simple to use as possible. Depending on any misuse of the forms and/or inventory list(s), the documents can be upgraded to counter the observed problems, for example by updating descriptions or building in more control on data entry.

Lastly, when working with unstructured data formats to which data gets appended by multiple users (e.g. a central inventory list in the case of DNA specimens), best practice is to work with version control. This allows restoration of a previous document version, if needed, and to log who administered what changes and when. In Google sheets and SharePoint environments, this is already in place. It is advised that a central document manager monitors the activity in the shared documents on a regular basis.

Level 2: Convert unstructured data to structured data as much as technically (and logistically) possible. 'Structured data' is data in a predefined format, complying to a data model, which is searchable (i.e. using queries) and easy to analyse. It is often stored in relational databases. Unstructured data is stored in a variety of formats and requires more work (e.g. data cleaning) to process and understand.

A structured data system ideally facilitates effective management of the DNA collection: knowing what there is (biological information, technical information), where it is (storage location) and in which state (quality estimate, volume), and what has happened with the archived sample throughout its lifecycle (loans, curation). When revising options, it is perfectly fine to (partly) switch to structured data systems and to decide to continue to (partly) work with unstructured data ([level 1](#)), if that is what is technically and logistically the best option for the DNA collection at hand.

It is recommended to document the decision-making process. Potential decisions could be:

- 1) We want a new structured digital management system: why?
- 2) We want to keep working with the current (un)structured digital management system, and not invest in (more) changes: why?
- 3) We want to start or to keep working with a(n) (un)structured data system at our institute. However, we intend to implement changes so that the system better fits the user needs (e.g. for managing DNA specimens): why? For example: the institute already has a LIMS or CMS, but these systems are not (yet) operational for DNA specimens.

Logging the decision-making process allows one to be reminded of the rationale behind a previous choice. This can be helpful in various situations, such as bringing new colleagues up to date on the system and its history, considering a new system or assessing a new question/opportunity. In this way, the collection manager(s) and other stakeholders can avoid repeating the decision process and instead build on previous actions. These decision histories

can also be invaluable when collaborating with other collection managers, especially when helping collections with very limited resources to find the most efficient processes possible.

When entering the process of converting unstructured data to structured data, it is recommended to first clearly outline user needs. For example, working with priorities like 'must-haves', 'should-haves' and 'nice-to-haves'.

Options for structured data systems depend on the user group and the purpose that the system was originally designed for. There is usually a trade-off in specialisation versus generalisation; more specialised systems perform very well for the function they were designed (e.g. LIMS, CMS, Custom Systems), while more generalised systems (e.g. ELN/ERNs, All-in-one Systems) can sometimes have functionalities you never use (yet still pay for) and other functionalities that work suboptimally for your situation, but the advantage of the digital management system allowing you to execute the full extent of actions within the same digital environment is great enough to be 'worth it'. Five types of digital systems are listed below. It is possible to have multiple systems in place that support the management of the DNA collection. For example, it is possible to have both a LIMS for the lab technicians and a CMS for the collection manager(s). Both log information on DNA specimens and they can be set-up so that relevant information between the databases is (automatically) exchanged (see [level 3](#)). It is even possible to combine a structured data system with an unstructured data system, for example, by having a CMS in which only the DNA collection manager(s) work(s) and having Google sheets per project in which the researchers/suppliers work.

If the digital management of the DNA collection changes when working towards (partly) structured data storage, the aligning DMP needs to be updated throughout the change process. Using whichever DMP tool best suits your collection/institute (e.g. [DMPonline](#), the [Data Stewardship wizard](#), [DMPtool.org](#) etc.), will allow thorough consideration of where data is hosted, who has access, documentation, digital security etc., all of which facilitates long-term thinking and more durable changes. If the digital management of the DNA collection is integrated into a larger system, supporting for example multiple laboratories or multiple Natural Science collections, the associates of the DNA collection could be involved in writing and/or revising the DMP, with special emphasis on the DNA specimen use-case throughout. It is also possible that the DMP is written by a staff member in the role of a data manager or IT operations, although in such a situation it is strongly advised to have sufficient input and feedback from the staff involved in the DNA collection when writing the DMP. For a DMP to be effectively implemented, all relevant stakeholders need to be 'on board'.

The five types of structured data systems that can currently be used for managing DNA specimens with both commercial and open-source examples included, are:

### **(1) Laboratory Information Management System (LIMS)**

This type of database and linked application/software are aimed at supporting lab technicians



and facilitating collaborative working, especially between larger groups. It is sample-centric and logs all activity on lab samples and reagents in the laboratory environment. A LIMS can be used in a solely genetic laboratory or in a combination of different types of laboratories (e.g. a genetic lab and an analytical or physico-chemical lab). The concept underlying a LIMS is that the lab workflow is carefully and thoroughly logged: receipt of samples; test assignment; result entry; calculations; storage of reagents and samples; instrument information; and so on. The LIMS facilitates standardised processes, such as repetitive testing according to a standard operating procedure.

- Commercial examples: [LabWare LIMS](#), [Freezerworks](#), [FreezerPro](#), ...
- Open Source examples: [Open LIMS](#), [BikaLIMS](#), [Clover \(plant labs\)](#), ...

## **(2) Electronic Lab/Research Notebooks (ELN/ERN)**

This type of database and linked application are aimed at supporting lab technicians and researchers that work in variable experimental settings (i.e. non-repetitive lab/research work, exploratory testing). These types of systems allow you to log all data relevant to being able to replicate an experiment: it allows you to log notes, images and workflows chronologically. This setting is an active project environment, and hence it is project-centric. This type of system is often very flexible, supports collaboration and has little learning curve. However, there are no standardised fields and therefore there is limited interoperability. Nowadays, most ELN/ERN systems also (partly) have LIMS functionalities.

- Commercial examples: [OneNote](#), [Benchling](#), ...
- Open Source examples: [eLabFTW](#), ...

## **(3) Collection Management Systems (CMS)**

This type of database and linked application/software are aimed at supporting collection managers. It is specimen-centric (or sample-centric) and logs all activity on specimens, often including all specimens of one institute (i.e. both DNA specimens and non DNA specimens). The concept underlying a CMS is that the data (e.g. occurrence data, creation data) and the life cycle (e.g. loans, treatments, storage location, exhibitions, ...) of the archived (DNA) specimens is carefully and exhaustively logged.

- Commercial examples: [EarthCape](#), [BRAHMS](#), ...
- Open Source examples: [CollectiveAccess](#), [DINA](#), [Specify](#), [Arctos](#), ...

## **(4) All-in-one Systems**

These types of systems are known by a variety of names: “Life Sciences Data Management Software”, “Workbench for biodiversity scientists”, “Data management and publishing platform”... These types of databases and application/software try to capture all data in the full research cycle, such as field data, collection data, lab data and analysis data. This type of system is aimed more towards the researcher community, however, it can also be effectively used by a collection curator if CMS-like functionality is included (i.e. logging storage location, volume,

loans, ...) or perhaps even lab technicians if there is some LIMS-like functionality available (i.e. logging SOPs and tests, instrument information, ...).

- Commercial examples: [BioloMICS](#), ..
- Open Source examples: [DataHub](#), [PlutoF](#), [TaxonWorks](#), ...

## **(5) Custom Systems**

If sufficient resources are available to do so, it is always possible to design a custom database. This frequently also involves designing an application for easy data entry and retrieval. The greatest benefits here are that it is wholly customised and that the creators have complete control of and insight into both the database and application. This brings with it some drawbacks: such a unique system/application generally requires more resources and a longer process to build and maintain, and often has a limited user base.

- Examples: [DaRWIN](#), ILVO-CMS,...

Level 3: Links to documentation and other data(bases) are in place. In the structured and/or unstructured Digital Management System(s), there are procedures that ensure linkage which is as stable as possible between each DNA specimen and information such as:

- Unpublished, internally stored sequence data
- Relevant field photos
- Permit documents
- Loan contracts
- SOPs (with the correct version at the time of usage)
- Other internal databases
- External taxonomic databases (such as Catalog of Life)
- External sequence databases (such as ENA)
- Published articles
- The Digital Extended Specimen (DES) of the specimen itself
- The Digital Extended Specimen (DES) of related specimens (such as the source material)

Whenever the links themselves are persistent by nature, for example DOI's of published articles or SOPs, and the DOI of the (future) Digital Extended Specimen ([Hardisty et al., 2022](#)), it suffices to have a data field in which the DOI-link is entered.

However, for all other types of links, different strategies can be put in place:

- 1) Clearly documenting all the dependencies, building procedures and protocols to make the dependencies as stable as possible. Documentation can be in the DMP as mentioned in [level 1](#) and [level 2](#).
- 2) A proactive checking procedure: this means appointing a responsible person and agreeing on a systematic approach/procedure to check if any of the links are broken.
- 3) A reactive fix protocol: this means appointing a responsible person who has to fix any broken links when they are noticed/reported.



**Importance:** Data associated with specimens is of the utmost importance for effective scientific research and reproducibility ([ISBER, 2023](#)). DNA specimens have information linked to them, logged at different stages in the research cycle before entering the collection and often obtained from different people, as well as during its time as an archived specimen of the DNA collection.

As the specimen itself is relatively small, all information has to be logged physically separately from the specimen, unlike certain larger specimens which often have larger labels and work with annotations on the specimen itself, such as herbarium specimens.

To manage DNA specimens, it is vital to be able to access this linked information in an efficient manner when needed, as well as log new information that could influence the (scientific) value of the DNA specimens. This creates a need for an accessible and searchable (or queryable) digital management system. Naturally, each DNA collection exists within a unique context and this needs to be taken into account to ensure that users can efficiently work with the chosen Digital Management System.

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific Key-category: “Digital management”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions.

*DiSSCo Flanders DNA working group experience (2024) with “Digital Management”.*

As far as can be ascertained from general discussion, most of the working group members were preoccupied during the timeframe of the DiSSCo Flanders project with fine-tuning their digital management. In practice, this is the very core of (DNA) collection management.

Most institutes obtain data from suppliers in an Excel or Google sheet form, provided by the genetic laboratory manager or DNA collection manager, or are working towards a standardised procedure to obtain field (and lab) data from suppliers.

There is one institute that makes use of a LIMS system (i.e. [Research Institute for Nature and Forest: LabWare LIMS](#)), where the genetic laboratory shares the database and software with the analytical laboratory.

In practice, most institutes still have centrally managed inventory lists in an unstructured format, often mostly with human-operated data quality control in the form of an online document system (with versioning). This is due to a variety of reasons: either the institute is implementing a new CMS that will better fit DNA collection management (e.g. [Meise Botanic Garden](#) and [Ghent University](#) research collections: [EarthCape](#)); they are initiating the CMS implementation from scratch, tailored specifically for including the DNA specimens use case (e.g. [Research Institute for Nature and Forest: CollectiveAccess](#); [Flanders Research Institute for Agriculture, Fisheries and Food](#): Custom System ‘ILVO-CMS’); they are adopting an All-in-one System (e.g. [Belgian](#)

[Coordinated Collections of Microorganisms: BioloMICS](#)); or they are reconfiguring the current CMS to fit the DNA collection (e.g. [Royal Belgian Institute of Natural Sciences](#) and the [Royal Museum for Central Africa: DaRWIn](#)).

The EAZA biobank makes use of [ZIMS \(Zoological Information Management System\)](#), a database and application that was purpose-built to collect and share data of animals kept in more than 1000 zoos globally. For the EAZA biobank a new biobank module in ZIMS was recently developed to accommodate the inclusion of tissue and DNA collections and their management.

**Return to [Table1](#).**

## 15. Unlocking the collection

**Category:** Unlocking the collection

**Level 1 goal:** The DNA collection is on GrSciColl and the institute's website.

**Level 2 goal:** A test dataset is in a public repository + there is a publishing strategy.

**Level 3 goal:** The complete DNA collection is in a public repository.

**Explanation:** This category is about bringing the relevant in-house (digital) information to the scientific community by publishing or “unlocking” the data, in a standardised format, in public repositories. In this way, the data of the DNA collection and the DNA specimens become FAIR (Findable, Accessible, Interoperable and Reusable); especially to external users of the data and specimens. Users can use the unlocked data of the specimens in (meta-)analyses, they can use the found DNA specimens of interest to start loan requests ("[16. Loans](#)"), or they can use the found DNA specimens to start new collaborations with experts that are linked to the collections. Unlocking the collection data digitally greatly facilitates physically unlocking the collections ("[16. Loans](#)"), as they become much more findable. Once found, however, to access the specimen(s) *physically*, users have to place a loan request.

Level 1: The DNA collection is described in [GrSciColl](#) and the institutional website.

The [Global Registry of Scientific Collections, or GRSciColl](#) ([Grosjean et al., 2022](#); [Grosjean et al., 2023](#)), is a comprehensive, community-curated repository of information on scientific collections. GRSciColl was initially developed by the Consortium of the Barcode of Life (CBOL) and hosted until 2018 by the Smithsonian Institution. In 2018, GRSciColl was adopted by GBIF and merged into its registry.

Alongside having the DNA collection on GrSciColl, it is also the goal to have the DNA collection actively mentioned on the institute's website, including extra information on the scope of the DNA collection and a link to its location on GrSciColl.

Level 2: A test dataset is in a public repository + a publishing strategy is in place.

Current available public repositories options for DNA collections are:

1. [GGBN portal](#): The Global Genome Biodiversity Network Portal for publishing of DNA specimens and collections. To be able to publish on the GGBN portal, the institute must pay a membership fee, see <https://wiki.ggbn.org/ggbn/Membership>.
2. [GBIF](#): The Global Biodiversity Information Facility to publish:
  - a. Occurrence data preceding DNA specimens.
  - b. Occurrence data derived from DNA specimens ([Abarenkov et al., 2023](#)).
  - c. Occurrence data derived from eDNA specimen: the eDNA converting tool (under development).
  - d. (DNA) collection inventory as a dataset.

To become a data publisher, your organisation must request endorsement from your

---

Supplementary material 2 of “The key to bringing DNA collections to the next level (Veltjen et al., 2024)”: Interactive format of the Key, including the guidance documentation

GBIF country node or directly from the GBIF secretariat.

3. [Genomic Observatories MetaDatabase \(GEOME\)](#) ([Deck et al., 2017](#)): metadata on biological samples, used for biodiversity inventories, population studies, and environmental metagenomics.
4. [ENA](#) ([Cummins et al. 2022](#)) or [GenBank](#) ([Schoch et al. 2020](#), [Sayers et al. 2021](#)) to publish nucleotide data. “Sequence Submissions” in [de Mestier et al. 2022](#) is a recommended read on how to organise submitting DNA sequences to the European Nucleotide Archive (ENA) and to NCBI GenBank. Interesting resources here are a) the ENA checklists: <https://www.ebi.ac.uk/ena/browser/checklists>, and b) the ncbi biosample packages <https://www.ncbi.nlm.nih.gov/biosample/docs/packages/>, that stipulate which information should be logged in the case of which type of source material.

In 2026, [DiSSCo](#) should be operational and also be an option for publishing (or “unlocking”) Natural Science specimens. It is expected that the DiSSCo Research Infrastructure will allow unlocking of DNA specimens via [ELViS, the European Loans and Visits portal](#) - a ‘one-stop-shop’ allowing users to request virtual access (data exports and digitisation-on-demand), specimen loans and (collection) visits.

It is expected that GBIF will also harvest occurrence data from DiSSCo, GeOME and GGBN databases.

By designing and executing the first (manual) test(s), the experience will allow one to write a strategy on how the collection will be (completely) unlocked in [level 3](#). This could include writing a procedure or workflow, if deemed worthwhile. In all cases, this strategy must include the designation of a responsible person who will follow up on the data publication of the (DNA) collection in the long term.

To effectively publish the DNA specimen data from the chosen digital management system (["14. Digital management"](#)), interoperability is key. The data of the institute's database should be mapped with the standards used by the repository, or repositories, of choice.

Level 3: The complete DNA collection is in a repository. The most efficient way to have this goal met at all times is via automated or human-induced systematic data publication from the chosen digital management system (see ["14. Digital management"](#)), if applicable.

**Importance:** Data publishing (of DNA collections) is important because it allows:

- The data (and indirectly the physical specimens), to be available for others:
  - To (re)use for new research.
  - To validate previously executed research.
- The data and the specimens to be citable
  - Allowing academic credit for collectors as well as data managers of Natural Science specimens.
  - Allowing more persistent linkage of publications (articles, datasets, ..) to the

- specimens.
- Allowing better tracking of the impact of Natural Science specimens.

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific category: “Unlocking the collection”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions.

*DiSSCo Flanders DNA working group experience (2024) with “Unlocking the collection”:*

Currently, as far as can be ascertained from general discussion, none of the DNA collections in the working group are unlocked except for the BCCM collections via an [in-house catalogue](#) and associated [GBIF-datasets](#). In terms of unlocking datasets, there has been a first experience using the GBIF eDNA converting tool ([Abarenkov et al., 2023](#)) to publish certain ILVO marine eDNA datasets, soon to be followed by various INBO eDNA datasets. This has been accomplished by working closely with the Belgian GBIF node.

**Return to [Table1](#).**

## 16. Loans

**Category:** Loans

**Level 1 goal:** Responsible person, Public statement on how-to-loan, Loan Agreement.

**Level 2 goal:** A FAIR loan policy + established incoming and outgoing loan procedures.

**Level 3 goal:** Lending policy and procedures are operational + creation of a Loan Agreement breach protocol.

**Explanation:** This category is about managing loans. For a loan request users first need to be aware of the existence of the specimens. This could be external users (i.e. external loans) by having the DNA collection or specimen data unlocked ("[15. Unlocking the collection](#)") or, by internal users (i.e. internal loans) in an internal Digital Management System ("[14. Digital management](#)"). An especially useful resource regarding most aspects of this category is de [Mestier et al. 2023](#).

DNA specimens are always destructively "loaned", meaning that they are subsampled and then sent to the requester. Part of the specimen is therefore lost and a smaller quantity of the DNA specimen remains which can be re-used for other potential purposes. In practice, a new aliquot is made from the DNA specimen or the DNA specimen is loaned in its entirety. The latter is sometimes the case if the DNA collection management invests in making duplicate specimens beforehand to anticipate loans and minimise freeze-thaw cycles of one replicate. It is not common practice to receive a "loaned" DNA specimen back after usage by the requester, as there is little control over the further quality (e.g. more freeze-thaw cycles, pipetting that can degrade the DNA further or even contamination or labelling mistakes), however it can be that any unusable or unused DNA specimen is returned to the collection. It is sometimes requested to destroy any remaining material after the requested usage is completed.

Level 1: Firstly, actively appoint a person who is responsible for handling loan requests for the DNA collection. This is for both incoming and outgoing loans of DNA specimens, meaning that this same person oversees all transactions. The responsible person can be a single individual in the institute who handles all loan requests, or a different person for each subcollection. This responsible person's activities align with the collection's current organisation of the 'Digital management' ("[14. Digital management](#)"). For example:

- The digital inventory should allow the responsible person to consult the available volume of the requested specimens.
- The responsible person should log any changes or new information for executed loans (e.g. remaining volume, number of freeze-thaw cycles, new sequence information, ..) in the digital inventory.
- The responsible person should archive relevant documentation of loan requests (i.e. the made agreements) in the centralised folder.

Secondly, write a publicly visible statement, for example on your institute's website, next to the

collection description (see [level 1 of “15. Unlocking the collection”](#)). This statement should clearly explain who to contact when wanting to request a loan, or if not yet publicly available, who to get in touch with to request more information (i.e. an information request) on the content of the DNA collection. Optionally, you can also specify in this public statement what information should be included in the loan or information request, to minimise emailing back and forth. It is recommended to have an agreed upon turnaround time for answering requests, either by communicating this directly next to the contact information, or by having a standardised ‘request received’ confirmation email that gives a time estimate with regard to handling of the request.

Thirdly, have a template Loan Agreement for (outgoing) DNA specimen loan requests. The template Loan Agreement for the DNA specimens can be self-created or adapted from a more generic Loan Agreement template provided by the organisation which the DNA collection is part of, such as an institute’s loan agreement template or a template provided by specific consortia or organisations. For example, DNA collections that are in institutes which are CETAF members are encouraged to use [the CETAF MTAs \(Material Transfer Agreements\)](#) and DNA collections in institutes which are GGBN members are encouraged to use the [GGBN MTAs](#). DNA collections that are not (yet) part of an overarching organisation specialised in (DNA) collections can adopt applicable clauses of the existing MTA template, if evaluated to be applicable to their vision. The goal of a Loan Agreement is to have the agreed upon terms and conditions of the usage of the requested DNA specimens clearly described in writing for all parties. For more in depth recommendations on Loan Agreements of DNA collections, consult [de Mestier et al. 2022](#): “Receiving molecular sample requests from external and internal researchers (outgoing loans) - Loan Agreement”.

If the goal [“Stability level 1”](#) is completed, make sure that both the public statement and the Loan Agreement align with the decisions made on financial responsibility.

Optionally, it is also possible to make a checklist for Loan Agreements provided by other institutes (incoming loans), meaning that you can make a list of clauses that must be present in the contract, as well as clauses that are red flags. If any clause is missing or not-aligned with the vision of your institute, negotiations should be held to reach a contract that both parties agree on. However, on the basis that all DNA collections are executing good practices, simply reading the contract, asking questions if clauses are unclear and trusting the providing party to have good intentions often suffice. For more information, read [de Mestier et al. 2023](#): “Requesting molecular material at third party institutions (incoming loans) - Completion of agreements required for the transfer of material”.

Level 2: Have a collection loan policy and procedures on the level of the institute that include the use-case of the DNA collection. The more FAIR the loan policy and procedures are, the better.

The loan policy should inform the borrower/user about the process, terms and conditions of a loan ([de Mestier et al. 2023](#)), including both outgoing and incoming loans. This could be in the



form of a general loan policy for the institute that is also applicable to the use case of DNA collections, or a separate loan policy that is solely for the DNA collection. Ideally, the loan policy is available via the institute's website (along with the public statement regarding contact people, [see level 1](#)), or, alternatively, is shared with potential loan requestors as early as possible in the communication. It is encouraged to publish the loan policy (with versioning) on a platform like [Zenodo](#) to increase the FAIRness of the document and act as a guide for DNA collections looking to write or edit their loan policy. Consult [de Mestier et al. 2023](#): "Receiving molecular sample requests from external and internal researchers (outgoing loans) - Loan Policy" for more detailed recommendations on what should be written in the loan policy considering outgoing loans (and a loan request form). A loan request form is also recommended, as it will ensure easier and quicker evaluation of the request - but it is not a prerequisite: information on the loan request can also be obtained via unstructured written requests, oral requests, ... In any case, the better documented loan communications are, the less chance there is of complications/discussions afterwards. Consult [de Mestier et al. 2023](#): "Receiving molecular sample requests from external and internal researchers (outgoing loans) - Policy for Incoming Material (incoming loan) for molecular work" for more detailed recommendations on what should be written in a good loan policy regarding incoming loans. A worthwhile statement to write in this policy is the DNA collection's vision on the finality of the destructive sampling: (when) do we allow DNA specimens to become consumed? And when don't we?

Alongside the overarching loan policy and often on an institutional level for all collections, there should be two written procedures: one for outgoing and one for incoming loans. Again, these could be general procedures or procedures specialised for the DNA collection. Both procedures should include a legal checkpoint (e.g. Nagoya) when evaluation of loan approval is carried out. Procedures can be internal documents or openly published documents; in both cases there needs to be clear versioning of the procedures.

For outgoing loans, consult [de Mestier et al. 2023: Chapter 2: Receiving molecular sample requests from external and internal researchers \(outgoing loans\)](#), including the informative Figure 1 workflow overview. New roles and responsibilities to appoint, including one responsible person handling the loan requests (level 1) are: responsibility for checking legal compliance, responsibility for approving the loan request, responsibility for handling the loan agreement, responsibility for preparing the material, responsibility for sending the material and logging good receipt, and responsibility for follow-up of the loan agreement. One person can be responsible for multiple of the described steps in the workflow, or each responsibility could be appointed to a different person. In [de Mestier et al. 2023](#): "Managing legal information, rights and restrictions related to specimens and samples", there are two relevant workflows visualised: Figure 3.2 'Authorising use where there are documentation gaps in provenance' and 3.3. "Giving permission to the use of items (items leaving, or being used in, the organisation)" that can be included in the outgoing loan protocol.

For the incoming loan procedure, please consult [de Mestier et al. 2023 Chapter 3: Requesting molecular material at third party institutions \(incoming loans\)](#), including the informative Figure 2



workflow overview. Here, it is key that all communication is logged as sent and all documentation is logged as signed and stored in an agreed upon centrally accessible folder; as well as clearly assigning roles and responsibilities at every point of the procedure. In [de Mestier et al. 2023](#): “Managing legal information, rights and restrictions related to specimens and samples”, there is one relevant workflow visualised: Figure 3.4 “Receiving permission to the use of items (eg. items being lent to the organisation)” that can be included in the outgoing loan protocol.

For both the incoming and outgoing loan procedures, having a structured data management system ([“15. Unlocking the collection”, level 2](#)) that has the functionality to log loans can provide better oversight. Most Collection Management Systems (CMS) allow logging of specific data on loans in a standardised way and allow you to track the status of the loan (e.g. which loans are still open etc.). However, as long as there is a clear procedure on who has to log what and where, and it is followed, this can be managed sufficiently in a simple folder system ([“15. Unlocking the collection”, level 1](#)). Good linkage to communication and contracts is key for collection management ([“15. Unlocking the collection”, level 3](#)).

If the goal [“Stability level 1”](#) is completed, make sure that the loan policy and procedures align with the decisions made on financial responsibility.

Level 3: The loan policy and procedures are operational + a Loan Agreement breach protocol is created.

For the loan policy and procedures to be operational, they must have been implemented and there is a strategy to optimise both documents (i.e. creating new versions). The ‘Loan Agreement breach protocol’ is a protocol that stipulates what to do when the Loan Agreement is not followed. This can be both in cases of internal and external reuse. The most likely breach of the Agreement is the lack of a user reporting back new publications, datasets, raw sequence data to be linked with the DNA specimen in the collection. There can be considerable time between sending the loan and having to link products to the DNA specimen record, which can make it more difficult to proactively ask users after a set time period. The protocol should state who is responsible for executing which steps and when. For example, if there are any consequences for not reporting back as agreed upon, and after what time period without a satisfactory response these consequences will be enforced.

**Importance:** Effective access worldwide to molecular collections, including DNA collections, considerably facilitates and improves the quality of investigations, highlighting the need for effective exchange of material and cooperation between institutions ([de Mestier et al. 2023](#)). A robust loan workflow is key to successful management of DNA collections and crucial to respecting the regulations in place ([de Mestier et al. 2023](#)). The loan workflow should ensure traceability of the DNA specimens, as well as further enrichment of the DNA specimens’ digital record as their new (data) derivatives are being linked to them.

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific Key-category: “Loans”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions.

*DiSSCo Flanders DNA working group experience (2024) with “Loans”.*

Currently, as far as can be ascertained from general discussion, the [BCCM collection](#) and the [EAZA biobank](#) provide public information on donations and loans, including a template Material Transfer Agreement. This pattern in the working group is similar to the missing statements on the existence of the DNA collection (see experiences of [”15. Unlocking the collection”](#)). As the majority of the DiSSCo Flanders working group collections are still “locked”, much of their reuse value is solely internal. Internal reuse is tractable by the same specimen identifiers being used in different projects. Even though most reuse is internal, external reuse does occur. The collections are found by researchers either via publications or networking activities. The loans are usually managed by the researchers, lab technicians and DNA collection managers themselves via email requests, sometimes with or without official Loan Agreements depending on the collection and project contexts. If there is no official contract, terms are stipulated in oral contracts or agreements via mail.

Return to [Table1](#).

## 17. Stability

**Category:** Stability

**Level 1 goal:** Financial responsibilities are outlined.

**Level 2 goal:** Estimate workload and optimise staffing.

**Level 3 goal:** The DNA collection has a stable budget and team.

**Explanation:** The goals in the stability category require close collaboration with representatives of all internal stakeholders ("[1. Involving internal stakeholders](#)"), wherein higher management levels play a key role.

Level 1: The financial responsibilities with regard to the DNA collection are clearly outlined. The decisions made here are documented. For more in-depth information on financial stability, please consult: [ISBER, 2023](#) chapter A3.2.2.3. Financial Sustainability.

Financial costs are diverse and typically include: provision of materials ~ consumables (i.e. vials, labels, pipettes, gloves), operational costs of the storage (e.g. electricity costs for freezers, costs for server space, maintenance costs and staff salaries), handling costs for archiving (i.e. staff time) and handling costs for loans (i.e. staff time, shipments). This should be well aligned to the relevant higher-level (research-group, institutional, ...) policies, as well as the (previously defined) [scope](#) of the DNA collection.

DNA collection management and the provision of DNA specimens are services that an institute provides to the scientific community. If an institute commits to the concept of hosting a DNA collection (see "[2. Having the DNA collection within the institute's vision](#)"), it should be able to (independently) provide sufficient funds for the day-to-day operation of the DNA collection.

To create room for optimization of processes or working on a backlog of tasks, supplementary funding can be looked into. The most durable supplementary funding option is to request financial contributions from users of the collection: whether from depositors adding their DNA specimen to the DNA collection or requesters using specimens from the collection. This could be both internal and external depositors and requesters. If an institute decides to go down this path, sufficient consideration must be given to:

- Working out a clear framework for billing externals.
- Investing in internal workflows so that funding for archiving or loan requests is foreseen in project proposals.
- Investing in clear communication of the (new) financial responsibilities.

Differentiation in pricing can be based on multiple factors such as the number of specimens to archive or requested for loan, travel distance, package size and any special transport conditions of a loan shipment, type of handling needed for archiving (e.g. if the DNA specimen has to be transferred to new vials or not), etc. Asking for financial contributions can potentially hamper the intent to archive (for depositors), as well as to use the collection (for requesters) and therefore a

risk analysis needs to be conducted to monitor the potential impact of such a decision related to the outlined purpose of the DNA collection in its [scope](#).

For more inspiration for supplementary funding, information under the challenge of “lack of financial resources” can be consulted [Corrales, Luciano, Astrin \(2022\)](#).

Level 2: Optimise staffing. This goal is fulfilled once staffing needs are clearly outlined, feedback of the stakeholders with the power to induce changes considering staffing has been formally given (e.g. by a concrete action plan) and this feedback has been approved by all internal stakeholders.

To outline the staffing needs effectively, the following estimates need to be made:

- Day-to-day operations: needs
- Day-to-day operations: overview of the responsibilities and time of the current staffing
- Day-to-day operations: needed expertise versus available expertise
- Day-to-day operations: stability of the staffing
- Backlog operations: needs
- Backlog operations: action plan

More in depth information on the six listed bullet points is given below. Actively document this assessment and communicate identified needs to management and colleagues (if not yet involved in the exercise from the start). The goal of outlining the staffing, responsibilities, stability and expertise should help optimise staffing by dividing responsibilities more evenly, assigning responsibility, investing in training of staff or investing in hiring new staff.

The day-to-day operations should be listed, which are likely to align with the categories of this tool (see [Table 1](#)), but can be different for every DNA collection:

- Support future depositors with Data Management Planning ("[8. Involvement of suppliers](#)")
- Quality management ("[9. Quality management](#)")
- Legal checkpoints (minimally executing or supporting 2 of the 5 suggested checkpoints specifically for the DNA collection: the pre-deposit and collection-action checkpoints: "[10. Due diligence & legal compliance](#)")
- Supporting depositors with physical archiving of DNA specimens ("[11. Physical storage](#)")
- Maintenance of the Digital Management System ("[14. Digital management](#)")
- Data cleaning of the records in the Digital Management System ("[14. Digital management](#)")
- Data publication of the (DNA) specimen records ("[15. Unlocking the collection](#)")
- Physically retrieving specimens for loans ("[16. Loans](#)")
- Handling loan requests, including administration and logging ("[16. Loans](#)")
- Managing the strategy of the DNA collection ("[2. Having the DNA collection within the institute's vision](#)", '[3. Outlining a scope](#)")
- Managing the stability of the DNA collection ("[17. Stability](#)")

- Community engagement (["18. Community"](#))

To evaluate if the current amount of staff hours is sufficient for the day-to-day operations of the DNA collection, one can outline which staff is currently appointed to, or responsible for, the tasks listed above and for how much of their time, for starters. After this baseline situation has become clear, an estimation can be made on how much time is needed for the task to be executed properly (working days/months). Comparing them side-by-side will facilitate identification of current problems in staffing. Estimating the workload will depend greatly on the size and the current organisation of the DNA collection. For example, the more unlocked the DNA collection is (["15. Unlocking the collection"](#)), the more loan requests are expected ([Corrales, Luciano, Astrin, 2023](#)) and the higher the estimate will be for sufficient time to be allocated for the specific tasks related to handling loans.

To estimate if the necessary expertise is present in-house for the outlined day-to-day operations, begin by focusing on as-yet unassigned tasks when listing the current responsibilities, tasks that require a lot of time with the current amount of staff hours, or tasks that the current personnel signal as being too complicated for their current skill set.

To estimate stability, a brief risk assessment can be carried out. Firstly, note which staff members who have a responsibility involving the management of the DNA collection are employed on a permanent contract and which are dependent on external funds (e.g. project funding). If staff members with a permanent contract are linked to the above responsibilities, the DNA collection is more stable. Secondly, note who can back up staff members in case they are unavailable for a prolonged period of time (cf. the hit-by-a-bus-exercise of [Huxley et al. 2020](#)). If there is absolutely no overlap in skills, the DNA collection is less stable than if there are back-up staff for certain (key) responsibilities. Thirdly, note how information transfer occurs when there is staff turnover. Development of a staff succession plan that includes procedures to ensure the transfer of knowledge, is a worthwhile investment which can significantly improve the long-term management of the collection ([Crop Trust 2020](#)).

Similarly to the assessment(s) regarding day-to-day operations, there should also be an assessment of the staffing needed to work through any task backlog, if applicable. Backlog operations are expected to be related to the following goals/operations:

- Ensuring that the complete DNA collection is legally obtained (["10. Due diligence & legal compliance"](#)).
- Bringing the complete DNA collection into its best-fit physical storage (["11. Physical storage"](#)).
- Having the complete DNA collection in a public repository (["15. Unlocking the collection"](#)). Namely, to fulfil this goal, it is also necessary to have the complete DNA collection digitally inventorized.

For backlog operations, there should be a strategy in place and an appointed responsible person to make sure the plan is being executed. Different strategies are possible when it comes

to staffing needs for backlog operations:

- Available staff allocates time.
- New temporary staff via short-term contracts and projects.
- New temporary staff such as student projects or interns.

For more in-depth information and inspiration on how to tackle the challenge of “Personnel” and the identified gaps of “Lack of personnel and dedicated staff”, and “Lack of training programmes” consult [Corrales, Luciano, Astrin \(2021\)](#) and references therein.

Level 3: The DNA collection has a stable budget and team to ensure qualitative functioning, at least until the next re-evaluation of the DNA collection, as strategized (["7. Reusing the Key"](#)). This means that both the efforts to achieve [level 1](#) and [level 2](#) have resulted in this goal being met.

**Importance:** One of the major challenges for (DNA) collections is long-term thinking and planning, often made more difficult by an environment of short-term research projects and an ever-changing political and organisational context. A (DNA) collection needs to involve its management as much as possible in long-term strategizing, while continuing to invest in stipulating the importance, experience and needs of the collection. Here, the two most limiting factors are typically finances and expertise. Outlining how a repository can maintain its financial sustainability is key to its survivability ([ISBER, 2023](#)). Investing in knowledge training for specialised roles can also significantly improve the efficiency of operations ([ISBER, 2023](#)).

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific Key-category: “Stability”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions.

#### *DiSSCo Flanders DNA working group experience (2024) with “Stability”:*

One of the main concerns for many of the staff associated with the DNA collections in the working group, is that there are often very limited man-hours available to (help) manage the DNA or biodiversity biobank collection properly. Many of the “curators” frequently have other prior commitments, for example, their collection-related responsibilities are combined with the roles of researcher, lab technician, data manager, ... This often results in the staff who take on (part of) the responsibilities being overwhelmed by a to-do list which is simply too big to manage, as well as having to continue to work inefficiently as a consequence of no time being allocated for process optimization. Despite this overall impression being shared amongst many of the participants, some remarkable experiences have been shared, showing that even at a slow pace and with limited staff, step-by-step changes are possible. Such experiences inspired this tool; breaking down the task of maturing the DNA collection into smaller goals. Furthermore, during the existence of the working group, we noted that active participation in the working group events (["18. Community"](#)) also by-passes a small aspect of the staffing issue due to the exchange of information and ideas, and that joining a community can help to boost morale and

keep focus on what is practically possible. Next to step-by-step changes and inclusion in a community, another good example present in the working group, that ensures stability, at least for a longer timeframe (e.g. 10 years) was outlining responsibilities for example in a Memorandum of Understanding (MOU) between partner DNA or biobank collections that includes details on staffing, finances, workflows, etc.

**Return to [Table1](#).**



## 18. Community

**Category:** Community

**Level 1 goal:** Become part of at least one community.

**Level 2 goal:** Active participation in co-creation.

**Level 3 goal:** Engagement in maintaining a community.

**Explanation:** Another challenge considering (DNA) collection management, or perhaps even all types of roles and responsibilities, is keeping up-to-date with relevant information while only having limited time (and staff). To benefit from the progress being made on a global scale, joining a community is crucial.

Level 1: Become part of at least one community that links best to your DNA collection (or wider concepts, such as biodiversity biobank/molecular collection/...). Communities are at different organisational levels, for example at the regional, national or international level, and they can focus on different themes linked to managing DNA collections, such as strategic management, physical management, digital management, data publication, data standards, genetic research, etc. The goals of joining a community are:

- 1) To share information, experiences and best practices,
- 2) To keep track of relevant projects, literature and training,
- 3) To communicate needs, questions and problems from your DNA collection,
- 4) To co-create reviews, handbooks, guidelines, data or tools (per demand of certain needs).

Examples of communities that may be useful for DNA collections associates:

- DF-DNA wg: The DiSSCo Flanders DNA collection working group. This working group brings together different institutes in Belgium that want to mature their DNA collections by sharing experiences, information and searching for solutions for shared unresolved problems.
- [BeBOL: Belgian Network for DNA Barcoding](#). This is a research network that provides an integrated and multidisciplinary molecular, systematic and DNA barcoding task force.
- [IBOL](#): International Barcode of Life. This is a project with the mission to extend the geographic and taxonomic coverage of the barcode reference library: the [Barcode Of Life Data System \(BOLD\)](#).
- [SPNHC](#): the Society for the Preservation of Natural History Collections. This is an international organisation devoted to the preservation, conservation and management of natural history collections.
- [GGBN](#): the Global Genome Biodiversity Network. This is an international network of institutions that share an interest in long-term preservation of genomic samples representing the diversity of non-human life on Earth.
- [Museum genomics](#): Museum Genomics Community Webpage aims to allow transparent discussions among museum genomics researchers ([Card et al., 2021](#)).



- [sedaDNA Scientific Society](#): sedaDNA stands for sedimentary ancient DNA. The sedaDNA scientific society aims to transmit information about current sedaDNA research, promote best practices and increase collaborations between research groups.
- [ESBB](#): the European, Middle Eastern & African Society for Biopreservation and Biobanking. Mission: advance biosharing for a better world through mobilising, inspiring and educating the biobank community.
- [ISBER](#): International Society for Biological and Environmental Repositories. ISBER is a global biobanking organisation that creates opportunities for networking, education, and innovation. ISBER provides a community for harmonising approaches to emerging challenges in repositories, as well as fostering ideas to create new solutions.
- [TDWG](#): Biodiversity Information Standards. A non-profit organisation and a community dedicated to developing biodiversity information standards.
- [OBPS](#): Ocean Best Practices: a global initiative aimed at improving and standardising practices for collecting, analysing and sharing ocean data and information ([Pearlman et al., 2019](#)).

Some communities ask for a member fee, others are free - make sure to check the requirements on how to become involved (~become a member) in the community. Some communities are already longstanding, while others are expected to be temporary (e.g. linked to a specific project).

The first step is deciding which community or communities to join. It is also often possible to try out a certain community for an agreed period of time, to evaluate if there is enough merit to join the community for the particular context of the DNA collection at hand. This agreement has to be made with the internal stakeholders (“[1. Involving internal stakeholders](#)”) of said DNA collection. Alternatively, it can be decided to start a local community cf. the DiSSCo Flanders DNA collection working group, to temporarily join forces and invest in knowledge sharing. If taking part in the latter, it should be made clear what the goal and scope is of the community and it should be thoroughly evaluated if there is merit in creating a new community rather than joining an existing one.

A second step is deciding which of the internal stakeholders (“[1. Involving internal stakeholders](#)”) will follow up the chosen community or communities. Here, it is especially important to actively budget time for community involvement, as it often happens that “more important” or “more urgent” work is prioritised over community involvement.

A third step is deciding how information will flow to and from the community, as well as how you plan to engage with the community.

- For information flowing from the community to the DNA collection, it should be decided whom to inform, when and how. For example, it can be planned to debrief all the internal stakeholders of the DNA collection at an annual internal stakeholder meeting.
- For information flowing from the DNA collection to the community, it is advised that listing the relevant needs, questions and problems of your specific DNA collection to the wider

communities is made a recurrent item on the yearly quality meeting of the DNA collection ("[9. Quality management](#)"). This also means that the appointed link (i.e. person or persons) between the DNA collection and the chosen community (or communities) is responsible for communicating the needs, questions and problems via forum, during meetings, via mail conversation; and that there is follow-up and reporting back to the internal stakeholders of the DNA collection if there are answers.

- Engagement can be reading documentation provided by the community, participation in events, reading the reports of the events, delivering input in surveys, co-creation, ... This will depend on how the chosen community organises the involvement of its members.

Level 2: Active participation in co-creation. This could be embedded in community ([level 1](#)) or standalone initiatives. With 'co-creation', we mean any form of collaborative creation. With 'collaboration', we mean any form of working with people across different institutes, linked to different DNA collections. Collaborative efforts are expected to give a wider variety of use cases or experience to draw from, in comparison to "solo" efforts.

Co-creation can be in the format of reviews, handbooks, guidelines, data or tools that enhance any of the aspects of DNA collection management (~categories in [Table 1](#)). Some examples of co-creation are:

- Original research and data. Examples are comparative quality research on DNA storage methods, quality research on DNA extraction methods, quality research on field sampling methods, ...
- Reviews: literature reviews on any of the above research topics, these can be in the format of handbooks or guidelines.
- Surveys: this could be participation in filling in surveys - or even better - participation in constructing or evaluating a survey.
- Strategic tools. For example you can co-develop this current tool, by:
  - Sharing the experience of your DNA collection management (maturation) with the user-community in the format of "experiences" linked to the goals that you've achieved. It is very valuable for other DNA collections in the process of maturation and re-evaluation to read the experiences, good and bad, as well as the rationale of the made decisions of DNA collections that have a similar context.
  - Adding information to any of the appendices of the tool, where relevant and missing in the current version.
- Technological tools such as scripts or applications. Here, co-creation can be active feedback on open source tools on Github for example.

For the interpretation of this goal in the context of this tool: during the current usage of the maturation tool, you invest at least once in co-creation. Keep in mind that the last step of this maturation tool is to make a re-evaluation strategy ("[7. Reusing the Key](#)"), and therefore, depending on the outlined strategy, this strategy potentially involves using the tool again after a defined period of time. This would mean that, when the DNA collection is re-evaluated and the

tool is being used again, a new co-creation action should be undertaken. This aligns the evaluation of the DNA collection with undertaking action to engage (more) in creation of information and tools that can help global DNA collection maturation (i.e. balancing out giving and taking).

Level 3: Engagement in maintaining a community. For DNA collections that are (on their way to become) fully matured within their context, the goal here is to take on some form of ongoing activity/engagement in maintaining a community. This could be by having someone linked to the DNA collection taking responsibility for a defined role in the community. For example, being part of an executive committee or a task force. Another option could be having someone from the DNA collection engaged in organising or hosting an event, maintaining a website, managing social media presence, sponsorship (i.e. having a new member under your wing), ... The community engagement can be limited in time (i.e. for a defined period of time only), or ongoing continuous. It is recommended that the decision on engagement in community maintenance is discussed with all the [internal stakeholders](#), that the outcome (i.e. community, responsible, type of role, time period) is documented, and that sufficient time is allocated so that obtaining this goal does not strain day-to-day operations (["17. Stability"](#)).

**Importance:** The main reasons to join a community are to avoid isolation and inefficient practices, and to generate extra motivation to work on maturing the DNA collection. In [Corrales, Luciano, Astrin \(2023\)](#) the strength of drawing parallels between different biobanking communities for the basic biobanking processes is put forward as a worthwhile investment for biodiversity and environmental biobanks, it is for example expected that improvements originating from biobanks related to the better-funded human domain can benefit biobanks in other domains.

**Experiences:** This section allows writing down experiences (synonyms in this context: “stories”, “use-cases”) when tackling this specific Key-category: “Community”. Sharing experiences aims to inspire. It is encouraged to add new experiences in new versions.

*DiSSCo Flanders DNA working group experience (2024) with “Community”.*

*DiSSCo Flanders DNA working group experience (2024):* The DiSSCo Flanders DNA collection working group is a positive example of how a local community can make a difference: the group has shared ideas throughout meetings and mailings in 2021-2024, and together we have designed this Key (to be tested in a next phase). The main reason a local working group was created was to create a safe space for the members to speak about doubts or “things that are not working well”. The working group has an overview document in which its scope and goal are stated. The group also drew up internal guidelines for guests, new participants and invited speakers, and a code of conduct for meetings. The working group has a shared document space (for slides, literature, meeting reports, ..) and mailing address. After each meeting, an informal report is written for the passive members (reading only), members who could not attend the meeting, and to serve as a collective memory of the group. We remain critical regarding the

ongoing necessity of the group and will evaluate the need for its continuation at the last meeting of the working group in the course of the DiSSCo Flanders project (October 2024).

**Return to [Table1](#).**

## References

Abarenkov K, Andersson AF, Bissett A, Finstad AG, Fossøy F, Grosjean M, Hope M, Jeppesen TS, Kõljalg U, Lundin D, Nilsson RN, Prager M, Provoost P, Schigel D, Suominen S, Svenningsen C & Frøslev TG (2023) Publishing DNA-derived data through biodiversity data platforms, v1.3. Copenhagen: GBIF Secretariat. <https://doi.org/10.35035/doc-vf1a-nr22>.

Access to Biological Collection Data task group (2007) Access to Biological Collection Data (ABCD), Version 2.06. Biodiversity Information Standards (TDWG). URL: <http://www.tdwg.org/standards/115>

Borisenko A, Young R, Hanner R (2024) A lab-centric, workflow-based data management system for environmental DNA research. Research Ideas and Outcomes 10: e120483. <https://doi.org/10.3897/rio.10.e120483>

Card D, Shapiro B, Giribet G, Moritz C, Edwards S (2021) Museum Genomics. Annual Review of Genetics 55(1):633-59. <https://doi.org/10.1146/annurev-genet-071719-020506>

CETAF Legislations and Regulations Group (2019) CETAF Code of Conduct and Best Practice for Access and Benefit Sharing, Annex 5: Practice Guidance. Available at: <https://cetaf.org/resources/best-practices/>.

Corrales C, Leliaert F, Astrin JJ (2023) Chapter 1. Biobank Management. In: Corrales C, Astrin JJ (Eds) Biodiversity Biobanking – a Handbook on Protocols and Practices. Pensoft, Advanced Books, Sofia, pp. 13-17. <https://doi.org/10.3897/ab.e101876>

Corrales C, Leliaert F, Forrest L, Fulcher T, Poczai P, Haring E, Krukenhauser L, Thines M, Mulcahy D, Mackenzie-Dodds J, Paz Martín M, Sanmartín I, Ruiz Y, Diéguez-Urbeondo J, Vondráček D, Poczai P, Astrin JJ (2023a) Chapter 8. DNA. In: Corrales C, Astrin JJ (Eds) Biodiversity Biobanking – a Handbook on Protocols and Practices. Pensoft, Advanced Books, Sofia, pp. 126-136. <https://doi.org/10.3897/ab.e101876>

Corrales C, Poczai P, Mulcahy D, Mackenzie-Dodds J, Haring E, Paz Martin M, Haston E, Vondráček D, Astrin JJ (2023b) Chapter 2. Metadata and Data Management. In: Corrales C, Astrin JJ (Eds) Biodiversity Biobanking – a Handbook on Protocols and Practices. Pensoft, Advanced Books, Sofia, pp. 18-23. <https://doi.org/10.3897/ab.e101876>

Corrales C, Luciano S, Astrin JJ (2023) Biodiversity biobanks: a landscape analysis. ARPHA Preprints. <https://doi.org/10.3897/arphapreprints.e103105>

Crop Trust (2020) Seeds for Resilience Genebank reviews. URL: <https://www.croptrust.org/work/projects/seeds-for-resilience/#c7192>. Accessed on: 2024-08-19.

Cummins C, Ahamed A, Aslam R, Burgin J, Devraj R, Edbali O, Gupta D, Harrison PW, Haseeb M, Holt S, Ibrahim T, Ivanov E, Jayathilaka S, Kadhivelu V, Kay S, Kumar M, Lathi A, Leinonen R, Madeira F, Madhusoodanan N, Mansurova M, O’Cathail C, Pearce M, Peasant S, Rahman N, Rajan J; Rinck G, Selvakumar S, Sokolov A, Suman S, Thorne R, Totoo P, Vijayaraja S, Waheed Z, Zyoud A, Lopez R, Burdett T, Cochrane G (2021) The European nucleotide archive in 2021. *Nucleic Acids Research* 50 (D1): 106-110. <https://doi.org/10.1093/nar/gkab1051>

Deck J, Gaither MR, Ewing R, Bird CE, Davies N, Meyer C, et al. (2017) The Genomic Observatories Metadatabase (GeOMe): A new repository for field and sampling event metadata associated with genetic samples. *PLoS Biol* 15(8): e2002925. <https://doi.org/10.1371/journal.pbio.2002925>

de Mestier A, Mulcahy D, Harris DJ, Korotkova N, Long S, Häffner E, Paton A, Schiller EK, Leliaert F, Mackenzie-Dodds J, Fulcher T, Stahls G, von Rintelen T, Martín MP, Lücking R, Williams C, Lyal CHC, Güntsch A, Aronsson H, Castelin M, Pielach A, Poczai P, Ruiz-León Y, Sanmartin Bastida I, Thines M, Droege G (2023) Policies Handbook on Using Molecular Collections. *Research Ideas and Outcomes* 9: e102908. <https://doi.org/10.3897/rio.9.e102908>

Droege G, Barker K, Seberg O, Coddington J, Benson E, Berendsohn WG, Bunk B, Butler C, Cawsey EM, Deck J, Döring M, Flemons P, Gemeinholzer B, Güntsch A, Hollowell T, Kelbert P, Kostadinov I, Kottmann R, Lawlor RT, Lyal C, Mackenzie-Dodds J, Meyer C, Mulcahy D, Nussbeck SY, O’Tuama É, Orrell T, Petersen G, Robertson T, Söhngen C, Whitacre J, Wieczorek J, Yilmaz P, Zetzsche H, Zhang Y, & Zhou X (2016) The Global Genome Biodiversity Network (GGBN) Data Standard specification. Database 2016. <https://doi.org/10.1093/database/baw125>

Forsdick N, Wold J, Angelo A, Bissey F, Hart J, Head M, Liggins L, Senanayake D, & Steeves T (2023) Journeying towards best practice data management in biodiversity genomics. *Molecular Ecology Resources*. <https://doi.org/10.1111/1755-0998.13880>

Grosjean M, Høfft M, Gonzalez ML, Hahn A, Robertson T (2022) GRSciColl: What is the Global Registry of Scientific Collections and How to Contribute? *Biodiversity Information Science and Standards* 6: e93875. <https://doi.org/10.3897/biss.6.93875>

Grosjean M, Robertson T, Gonzalez ML, Høfft M, Hahn A (2023) Ongoing Work with the Global Registry of Scientific Collections. *Biodiversity Information Science and Standards* 7: e110932. <https://doi.org/10.3897/biss.7.110932>

Güntsch A, Groom Q, Ernst M, Holetschek J, Plank A, et al. (2021) A botanical demonstration of the potential of linking data using unique identifiers for people. PLOS ONE 16(12): e0261130. <https://doi.org/10.1371/journal.pone.0261130>

Guralnick R, Conlin T, Deck J, Stucky BJ, Cellinese N (2014) The Trouble with Triplets in Biodiversity Informatics: A Data-Driven Case against Current Identifier Practices. PLoS ONE 9(12): e114069. <https://doi.org/10.1371/journal.pone.0114069>

Harati MD, Williams RR, Movassaghi M, Hojat A, Lucey GM, Yong WH (2019) An introduction to starting a biobank. Methods in molecular biology (Clifton, N.J.), 1897: 7–16. [https://doi.org/10.1007/978-1-4939-8935-5\\_2](https://doi.org/10.1007/978-1-4939-8935-5_2)

Hardisty AR, Addink W, Glöckler F, Güntsch A, Islam S, Weiland C (2021) A choice of persistent identifier schemes for the Distributed System of Scientific Collections (DiSSCo). Research Ideas and Outcomes 7: e67379. <https://doi.org/10.3897/rio.7.e67379>

Hardisty AR, Ellwood ER, Nelson G, Zimkus B, Buschbom J, Addink W, Rabeler RK, Bates J, Bentley A, Fortes JAB, Hansen S, Macklin JA, Mast AR, Miller JT, Monfils AK, Paul DL, Wallis E, Webster M (2022) Digital Extended Specimens: Enabling an extensible network of biodiversity data records as integrated digital objects on the Internet. BioScience72: 978–987. <https://doi.org/10.1093/biosci/biac060>

Huxley R, Quaiser C, Butler CR, Dekker RW (2020) Managing Natural Science Collections: A Guide to Strategy, Planning and Resourcing (1st ed.). Routledge. <https://doi.org/10.4324/9780429426599>

ISBER (2023) ISBER Best Practices: Recommendations for Repositories, Fifth edition. Editor-in-Chief: Snapes E. 195 p. Free download: <https://www.isber.org/page/BPR>

Luger T, Maes S, Milkers A, Bergmans B (2021) Een collectieplan schrijven? Dat doe je zo! Editor: Versloot A. 88 p. OBT, Leiden. Can be downloaded here: [Een collectieplan schrijven? Dat doe je zo! | Publicatie | Rijksdienst voor het Cultureel Erfgoed](#)

Norton, B., Woodburn, M., Grant, S., Jones, J., Webbink, K., Vincent, S., Trekels, M., Buschbom, J., Baskauf, S. J., & Biodiversity Information Standards (TDWG). Latimer Core (LtC) GitHub Repository [Computer software]. <https://github.com/tdwg/ltc> (accessed 2024-04-09).

Pearlman J, Bushnell M, Coppola L, Karstensen J, Buttigieg PL, Pearlman F, Simpson P, Barbier M, Muller-Karger F, Munoz-Mas C, Pissierssens P, Chandler C, Hermes J, Heslop E, Jenkyns R, Achterberg E, Bensi M, Bittig H, Blandin J, Bosch J, Bourles B, Bozzano R, Buck JH, Burger E, Cano D, Cardin V, Llorens MC, Cianca A, Chen H, Cusack C, Delory E, Garelló R, Giovanetti G, Harscoat V, Hartman S, Heitsenreth R, Jirka S, Lara-Lopez A, Lantéri N,



Leadbetter A, Manzella G, Maso J, McCurdy A, Moussat E, Ntoumas M, Pensieri S, Petihakis G, Pinardi N, Pouliquen S, Przeslawski R, Roden N, Silke J, Tamburri M, Tang H, Tanhua T, Telszewski M, Testor P, Thomas J, Waldmann C, Whoriskey F (2019) Evolving and Sustaining Ocean Best Practices and Standards for the Next Decade. *Frontiers in Marine Science* 6 URL: <https://doi.org/10.3389/fmars.2019.00277>

Petersen, M., Glöckler, F., Kiessling, W., Döring, M., Fichtmüller, D., Laphakorn, L., Baltruschat, B., and Hoffmann, J.: History and development of ABCDEFG: a data standard for geosciences, *Foss. Rec.*, 21, 47–53, <https://doi.org/10.5194/fr-21-47-2018>, 2018.

Santi I, Casotti R, Comtet T, Cunliffe M, Koulouri Y, Macheriotou L, Not F, Obst M, Pavludi C, Romac S, Thiebaut E, Vanaverbeke J, Pade N (2021) European Marine Omics Biodiversity Observation Network (EMO BON) Handbook. (Version 1.0). Paris, France, EMBRC-ERIC, 63pp. DOI: <http://dx.doi.org/10.25607/OBP-1653>

Sayers EW, Cavanaugh M, Clark K, Pruitt KD, Schoch CL, Sherry ST, Karsch-Mizrachi I (2021) GenBank. *Nucleic acids research* 49 (D1): 92-96. <https://doi.org/10.1093/nar/gkaa1023>

Schoch CL, Ciuffo S, Domrachev M, Hotton CL, Kannan S, Khovanskaya R, Leipe D, Mcveigh R, O'Neill K, Robbertse B, Sharma S, Soussov V, Sullivan JP, Sun L, Turner S, Karsch-Mizrachi I (2020) NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database* <https://doi.org/10.1093/database/baaa062>

Ståhl G, Aleixo A, Hyvärinen M-T, Kuusijärvi A, Myllys L, Väre H, Sihvonen P, Luttinen A, Kröger B, Oinonen M, Schulman L, Juslén A (2021) The Genomic Resources Collection Policy of the Finnish Museum of Natural History. *Research Ideas and Outcomes* 7: e61664. <https://doi.org/10.3897/rio.7.e61664>

Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, et al. (2012) Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. *PLoS ONE* 7(1): e29715. <https://doi.org/10.1371/journal.pone.0029715>

Woodburn M, Buschbom J, Droege G, Grant S, Groom Q, Jones J, Trekels M, Vincent S, Webbink K (2022) Latimer Core: A new data standard for collection descriptions. *Biodiversity Information Science and Standards* 6: e91159. <https://doi.org/10.3897/biss.6.91159>