Digital Sequence Information: An Evidence Review

Final Report

14 August 2020

ICF makes big things possible

ICF is a global consulting and technology services provider with more than 7,000 professionals focused on making big things possible for our clients. We are policy specialists, social scientists, business analysts, technologists, researchers, digital strategists and creatives. Since 1969 government and commercial clients have worked with ICF to overcome their toughest challenges on issues that matter profoundly to their success. Our five core service areas are described below. Engage with us at icf.com.

Research + Analyse

Our teams delve deep into critical policy, industry and stakeholder issues, trends, and behaviour. By collecting and analysing data of all kinds, we help clients understand the current landscape clearly and plan their next steps wisely.

Assess + Advise

With equal parts experience and dedication, our experts get to the heart of the issue—asking all the right questions from the start. After examining the results and evaluating the impact of research findings, we counsel clients on how to best navigate societal, market, business, communications, and technology challenges.

Design + Manage

We design, develop and manage plans, frameworks, programmes, and tools that are key to each client's mission or business performance. These solutions often stem from our analytics and advice.

Identify+Implement

Our experts define and put into place the technology systems and business tools that make our clients' enterprises more effective and efficient. We deploy standard or customised methodologies based on the business context.

Engage

Realising the promise of the digital revolution requires foresight and heightened understanding. Both are baked into the solutions-focused engagement work that runs through all we do.

Digital Sequence Information: An Evidence Review

Final Report

A report submitted by ICF Consulting Services Limited
in association with

Elta Smith, University of Strathclyde

Date: 14 August 2020

Job Number J30302432

ICF Consulting Services Limited
Riverscape
10 Queen Street Place
London
EC4R 1BE

T +44 (0)20 3096 4800

www.icf.com

Document Control

Document Title	Digital Sequence Information: An Evidence Review
Job No.	J3030243230302432
Prepared by	Mar Maestre (ICF), Simona de Paolis (ICF), Stephanie Switzer (University of Strathclyde), Elta Smith (Independent)
Checked by	Rupert Haines (ICF), Elisa Morgera (University of Strathclyde)
Date	14 August 2020

Contents

Executive summary

Introduction

The Convention on Biological Diversity (CBD), adopted in 1992, is an international treaty that recognises the sovereignty of states over their natural resources and includes the authority to determine access to genetic resources. One of the CBD's primary objectives is to ensure 'the fair and equitable sharing of the benefits' derived from the utilisation of genetic resources and traditional knowledge associated with genetic resources. Access to and utilisation of genetic resources is currently covered under the Nagoya Protocol (NP), adopted by the CBD parties in October 2010, and entered into force in 2014.

Digital sequence information (DSI) refers to digital representations of certain aspects of genetic resources such as nucleotide or amino acid sequence data, and related information that is downloadable from databases around the world. DSI is regularly used as part of new research without the need for any new access to physical material. Divergent positions on whether DSI should be included in the NP have emerged over the past few years. These issues, including how benefits from DSI may be shared, are being discussed in the context of the CBD. To date, there is no agreement on the most appropriate way forward.

In this context, ICF with partners Elta Smith and the University of Strathclyde's Centre for Environmental Law and Governance, were commissioned by Defra to conduct a study to bring together evidence that can help inform CBD negotiations on DSI.

The study was based on a rapid review of key literature and a consultation with 14 UK stakeholders. The study was completed between March and July 2020. It addressed the following questions:

- What is DSI?
- How is DSI generated?
- How is DSI accessed?
- What are the characteristics of the DSI stakeholder landscape?
- What are the options for DSI benefit sharing?[1]

DSI evidence review

DSI is an umbrella term adopted in December 2016 by the CBD[2] (CBD, 2016) that refers to (amongst other terms): "genetic sequence data", "nucleotide sequence data", "nucleotide sequence information", and "genetic sequences". The precise scope and meaning of the term 'DSI' is debated, so the term is currently used as a 'placeholder' in international discussions until a final agreement on the term and its scope is reached. The definition(s) adopted in international policy discussions will have implications for how (if any) benefit-sharing options may be introduced and governed.

A study commissioned by the CBD (Houssen et al., 2020) developed four groupings to define DSI, based on technical and scientific distinctions that relate to the distance between the

[1] The study does not explore the question of whether benefits should be shared or not. Rather, it explores what the potential benefit sharing policy options could be.

[2] Decisions XIII/16 and NP-2/4 taken during the thirteenth meeting of the Conference of the Parties to the Convention on Biological Diversity (CBD), and the second meeting of the Conference of the Parties serving as the meeting of the Parties to the Nagoya Protocol on Access and Benefit-sharing, respectively.

genetic material itself and the degree of analysis or processing of information about that material (the flow of data and information). ‘Group 1’ represents the narrowest concept, including only DNA and the transcription process associated with it. ‘Group 4’ represents the broadest definition, including all of the underlying genetic resources used, and all data and information generated (extending to behavioural data, information on ecological relationships and traditional knowledge).

DSI is generated by transforming the DNA of a genetic resource – whether obtained from a natural source or developed artificially – through ‘basic’ research. DSI can undergo different processes to get to its final use. The same type of sequence can be used for research and conservation or applied in commercial research. DSI can be seen as a tool that allows users to obtain something – which may be the diagnosis of new pathogens in animals or plants, food safety screening, investigating the relationships within and between species, or to support the development of commercial products (Sirakaya, 2019).

DSI is an easily shareable resource and requires multiple users to process high volumes of digital data in order to maximise its societal use-value. The large volumes of data being accessed, shared and used on a daily basis mean that DSI databases and users often cannot identify where a resource originally came from; who they shared it with or received it from; or at what point value was added.

Some benefit sharing options would require tracing the use of genetic resources back to its origin, identifying the parties engaged in creating the DSI output, and/or locating where value was added in the process. To track and trace a resource, one must be able to trace back to origin upon commercial use, or to track the information of how it was shared and used along the value chain. However, robust traceability of the broad array of DSI is considered problematic, hence these options would be very challenging to implement,

DSI is used both in basic research and to create new products or processes. It is difficult to differentiate between commercial and non-commercial DSI users aiming to develop new products (e.g. new vaccines) or processes (e.g. new technologies to process DSI more efficiently). The main sectors that use DSI in the UK are:

- The taxonomy and conservation sectors, which undertake basic research and include database managers.
- The biotechnology field, which provides next-generation sequencing technologies that allow other sectors to develop new products and processes.
- Other sectors including healthcare, agriculture technology, and industrial biotechnology.

Monetary and non-monetary benefits are accrued through generating, sharing and processing DSI. As DSI value comes from the processing of large volumes of data multiple times by multiple stakeholders, it is challenging to determine with legal certainty the extent of the value added by each party. Furthermore, assessing the total value of DSI is difficult, even when it is utilised in final products. As a result, stakeholders have different perspectives on how benefits should be shared in a fair and equitable manner.

Benefit sharing options

At an international level, the first time that stakeholders discussed ‘ways forward’ on DSI with a particular focus on benefit-sharing options, was during the Global Dialogue on Digital Sequence Information on Genetic Resources, that took place in South Africa in 2019.

This current study took all options discussed during this event as the starting point and has developed them further based on desk research and the experience of our technical advisors at the University of Strathclyde. The options discussed are:

- Option 0: Exclusion of DSI from the scope of the CBD and Nagoya
- Option 1: Nagoya – bilateral. This option would effectively embody the *status quo* under the Nagoya Protocol.
- Option 2: Open Access – bilateral. Under this option, benefit-sharing would not be specified by the Nagoya Protocol, but rather terms and conditions would be attached to the origin country of the DSI used or uploaded. This would require country tags to accompany uploaded DSI so users can direct monetary benefits back to the provider country.
- Option 3: Multilateral. The next set of options would be agreed by all CBD parties, with monetary benefits being accrued through a multilateral fund for subsequent distribution among Parties or to particular projects to support the objectives of the CBD. The sub-options could be:
  - Option 3(a): Subscription fee tied to database access, generating funds for monetary benefit-sharing. Such a fee could be payable upon access to databases hosting DSI and chargeable on a one-off basis. As an alternative, a subscription fee could be introduced.
  - Option 3(b): A tax, fee, levy could be considered as a way of generating funds for monetary benefit-sharing linked to DSI
  - Option 3(c): Payment upon commercialisation, likely attached to intellectual property such as the grant of a patent - derived from a successful innovation based on DSI.
  - Option 3(d): Voluntary donations, sought from a range of actors, from states to other actors (including the biotech sector), which could then be disbursed multilaterally in accordance with defined criteria focusing on conservation and sustainable use.
- Option 4: Open Access – Capacity Development. Under this option, which to an extent reflects the current *status quo*, benefit-sharing obligations would be required for non-monetary benefits (such as capacity development or technology transfers), attached to the use of DSI. Some stakeholders already engage in this type of non-monetary benefit sharing activities, but under this option, these arrangements would be mandatory.

Stakeholder consultation

The study team completed a multicriteria assessment of the options based on information gathered through stakeholder consultation. The consultation design took into account the following preferences, as established through study project management discussions, which were:

- Maximise the number of potential benefit-sharing options discussed with stakeholders.
- Prioritise gathering information on the potential options rather than the status quo.
- All interviewees should discuss the same options, to avoid evidence gaps.

There was a trade-off between the number of options that could be discussed and the number of questions that could be asked for each option. Based on that, and the preferences stated above, the approach taken maximised the number of options being discussed.

The final selection of options from the long-list presented above were:

- Option 3(a): Multilateral - subscription fee tied to database access
- Option 3(c): Multilateral - payment upon commercialisation

- Option 3(b) and option 3(d) combined: Multilateral - tax, fee or levy; and/or voluntary donations
- Option 4: Open Access – Capacity Development

Stakeholder views were sought for each option according to the following assessment criteria:

- Viability: technical and legal viability
- Impacts: anticipated costs, benefits and overall competitiveness
- Support: overall support for the option

Stakeholder selection was based on a mapping of the key UK stakeholder groups involved in generating, sharing and using DSI. These were the UK stakeholder groups that will likely be impacted by any change in the DSI policy landscape. The types of stakeholders consulted were:

- DSI regulators and related government agencies.
- DSI data repositories.
- Commercial DSI users, such as pharmaceuticals, agriculture and food industry, biotech, etc.
- Non-commercial users of DSI, such as universities and research institutions.
- Other experts, e.g. patents/IP.

Assessment of the benefit sharing options

The study assessed the options based on the criteria presented above. A precise assessment of the impacts of the proposed options was not possible in this study, in part because all the options presented require significant further specification and in part due to the lack of agreement on a definition of DSI and limited available information on the value of DSI use for individuals, organisations and society. Nonetheless, some general conclusions were drawn.

While none of the options assessed in this study were considered to address all the needs and concerns of stakeholders consulted in relation to accessing and using DSI, consultees suggested that combinations of these options could be more favourable.

The majority of those consulted believed that non-monetary benefits (Option 4) should be part of any agreement, but also that this option could not be considered a viable option on its own. An absence of monetary benefit-sharing from any potential solution would also be unlikely to satisfy the demands of certain biodiversity-rich countries that benefits arising from DSI are shared fairly and equitably.

Further, all consultees considered that any option discussed must be in the form of a multilateral approach to benefit-sharing, agreed by all countries, which prioritises open access to DSI. Consultees were not in favour of adapting the bilateral ABS regime currently applicable under the NP to include DSI. Further, any funds raised through monetary benefit sharing should clearly be destined to support conservation, in line with CBD objectives.

With regards to the multilateral options discussed, the option of payment upon commercialisation (3c), while perceived as fair, was seen as very challenging to implement. The options of payment on access (3a) and a tax, fee or levy (3b and 3d combined) were the least supported options as interviewees felt that payment levels seemed arbitrary and could lead to a large increase in costs for some organisations as well as hinder research and innovation. All these options were considered to be very difficult to implement.

1 Introduction

The Convention on Biological Diversity (CBD), adopted in 1992, is an international treaty that recognises the sovereignty of states over their natural resources and includes the authority to determine access to genetic resources. One of the CBD's primary objectives is to ensure 'the fair and equitable sharing of the benefits' derived from the utilisation of genetic resources and traditional knowledge associated with genetic resources. Access and utilisation of physical genetic resources is currently covered under the Nagoya Protocol (NP), adopted by the CBD parties in October 2010, and entered into force in 2014.

Digital sequence information (DSI) refers to digital representations of certain aspects of genetic resources such as DNA or amino acid sequences often available from public databases. DSI is regularly used as part of new research without the need for any new access to physical material. Sequencing and digital storage/sharing of sequence data for research purposes has been taking place for over 30 years. Currently, DSI is not within the scope of the CBD or Nagoya. However, divergent positions on whether DSI should be included in the NP have emerged over the past few years. This has led to concern from some stakeholders that the benefits of genetic material are not being shared fairly and equitably.

These issues, including how benefits from DSI may be shared, are being discussed in the context of the CBD. To date, there is no agreement on the most appropriate way forward.

In this context, ICF with partners Elta Smith and the University of Strathclyde's Centre for Environmental Law and Governance, were commissioned by Defra to conduct a study to bring together evidence that can help inform CBD negotiations on DSI. The study was based on a rapid review of key literature and a consultation with UK stakeholders, which were used to respond to the following research questions:

- What is DSI?
- How is DSI generated?
- How is DSI accessed?
- What are the characteristics of the DSI stakeholder landscape?
- What are the options for DSI benefit sharing?[3]

The report is structured as follows:

- Study approach (section 2)
- Defining DSI (section 3)
- DSI generation, access, and valuation (section 4)
- DSI stakeholders in the UK (section 5)
- Benefit sharing options (section 6)
- Assessment of the benefit sharing option (section 7)
- Conclusions (section 8)

[3] The study does not explore the question of whether benefits should be shared or not. Rather, it explores what the potential benefit sharing policy options could be

2 Study Approach

This section summarises the study approach, including the method for the desk research and stakeholder consultations.

2.1 Desk review

A targeted approach was followed to identify a set of priority documents to be reviewed. The documents were identified from the following sources:

- Core literature identified in the ITT.
- A selection of priority references drawn from the core literature.
- Other key literature known to Defra and the study team.
- A rapid and targeted search of journal databases and a standard web search engine (e.g. Google) to identify additional scientific and grey literature that may not be included in the initial selection above.

The literature was selected based on the relative importance of the document to answer the research questions (section 1). The desk research was undertaken during March and April 2020. Over 30 documents were reviewed, including:

- Studies on DSI commissioned by the Executive Secretariat of the CBD.
- Review of the submissions to the CBD by stakeholders in 2017/2018 and 2019/2020.
- Over 20 recent (2015 to 2020) academic articles on the topic.
- Other relevant documents.

2.2 Stakeholder consultation

Stakeholder consultation followed the desk review and consisted of 14 interviews with UK DSI stakeholders. Data collection was focused on an assessment of selected options for DSI benefit sharing.

The interview design took into account the following, as established through study project management discussions:

- Maximise the number of potential benefit-sharing options discussed with stakeholders.
- Prioritise gathering information on the potential options rather than the status quo.
- All interviewees should discuss the same options, to avoid evidence gaps.

There was a trade-off between the number of options that could be discussed and the number of questions that could be asked for each option. Based on that, the approach taken maximised the number of options being discussed. Stakeholder opinion was sought for a limited number of option assessment criteria. The criteria included were viability (including technical feasibility and legal coherence and conformity), impacts (including costs, benefits, competitiveness), and overall support for each option.

The stakeholder selection was based on a mapping of the key stakeholder groups in the UK involved in generating, sharing and using DSI (section 5). These are the UK stakeholder groups that will likely be impacted by any change in the DSI policy landscape. The organisations within the sectors were identified based on the stakeholder's:

- use and reliance on DSI;
- knowledge and insight of developments in the UK to provide context to the major stakeholder groups likely to be impacted by changes in the DSI policy landscape;
- previous engagement with Defra.

The final set of stakeholders interviewed are summarised in the Table 2.1 below.

Table 2.1 Stakeholders interviewed

Stakeholder group	Number of interviews
DSI regulators and related government agencies in the UK	1
DSI data repositories: representing at least one large and one smaller database	3
Commercial DSI users, such as pharmaceuticals, agriculture and food industry, biotech, etc.	3 representing the pharmaceuticals industry 1 representing the agriculture industry 2 representing the bioindustry
Non-commercial users of DSI, such as universities and research institutions	4
Other experts, e.g. blockchain, patents/IP	1

2.3 Study limitations

The following limitations to this report should be recognised:

- Lack of available information – there is ongoing global debate about core DSI issues – such as an appropriate definition – and understanding of the potential options is only recently emerging. The topic has only relatively recently been the subject of academic research. As such there is limited consistent and reliable published information on DSI, its use, access and value.
- The team used a targeted search of grey materials to fill in evidence gaps, particularly for the section that maps and describes the key DSI stakeholders (section 5). This evidence is limited for some of the sectors described.
- The stakeholder consultation only collected a limited level of information on the baseline situation – through an introductory question. Meaningful further elaboration of the baseline, beyond that completed with the desk research, was not possible.
- The stakeholder consultation sample was balanced across different stakeholder groups (as shown in Table 2.1). However, it draws on a small, non-random sample and may therefore be subject to biases.
- Section 6 provides a brief overview of each available benefit sharing option in general terms, based on desk research only. These descriptions were shared with stakeholders during the consultation phase.
- The options were described in general terms, given the lack of consensus in the ongoing policy discussions on the topic. This made it difficult to collect specific data on the potential impacts of the options.

3 Defining DSI

This section responds to the first research question: How is DSI defined? What are the predominant definitions and what are the main points of different between them?

The section first introduces what DSI refers to and the debates around its concept. It then moves on to briefly explain why DSI is currently not part of the Nagoya Protocol; and finally describes what the policy implications of the terminology debate on DSI have for any potential benefit sharing agreement to come.

3.1 Background to DSI

The Convention on Biological Diversity (CBD) is an international treaty that recognises the sovereignty of states over their natural resources and includes the authority to determine access to genetic resources, adopted in 1992. The CBD has three objectives; biodiversity conservation, sustainable use of biological resources, and fair and equitable sharing of the benefits derived from the utilisation of genetic resources.

'Genetic resources' are defined by Article 2 of the CBD as "genetic material of actual or potential value" while the term 'genetic material' is defined as "any material of plant, animal, microbial or other origin containing functional units of heredity" (CBD and WHO, 2018).

The NP on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization (ABS) is an international agreement adopted in October 2010 as a supplementary agreement to the CBD. Its purpose is to ensure the sharing of benefits arising from the utilization of genetic resources in a fair and equitable way. The NP also covers traditional knowledge associated with genetic resources that are covered by the CBD and the benefits arising from its utilisation. The NP sets out core obligations for its contracting Parties to take measures in relation to access to genetic resources, benefit-sharing and compliance. It entered into force on 12 October 2014. The UK is a Party to the Nagoya Protocol, and thus has a legal responsibility to implement it.

Access to and utilisation of genetic resources (including plants, animals and microorganisms) is covered by the NP. However, DSI – which is the digital representation of certain aspects of genetic resources such as DNA or amino acid sequences – is not. The issue of access and benefit sharing from DSI on genetic resources was introduced in 2016 in decisions within the CBD and the NP.[4] Some stakeholders have expressed concerns that DSI could undermine the ABS system established by the CBD and NP, while others claim that benefit sharing from the application of DSI is not being realised. In response, the CBD initiated a science and policy-based process on this topic.[5]

DSI is also being discussed in other international fora. The Food and Agriculture Organisation (FAO) is currently discussing DSI in the context of the ITPGRFA; the WHO in the context of the PIP Framework; and the United Nations Convention on the

[4] COP decision 14/20 and COP-MOP decision NP-3/12.

[5] COP decision 14/20 established a science and policy-based process on digital sequence information on genetic resources for further work on this topic. The process entails the submission of views and information, the commissioning and peer review of four studies as well as work by an ad hoc technical expert group (AHTEG).

Law of the Sea in the context of developing the International Instrument on Biodiversity Beyond National Jurisdiction.

3.2 Definitions and main debates

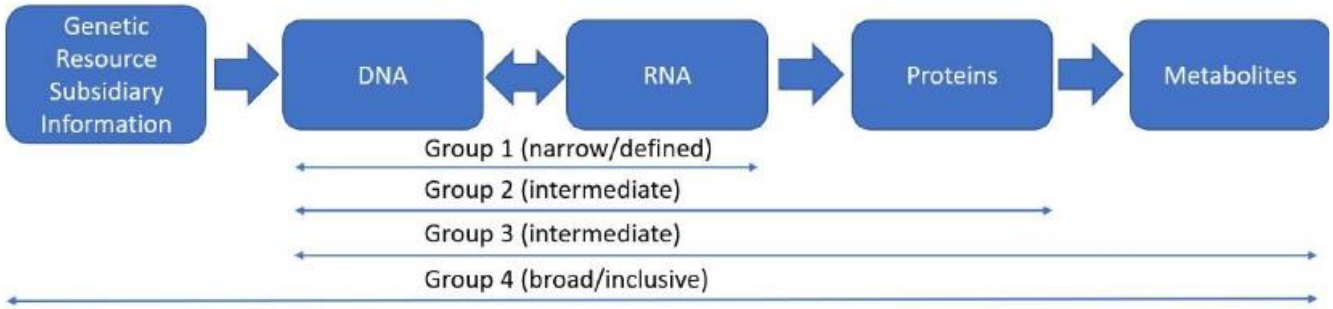
DSI is an umbrella term used to refer to DNA sequence data and related information that is downloadable from databases around the world (Hiemstra et al., 2019). These databases contain DNA sequences, RNA sequences, and protein sequences, as well as annotations, metadata and other relevant information necessary to use and process DSI (Laird et al., 2018). DSI can encompass, amongst other terms: “genetic sequence data”, “nucleotide sequence data”, “nucleotide sequence information”, and “genetic sequences”. The precise scope and meaning of the term DSI is debated, however it has been adopted as a ‘placeholder’ in international discussions until a final agreement on the term and its scope is reached.

In 2018, discussions at the Ad Hoc Technical Expert Group (AHTEG) of the CBD (AHTEG, 2018) on DSI did not come to an agreement on the definition for DSI. At this meeting experts identified various types of information that may be included as part of the DSI definition:

1. The nucleic acid sequence reads and the associated data.
2. Information on the sequence assembly, its annotation and genetic mapping. This information may describe whole genomes, individual genes or fragments thereof, barcodes, organelle genomes or single nucleotide polymorphisms.
3. Information on gene expression.
4. Data on macromolecules and cellular metabolites.
5. Information on ecological relationships, and abiotic factors of the environment.
6. Function, such as behavioural data.
7. Structure, including morphological data and phenotype.
8. Information related to taxonomy.
9. Modalities of use.

To help clarify this further, AHTEG commissioned the study “Digital Sequence Information on Genetic Resources: concept, scope and current use”. The study developed four groupings based on technical /scientific distinctions that relate to the distance between the genetic material itself and the degree of analysis/processing of information about that material (the flow of data and information). Figure 3.1 shows this flow and the groupings (Houssen et al., 2020).

Figure 3.1 Definition groupings based on DSI flow of data and information



Source: Houssen et al., 2020

The four proposed groups range from a narrow scope and proximity to the genetic resource to a broad form of inclusion:

- Group 1, which is the narrowest concept, including only DNA and the transcription process associated with it.
- Group 2, including intermediary steps that include transcription and translation processes, including protein sequences.
- Group 3, which extends the definition to also include metabolites and biochemical pathways, thus comprising information associated with transcription, translation and biosynthesis.
- Group 4, the broadest definition that includes all the underlying genetic resources used, and all data and information generated. It also extends to behavioural data, information on ecological relationships and traditional knowledge (including upstream and downstream subsidiary information).

According to the study, "scientifically, groups 1-3 are all based on the molecular structure of macromolecules and small molecules, the information they carry and information associated with their acquisition. Group 4 also includes information that is not related to molecular structure or information associated with their acquisition" (Houssen et al., 2020: 32).

Submissions to the CBD from organisations and countries in 2018 and 2019 indicate that there is no agreement on the definition of DSI. Some advocate for the broadest definition (e.g. the submission from the African Group of Negotiators), whereas others support a narrower definition (e.g. the submission from the Australian government). These differences in terminology reflect divergent concepts of what should or could fall within the scope of the regulatory framework on ABS and related national laws (Hiemstra et al., 2019).

Different international fora also utilise different definitions. For example, Welch et al. (2017), in a study commissioned by the ITPGRFA, elected to use the term "sequence data". Another background paper for the FAO (Manzella, 2016), uses the term genetic information (processed sequenced data), differentiating between 'data' (group 1 within the definition groupings above) and 'information' (group 2 or 3). Finally, the WHO uses the term genetic sequence data defined as 'nucleotides found in a molecule of DNA or RNA…contain[ing] the genetic information that determines the biological characteristics of an organism or a virus" (WHO, 2018: 1).

The different definitions highlight that debate on terminology focusses on the extent to which a DSI definition should include processing activities performed with that data (DNA/RNA) to generate information in different formats, media, and shapes, by data producers, curators, and users (also called 'information'). As explored in section 3.4, the different definitions have policy implications as to what benefit sharing options are more appropriate.

3.3 The debate around DSI in the context of the Nagoya Protocol

The NP is based on the fundamental principles of ABS enshrined in the CBD. These principles are based on potential users of genetic resources obtaining the prior informed consent of the country in which the genetic resource is located before accessing the resource, and negotiating and agreeing on the terms and conditions of resource access and use through the establishment of mutually agreed terms. Such an agreement includes the sharing of benefits arising from the use of the resource with the provider, as a prerequisite for access to the genetic resource and its use. Countries acting as providers of genetic resources should reciprocally provide fair and non-arbitrary rules and procedures for access to their genetic resources (Aubrey, 2019).

The way DSI is stored and managed has seen rapid changes since 2015. The rise of technologies to generate and sequence DSI has led to a decrease in processing costs, and an increase of data sharing to databases. Stakeholders can download DNA sequences from public databases and reconstruct the DNA, then utilise it as one might a gene extracted directly from an organism. This means that anyone, including researchers and companies, can and do use publicly accessible data for commercial and non-commercial purposes. The open access nature of DSI has helped facilitate these arrangements and a rapid increase in the use of DSI.

These changes have led to the ongoing debate around how DSI should be governed and concerns about how DSI related benefits are being shared. Lawson and Rourke (2016) argue that open access to DSI can undermine the NP, by making data available without any binding terms and conditions. In addition, benefit-sharing obligations have previously been linked to the physical genetic resource (or raw material). With open access DSI, this is also undermined, as there is no need to use the physical resource anymore, since its content can be processed and exchanged in its own right. Some CBD parties agree with this and argue that DSI should fall within the scope of the NP (e.g. Ethiopia, Kenya or South Africa). Some of these countries have domestic measures in place that cover ABS for DSI but prefer international legal measures to share DSI benefits more equitably.

Alternatively, many stakeholders believe that including DSI within the NP would have a negative effect on the use and benefits derived from DSI, both for commercial and non-commercial purposes. In their view, inclusion would affect the existing open access arrangements for DSI (see section 4 where DSI access is discussed).

DSI researchers and users claim that open access is required for non-commercial scientific research. For example, a submission from UK non-commercial

stakeholders[6] explains that the ongoing DSI open access model "supports scientific research and environmental management globally; […] supports scientists in all countries, noting that research results become more reliable with increasing amount of data on which they are based". Their concern is that including DSI under the NP could undermine this open access model. Other CBD parties have maintained a similar position, claiming that DSI falls outside the scope of the NP (e.g. the European Union and Japan). The European Union considers that DSI is not equivalent to genetic resources, and as such, it should not fall under the NP, nor be subject to PIC[7].

Laird et al. (2020) argue that the NP model, based on bilateral transactions, is a poor fit with current scientific practices when applied to DSI. However, they also argue that "not to capture DSI would mean leaving a massive loophole in the ABS endeavour… [since] users might simply sidestep benefit-sharing obligations by digitising genetic resources and synthesising [required components using openly available DSI]" (Laird et al., 2020: 1200). Rourke et al. (2020) offer numerous examples of this occurring.

Vogel et al. (2019) explain that while this DSI loophole exists, the CBD objective of fair and equitable benefit-sharing can be undermined and frustrating the other two CBD objectives of sustainable use and conservation. The loophole can also result in individual countries imposing more restrictive access policies, with a knock-on effect on scientific innovation. This would also put at risk the first two CBD objectives and, more generally, it can put public goods such as food security at risk (Sirakaya, 2019). These implications are further discussed in section 4.

3.4 Implications for adopted definition(s)

This section has introduced the debate around DSI terminology. The different definitions have policy implications for how benefit sharing options will be governed. The main implications that the different definitions have for benefit sharing can be summarised as follows:

- As the definition of DSI becomes broader in context (from group 1 to group 4) and it is more the result of cognitive processes, the number of entities that might claim rights or have to share benefits will change. This will have policy implications for how benefit sharing options are governed.
- The definition also has implications for the process that needs to be governed under the benefit sharing agreement.
  - A narrow definition (e.g. group 1 definition) could ease the implementation of a benefit-sharing option as, to share benefits in a fair and equitable manner, one needs to be able to determine who contributed what to the final output that contains DSI. .
  - However, others argue that using a narrow definition (e.g. group 1 definition) would not allow natural resource-rich countries to benefit from a potential benefit-sharing agreement, as they do not always have the technology to exploit their knowledge and resources.
- The debate on terminology and inclusion implies that if DSI is accepted as part of the NP (i.e. PIC and MAT are required), it will complicate the legal use of DSI. On

[6] The Natural History Museum, Royal Botanic Gardens Kew, and Royal Botanic Garden.

[7] https://www.cbd.int/dsi-gr/2019-2020/submissions/

the other hand, if DSI is not part of the NP, a new agreement would need to be designed to ensure benefit sharing. Regardless of the outcome, clarity on the definition will allow stakeholders to have legal certainty when and if a benefit sharing agreement is reached.

This report uses DSI as an umbrella term, encompassing the different definitions, unless otherwise stated.

4 DSI generation, access and valuation

This section responds to the research questions: How is DSI generated? And how is it accessed?

Section 3 discussed the definition of DSI and the ongoing debates around its terminology. While there is no agreement on the definition, parts of this section (sections 4.2 and 4.3) refer to DSI as nucleotide/genetic sequence data (NSD), nucleotide sequence information, or genetic sequences. This aligns with the definition for group 1 as shown in Figure 3.1. DSI and NSD are used interchangeably. If it is decided to use a different definition for DSI, more research will be needed with regards to DSI access, as there is almost no literature available discussing access to DSI or DSI databases using a broader definition.

The section first introduces how DSI is generated and used, including the technologies and processes being used to work with and generate DSI; it then summarises the databases where DSI is stored and how users access DSI through them. It concludes with a summary of the challenges for DSI in terms of traceability, due to its digital nature, and what this implies for benefit-sharing.

4.1 Process (from genetic material to DSI information) and technologies used

The CBD (Art. 2) defines genetic resources as *'any genetic material of actual or potential value',* with genetic material meaning *'any material of plant, animal, microbial or other origin containing functional units of heredity'* (CBD, 1992).

DNA is the basic hereditary material found in all organisms. In order to become DSI, DNA is extracted from the organism and then read or decoded so that the order of its nucleotides becomes known. This results in a code comprising the letters A, T, C and G (DNA sequencing).[8] The code is processed (transcribed), resulting in an RNA copy. This copy is then translated into a protein comprising a specific sequence of amino acids (determined by the DNA sequence). Therefore, once the DNA/RNA has been sequenced, the protein sequence also becomes known. In summary, an RNA copy is made from the DNA – this is the process of transcription. The RNA sequence is converted into proteins – the process of translation – and finally into metabolites – that is, biosynthesis (Houssen et al., 2020).

DNA becomes DSI through the entry of this information into databases. At the most basic level, a DSI database can contain a single entry for each gene, including the specific order of DNA letters (i.e. the sequence) for that gene. Databases can also contain additional supporting information about the gene, including phylogenetic information about the source species and citation information related to gene discovery. If the full genome sequence of the source organism is known, the database entry may also contain contextual data, such as the genome map (Houssen et al., 2020).

[8] Each organisms' genes are coded using a specific combination of these letters (A, T, C, and G) - an average gene comprises for the sake of argument around 1000 letters.

In its broadest form (e.g. group 3 or group 4 – see section 3.2), DSI can be defined as both raw data and the product(s) developed after using sequencing technologies (Houssen et al., 2020).

DSI can undergo different processes to get to its final use. For example, analysing environmental DNA from water samples collected at watering sights in tropical forests can enable the detection of endangered mammal species, thus providing insight into their survival rates. Sequencing various genomic regions of bees and comparing them to those available in online databases is providing insights to threats to bee populations. Sequencing samples of viral pathogens in disease outbreaks such as Ebola has revealed critical insights into the origins of the infection and the evolution and transmission of the disease (CBD, 2018). The same type of sequence can be used for research and conservation or applied in commercial research. DSI can therefore be seen as a tool that allows users to obtain or do something. This can include diagnosis of new pathogens in animals or plants, food safety screening, investigating the relationships within and between species, or supporting the development of commercial products (Sirakaya, 2019).

The exchange of DSI is taking place in an increasingly globalized research context. Advances in information technologies, including large storage capacity, powerful data manipulation techniques, and graphical capabilities have significantly increased the knowledge and capacity to process DSI (Heinemann et al., 2018). For example, databases such as the European Molecular Biology Laboratory, GenBank, Sequence Read Archive and the DNA Data Bank of Japan have almost doubled in size in the last few years. These databases are repositories of quadrillions (i.e. more than 10 to the 15^{th} power) of nucleotides of DNA sequences, collected from over 300,000 organisms (Laird and Wynberg, 2018).

How DSI is processed and used is evolving quickly. Bioinformatics has emerged as a new scientific discipline to process large amounts of data, combining biology, computer science, information engineering, mathematics and statistics to analyse and interpret DSI and biological data. Researchers use bioinformatics programmes to compare sequences from different organisms, and predict gene function, thus increasing the sum of knowledge regarding an organism.

Another discipline is the field of synthetic biology, which combines science, technology and engineering to facilitate and accelerate the understanding, design, redesign, manufacture and/or modification of genetic materials, living organisms and biological systems. Although most activity in synthetic biology is concentrated in basic research, synthetic biology also relies on DSI to write new genetic information for commercial uses, such as the “production of chemicals, biofuels, medicines, plastics, polymers and rubbers as well as plant feedstocks for microbe consumption” (Karger, 2018). In essence, synthetic biology enables the synthesis of a ‘physical’ DNA fragment or molecule based on a digital sequence (i.e. DSI), which can then be used in the same way as if it had been extracted and copied from an organism using conventional molecular biology.

Digital technologies permit data storage, distribution and analysis with low or zero marginal costs, making it easier to process and use genetic resources without possessing the physical material. Currently, scientists can sequence all or part of hundreds or thousands of genetic resources (e.g. plant samples) originating from various sources (Smyth et al., 2019). This process can yield vast amounts of data: for

example, genomics alone is predicted to soon exceed other "big data" science in size and complexity (Heinemann et al., 2018). Sequencing technologies have become faster, cheaper, and more accurate in recent years. The falling cost of technology has made access to sequencers possible for many researchers, either in their own lab, or via larger-scale sequencing facilities. This can be seen clearly in the 'Cost per Genome' data generated by the National Human Genome Research Institute (NHGRI)[9], which shows the costs falling from £80 million (USD100 million) in 2001 to below £800 (USD 1,000) in 2019 (Houssen et al., 2020).

In addition to the increased amount of data being shared, distinctions between academic, governmental, and industry research using genetic sequences are blurred as partnerships increase and academic institutions are developing 'commercialisable' products as digital sharing increases (Bagley et al., 2019). Networks of researchers from diverse institutional homes (e.g. industry, government, academia, and community laboratories) are working together in a system of "open innovation" in which users add incremental value through data and knowledge along a chain. In this context it is challenging, if not nearly impossible to determine who increased value and at what point in that chain (Laird and Wynberg, 2018).

The combination of large increases in the amount of sequence data available and being processed at any given time, together with the growth in partnerships between commercial and academic DSI users, makes it increasingly difficult to trace the lines of who accesses DSI, when and where DSI value is being added, and by whom (Houssen et al., 2020).

4.2 Access to DSI

DSI is generally accessed directly through nucleotide sequence databases[10], but it can also be accessed through journal articles, supplementary files attached to published papers, collections, patents and patents applications, synthesis companies, foundries, or genetic parts registries.

Databases can be public or private. There are approximately 1,700 online, publicly accessible DSI databases. The number of private DSI databases is unknown. The best-known public database is the INSDC. This is comprised of three global databases, based in the United Kingdom, Japan and the United States. These partners "capture, preserve, share and exchange a comprehensive collection of nucleotide sequence and associated information" (Laird and Wynberg 2018: 10) for the scientific community, and develop new services to handle the changing landscape of data types. It is INSDC policy to enforce free, unrestricted access to all data (Rohden et al., 2020).

[9] https://www.genome.gov/

[10] Sections 4.2 and 4.3 refer to DSI as nucleotide/genetic sequence data (NSD) or nucleotide sequence information or genetic sequences. This is the narrow definition (group 1) or data, from the discussion in Section 3. For this section DSI and NSD are used interchangeably, as the team did not find any sources discussing access to DSI or DSI databases using a broader definition.

Box 1 The International Nucleotide Sequence Database Collaboration

The INSDC is an international collaboration between GenBank in the US, the European Nucleotide Archive in the UK, and as of the early 1980s, the DNA Data Bank of Japan. The partners share and exchange the collection of nucleotide sequence and associated information. These three databases (Japan, US and UK based) provide the scientific community around the world with a complete, high-quality, reliable, open, and free infrastructure for NSD. The three INSDC partners "mirror" (exchange) all NSD in their databases every 24 hours to maintain an up-to-date copy of all published NSD for global use.

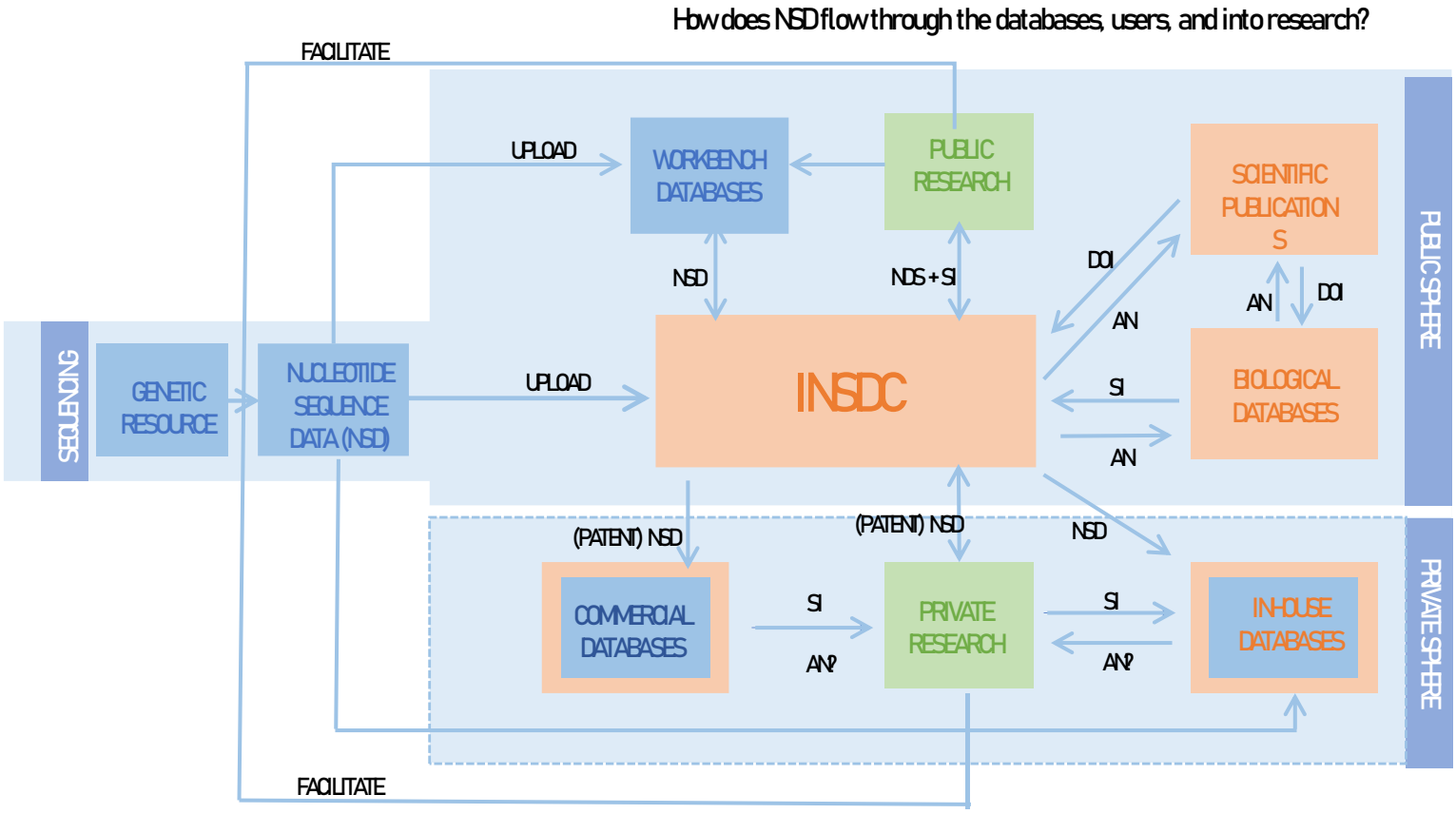
The combined costs across all three INSDC databases are estimated at £40-60 (USD 50-60) million annually. The public NSD databases that use and download NSD from INSDC agree to and depend on the INSDC's use policy. INSDC has around 10 to 15 million users.

95% of NSD databases directly link to or download NSD from the INSDC. The remaining 5% of NSD databases allow direct NSD submissions but require the use of unique identifiers – accession numbers – which are generated by the INSDC and so are inherently connected to the infrastructure. NSD databases rely on the INSDC and use accession numbers to enable traceability through the database landscape

Source: Rohden et al., 2020

Figure 4.1, from the "Combined Study on Digital Sequence Information in public and private databases and traceability" (Rohden et al., 2020), shows a simplified overview of the technical infrastructure for accessing and sharing DSI developed over the past four decades.

Figure 4.1 DSI flow through databases and DSI users at each stage



Legend: AN – Accession Number; DOI – Digital Object Identifier; NSD - Nucleotide Sequence Data; SI – Subsidiary Information

Source: Rohden et al., 2020

The INSDC is the core database infrastructure for publicly available DSI. The INSDC enables scientists to submit their DSI and receive a unique identifier or accession number. The database was set up under the principles of open access to enable scientific reproducibility and support scientific integrity. Open access refers to "free and unrestricted access" with "no use restrictions", and data "permanently accessible", according to INSDC policy (Brunak et al. 2002). In 2016 INSDC stated that "the core of the INSDC policy is maintaining public access to the global archives of nucleotide data generated in publicly funded experiments. A key instrument for this is submission as a prerequisite for publication in scholarly journals […] The database also provides training, technical assistance, free software tools, and tutorials" (Crochane et al. 2016: D50).

Access to DSI also occurs through other databases, public or private. Based on their level of precision, databases may be classified as "primary" (containing raw data) or "secondary" (containing curated and analysed data). Databases may also be specialised or comprehensive, with the latter type containing a wider range of data types from numerous species.

The types of databases can be described as:

- Knowledge hubs (public access): These databases provide aggregated knowledge on a specialized topic, by collecting and rearranging information from

other databases and already existing scientific publications. Within these, a number of larger databases are tightly integrated with the INSDC. They are often funded by the same public sources and have members of the INSDC or connected institutions on their steering board. Their main function is to rearrange and curate data from the INSDC, connect it with information obtained from scientific publications and make it publicly available.

- Bioinformatic tools (public access): provides a tool for researchers to analyse/process either their own research data or the data/information provided by the (knowledge hub) database. For example, the database RAID (RNA Interactome Database) has the tool PRIdictor ("Protein RNA Interaction Predictor") on its webpage, which predicts the interaction between RNA molecules and proteins. Here, a researcher can submit the data of an RNA molecule and the amino acid sequence of a protein and obtain a prediction of how these two molecules will bind to each other.
- Restricted databases (private access), often described as "in-house databases", privately store their generated or acquired DSI data for internal use only. Most databases will download the data from INSDC but the results of any further sequencing will be private.
- "Commercial" databases (private access), in which access to the stored DSI is only possible after payment. Commercial databases use, process, and analyse data in order to create more curated (value-added) sequences and information, as well as developing bioinformatic tools to use, process and analyse DSI. Commercial patent DSI databases are good examples of commercial databases and are very frequently used by companies.

DSI data flowing into databases is exponentially increasing, with some of the most widely known databases having almost doubled in size in the last few years (reaching repositories of quadrillions of sequences) (Laird and Wynberg, 2018). This had led to more people sharing, and processing DSI data, and, to more knowledge being developed and made available.

4.3 DSI value generation

DSI value is generated through its production and further processing. The rapid emergence of new sequencing technologies has lowered the associated costs of this activity and made sequencing capabilities available to more researchers (Houssen et al., 2020; Laird and Wynberg, 2018). This has led to the emergence of new processing technologies (e.g. bioinformatics) and new uses for DSI (see section 5 for a more detailed description of the potential uses of DSI). These emerging technologies present opportunities for the development of new products (e.g. pathogen resistant crops, pharmaceuticals), and processes (e.g. production of bioethanol, production of bacterial or fungal treatments) (section 5 includes examples of how DSI is used in different sectors), that use DSI and can generate monetary and non-monetary benefits.

Understanding how DSI value is generated and shared is relevant to operationalise a potential benefit-sharing policy option. However, there are challenges to quantify the 'value' DSI has and where it originated. DSI related products, processes and technologies often involve a combination of large sets of data that come from multiple

countries and different types of organisms (or genetic resources). Further, when stakeholders process DSI, even when there is information on the country of origin of the genetic material, such information is not always conclusive. This is because some genetic resources in different parts of the world may contain identical types of data, so users can find a similar sequence somewhere else. They may also decide not to use it at all if access is restricted. The implication being that, sometimes, if users do not have access to the data they require in a specific country, they can find it from another country (Laird and Wynberg, 2018).

One researcher quoted in Laird and Wynberg (2018: 49) explains: "a sequence on its own does not have real value. Value begins with identification of a valuable trait, a characteristic of an organism that is of interest like drought-resistance, fungal resistance, or a slug whose slime's stickiness helps close surgical wounds…". Further, DSI value can often be found in the aggregate, rather than individual sequences: "… an individual sequence or 'part' has more value in a library where it can be screened with other sequences to find the connections between a particular trait and its function and use in other things… As a result, the value of an individual sequence from a species may be very difficult to quantify." (Welch 2017, as cited in Laird and Wynberg 2018:49).

Value can also be generated with patents. However, as there is no legal definition of DSI, nor a clear difference between what is data and what is information, acquiring a patent for a DSI-related output is not easy. In most jurisdictions, patents on gene sequences are not permitted, but once functionality is added, and can be demonstrated, patents can be issued. The implications of the value added for DSI through patents has not been explored in the literature, but it will be relevant when considering benefit sharing models and who has rights over DSI (Ruiz Muller, 2018).

4.4 Traceability challenges

Operationalising the benefit-sharing aspects of the NP in respect of DSI arguably would require a system of tracking and tracing to be introduced so that the value extracted from utilisation of a resource may be linked back to the relevant provider country. Otherwise, determining the type and method by which to share the benefits can be challenging. The process of tracking and tracing is not an unequivocal one; it can encompass systems that trace back to origin upon commercial use, or to the tracking of information and use along the value chain. The data which may be relied upon to track and trace includes geographical coordinates, environmental contexts and information about collections from which sequences might have come. However, because of how DSI is generated, accessed and used, and because its value lies in large volumes of data being shared daily among multiple stakeholders, robust traceability of the broad array of DSI is considered problematic.

Having access permits (e.g. PIC/MAT) could allow one to trace the data to its country of origin. But with DSI, according to Rohden et al. (2020), this is not always possible. The traceability of certain forms of DSI may be a possibility. The INSDC, for example, operates on an open access basis and has an existing traceability system since 2011, where all NSD submitted to the INSDC is required to include information on country of origin and it is given an individual accession number (Rohden et al., 2020). However, the system relies on individual due diligence when submitting to the INSDC

(there is no monitoring mechanism), therefore only 16% of analysed NSD has a country tag (Rohden et. al., 2020).

Rohden et al. (2020) identify the following challenges related to tracing DSI:

- Existing traceability of the data depends on submitter diligence - the INSDC system was built to enable scientific integrity and transparency with a primary focus on publication and scientific exchange (not monetary benefit-sharing).
- The types of data stored in public databases are heterogeneous and out of CBD scope (e.g. human NSD, which accounts for 12% of data stored) - there is no "blanket" solution for traceability of the entire INSDC database as not all data stored would be required to be traceable under CBD. Significant intellectual, technological, and regulatory effort would need to be made to address this heterogeneity.
- There are more than 1,700 biological databases that are inter-connected and exchange data daily. Behind the scenes, different types of data are converted, transformed, and exchanged in many and multiple directions. Regulating this infrastructure is difficult.
- The volume of data is exponentially growing. Any long-term regulatory scheme would need to be prepared for "big data" interventions and the accompanying IT investments.
- INSDC efforts to include the origin of data and other contextual information have increased in the past decade, making traceability to provider countries easier. However, earlier records did not include information on the origin of the DSI data, so records are incomplete. Compliance with INSDC traceability regulations is entirely at the discretion of data submitters: database managers receive a new submission every six minutes on average, so other kinds of enforcement on individual submissions are mostly impractical.
- While the INSDC tracing system is helpful because every entry has a unique identifier, biology itself is more complicated. There are millions of repetitive NSD entries or parts of entries in the databases making it difficult to attribute all entries to a specific sovereign state. This means that since datasets can include a combination of different types of data and comparisons between them, it can be difficult to determine where the multiple datasets come from, increasing in complexity as these are processed over time. There are also many shared and homologous sequences; often, if a sequence from one country is not available, it can be found somewhere else. Accordingly, issues of valuation are intractably linked to the challenge of traceability.

The debate over the definition of DSI also requires consideration in respect of these issues, as a broader scope of what is considered DSI beyond NSD will raise more significant traceability issues.

The AHTEG meeting in 2018 discussed whether the ability to trace DSI data may be improving with new technological developments (e.g. blockchain). However new findings suggest that the use of blockchain to trace DSI is not feasible with the current technology (AHTEG, 2018). Significant concerns have also been raised with respect to the energy-intensive nature of blockchain, resource use and the potential to produce significant climate change emissions (Rohden et al., 2020).

Blockchain technology is being developed and applied for human patient NSD and accompanying patient health information, enabling patients to control access to their private genetic data. Technically, this could be applied to non-human NSD if developments in the field of blockchain continue. However, it could only work for newly generated NSD, as it would need to establish a private, standalone system outside of the INSDC and the public databases. It would also need intensive financial investments and upkeep, and it is debatable whether the benefits could exceed the costs. Other restricted access models from the publishing or media world (e.g. Spotify or Netflix) target only passive access of the user (e.g. listening) rather than the interactive "hands-on" use required by scientists using NSD (Rohden et al., 2020).

4.5 Implications of DSI generation, access and use for stakeholders and benefit-sharing options

This section has explained the process for generating DSI, and the different technologies that are used to process it. It has discussed the main ways stakeholders access DSI and how value is generated in this process. This has highlighted the nature of DSI as an easy to share resource, requiring multiple users processing high volumes of data in a digital manner for the value of DSI to be maximised. This complexity means that traceability is almost impossible.

The current bilateral approach to benefit-sharing lies in the ability to trace the use of genetic resources back to its origin, to identify the parties engaged in creating the DSI output and where value was added in the process. This means users of genetic resources must be able to identify throughout the supply chain those stakeholders from whom they received the material and those to whom they supplied it.

The policy implications of this for benefit sharing options are:

- DSI value comes from its complexity. Aiming to understand who generates the value and at what stage it has been generated is generally seen as unfeasible.
  - The use of patents and IP could be explored to link DSI to its use. However, these have not been fully explored in the literature and their implications are not known.
- Access to DSI is similar for most users. INSDC is the initial point where data is downloaded and processed and where some users upload their data after.
  - Private and public databases exist, and both would need to be part of any benefit-sharing agreement.
- DSI is increasingly used by multiple stakeholders working together. This means that unpacking the stakeholders that may have been involved and/or what the purpose of the DSI use was becomes challenging. As DSI does not have a legally binding definition yet, the implications of this are not clear.
  - It is challenging to determine with legal certainty the extent of the value added by each party in the process of generating DSI and who is using it for commercial or non-commercial purposes.
  - More information on DSI use will be required to better understand the implications.

- The current information structure in databases does not acknowledge nor distribute the value of the origin of the data. Guaranteeing traceability under INSDC – as the largest database - is currently hindered by the fact that INSDC does not monitor its entries (neither uploaded DSI nor the associated traceability attached to it).

5 DSI stakeholders in the UK

This section identifies and explains how DSI is used in specific sectors in the UK. The section provides an overview of the key stakeholder groups involved in generating, sharing and using DSI. It identifies the main UK stakeholder groups that will likely be impacted by any change in the DSI policy landscape and describes the sectors these stakeholders work in, the trends within those sectors,[11] and summarises submissions made to the CBD by UK stakeholders expressing their opinion on the objectives of the CBD towards DSI.

5.1 Groups of stakeholders

DSI is often the result of sequencing technologies using digital data that originated from genetic resources. The key actors are countries that provide genetic resources, public collections that store and grant access to DSI, and the users of DSI.

DSI benefits, both monetary and non-monetary, accrue across the process of accessing, using and generating DSI. As DSI value comes from the processing of large volumes of data multiple times by multiple stakeholders, it is challenging to determine with legal certainty the extent of the value added by each party. Furthermore, assessing the total value of DSI is difficult, even when it is utilised in final products. As a result, stakeholders have different perspectives on how benefits should be shared in a fair and equitable manner.

5.1.1 Provider countries of genetic material

A 'provider country' is any country that provides genetic resources for utilisation by a 'user'. Some provider countries, however, are so rich in biodiversity that they are also referred to as 'megadiverse'[12] countries. Every country has sovereign rights over its natural resources and can establish legislation regarding who holds rights to genetic resources within their country, who has the authority to grant access to those resources and how traditional knowledge associated with genetic resources is governed. However, many countries lack the resources or capacity to transform their genetic resources into DSI and use it in this advanced form, including some of those classified as megadiverse.

In this context, most megadiverse countries argue for DSI to be included in the NP, as they claim that open access to data does not allow them to fully realise the benefits that can be derived from the genetic resources in their countries (Aubry, 2019). Fifteen megadiverse countries have already developed domestic measures (PIC, MAT or

[11] This includes information on the economic value of the sector. As the sectors – and in particular the components of them reliant on DSI – are not well represented in the Standard Industrial Classification (SIC) of economic activities, partial data on economic value are drawn from multiple sources. The sectoral definitions used in these sources do not directly map onto the DSI sectors, and hence the values provided should not be taken as 'accurate', but rather given an indication of magnitude. The actual data presented may be over or underestimates of the sector value, as indicated when presented.

[12] The term megadiverse country refers to any one of a group of nations that host the majority of Earth's species and high numbers of endemic species (a megadiverse country must have at least 5,000 species of endemic plants and must border marine ecosystems). Conservation International identified 17 megadiverse countries in 1998, many of them located in, or partially in, tropical or subtropical regions.

permits) to address benefit-sharing arising from commercial and non-commercial use of DSI and eighteen countries have plans to introduce these (Bagley et al., 2019).[13]

5.1.2 Public collections (databases)

Databases are used to store DSI and database developers can provide open access to this information (section 4.2). Databases can be public or private, and there are thousands of them. The INSDC is a public database providing a well-established and broadly adopted data infrastructure for the open sharing of sequences. Other databases use the INSDC infrastructure to access DSI (Rohden et al., 2020).

Public and private database developers are also users of DSI. For example, research institutions in the taxonomy and conservation sector develop and host public collections to create and sustain biodiversity data (for both animals and plants). At the same time, the taxonomy and conservation sectors use DSI for biodiversity conservation and the sustainable use of natural resources.

Database developers argue that open access to data is a critical form of DSI benefit-sharing (INSDC, 2019). Countries hosting databases, particularly the three country hosts (UK, US, and Japan) for the INSDC, provide funding, expertise, and technological capacity to third parties to store, analyse and manage public databases (NHM et al., 2019). As such, they argue that the existing infrastructure allows for equitable and fair non-monetary benefit sharing. According to CBD submissions and the Laird and Wynberg (2018) fact-finding study, most researchers and database managers agree with the view that DSI should remain open access and the existing non-monetary benefit-sharing scheme (consisting largely of capacity building measures) should remain.

5.1.3 Users of DSI

DSI is used both in basic research and to create new products or processes and it is difficult to differentiate between commercial and non-commercial DSI users aiming to develop new products (e.g. new vaccines) or processes (e.g. new technologies to process DSI more efficiently) (Rohden et al., 2020; Laird and Wynberg, 2018; von Kries and Winter, 2015). This is due in part to increasing collaborations between universities and companies; furthermore, an individual researcher can use their work both for commercial and non-commercial purposes (Laird and Wynberg 2018).

DSI is used across a range of different sectors and for different purposes. The following subsections describe the major UK sectors that use DSI, based on desk research. The sectors were identified based on the selection made by Houssen el at. (2020: 25). The descriptions include a summary of trends for each sector and examples of how DSI is used, highlighting potential DSI value and the technologies that enable DSI use in that sector:

- The taxonomy and conservation sector (section 5.2) undertakes basic research and includes database managers.

[13] COP decision 14/20 requested a "Fact-finding Study on How Domestic Measures Address Benefit-sharing Arising from Commercial and Non-commercial Use of Digital Sequence Information on Genetic Resources and Address the Use of Digital Sequence Information on Genetic Resources for Research and Development".

- The biotechnology (section 5.3) field, of which synthetic biology is a subset of, provides next-generation sequencing technologies that allow other sectors to develop new products and processes.
- Other sectors that use DSI include healthcare (section 5.4), agriculture technology (section 5.5), and industrial biotechnology (section 5.6).

The sectors that use DSI are an important part of the UK economy. Annex 1 summarises the economic significance of these sectors, in terms of Gross Valued Added (GVA), turnover, and employment. The data are taken from the UK Annual Business Inquiry 2018 (ONS, 2019) and include the most detailed figures possible for relevant sub-sectors. The Standard Industrial Classification (SIC) codes included under each sub-sector are included in the table. The team has also included economic data from sector specific reports for some of the sectors as indications of the size of the sectors in the UK economy. Subsequently, the figures included below are best estimates and do not accurately reflect the GVA and employment of each sector.

5.2 Taxonomy and conservation[14]

5.2.1 Overview of the sector

Conservation is the management and protection of biodiversity. Taxonomy refers to the theory and practice of describing, naming and classifying living things in hierarchical order, and such work is essential to the fundamental understanding of biodiversity and therefore key to its conservation (Knowles et al., 2008). Access to DSI is important for this sector as it facilitates the rapid identification of new species through genetics and the classification of species already identified according to their biological function in the natural world (Shipman, 2012).

The main actors in this sector are universities, botanic gardens, museums, and culture collections (i.e. a collection of microorganisms, such as bacteria). In the UK the key actors are the Natural History Museum (NHM), the Royal Botanic Gardens Kew (RBG Kew) and the Royal Botanic Gardens Edinburgh (RGBE).

Stakeholders in the sector use DSI for research purposes, including taxonomic identification and systematics.[15] Actors within the sector maintain collections (or databases) of living and preserved plant material, animal specimens, and micro-organisms. These collections are used for education and training, conservation, and taxonomic identification. Collections can also facilitate acquisition of genetic resources for other organisations (e.g. research institutes, universities and companies) and help users of genetic resources identify particular specimens. Finally, DSI sequencing allows taxonomists and conservationists to identify and compare different organisms across the world in a rapid manner. As such, the work of the sector is heavily reliant on open access to DSI data (House of Lords, 2008).

Rohden et al. (2020), observe that the increasing availability of DSI sequences can enable actors in this sector to monitor biodiversity in a more efficient manner. In addition, a case study published in Rhoden et al. (2020) highlights how DSI has been

[14] Taxonomy and conservation can be described as two different sectors. Taxonomy is required for conservation, but the sectors can exist independently of each other. The actors identified here work in both fields.

[15] Biological systematics is the study of relationships and classification of living things. Taxonomy is a branch of systematics that studies the identities of living things and assigns names to organisms.

used to evaluate marine ecosystems and climate change impacts by observing and analysing changes in the biological composition of ocean regimes.

5.2.2 Trends in the sector

The main UK stakeholders are world leaders in this sector. The NHM collection has more than 80 million specimens. As of 2018 more than 8 billion items of scientific data had been downloaded from these collections (NHM, 2018). RGB Kew estimates that it has approximately 8 million specimens of plants and seeds. It also manages approximately 48,000 samples of DNA in the DNA and tissue bank collection from the Millennium Seed Bank Partnership[16] (Oxford Economics, 2019). The RGBE has a collection of 13,5000 living plant species and 3 million preserved herbarium specimens.

It was not possible to provide an estimate of the potential value related to DSI. There are estimations of the size of the sector (or parts of the sector) in the UK. However, these do not provide estimates of the value of DSI. An Oxford Economics (2019) study estimated that "Kew's scientific activities in 2018/19 could deliver a £76.3 million boost to the UK's economic output in the long term". Similarly, a study by Hill et al. (2016) estimated that the RBGE's scientific contribution to the global economy is £81 million GVA between 2015 and 2016 but did not isolate the value attributable to DSI.

5.2.3 Stakeholder submissions to the CBD

UK stakeholders in the taxonomy and conservation sector submitted joint statements to the CBD in 2017 and 2019. The NHM, RBG Kew and RBGE explained that the analysis of DSI data is critical to taxonomic research, and that outputs of this research are for non-commercial scientific purposes and therefore do not generate monetary benefits.

The existing DSI open access model is beneficial for this sector as it supports scientific research globally. Moreover, non-monetary benefit sharing (such as capacity building) is already occurring between non-commercial biodiversity research organisations and provider countries (NHM et al., 2019). UK stakeholders recognise that not all CBD Parties have equal capacity to exploit DSI to its full potential, due to a lack of resources in many countries, and reiterate their commitment to capacity building through training and joint research. However, UK stakeholders in the sector remain concerned about any potential restriction on the use and availability of DSI, which is used globally for research. These stakeholders worry that any restrictions on access to data may damage biodiversity research and conservation efforts, by imposing bureaucratic barriers, making it more difficult to share data in the way that it is done now.

UK stakeholders in this sector argue that changes to the current database infrastructure (particularly to the INSDC) are neither practical nor appropriate, as the database policies were developed to avoid both "barriers to downloading the sequence data and applying conditions on their use" (NHM et al., 2019: 2).

[16] The Millennium Seed Bank Partnership is the largest ex-situ plant conservation project in the world. It is based at Kew, and currently holds seeds for 10 per cent of the world's plant species.

5.3 Biotechnology

5.3.1 Overview of the sector

Biotechnology can be described as a set of cross-disciplinary technologies that use genetic resources to design new biological systems. Synthetic biology is an emergent multidisciplinary area of research within biotechnology that includes a broad range of methodologies from various disciplines (e.g. genetic engineering, and molecular biology, engineering, systems biology, bioinformatics etc). It is defined by AHTEG as "a further development and new dimension of modern biotechnology that combines science, technology, and engineering to facilitate and accelerate the understanding, design, redesign, manufacture and/or modification of genetic materials, living organisms and biological systems" (CBD, 2015).

DSI is core to these fields, and through its use, products and processes have been developed that can be applied in other sectors. The outputs from this sector generate both monetary benefits, in terms of producing products of commercial value, such as biofuels or bioplastics, and non-monetary benefits, such as inputs to the discovery of new flu vaccines (Keasling, 2013). For example, the Open Plant Synthetic Biology Research Centre, a joint initiative between the University of Cambridge, John Innes Centre and the Earlham Institute, and part of the UK Synthetic Biology for Growth programme is working on 'smart' breeding of crop systems, to address climate change and other urgent threats and challenges.[17]

These fields have applications across three UK sectors: healthcare, agriculture, and industrial biotechnology.

- Healthcare (section 5.4) refers to healthcare products such as medicinal or diagnostic products, or a vaccine that consists of, or has been produced in living organisms, as well as processes such as gene and cell therapies.
- Agriculture technology (section 5.5) includes techniques to modify existing genetic resources (e.g. seeds) to transfer specific traits, such as resistance to pests. The crops may be used for animal feed, food, biomaterials or energy production.
- Industrial biotechnology (section 5.6) makes products in sectors such as chemicals, food and feed, detergents, paper and pulp, textiles and bioenergy (such as biofuels or biogas).

5.4 Healthcare

5.4.1 Overview of the sector

DSI is commonly used in the healthcare sector for research and development of pharmaceuticals and the discovery of new pathogens and their cure. Most national and international stakeholders recognise the growing importance of access and use of DSI for health purposes, as DSI allows for the development and testing of new treatments (e.g. HIV test kits, drugs that have predictive elements, and enabling smaller and more targeted trials), and for the genomic study of some pathologies to uncover their origins and potential cure (DIT, 2020a/b; IBLF, 2018).

[17] www.openplant.org

The use of DSI in the healthcare sector yields important outputs both in terms of pharmaceuticals and medical research and diagnosis. Estimates indicate that 20-25% of the pharmaceuticals market is derived from genetic resources, with almost two out of three antibacterial agents coming from this source. DSI can also be used to diagnose and prevent diseases, such as food-borne illnesses, as DSI sequencing has the potential to increase the rapidity of diagnostic times and identification of new pathologies (Houssen et al., 2020).

In the UK, the subsectors that use DSI include (PWC, 2017):

- Development and manufacture of pharmaceuticals products.[18]
- Life sciences research, including those engaged in conducting research and experimental development within fields such as biotechnology, medicines and medical diagnostics and devices. These firms often operate at the early pre-commercialisation stage.[19]

5.4.2 Trends in the sector

The UK is among the global leaders in healthcare R&D. In the UK, this sector contributes to pharmaceutical R&D; products for dementia and oncology, and addressing global health challenges such as antimicrobial resistance, HIV/AIDS, and malaria (PWC, 2017). A PWC study commissioned by industry actors[20] estimated that, in 2015 the sector contributed £30.4 billion[21] in UK GDP, supported 482,000 jobs (140,000 direct employees of firms, 196,000 jobs in the supply chain and 146,000 supported through induced effects), and contributed £8.6 billion in taxes (PWC, 2017). It has more than 5,800 companies and the largest pipeline of biotech products in Europe (DIT, 2020c).

Pharmaceutical manufacturing is the largest contributor in terms of GVA (£15.7bn, 52% of total life sciences GVA) (PwC, 2017) and turnover (£33.4bn or 45%). In terms of research and development within the healthcare industry, new subsectors are emerging. Biopharmaceutics, linked to biotechnology, enables the development of personalised medicine, that is, tailoring medicine to the specific needs of the individual biology of patients. Personalised molecule drugs have the potential to lead to safer and more effective treatments. Moreover, biotechnology-enabled diagnostics promise to reduce the time taken to diagnose diseases and increase the quality and efficiency of prescription of new treatments (IBLF, 2018). The Novo Nordisk Research Centre Oxford made a £100 million investment in biotechnological health research to support academic-industrial collaborations, human genetics and big data to conduct world-class discovery and translational research (DIT, 2020b).

Life sciences research as a subsector provides the lowest monetary contribution within the healthcare sector, with a GVA of £3.3bn and employing 26,000 people

[18] SIC codes: manufacture of basic pharmaceutical products, manufacture of pharmaceutical preparations, wholesale of pharmaceutical products, manufacture of other chemical products not elsewhere classified, manufacture of other inorganic basic chemicals, manufacture of other organic basic chemicals

[19] SIC codes: Research and experimental development in biotechnology, other research and experimental development on natural sciences and engineering, other professional, scientific and technical activities (not including environmental consultancy or quantity surveying) not elsewhere classified, other human health activities

[20] The report was commissioned by the Association of the British Pharmaceutical Industry, UK Bioindustry Association, British In Vitro Diagnostics Association and the Association of British Health Tech Industries.

[21] These figures include the subsectors mentioned above, and the subsector 'development and manufacture of medical technology devices' (e.g. Beckton Dickinson, Olympus Keymed)

(PwC, 2017). However, the non-monetary contribution of life sciences research is substantial, providing scope for innovation in human health. Within the life sciences research subsector, the genomics segment provided a total turnover estimated at £1.9bn and included 2,400 employees (OLS, 2019), and within this segment, sequencing consumables and instruments is the largest activity with a total employment of 1,600 people and a generation of £1.7bn in turnover (OLS, 2019).

UK researchers and institutions have pioneered major scientific breakthroughs including the discovery of the structure of DNA and advancement in neuroscience and medical imaging, and developments in stem cells and in vitro fertilisation (DIT, 2020c). Moreover, the UK government made a commitment to increase R&D spending to 2.4% of GDP by 2027 (DIT, 2020b). Another important subsector increasingly using DSI is gastroenterology, aiming to improve the understanding of the functions of the gut microbiome. In the past five years, experts and clinical researchers have highlighted the links between the microbial composition of our guts and a wide range of illnesses (from irritable bowel disease, to cancer and depression). The UK has significant expertise in gut microbiology, which is attracting new investment in the sector (IBLF, 2018).

5.4.3 UK Stakeholder submissions to the CBD

The Wellcome Sanger Institute (a UK genomics and genetics research institute) submitted their opinion to the CBD on any potential changes to the DSI policy landscape.

The Institute strongly disagrees with the proposal to include DSI in the scope of the CBD and the NP. The Institute acknowledges that countries should share equitably the benefits of research and development derived from any genetic resources, but states that the inclusion of DSI in the NP would not achieve this goal, rather it would hinder global cooperation in scientific research. The Institute writes that open access to DSI is fundamental to responding promptly and efficiently to health crises. Any added administrative burden of agreements such as MAT or PIC would hinder the rapidity of the response. In their opinion, the potential impact of the inclusion of DSI within the scope of the NP would be damaging for the healthcare sector, imposing barriers to research and innovation. For example, they explain that this would impact efforts to collate DSI from around the world to track the emergence of drug resistance. Having to interpret and comply with domestic measures to access DSI may end up dissuading researchers from sharing and using sequences.

The Royal Society of Biology (RSB) expressed in 2017 that hampered access to DSI would seriously challenge national and international biosecurity and public, animal and plant health responses by impeding international research and surveillance activities relating to existing and emerging global health threats. They explain that surveillance capacity, for example, is based on constant open access of data. Timely sharing of information on pathogen samples and DSI in public databases is key for accurate epidemic risk assessment and rapid responses. Any barriers imposed to the existing mechanisms (such as new benefit sharing protocols or additional agreements that need to be signed) will disrupt the ongoing global health efforts to prevent health crisis (RBS, 2017).

5.5 Agriculture technology

5.5.1 Overview of the sector

Agricultural technology, or agri-tech, encompasses all the activities involved with technology, innovation and sustainability developments for agriculture. This includes innovations in biological, environmental and chemical sciences and the implementation of emerging technologies such as satellite imaging for agricultural production (DIT, 2020a). Within the sector there are many practices that make use of genetic resources and DSI, including plant breeding and crop modification research (to increase productivity and sustainability), genetic modification of livestock, support activities for plant production and veterinary pharmaceuticals.

The agri-tech sector actors that rely on DSI for research are plant breeding companies, companies that develop crop protection products (including chemicals and plant breeding for resistance to pests and diseases), and animal breeding companies. Some stakeholders are also using DSI to improve food safety (e.g. to monitor foodborne pathogens along the food chain).

DSI is used to develop new plant varieties for specialised food and feed, via marker assisted selection or other processes. In addition, the industry uses DSI for selective breeding to develop crops or livestock that are resistant to specific pathogens or parasites. Another use of DSI in the industry is the development of living modified organisms (LMO)[22]. For example, one can introduce in a plant a bacterial gene that has been modified to resist parasites. This can potentially contribute, among other things, to food security and nutrition (Heinemann et al., 2018).

5.5.2 Trends in the sector

The UK is a market leader in agricultural technologies (DIT, 2020a). In terms of agricultural technology research, in recent years, the UK government has established of four Centres for Agricultural Innovation, supporting the development of technologies to improve the productivity and sustainability of the agriculture sector (DIT, 2020a). For example, the Norwich Research Park has four research institutes that engage in agri-tech research: the John Innes Centre (JIC), and The Sainsbury Laboratory (TSL), at the forefront of plant sciences; the Institute of Food Research, dedicated to post-farm gate agri-tech; and The Genome Analysis Centre, that sequences and decodes genomes specifically of relevance to agriculture. In addition, some of the world's leading agri-tech companies are based in the UK, such as Syngenta, Genus, Aviagen, JCB, New Holland and Velcourt (BIS, 2013).

Recent estimates suggest that the overall agri-tech sector accounts for a total turnover of £56.8bn and GVA of £14.3bn, employing 543,000 people (SQW, 2016). However, only a fraction of the sector concerns activities related to DSI use. These mostly include plant breeding, production support activities (such as fertilisers and crop protection products), and animal breeding. As such, the fraction of the agri-tech sector which relies on DSI is not adequately represented in standard economic statistics. Multiple studies have been undertaken in recent years to establish estimates. While

[22] "Living modified organisms" are any living organisms that possess a new combination of genetic material obtained through the use of modern biotechnology; they are a subset of genetically modified organisms.

these studies define the sector in different ways, none of them is solely focussed on agri-tech subsectors that use DSI.

The UK plant and animal breeding subsectors constitute a small but important contribution to the agri-tech sector, particularly in terms of innovation and R&D, which relies on DSI as their main input. There are limited figures available for the other subsectors identified in agri-tech that use DSI.

For the plant breeding sector, it is estimated that UK plant breeders make use of 30-40% of genetic material from third parties to develop new plant varieties (Barnes et al., 2016). A report by SQW (2016) for the UK Department for Business and Innovation Skills completed with data from 2013 estimated that the plant agri-tech sector contributed £2.7bn in turnover and £0.7bn in GVA, 8,400 direct jobs, representing approximately 5% of the overall agri-tech sector. A report by Barnes et al. in 2016 estimated that the turnover for the plant breeding subsector was between £200 and £230 million. However, figures from the latest Annual Business Survey (ONS, 2019) suggest a turnover of £89 million (Annex 1). This latter figure is based on the 5-digit SIC subdivision, which is likely to underestimate the actual size of the plant breeding subsector. This may explain the considerable divergence between the two figures.

This agri-tech subsector has potential for indirect effects in the agri-tech industry in terms of production and processing. The majority of the food chain, whether crops or animals, is underpinned by plant breeding, so improvements in crops, yields and quality of plants will have an impact on supply costs and production (Barnes et al. 2016).

The other relevant subsectors in agri-tech are:

- Crop production support, in the development of crop protectors and fertilisers. It is estimated that the crop production subsector, using data from the 2018 Annual Business Survey, contributes £648 million in GVA and £2,471 million in turnover (Annex 1).
- Animal breeding, which is primarily conducted by private companies supplying livestock to multinational markets. However, data available on the UK actors working in the sector is limited. SQW (2016) mentions that the animal agri-tech subsector had £5.8bn in turnover and £1.1bn in GVA, and contributed 20,900 direct jobs, representing approximately 10% of the overall agri-tech sector (SQW, 2016).

5.5.3 UK Stakeholder submissions to the CBD

A range of stakeholders submitted a joint statement to the CBD (i.e. public and private sector organisations, academic and scientific institutions, data repositories and collections representing a broad range of stakeholders) many of whom represent the agri-food and biotech industry from the EU and worldwide. The statement argues that DSI is a critical tool in the use of genetic resources that is leading to many societal benefits (e.g. food and nutrition security). As such, including DSI within the NP will impose barriers to existing open access and information sharing approaches, perhaps damaging the exchange of information to support conservation, protection and sustainable use of biodiversity. By changing the system of DSI access, it will create an environment of legal uncertainty, increasing the time needed to use DSI, the costs

and adding administrative burdens, potentially stifling innovation and growth in the sector.

5.6 Other industrial uses: Industrial biotechnology

5.6.1 Overview of the sector

The pharmaceutical, and agribusiness and agri-food technology industries are the major industry subsectors that use and benefit from DSI currently. However, biotechnology has applications for other industrial uses, beyond healthcare and agriculture. These subsectors also use DSI and may be affected by any changes in the way DSI is accessed.

Industrial biotechnology covers other industrial uses, such as innovations along the food value chain to reduce waste or the environmental footprint (e.g. use of waste in energy production), environmental uses, cosmetics, and detergent manufacture, among others.

The UK is among the global leaders in industrial biotechnology, together with the US and Japan. The sector is not consolidated, with more than 80% of businesses in the industry being SMEs, and no major players having a greater than 5% market share (IBLF, 2018). A consortium of biotechnology actors in the UK came together to launch the 'National Industrial Biotechnology Strategy to 2030'.[23] It reflects the key organisations in the sector:

- Crossing Biological Membranes Network, a BBSCR NIBB;
- Network in biocatalyst discovery, development and scale up;
- Industrial Biotechnology Leadership Forum (IBLF);
- Biotechnology and Biological Sciences Research Council; and
- BioPilots UK.

The UK is also a leader in using DSI sequencing to produce flavours and fragrances (a potential global market worth £22 billion); non-caloric sweeteners (£2.4 billion); and functional food ingredients, such as prebiotics and soluble fibre (£16 billion) (IO, 2018).

5.6.2 Trends in the sector

Between 2014-2016, employment growth in industrial biotechnology outpaced other industries by more than 10% per year, with a median income at ~£20,000 above the national average. As of 2018, more than 14,000 full employment positions were registered in the sector, with high potential for growth (IBLF, 2018).

According to an IBLF estimation, the annual turnover in the industrial biotechnology sector was £1.8bn in 2010 (IBLF, 2018). The direct economic effect in 2013/14 accounted for a £2.9 billion turnover, 8,800 jobs and a GVA of £1bn (Chambers et al., 2015). In 2014, the figures already show an economic contribution of £17.2bn in turnover and £7.4bn GVA (Bauen, 2016). As of 2018, more than 14,000 full-time

[23] https://www.bioindustry.org/uploads/assets/uploaded/d390c237-04b3-4f2d-be5e776124b3640e.pdf

employment positions were registered in the sector, with high potential for growth (IBLF, 2018).

Industrial biotechnology has the potential to disrupt markets in the fields of industrial gas manufacturing, soap and detergent manufacturing, fertilisers, perfume and cosmetics, plastics manufacturing and organic based chemicals worth more than £27 million combined (IBLF, 2018). Indeed, industrial biotechnology has the capacity to meet the increasing demand for sustainable and environmental solutions for the economy in these fields. Moreover, there are tangible downstream effects deriving from its outputs in pharmaceuticals, chemicals, healthcare, personal care products and energy sectors. Chambers et al. (2015) estimated that the portion of those sectors that rely on industrial biotechnology products accounts for a turnover of £34bn, £4.4bn in GVA and approximately 63,500 jobs.

5.6.3 UK Stakeholder submissions to the CBD

The UK bioindustry association (BIA) submitted their opinion to the CBD in 2019. They argue that DSI should not be part of the NP as this could undermine the CBD principles of ABS, by increasing compliance challenges to the industry.

BIA lists a series of challenges that the implementation of the NP still faces and argues that solving these challenges should be the focus of CBD parties, rather than expanding the scope of the NP to include DSI. Further, they explain that the NP, as it currently stands, is already having a negative impact on research and development on some of its members, due to legal uncertainties as to the exact nature and scope of the obligations to be fulfilled.

They argue that additional compliance barriers to companies and limiting open access to DSI will have a negative impact on the industry, especially on SMEs and be a detriment to innovation, which is the main reason why the sector has been thriving. They further expand on the DSI traceability challenges and the difficulty of changing the current database infrastructure.

6 Benefit Sharing Options

In late 2019, the first Global Dialogue on Digital Sequence Information on Genetic Resources was held in South Africa. This marked the first of two such dialogues between stakeholders on 'ways forward' on DSI, with a particular focus on benefit-sharing options. This section sets out a long list of potential options for benefit-sharing which take the discussions at the South Africa meeting as a point of departure and develops them further.

Some of the disagreements that has arisen regarding the relationship between DSI and the NP and the CBD are not merely technical. Rather, they reflect the range of different political, legal, and ethical views that prevail on these issues , as well as global power inequities which are reflected (subject to varying interpretations) in international law objectives and principles (notably the CBD objective of "fair and equitable" benefit-sharing). According to Laird et al. (2020), "there are clear inequities between the global North and South in research funding, control over resources and data, benefit sharing, and other issues that must be addressed."

Added to this is the fact that the prevailing bilateral construct of the NP and the CBD is increasingly unable to play 'catch up' to the way in which scientific research and innovation is conducted with open sharing of DSI via databases arguably guiding norm and a common practice. Databases in this context are generally open access - in essence, in the public domain and free – or open source, to which some conditions may be attached such as user agreements (see Laird and Wynberg, 2018).

While provider countries of genetic resources, typically located in the global South, have tended to frame the issue of DSI as one of fairness and equity, there is also an acknowledgement by all parties of the importance of scientific research and scientific cooperation to the achievement of the CBD's objectives. Indeed, and as noted throughout this report, significant benefits are generated from DSI in respect of conservation science, planning and management (Laird and Wynberg, 2018; Laird et al. 2020). A new approach is required to deal with the extensive and at times conflicting range of views surrounding these issues.

With the above context in mind, this report discusses below a range of potential policy options with respect to this issue of DSI. The first option discussed considers a 'do nothing' approach, followed by a discussion of potential mechanisms – both bilateral and multilateral – for monetary benefit-sharing, while the final option considers the potential for non-monetary benefit-sharing by means of capacity building and development. A number of guiding principles are set out to govern such an exercise.

6.1 Guiding principles

The study team identified the following principles to guide any process for benefit-sharing, based on the experience of the team's technical advisors at the University of Strathclyde:

- Solutions must take account of and be consistent with international law.
- Solutions must be able to contribute more generally to the effectiveness of the international biodiversity regime.
- Accordingly, solutions must contribute to the achievement of the objectives of the CBD, at a minimum.

- Solutions must not undermine and should contribute to achievement of the UN Sustainable Development Goals.
- Solutions must promote legal certainty and predictability, with a view to fostering international (scientific and other) cooperation – the current 'non-position' is no longer tenable.
- Financial security must be ensured in respect of monetary benefit-sharing as a guarantee of feasibility of the proposed solution.
- The needs and experiences of relevant actors, including safeguarding the norm of 'open science', must be taken account in the design of any solutions.
- Solutions can only be arrived at after an iterative process of dialogue with key stakeholders.
- Criteria for disbursement of funds must be designed in advanced and arrived at on the basis of a collaborative process of dialogue.

6.2 A long list of potential options

This section provides a description and opinion, drawing on the available literature, on the following options:

- Option 0: Exclusion of DSI from the scope of the CBD and Nagoya
- Option 1: Nagoya – bilateral
- Option 2: Open Access – bilateral
- Option 3: Multilateral options (under the CBD)
  - Option 3a: Subscription fee tied to database access
  - Option 3b: Tax, fee, levy
  - Option 3c: Payment upon commercialisation
  - Option 3d: Voluntary donations
- Option 4: Open Access – Capacity Development

6.2.1 Option 0: Exclusion of DSI from the scope of the CBD and Nagoya

Description: This may be considered as akin to a 'do nothing' approach. It would result in efforts to resolve the issue of DSI under both the CBD and NP not being pursued further. While this may appear beneficial in that current scientific practices could continue unabated, it is also the case that such a position would arguably "maintain an uncertain position or the perpetuation of disagreements" (Morgera, 2018). It would also have the potential to lead to a breach of the third objective of the CBD on fair and equitable benefit-sharing, thereby potentially jeopardising the position of the CBD. Leaving the issue unresolved in this way would further undermine multilateral processes and would run contrary to the international legal principle of good faith. It may also result in more restrictive domestic laws governing access and benefit-sharing, which could in turn hinder scientific discovery and biodiversity conservation as well as undermine other public goods (Sirakaya, 2019; see also section 3.3).

- **Opinion**: This is not a recommended option. The 'do nothing' approach risks undermining the position of the CBD at a time when biodiversity is under significant

threat. In the words of Laird et al (2020), “to not capture DSI would mean leaving a massive loophole in the ABS endeavour” (sic) with continuing disagreements over ABS posing grave threats for the ongoing CBD discussions on the post-2020 global biodiversity framework. As noted above (section 6.1), solutions must promote legal certainty and predictability, with a view to fostering international (scientific and other) cooperation. The current ‘non-position’ is hence not tenable as a permanent ‘solution’, a point developed further in the discussion of Option 4 below.

6.2.2 Option 1: “Nagoya – bilateral”

Description: This option would effectively embody the *status quo* under the NP. The NP mandates that prior informed consent (PIC) from the country of origin must be gained in order for genetic resources to be accessed. The NP further mandates that benefits arising from utilisation together with subsequent applications and commercialisation must be shared fairly and equitably on the basis of mutually agreed terms (MAT). The bilateral model inherent in the NP is predicated on the notion of a single resource generating benefits that are then shared with the provider country on MAT. Extending this to DSI would require MAT to specify provisions related to the use of DSI and also set out applicable benefit-sharing obligations (see ABS Capacity Development Initiative, 2019). While, as noted in section 3.3, certain countries already include DSI within PIC and MAT (see also Bagley et al., 2020), it is unclear to what extent such obligations can support non-monetary benefit-sharing in current scientific practices. As noted by Laird et al (2020), “research practices and concepts of ethics and benefit sharing associated with DSI that have evolved in recent decades within the scientific community emphasize openness, transparency, networks, and free exchange. By contrast, ABS is a transactional mechanism that restricts access to genetic resources so that their use can be exchanged for benefits between identified users and providers of these resources.” In addition , parallel uploading of MAT and DSI is not currently available in respect of the templates utilised by databases hosting DSI (ABS Capacity Development Initiative et al., 2019).

More fundamentally, the current bilateral system under the NP, when applied to DSI, raises significant questions as to the valuation of monetary benefits. A sequence may have generally little value on its own. Rather, value lies in its ability to be screened against other sequences to understand differences and commonalities in the traits and functions of different genes (Welch et. al., 2017).

In the vast majority of cases, sequences will be common across many different countries (Laird and Wynberg, 2018), presenting further difficulties in attributing ownership. In addition, current scientific processes often involve multiple actors in the creation of value from DSI. As noted by Laird and Wynberg (2018), “diverse networks of researchers from industry, government, academia, and community laboratories commonly span the globe in a system of “open innovation” in which users add incremental value through data and knowledge sharing along a chain that involves multiple databases and gene sequence.” The practical import of multiple users adding value incrementally is that it is difficult to see how a bilateral model predicated upon the current ABS regime under the Protocol could operate effectively. When, for example, would benefit-sharing obligations cease?

Opinion: Difficulties with the application of the current bilateral model of the NP to DSI presents problems for ensuring fairness and equity in benefit-sharing. While a range of countries have implemented domestic measures applying PIC/MAT to DSI (Bagley et al., 2020), issues with valuation, combined with traceability issues (section 4.4) mean that application of the current bilateral, transactional model of ABS to DSI is problematic (Laird et al., 2020). Leaving issues of fairness and equity in benefit-sharing in respect of DSI to be resolved bilaterally would not only result in a 'non-solution' but could, as noted above, result in countries at a domestic level imposing more restrictive access policies, with a knock on effect on scientific development and innovation (section 3.3).

6.2.3 Option 2: "Open Access – bilateral"

Description: Under this option, benefit-sharing would not be specified by bilateral MAT but rather through terms and conditions defined on a sectoral or "regime" basis (see ABS Capacity Development Initiative, 2019). In other words, terms and conditions would be attached to the use of DSI obtained from databases with the user informed of such terms when accessing DSI via a database. This would require country tags to accompany uploaded DSI so users can direct monetary benefits back to the provider country. Users could be made aware of any such terms and conditions of use via, among other things, click-through-agreements which require acceptance before access to a sequence/database is granted. Such agreements can also provide a record that the terms were accepted by the user (Perron-Welch, 2018). DSI under such a system would be open access, with benefit-sharing obligations triggered later, likely at the point of commercialisation with benefits to be shared with the provider country. Other types of user agreement may also be possible (Laird and Wynberg, 2018).

Opinion: Click-through agreements requiring users to confirm their consent to certain conditions when accessing DSI via specific databases already exist. It is, however, not very clear to what extent these conditions are actually abided by (Laird and Wynberg, 2018). While there may be ways to ensure greater prominence and moreover compliance with click-through or similar types of agreements (Laird and Wynberg, 2018), these would need to be accompanied by enhanced traceability mechanisms to assist with enforcement. More generally, a bilateral system, even if implemented through terms and conditions agreed to at the point of access to a database, would likely encounter the same issues as discussed under Option 1. A bilateral system for DSI, for example, raises significant questions as to the valuation of monetary benefits, particularly given the presence of shared and homologous sequences (Laird and Wynberg, 2018). More generally, and as stated above, application of the current bilateral, transactional model of ABS to DSI is something of a poor fit with current scientific practices (Laird et al., 2020), where value is rarely generated from a single sequence but from the ability, among other things, for it to be screened against others. Hence, while this option does have the advantage of allowing DSI to be accessed for free, it nevertheless does not address the significant issues with valuation associated with how science is actually done. The failure to address sufficiently questions of valuation also has a consequential effect on the fairness and equity of any resulting monetary benefit-sharing arrangement.

Synthesis: Options 1 and 2 discussed above encompass DSI within the context of a bilateral ABS framework. However, a range of difficulties surrounding the application

of such bilateralism to DSI have been identified. These difficulties have been discussed extensively in the literature (e.g. Laird et al., 2020; ABS Capacity Development Initiative et al., 2019; Vogel et. al., 2017).

The difficulties of a bilateral approach can be traced to a number of issues, some of which are listed below:

- Problems with valuation: Monetary benefits are difficult to value on a bilateral basis, in part because value rarely lies in a single accession but rather more typically manifests in the potential for a sequence to be screened against others (Aubry, 2019).
- Problems with operationalisation of a bilateral system: Many sequences are not unique to a particular country. Indeed, conserved sequences, which "are similar or identical sequences in DNA, RNA and proteins or polysaccharides occurring across species, or within different molecules produced by the same organism" are very common (Laird and Wynberg, 2018). This, alongside the availability of *ex situ* specimens and public databases, means that bilateral arrangements are increasingly difficulty to operationalise (Laird and Wynberg, 2018).
- Low monetary benefits: Where a sequence is shared between jurisdictions, the possibility of jurisdiction shopping results in oftentimes very low monetary benefits (Ruiz Muller, 2018; Ruiz Muller, 2010).
- Fairness, equity and support for conservation and sustainable use: Low monetary benefits raise issues of equity, which has led to questions as to whether the current bilateral ABS approach embodied by the Nagoya Protocol does in fact support fairness and equity in benefit sharing and overall conservation and sustainable use of genetic resources (Ruiz Muller 2018; ; Laird et al. 2020).

The next set of options examines whether multilateral – as opposed to bilateral - options might exist for monetary benefits generated from the use of DSI. Such monetary benefits could be disbursed to Parties via a multilateral fund or used to fund particular projects aimed at supporting the objectives of the CBD (Bagley, 2016). Under such an arrangement, DSI would be (largely) available on an open access basis, with the norm of open science prevailing and monetary benefits generated and distributed fairly and equitably through multilateral processes.

The options explored below take as their point of departure the discussions of participants at the first Global Dialogue on Digital Sequence Information on Genetic Resources held in South Africa in 2019 (ABS Capacity Development Initiative et al., 2019). However, the options discussed in the South Africa meeting are refined and expanded upon here, to provide a potentially more comprehensive list of possible approaches going forward. In addition, the potential mechanisms for accrual of funds are discussed, together with possible spaces for the governance of such funds, as well as how monetary benefits might be disbursed. While the primary focus of this assessment is on monetary benefit-sharing, an option (option 4) of non-monetary benefit-sharing is also considered. Finally, these options, while presented separately, should not be thought of as mutually exclusive (ABS Capacity Development Initiative et al., 2019) and could be combined or structured as part of a phased, stepwise approach.

6.2.3.1 Option 3 (a) Subscription and/or access fee tied to database access and established under CBD

Description: An obvious way to generate funds for monetary benefit-sharing tied to DSI is to charge a fee for access to DSI. Such a fee could be payable upon access to databases hosting DSI and chargeable on a one-off basis. An alternative to this would be the introduction of a subscription fee for database access.

Monies generated from this could be disbursed in a number of ways. The first would involve any fees generated being redistributed directly to States as 'provider countries'. There would hence be an attempt to link provider country deposits to disbursement of funds. The criteria for this would need to be designed in advance and it is not altogether clear that megadiverse countries would benefit since, under the International Nucleotide Sequence Database Collaboration (INSDC) at least, the bulk of country-tagged deposits are from the US, Japan, China and Canada (Rohden et al., 2020). Other possibilities to 'match' payments to provider countries include the use of a mechanism to determine geographical distribution of genetic resources such as the International Barcode of Life (Ruiz Muller, 2018).

An alternative would see applicable funds distributed by a multilateral fund, with disbursement according to criteria identified in advance through a deliberative international process and in accordance with CBD objectives, including conservation and sustainable use. The fund could be established under the CBD, with funds disbursed via project-based funding, perhaps managed by the Global Environment Facility (GEF) (ABS Capacity Development Initiative et al., 2019), which is the existing financial mechanism for implementation of the CBD.

Under this benefit-sharing mechanism, there would be no attempt to track and trace, and disbursement would not be 'matched' to country deposits, thereby reducing transaction costs (Tsioumani, 2019). One advantage of locating the multilateral benefit-sharing mechanism under the CBD is its ability to rely on an existing treaty structure with almost universal membership. The CBD also provides the objectives and international obligations under which an exercise could be structured. While the NP could also offer such a forum via the global multilateral benefit-sharing mechanism foreseen in Article 10 (Vogel et al., 2019), its limited membership combined with continued controversy over its temporal application make it less of an attractive prospect than the CBD. A similar preference for any multilateral fund to be situated at the CBD also extends to the other options (Option 3(b) to 3(d)) discussed below. Finally, other possibilities exist beyond the GEF in respect of the mechanism by which funds might be disbursed, and experiences with other multilateral funds could be drawn upon in assessing the pros and cons of particular options.

Opinion: While a number of specialised databases do require payment for access (Winter, 2013), public databases such as those within the INSDC are free. Indeed, it is difficult to reconcile the core mission of the INSDC to provide, "free, unrestricted access to all of the data records in their database" (Laird and Wynberg, 2018) with the introduction of a payment scheme for access. The practical issue of the fact GenBank is hosted by the US – and indeed, funded by US taxpayers (ABS Capacity Development Initiative, 2019) – is an accompanying concern. In addition, any payment or subscription fee would need to differentiate between different categories of user(s) including, for example, commercial and non-commercial users (see generally Prathapan et al., 2018). In practice, the demarcation line between these and

other categories of user can be difficult to draw. The design of any such payment scheme would need to be sufficiently nuanced to avoid any particularly detrimental impact upon researchers and companies from the Global South. Most managers of databases, and indeed researchers, are not in favour of the adoption of a fee-based system (Laird and Wynberg, 2018).

While it is difficult to envisage the introduction of a system for payment upon access to public databases such as those under the INSDC, more delineated sectoral approaches may be more of a possibility. Indeed, a subscription system based upon upfront payments was proposed as part of the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) inter-sessional process (Morgera, 2016). In addition, a sectoral approach might be better placed to take account of the economics of different sectors and to respond to the needs of different categories of users. However, a subscription fee would only work with respect to a relatively closed system of DSI; for example, if a sequence can be found elsewhere for free – as it often can – there is little incentive for potential users to pay for it. For databases that host DSI potentially accessible elsewhere, subscription fees would likely be best operationalised where there is a value-added component available from the database. However, this would likely require development of a range of bioinformatics tools, which in itself would require expenditure, thereby posing difficult questions for the economics of such a system (Rohden et al., 2019).

It is difficult to see how a subscription fee or a fee payable upon access to DSI could be operationalised in practice to encompass the large range of DSI in scope for the following reasons:

- First, payment of a fee is fundamentally at odds with the underpinning ethos of the likes of the INSDC. It would also put the strain of benefit-sharing primarily on scientists (ABS Capacity Development Initiative et al., 2019). While sectoral approaches appear to offer more potential, any such scheme would need to be relatively closed or offer a value-added component, which may raise question marks in respect of the economics of such a system.
- Second, the application of a fee for access/ a subscription fee would need to be able to differentiate between different categories of user which could prove difficult in practice. Such differentiation is particularly important to ensure that any payment for access to DSI does not undermine certain objectives of the CBD, such as biodiversity conservation and sustainable use. In essence, concerns would arise if any requirement for payment were to introduce financial barriers to such work.
- Third, the significant public funding already given to public databases needs to be considered. The INSDC is publicly financed by the US, EU and Japan and accordingly, it is not clear to what extent adding another layer of cost would receive support from such key stakeholders. Perhaps most fundamentally, any introduction of a fee for access to DSI would require close dialogue with the range of stakeholders along the DSI value chain, with databases arguably central to any process of dialogue.
- Finally, the possible movement of databases to non-Parties, most notably the US, needs careful consideration. Dialogue with the range of stakeholders involved in the DSI value chain would hence be a fundamental starting point to the introduction of any such subscription or access fee, with a stepwise sectoral approach arguably the most appropriate way forward should this option be pursued.

6.2.3.2 Option 3 (b) Tax, fee, or levy established under the auspices of the CBD

Description: A tax, fee or levy could be considered as a way of generating funds for monetary benefit-sharing linked to DSI (e.g. Laird et al., 2020). A number of proposals have been advanced in this regard, including a mandatory levy or tax on products and processes associated with the 'creation' of DSI. This would potentially cover the purchase of equipment such as sequencing machines (ABS Capacity Development Initiative, 2019). It is to be noted, however, that the introduction of such a tax, fee or levy could potentially have a greater impact on researchers than commercial entities and so careful thought would be needed to avoid particular negative externalities (ABS Capacity Development Initiative et al., 2019).

Another possibility to generate funds for monetary benefit-sharing would be to expand the scope of any tax, fee or levy beyond products and processes associated with the 'creation' of DSI. This could, for example, involve the application of a small fixed fee to a range of biotrade products, an idea which derives from the literature on traditional knowledge (e.g. Ruiz Muller, 2013) or a more general biodiversity tax, based on the concept of applying a small charge to any activity, product or process derived from a genetic resource (Winter, 2013). A further option would be to extend the tax/levy to include a wide range of activities with the potential to harm biodiversity (Laird et al., 2020). This could include, for example, mining and logging activities. Extending coverage of such a tax/levy in this way would ultimately result in the creation of a Pigovian-type tax. A flat rate charge on products derived from DSI, adapting the model of Brazil (Brown, 2017), is yet another possibility.

Objections to a more broadly specified tax, levy or fee include the contention that activities such as mining are already regulated and, at least in some cases, taxed in the state of activity. Clearly, however, legitimate questions may be raised over the effectiveness of such regulation and taxation given the continuing loss of biodiversity globally. For example, the 2019 Global Assessment of Biodiversity and Ecosystems Services indicated that current negative trends in biodiversity and ecosystems will undermine progress in respect of 80% of the targets under the Sustainable Development Goals related to poverty, hunger, health, water, cities, climate, oceans and land.

Other potential objections to a tax, fee or levy applied to a broader range of products and processes than those directly linked with DSI relate to the legitimate question as to scope. In essence, the criticism may be made that it would be a step too far to include products and processes not necessarily linked to DSI. In response, it could be stated that a biodiversity levy or tax, however constructed, would not be another revenue raising tax, but would be used to fund projects directly related to the conservation or sustainable use of biodiversity, thereby contributing to relevant Sustainable Development Goals. Positioning the taxes/fees in this way may help such proposals to gain more by way of appeal, particularly because the current bilateral, transactional ABS regime when applied to DSI has arguably drawn attention away from the biodiversity challenges the world is currently confronting (Laird et al., 2020).

Opinion: The specific design of the measure will be a crucial determinant of its success. Taxes, charges and levies would undoubtedly need to be implemented domestically with funds generated sent to a multilateral entity for disbursement. As noted above, there are significant advantages to locating such a multilateral benefit-sharing mechanism at the CBD: because it enjoys almost universal membership and

because it is the CBD that provides the objectives and international obligations to structure such an exercise. Similar considerations in terms of fund disbursement and a potential host organisation to that discussed under Option 4 (a) would hence apply.

The fact that a tax, charge or levy will, however, require domestic implementation may present difficulties, both from a practical perspective given the likely costs involved, as well as in respect of addressing concerns regarding the impact of any such tax, fee or levy upon research. In addition, any such tax, fee or levy would need to take into account a range of country situations, and would need to take a nuanced, sectoral approach. While a properly designed tax/levy should in principle provide for the financial sustainability of such a system, the practical operation of such a scheme, together with its impacts upon particular sectors, needs to be carefully considered. Issues pertaining to enforcement would also need to be resolved (ABS Capacity Development Initiative et al., 2019).

6.2.3.3 Option 3 (c) Payment upon commercialisation with disbursement from a multilateral fund established under the CBD

Description: A number of commentators have proposed the payment of royalties attached to commercialisation – likely attached to intellectual property such as the grant of a patent - derived from a successful innovation based on DSI (Ruiz Muller, 2018). Disbursement of such royalties would be through a multilateral fund established under the CBD (see discussion above Option 3 (a)) and could potentially be directed back to the country of origin, or, given the prevalence of shared and homologous sequences, used to support projects aimed at biodiversity conservation or sustainable use. Similar considerations in terms of fund disbursement to that discussed under Option 3 (a) would apply.

Opinion: One potential issue that would need to be overcome with respect to a charge on commercialisation is that, simply put, commercialisation of products takes time. Relying on commercialisation to produce funds for monetary benefit-sharing means that little by way of monetary benefits are likely to be produced, at least in the short term. This has certainly been the case with the ITPGRFA which is currently reviewing the operation of its standard material transfer agreement (SMTA) to enhance monetary benefit-sharing due to a long and uncertain timeframe for commercialisation (Ruiz Muller, 2018). While the issue identified here assumes that the temporal scope of any such commercialisation agreement is prospective, it is difficult to foresee its retrospective application. Clearly, if financial security is to be a guiding principle for any system of monetary benefit-sharing, this would require careful consideration and hence this option may need to be combined with other potential mechanisms for accrual of funds in order to ensure such stability.

6.2.3.4 Option 3 (c) Voluntary donations

Description: Voluntary donations are another way in which funds could be raised for a multilateral benefit-sharing mechanism. Such donations could be sought from a range of actors, from states to other actors (including the biotech sector). These donations could then be disbursed multilaterally in accordance with defined criteria focusing on conservation and sustainable use. The Treaty objectives of the CBD could be drawn on to establish the fund, with funds potentially disbursed via the existing facility of the GEF. As previously discussed above, the CBD has almost universal

membership, it would be a particularly attractive venue to host such a fund. Voluntary donations could also follow the model of the World Health Organization (WHO) Pandemic Influenza Preparedness (PIP) Framework Partnership Contribution (Rourke, Phelan and Lawson, 2020), with a questionnaire sent to large users, for example, of the INDSC to seek donations to fund activities focused on conservation and sustainable use, as well as capacity building initiatives. Governments could also be asked to contribute on a voluntary basis.

Opinion: Voluntary contributions are perhaps the least controversial option for the multilateral accrual of funds for monetary benefit-sharing. More generally, the multilateral system is heavily reliant upon voluntary contributions, to varying degrees of success. Indeed, one concern that might arise from reliance upon voluntary contributions is the fact that such a system may do little to guarantee financial stability given the unpredictability of voluntary funding sources. International institutions and funds are notoriously underfunded. Accordingly, while voluntary donations are the least controversial option, they are perhaps the most uncertain. Even if funding is secured, it could potentially pull funding away from other existing funds or institutions. If financial stability is to be a core guiding principle in respect of the establishment of any multilateral fund for benefit-sharing in respect of DSI, clearly more consideration needs to be given to how such a scheme could operate to ensure adherence to this key principle where voluntary contributions are concerned. Accordingly, and taking into account the principle of financial stability, voluntary donations may need to be combined with the other options discussed here.

6.2.4 Option 4: Open Access – Capacity Development

Description: Under this option, which to an extent reflects the current *status quo*, no monetary benefit-sharing obligations would be attached to the use of DSI and the current open access model would prevail. Proponents of this option might stress that significant non-monetary benefits already exist with respect to the use of DSI (section 4.5) and that the publication of sequences in itself represents a significant form of benefit as it allows anyone to use that information. However, there is a significant disparity in the ability of researchers to access and use DSI and accordingly, capacity development for the users of DSI, specifically those in the Global South, would be introduced pursuant to this option as a non-monetary benefit-sharing measure.

Opinion: As noted by participants at 2019 South Africa meeting, while this option largely reflects the current default situation, it is unlikely to "satisfy CBD Parties that are demanding that benefits arising from the use of DSI be shared fairly and equitably" (ABS Capacity Development Initiative et al., 2019). Indeed, in keeping with the current *status quo*, this option is likely to offer something of a non-solution and is unlikely to attract significant support from key stakeholders as a 'full' answer to the issues raised by DSI. However, it could prove an attractive option (Laird and Wynberg, 2018) if introduced in conjunction with one or more of the other options outlined above in respect of multilateral monetary benefit-sharing. Non-monetary and monetary benefits are not mutually exclusive, since the former requires funding. Should this option be implemented, provider countries would still be free to set down their own non-monetary benefit requirements for users under PIC/MAT for physical samples (which is not included within the scope of this study).

6.3 Synthesis

The analysis of the study team completed for each of the potential options demonstrates that a multilateral approach to monetary benefit-sharing which prioritises open access to DSI is clearly preferable either a 'do nothing' approach or to the adaptation of the current bilateral ABS regime currently applicable under the NP and/or CBD. Furthermore, while there is an overlap between non-monetary and monetary benefit-sharing, it may be argued that an absence of monetary benefit-sharing from any potential solution would be unlikely to satisfy the demands of certain biodiversity-rich countries that benefits arising from DSI are shared fairly and equitably (see also ABS Capacity Development Initiative et al., 2019).

There is no one size-fits-all solution in respect of how funds for a multilateral benefit-sharing mechanism might be accrued. Options for monetary benefit-sharing may need to be combined, so as to ensure financial stability of the system. Other options not currently explored in the literature may also need consideration. Possibilities in this regard involve the creation of a biodiversity bond, perhaps drawing on the various innovative funding mechanisms within the global health arena such as the Vaccine Alliance's innovative International Finance Facility for Immunisations (IFFIm) which allows long-term financial pledges to be made by donor states with 'vaccine bonds' – in essence, debt instruments – then sold to investors (Klock, 2016). An underpinning advantage of such a mechanism is that it allows funding to be raised in the short term, with donors committing to funding over the longer term (Gostin, 2014).

In terms of the disbursement of funds, a multilateral system for monetary benefit-sharing is likely to be preferred. This is in part because of the fundamental challenges that undermine bilateral approaches. As outlines above, the CBD is arguably the most appropriate forum for the creation of any such mechanism. Disbursement of funds would need to be in accordance with predefined criteria agreed to by applicable stakeholders. Any such criteria would, as a minimum, need to contribute to the CBD's objectives and also to achievement of relevant SDGs. In addition, careful consideration should be given to the design of governance structures so as to ensure equity as well as responsiveness to both the needs of beneficiaries and take account of any changes in scientific practices. In this regard, regular dialogues with stakeholders, so as to take account of the implications of scientific and technological developments for benefit-sharing and the fund, would arguably be required as a minimum.

7 Assessment of the Benefit Sharing Options

7.1 Options discussed

Based on the options introduced in section 6, it was agreed there would be a focus on a short-list of options for the stakeholder consultation process. As discussed in section 2, the priorities were to maximise the number of options discussed during the interviews, which reduced the number of criteria against which interview questions were asked, that is, the scope and detail of the interview questions were reduced compared to the original framework proposed in order to reduce the time required to discuss each option.

The options were assessed considering the conclusions from the latest AHTEG meeting (AHTEG, 2020):

- In clarifying the scope of DSI, AHTEG agrees that groups 1-3 (see DSI definitions, section 3) can be considered DSI, whereas Group 4 (that includes information not related to molecular structure or traditional knowledge) is not DSI. There is no agreed DSI definition.
- The AHTEG discussed implications on the different definitions and agreed that these would depend on the nature of the benefit-sharing approach.
- Experts agreed that “any approach to benefit-sharing should provide legal certainty, incentivise the use of digital sequence information and decrease unnecessary burden in monitoring, tracing, and tracking requirements”.

The final selection of options and criteria discussed are summarised in Table 7.1 (options and criteria considered but excluded are summarised in Annex 2).

Table 7.1 Summary of options and criteria selected

Questionnaire framework
Options discussed
Multilateral Monetary Benefit-sharing – data access fee
Multilateral Monetary Benefit-sharing – payment upon commercialisation
Multilateral Monetary Benefit-sharing – tax, fee or levy
Non-monetary Benefit-sharing
Assessment criteria
Viability: questions on the technical and legal viability
Impacts: questions on anticipated costs, benefits and overall competitiveness
Support: comparative question on overall support for each option

7.2 Summaries of the policy options assessed

The following presents short summaries of the options that were discussed during stakeholder interviews (and were presented to the interviewees ahead of the interview).

7.2.1 Option 1: ‘Multilateral Monetary Benefit-sharing’

Monetary benefits are generated through database access fees, payments tied to commercialisation and/or taxes or levies. Parties receive the benefits through distribution via a multilateral process (e.g. fund). The funding mechanism could be managed by the Global Environmental Facility (GEF) under the CBD. The degree of

open access for users depends on the stage at which payments are required. Sub-options could include one or more of the following:

7.2.1.1 Option 1a: Database access fees (one-time or subscription-based)

Database access or subscription fees could be introduced. The level of fee could be based on the type of users or the intended use, in order to address the needs of different categories of users. Fees could be distributed in different ways:

a Benefits could be directed at specific provider countries by linking provider country deposits/DSI to disbursement of funds or through 'matching payments' by using a mechanism to determine geographical distribution of genetic resources such as the International Barcode of Life.

b Benefits could be distributed through a multilateral fund amongst a larger number of Parties according to pre-established criteria (e.g. where the funds are used to support biodiversity conservation or sustainable use). There would be no attempt to match payments to provider countries and funds.

7.2.1.2 Option 1b: Payment upon commercialisation. For example, royalties could be paid in relation to the grant of a patent for a successful innovation based on DSI.

Under either of these sub-options:

a Benefits could be directed at specific provider countries by linking provider country deposits/DSI to disbursement of funds or through 'matching payments' by using a mechanism to determine geographical distribution of genetic resources such as the International Barcode of Life.

b Benefits could be distributed through a multilateral fund amongst a larger number of Parties according to pre-established criteria (e.g. where the funds are used to support biodiversity conservation or sustainable use). There would be no attempt to match payments to provider countries and funds.

7.2.1.3 Option 1c: A tax, fee or levy:

A tax, fee or levy may be placed on one or more parts of the DSI value chain:

a Products and processes associated with the 'creation' of DSI (e.g. sequencing machines).

b A broader range of 'biotrade' products and/or processes involving the use of genetic resources (building on approaches suggested for traditional knowledge).

7.2.2 Option 2: 'Non-monetary Benefit-sharing'

Capacity development is provided for DSI 'creators' and users in the Global South to address the disparity in the ability of researchers to access and use DSI in many provider countries.

Capacity development could include:

a Education and/or information exchange (e.g. through courses and workshops, exchange visits, sharing of lessons learnt, intercultural dialogue with indigenous groups and local communities regarding traditional knowledge).

b Collaboration in scientific research and technology transfer, including through regional networks.

c Support for development of scientific infrastructure, including through regional approaches (for example, CGIAR centres).

This option would require investment to secure these types of arrangements. This could happen via a secure multilateral fund targeting researchers, for example, in megadiverse countries or in countries with less capacity or access to technology.. The latest AHTEG meeting (AHTEG, 2020) also outlined numerous potential options for capacity building.

7.3 Assessment of the proposed options

This section provides an overall assessment of the options by drawing on the data collected in the interviews.

7.3.1 Multi criteria assessment

The options assessment is based on the following criteria:

- **Viability**: Will the option work in practice? In particular, from a technical and legal perspective:
  a. Technical feasibility:
     - What are the anticipated operational challenges to implementation?
     - Are these expected to be minimal (i.e. resolved with minor or no changes to the way DSI is currently used/accessed, moderate (i.e. requiring modest but not major changes to the way DSI is currently used/accessed, for example changes or adjustments within existing systems) or major (i.e. requiring significant changes to the way DSI is currently used/accessed such that, for example, new systems would need to be put in place)?
  b. Legal coherence and conformity:
     - Is the option viable under existing UK law or would it require amendments or new laws to be enacted?
     - Does it align with UK obligations as a party to international agreements or participation in international systems (e.g. standards)?
     - Are there potential challenges for those accessing or using DSI in terms of legal certainty (i.e. can users be confident that they are able to comply with the likely requirements of the option)?
     - Are there any anticipated challenges to the enforceability of the option (i.e. would authorities be able to use legal and administrative means to encourage or compel compliance)?
- **Impacts:** What are the competitiveness implications for those who access and use DSI? What are the other potential benefits and/or costs (for providers and users of DSI)?
  a. Costs**:**
     - What are the types and level of anticipated costs associated with the option (e.g. will it be costly to implement, manage, or participate in)?
     - How/how much (e.g. type and scale of costs)?
     - Who is likely to bear these costs?

b. Benefits:
   - What are the types and level of anticipated benefits associated with the option?
   - How/how much (e.g. type and number of Parties who benefit, level of benefit)?
   - Who is likely to benefit from the option?

c. Competitiveness*:*
   - What impact does the option potentially have in terms of industry's ability to operate or innovate? For example, would it enable businesses to produce more or higher quality products and services?
   - In what ways does the option potentially stimulate (or hinder) research? For example, in what ways does the option facilitate the introduction and/or dissemination of new methods, technologies or products?

- **Support**: What is the overall level of support for the option?

  This criterion represents a general reflection of the extent to which the option addresses the needs and concerns of specific groups or individuals in relation to its potential viability, effectiveness, and impacts. Support will be registered on a scale of low/medium/high with a description of the reasons for this score.

7.3.2 Sectors consulted

The stakeholder consultation categorised stakeholders into two types of DSI actors, commercial users, and non-commercial, to facilitate the analysis, as many sectors share similar characteristics that relate to the impact of the proposed options. In doing so, unnecessary repetition of impacts is avoided, and the analysis is simplified to better understand the direction of change of the proposed options. DSI data repositories stakeholders are part of the non-commercial actors category, while opinions from other experts and regulators have been used to explain the nuances of assessment.

Table 7.2 Characteristics of user groups

User	Characteristics
Non-commercial	High dependence on DSI, both in volume and in importance to sector for their area of business
	DSI used for non-commercial research and other purposes (e.g. education)
	Often involved as intermediaries in sharing DSI between organisations in each sector and across sectors
	Access databases often
	Often have access to fewer resources
Commercial	Dependence on DSI varies by sector
	Often download an entire database for use later
	DSI relatively substitutable (i.e. multiple databases may contain the data required)
	Fewer and less well-established access and benefit-sharing practices
	DSI use most likely to be for commercial purposes; commercialisation of products derived from DSI can take a very long time

7.3.3 Ranking the impact of the proposed options

The options assessment is summarised in Table 7.4. Each option has been assessed according to the criteria set out in Section 7.3.1 using the scoring system outlined in Table 7.3.

The scale and exact nature of the expected changes under the proposed options are unknown, and therefore the assessment is based on a qualitative judgement of impacts, according to the likely direction of change for a given option and the likely intensity of that change.

Table 7.3 Options assessment rating

Rating	Description
	May have major impact
	May have moderate impact
	May have minimal impacts
	Impacts uncertain – stakeholders disagreed or there is not enough data to assess the option

Table 7.4 Assessment of options: summary table

	Stakeholder	Option 1a	Option 1b	Option 1c	Option 2
Viability[24]					
Technical	All	Infrastructure, legal, and administrative changes expected. Viability depends on how the option is enforced.	Depends on the point in the value chain when payment is introduced. Potential requirements of audit documents for collaborators.	Unclear. Opinions presented vary widely.	Potentially minimal. Organisations already do this but depends on the obligations.
Legal	All	Can lead to major legal burden	Potentially major; depends on how it is operationalised	Unclear. Opinions presented vary widely.	Potentially minimal. Organisations already do this but depends on the obligations.
Impacts					
Costs	Non - commercial				
		Legal, administrative, financial, time scale	Not applicable – although it impacts collaborations with others	High (e.g. sequencing machines) to none	Time, financial, administrative, management changes expected
	Commercial				
		Legal, administrative, financial, time scale	Legal, licensing, administrative. Can lead to uncertainty	High (e.g. sequencing machines) to none	Potentially minimal. Most companies already do this, but it depends on the obligations.
Benefits	Non - commercial				
		It could offer legal certainty if all countries agree	It depends, if done with clarity and simple it can provide legal certainty and protect biodiversity	Unclear	Major benefits expected and any option discussed should include this.
	Commercial				
		None unless databases upload new and well curated data	Given the complexity of operationalising it, no benefits	Unclear	Major benefits expected and any option discussed should include this.
Innovation	Non - commercial				
		Detrimental to research and conservation. Unfair unless there is a tiered system	It could be detrimental for small organisations	Unclear	Potentially positive, but uncertainty as to how it works.
	Commercial				
		Likely to hinder collaboration and innovation	It depends. It could hinder small organisations' capacity	Unclear	Potentially positive, but uncertainty as to how it works.

[24] The viability criteria has not been broken down by stakeholder type as the opinions across different types were consistent

7.3.4 Summary of the assessment of the options

Option 1a: Multilateral Monetary Benefit-sharing' – database access fees

All types of interviewees agreed in their assessment of this option: they do not support it, particularly when compared to the status quo or the other options. This option, from their perspective, goes against current open data policies that are in place for many existing DSI databases, and against wider global trends for increasing open data.

It is expected that the option will reduce access to DSI data. This may restrict and hamper research, reducing the private and societal benefits from DSI use. Reduced access may occur in a number of ways; examples provided include:

- Discouraging researchers from sharing their own data, which would reduce the overall level of available DSI and hence impede scientific use.
- Increasing access costs may result in users not being able to afford access, potentially disenfranchising a range of users.
- A risk that data is no longer shared openly, in contrast to current open access policies, through which users can access and share their data at any point.

Interviewees indicated a preference towards maintaining the status quo and not having to pay to access data. However, most interviewees indicated that, if this option were to be implemented, they would consider paying for access. Some interviewees explained they would prefer to pay for new data to be uploaded (but not for existing data). If they were to accept an access fee option, most interviewees indicated a preference for a 'reasonable' flat rate access fee, irrespective of how many times one accesses data or downloads it.

Technical and legal viability

Depending on how the option is operationalised, interviewees could require technical changes in their administration systems. When comparing the option to the current system used by journal articles, interviewees agreed that this would be feasible. Alternative ways of operationalizing the option may lead to complex changes for any organisation. As experts discussed in the AHTEG meeting (AHTEG, 2020), traceability of DSI to provider countries, and developing an infrastructure to monitor and report on its use, is likely to be complex and expensive.

However, interviewees noted that there are currently thousands of databases available (see also section 4.2). This option would need to guarantee all databases are part of the agreement, or that those under a paywall offer value added to what is already available (e.g. better curated data, new sets of sequences) (see also section 4.3). Otherwise, those databases requiring a fee may not be used. Questions raised by interviewees included:

- Would an access fee only be applied to some databases, for example those established by provider countries?
- How can it be guaranteed that provider country data is useful and not freely available somewhere else?
- Would the access fee only apply to data that is not currently available for free, or would the fees apply to all existing data?
- Would there be a flat rate, or would they pay per download?
- On what basis would the 'cost' of data be assessed, and access fees set?

Impacts

Costs would likely be high: in addition to the fees, new infrastructure would need to be developed, and there would be associated administrative and legal costs. Depending on how the option is implemented, it could lead to major long-term costs for all organisations in terms of time and financial costs. Beyond potential legal certainty, assuming all databases available are part of the agreement, there are no benefits perceived for the option.

The option may hinder research on conservation, biodiversity, and public health issues among others, thereby having a negative societal impact. The option is not considered a fair option, and to ensure the level of access fee imposed on different user groups is fair, several interviewees suggested the following be considered:

- Tiers of payment by type of user (i.e. commercial / non-commercial). This was raised as commercial users often download the database once and store it, while non-commercial users tend to download data more often.
- Tiers of payment by capacity to pay, for example, by company size. Most interviewees thought that this option would otherwise impact small organisations disproportionately.

Option 1b: 'Multilateral Monetary Benefit-sharing' - Payment upon commercialisation

Stakeholders agreed that this option seemed fair, compared to the other options. The option has support from most interviewees - assuming the payment is placed fairly on the value chain, and fees are proportionate to the benefits obtained (i.e. a percentage of profit) and/or DSI used. However, the option is seen as very complex to implement.

Identifying the fee to be paid poses challenges:

- How to 'value' how much DSI was used to develop a specific product? (section 4.3).
- How much is 'fair' to pay? As profit margins differ greatly between sectors and companies.

Technical and legal viability

This option may require identifying who is responsible for the 'commercial' aspect to ensure legal certainty, leading to traceability challenges. Enforcement of the option given the traceability difficulties (section 4.4) would be extremely complex. It is considered almost impossible to operationalise from a policy perspective. This is linked to two challenges:

- Defining commercialisation: How do you define commercialisation? When do you introduce fees? Are there different rates, depending upon if the product is a global public good (e.g. vaccinations) or not? Do you pay royalties as a percentage of your profit or do you pay a fee when you patent (even if there is no profit)?
- Traceability: Can you trace all users of DSI and regulate them or will it be based on honesty? At what point do you pay and who is responsible for determining its definition and judging when a user has reached it? How would it be enforced? Could it be enforced via disclosures to patent offices where the commercialisation process results in a patent application?

For non-commercial users this option would require minor changes and it could be viable. However, should this option require a traceability system so commercial users know the fee to be paid when they commercialise the product, it could also require non-commercial users to keep audit trails of their collaborations. This would impact

them negatively by increasing the administrative burden related to traceability and reduce legal certainty.

Depending on how the option is implemented, commercial users may require minor to major changes. The two main challenges raised were:

- Defining commercialisation: being able to determine 'when' to pay and 'for what', potentially leading to legal uncertainty.
- Traceability issues, and the viability challenges related to that.

Impacts

Issues around the potential costs and benefits related to this option were raised:

- Depending on how the option is operationalised, it could help companies have clarity upfront of what the costs would be and protect their intellectual property if patents are part of the process.
- Most non-commercial users would not be impacted by this option.
- It could make collaborations between non-commercial users and commercial users more difficult. This is particularly the case, as it is sometimes difficult to identify if the research carried out would have a commercial end or not.
- Increased costs are perceived as creating barriers to commercialisation and to accessing and sharing data and collaborating, impacting negatively smaller or less profitable organisations (e.g. start-ups).
- Depending on the tracking requirement of the option, data repositories and commercial users would require moderate to major changes in the way they operate currently.
- If all parties agree, it could provide legal certainty in the long term and the ease of only negotiating with one party rather than multiple databases / countries.

Regarding benefit-sharing, the option could protect biodiversity, generate profit for provider countries, and avoid imposing additional costs on non-commercial users. However, the redistribution of benefits would, potentially, take a long time. DSI product development often takes many years, so funding would only be generated in the long term.

Option 1c: 'Multilateral Monetary Benefit-sharing' - A tax, fee or levy

Many interviewees found the option difficult to understand and to conceive how it could realistically be implemented – most considered it almost impossible to implement. From the two options presented (see section 7.2.1.3):

- A tax on products and processes associated with the 'creation' of DSI had very little support from interviewees.
- A fee on a broader range of 'biotrade' products and/or processes such as voluntary schemes (like Rainforest Alliance or Fairtrade labels) had more support from interviewees.
  - However, certain industries (e.g. pharma) may not be allowed to use this 'biotrade' type of scheme.
  - Further, some mentioned not all companies would follow an 'honour'/voluntary system if the scheme was voluntary, as opposed to compulsory. This could lead to not collecting sufficient funds to be considered a fair redistribution of benefits.

Technical and legal viability

Neither of the options seem to pose major operational challenges for the organisations. However, interviewees raised the challenges of implementing either of the options, particularly a tax (1.c.a). When discussing this option all interviewees were sceptical about it. Some of the questions raised were: who would tax (countries or CBD)? What will the tax be based on? How would it be enforced?

Impact

In terms of impact, the option could have the same negative impacts as Option 1a (access fees), as costs may end up being distributed arbitrarily across DSI users. Another concern was that the tax could indirectly impact other sectors (e.g. sequencing machines are not only used to sequence genetic resources) or have a disproportionate cost for some organisations.

Option 2: Non-monetary benefit sharing

All interviewees agreed that this option is very important. However, the option was not considered viable on its own because benefit-sharing activities would need to be funded. Hence it would need to be merged with another 'monetary' option.

Stakeholders agreed that many organisations, though not all, already engage in non-monetary benefit sharing in an ad hoc manner. There is support for better coordination of existing efforts and a potential to scale them up.

Stakeholders support the idea of having a coordinating institution that acts as a 'matchmaker' between knowledge needs and capacity to offer. This institution would have to guarantee that the benefit-sharing support being given targets those working in conservation and biodiversity. Some questions were raised:

- How would the option be operationalised?
- Will organisations be expected to fund training activities, or would DSI users be required to engage in non-monetary benefit sharing?
- How would it be monitored and enforced? How would it be guaranteed that all organisations using DSI are engaging in non-monetary benefit sharing? How would these actions be quantified so as to determine where an organisation's benefit sharing obligations had been met?

7.4 Support for the options

This section summarises the points raised by interviewees in relation to the level of support expressed for each of the options. Stakeholders found it difficult to assess the merits of the options given prevailing uncertainties over the definition of DSI and the lack of specificity available for the benefit-sharing options. Indeed, this echoes the findings of the recent AHTEG meeting in which participants noted that, 'discussion on potential implications for the different groups concerning measures governing access, benefit-sharing and compliance was of a preliminary nature, and it was noted that this issue would benefit from further discussion' (AHTEG, 2020).

None of the options address all the needs and concerns of interviewees in relation to accessing and using DSI. Some interviewees suggested combinations of these options. Nevertheless, all the options were considered to have potentially major legal implications. A summary of the interviewees' level of support for the options is included in Table 7.5 below.

Table 7.5 Summary of the options' support

	Support	Comments
Option 2	Highest supported option by most interviewees.	Not considered viable on its own and needs to be combined with a monetary option.
Option 1b	Perceived as the fairest option by many interviewees.	Operationalising this would be almost impossible for all stakeholders (data repositories, users, regulators), and enforcement will end up being based on users' honesty. This option could be supported and viable, if there were to be agreement as to when commercialisation starts.
Option 1c	A few interviewees supported sub-option 1c: biotrade' in a voluntary manner Sub-option 1c: a tax or fee on products and processes associated with the 'creation' of DSI had no support.	The option was considered to be difficult to implement and arbitrary in the way payment was structured.
Option 1a	Very little support for this option. The option was perceived slightly more favourably were it to require a one-time access fee only.	Payment levels seemed arbitrary and could lead to a large increase in costs for some organisations and hinder research and innovation.

Interviewees indicated that for any of these options to work, there would need to be agreement across all CBD members – otherwise there would not be legal certainty and it would defeat the purpose of the agreement. There was also a strong sense amongst interviewees that any option that poses risks to ongoing open access policies or to conservation research or practices, should not be pursued (e.g. option 1.a.).

All interviewees strongly rejected the concept of removing open access (with an emphasis on 'no-monetary cost' access) as it will hinder research, innovation and collaboration. It was also broadly agreed that any option that requires traceability for benefit-sharing or payments is technically not viable.

The findings from the stakeholder interviews echo the recent discussion of the AHTEG which 'highlighted the importance of having legal certainty regarding usage of digital sequence information for all sectors (…) therefore any approach to benefit-sharing should provide legal certainty, incentivize the use of digital sequence information and decrease unnecessary burden in monitoring, tracing, and tracking requirements' (AHTEG, 2020). They also seem to support another finding of the recent AHTEG meeting (AHTEG, 2020) to the effect that sector-based approaches may be necessary.

8 Conclusions

This study provides a review of the state of play on DSI and implications for benefit-sharing in line with the CBD and the NP, including the financial, social and economic impacts of DSI. It includes an assessment of mechanisms which could be used to achieve fair and equitable sharing of the benefits arising from its access and use.

The report addressed the following questions, based on a rapid review of key literature and consultation with UK stakeholders:

- What is DSI?
- How is DSI generated?
- How is DSI accessed?
- What are the characteristics of the DSI stakeholder landscape?
- What are the options for DSI benefit sharing?

DSI is a controversial topic. The definitions debated in international policy discussions will have implications for how benefit-sharing options will be introduced and governed. In particular, the scope of the definition adopted will have major implications for any potential changes in the DSI policy landscape. Debates regarding terminology imply that if DSI were accepted in the future as part of the NP (i.e. PIC and MAT are required), it would complicate the legal use of DSI. Further, international policy discussions are debating alternative policy options if DSI remains out of scope of the NP, to ensure benefit sharing. Regardless of the outcome, clarity on the definition will allow stakeholders to have legal certainty if and when a benefit-sharing agreement is reached.

The study team selected a short list of policy options for the assessment based on recent discussions, and the way in which DSI is generated, accessed, shared, and used. DSI is an easily shareable resource and requires multiple users to process high volumes of digital data in order for the beneficial societal use-value of DSI to be maximised. DSI traceability is perceived as extremely complicated, since DSI users would be required to identify throughout the supply chain those stakeholders from whom they received the data and those to whom they supplied it. This traceability challenge implies that benefit-sharing options that require the ability to trace the use of genetic resources back to its origin and identify the parties engaged in creating the DSI output and where value was added in the process would be very challenging to implement.

A precise assessment of the impacts of the proposed options was not possible in this study, in part because all the options presented require significant further specification and in part due to the lack of agreement on a definition for DSI as well as information available regarding the value of DSI use for individuals, organisations and society. Nonetheless, some general conclusions can be drawn.

Assessment of the potential options demonstrates that a multilateral approach to benefit-sharing which prioritises open access to DSI is clearly preferable for the UK DSI users consulted in this study. Consultees were generally not in favour of adapting the bilateral ABS regime currently applicable under the NP/CBD to include DSI.

The option of payment upon commercialisation (1b), while perceived as fair, was seen as very challenging to implement. The options of payment on access (1a) and a tax, fee or levy (1c) were the least supported options as interviewees felt that payment levels seemed arbitrary and could lead to a large increase in costs for some organisations and hinder research and innovation. These options were also considered to be very difficult to implement.

While none of the options assessed in this study address all the needs and concerns of stakeholders consulted in relation to accessing and using DSI, they suggested combinations of these options could be more favourable. The majority of those consulted agreed with the fact that non-monetary benefits (Option 2) should be part of any agreement, and this option could not be considered a viable option on its own. An absence of monetary benefit-sharing from any potential solution would also be unlikely to satisfy the demands of certain biodiversity-rich countries that benefits arising from DSI are shared fairly and equitably.

Ultimately, policy changes for DSI benefit-sharing should be multilateral and agreed by all countries, otherwise they would not be effective. Further, any funds raised through monetary benefit sharing should clearly be destined to support conservation, in line with CBD objectives.

References

ABS Capacity Development Initiative, the South African National Department of Environment, Forestry and Fisheries and Norwegian Government (2019). First Global Dialogue on Digital Sequence Information on Genetic Resources. Pretoria, South Africa.

Ad Hoc Technical Experts Group on Digital Sequence Information on Genetic Resources (AHTEG) (2018). Report of the Ad Hoc Technical Expert Group on Digital Sequence Information on Genetic Resources. Convention on Biological Diversity (CBD). CBD/DSI/AHTEG/2018/1/4

Ad Hoc Technical Experts Group on Digital Sequence Information on Genetic Resources (AHTEG) (2020). Report of the Ad Hoc Technical Expert Group on Digital Sequence Information on Genetic Resources. Convention on Biological Diversity (CBD). CBD/DSI/AHTEG/2020/1/7.

Hill, G., Webb, D., Sheppard, S. (2016). Royal Botanic Garden Edinburgh: Economic Impact Assessment. United Kingdom. Available at: https://www.rbge.org.uk/media/5468/economic-impact-assessment-oct-2016.pdf.

Aubrey, S. (2019) The Future of Digital Sequence Information for Plant Genetic Resources for Food and Agriculture. Frontiers in plant science 10 (2019): 1046.

Bagley, M., Karger, E., Ruiz Muller, M., Perron-Welch, F. and Thambisetty, S. (2020). Fact-finding Study on How Domestic Measures Address Benefit-sharing Arising from Commercial and Non-commercial Use of Digital Sequence Information on Genetic Resources and Address the Use of Digital Sequence Information on Genetic Resources for Research and Development. [online] Montreal, Canada: Convention on Biological Diversity (CBD) AHTEG, pp.1–66. Available at: https://www.cbd.int/doc/c/428d/017b/1b0c60b47af50c81a1a34d52/dsi-ahteg-2020-01-05-en.pdf.

Barnes, A.P., Ferreira, J., Revoredo-Giha, C.R.G, Hoad, S., Hoebe, P., Burnett. F. (2016). The UK Plant Breeding Sector and Innovation (CT-RES-042). Report for the Intellectual Property Office. HMSO, London.

Bauen, A., Chambers, G., Houghton, M., Mirmolavi, B., Nair, S., Nattrass, L., Phelan, J. and Pragnell, M. (2016). Evidencing the Bioeconomy - An assessment of evidence on the contribution off, and growth opportunities in, the bioeconomy in the UK. Capital economics.

Brown, M. ‘New Brazilian Law on Genetic Heritage gives one year to companies to report on their past activities having used Brazilian genetic heritage’ (Lexology, 2017) Available at: https://www.lexology.com/library/detail.aspx?g=3f8fb766-b4f0-437d-80ee-ae2ee742f360.

Brunak, S., Danchin, A., Hattori, M., Nakamura, H., Shinozaki, K., Matise, T., Preuss, D. (2002) Nucleotide Sequence Database Policies Available at: http://www.insdc.org/policy.html.

Chambers, G., Evans, A., Phelan, J. and Pragnell, M. (2015). Biotech Britain. [online] BBSRC. Available at: https://bbsrc.ukri.org/documents/capital-economics-biotech-britain-july-2015/.

Chapman, T., Trivedi, C. and Miles, S. (2018). “Capturing, protecting and restoring plant diversity in the UK: RBG Kew and the Millennium Seed Bank”. Plant Diversity, 40.

Cochrane, G., Karsch-Mizrachi, I., Takagi, T., and International Nucleotide Sequence Database Collaboration (2016). “The International Nucleotide Sequence Database Collaboration”. Nucleic Acid. 40 Res, 2016. 44(D1): p. D48-50, DOI: 10.1093/nar/gkv1323

Convention on Biological Diversity (CBD) (1992) Convention on Biological Diversity. The Text. Available at: https://www.cbd.int/doc/legal/cbd-en.pdf

Convention on Biological Diversity (CBD) (2016). Decision adopted by the conference of the parties to the Convention on Biological Diversity. Available at: https://www.cbd.int/doc/decisions/cop-13/cop-13-dec-17-en.pdf

Convention on Biological Diversity (CBD) (2018). Synthesis of views and information on the potential implications of the use of Digital Sequence Information on genetic resources for the three objectives of the Convention and the objective of the Nagoya Protocol. *Note by the Executive Secretary. Addendum. Case studies and examples of the use of Digital Sequence Information in relation to the objectives of the Convention and the Nagoya Protocol.* Available at: https://www.cbd.int/doc/c/916b/1ee4/9dcf0ef617da984b6d107500/dsi-ahteg-2018-01-02-add1-en.pdf

Department for Business Innovation and Skills (BIS) (2013). A UK Strategy for Agricultural Technologies, July 2013. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/227259/9643-BIS-UK_Agri_Tech_Strategy_Accessible.pdf.

Department for International Trade (DIT) (2020a). Agricultural technology in the UK - great.gov.uk international. [online] www.great.gov.uk. Available at: https://www.great.gov.uk/international/content/about-uk/industries/agricultural-technology/.

Department for International Trade (DIT) (2020b). Choose UK Life Sciences.

Department for International Trade (DIT) (2020c). Healthcare and Life Sciences in the UK - great.gov.uk international. [online] www.great.gov.uk. Available at: https://www.great.gov.uk/international/content/about-uk/industries/health-and-life-sciences/.

Department for International Trade (DIT) (2020d). UK agritech strengths. [online] GOV.UK. Available at: https://www.great.gov.uk/international/content/about-uk/industries/agricultural-technology/uk-agritech-strengths/.

Food and Agriculture Organization of the United Nations (FAO) (2019). “Digital Sequence Information” on Genetic Resources for Food and Agriculture and its Relevance for Food Security. Commission on Genetic Resources for Food and Agriculture. Available at: http://www.fao.org/3/my588en/my588en.pdf.

Gostin, L. (2014) Global Health Law. Harvard University Press.

Heinemann, J.A., Coray, D.S., Thaler, D.S. and Ad Hoc Intergovernmental Technical Working Group on Aquatic and Genetic Resources for Food and Agriculture (2018). Draft Exploratory Fact-Finding Scoping Study on “Digital Sequence Information” on Genetic Resources for Food and Agriculture. [online] Rome: Commission on Genetic Resources for Food and Agriculture (FAO). Available at: http://www.fao.org/fi/static-media/MeetingDocuments/AqGenRes/ITWG/2018/Inf10e.pdf.

Hiemstra, S.J., Brink, M. and van Hintum, T. (2019). Digital Sequence Information (DSI) Options and impact of regulating access and benefit sharing - stakeholder perspectives. [online] Wagening: Centre for Genetic Resources, the Netherlands (CGN), Wageningen University & Research. Available at: https://library.wur.nl/WebQuery/wurpubs/fulltext/470286.

House of Lords (2008). Systematics and Taxonomy: Follow-up. [online] House of Lords, UK Parliament. Available at: https://publications.parliament.uk/pa/ld200708/ldselect/ldsctech/162/162.pdf.

Houssen, W., Sara, R. and Jaspars, M. (2020). Digital sequence information on genetic resources: concept, scope and current use. [online] Montreal, Canada: Convention on Biological Diversity (CBD) AHTEG, pp.1–49. Available at: https://www.cbd.int/doc/c/fef9/2f90/70f037ccc5da885dfb293e88/dsi-ahteg-2020-01-03-en.pdf.

Industrial Biotechnology Leadership Forum (IBLF) (2018). Growing the UK Industrial Biotechnology Base. [online] Available at: https://www.bioindustry.org/uploads/assets/uploaded/d390c237-04b3-4f2d-be5e776124b3640e.pdf.

Intellectual Property Office (IPO) (2012). Agrifoods - a brief overview of the UK agrifood patent landscape. [online] Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/312605/informatic-agrifood.pdf.

International Nucleotide Sequence Database Collaboration (INSDC) (2019). Response from International Nucleotide Sequence Database Collaboration (INSDC) to CDB call for views and information on Digital Sequence Information on Genetic Resources. Available at: https://www.cbd.int/abs/DSI-views/2019/INSDC-DSI.pdf.

Kamau, E.C. and Winter, G., 2013. Knowledge commons, intellectual property and the ABS regime GER DWINT ER. In Common Pools of Genetic Resources (pp. 317-334). Routledge.

Karger, E. (2018). Study on the use of Digital Sequence Information on genetic resources in Germany. [online] Regensburg: Institut fur Biodiversitat (IfB). Available at: http://www.biodiv.de/fileadmin/user_upload/PDF/Projekte-aktuell/DSI-Study.pdf.

Klock, K. A. 'International Public-Private Partnerships as Part of the Solution to Infectious Disease Threats: Operational, Legal, and Governance Considerations' in Sam F Halabi (eds), Global Management of Infectious Disease After Ebola (OUP, 2016) Chapter 9.

Keasling, J. (2013) Why Synthetic Biology Is the Field of the Future. Nova Newsletter. Available at: https://www.pbs.org/wgbh/nova/article/why-synthetic-biology-is-the-field-of-the-future/.

Kries, C.V., & Winter, G. (2015). Defining commercial and non-commercial research and development under the Nagoya Protocol and in other contexts.

Laird, S. and Wynberg, R. (2018). Fact-finding and scoping study on digital sequence information on genetic resources in the context of the convention on biological diversity and the Nagoya Protocol. Note by the Executive Secretary. [online] Convention on Biological Diversity (CBD) - AHTEG. Available at: https://www.cbd.int/doc/c/b39f/4faf/7668900e8539215e7c7710fe/dsi-ahteg-2018-01-03-en.pdf.

Laird, S. Wynberg, R. Rourke, M., Humphries, F., Ruiz Muller M., and C. Lawson (2020), 'Rethink the expansion of access and benefit sharing' Science 367 (6483), 1200-1202; DOI: 10.1126/science.aba9609.

Lawson, Charles and Rourke, Michelle, Open Access DNA, RNA and Amino Acid Sequences: The Consequences and Solutions for the International Regulation of Access and Benefit Sharing (October 5, 2016). Griffith University Law School Research Paper No. 16-12. Available at SSRN: https://ssrn.com/abstract=2848136 or http://dx.doi.org/10.2139/ssrn.2848136.

Manzella, D. (2016). The global information system and genomic information: transparency of rights and obligations. Available at: www.fao.org/3/a-bq620e.pdf.

Morgera, E., (2016) 'Study on Experiences Gained with the Development and Implementation of the Nagoya Protocol and Other Multilateral Mechanisms and the Potential Relevance of Ongoing Work Undertaken by Other Processes, Including Case Studies' (1-3 February 2016) UN Doc UNEP/CBD/ABS/A10/EM/2016/1/2.

Morgera, E. (2018) 'Fair and Equitable Benefit-Sharing in a New International Instrument on Marine Biodiversity: A Principled Approach Towards Partnership Building?' (2018) 5 Maritime Safety and Security Law Journal 48

Muller, M.R. (2018). Innovation and IP Access to Genetic Resources and Benefit Sharing 25 Years on: Progress and Challenges. [online] International Centre for Trade and Sustainable Development (ICTSD). Available at: https://www.voices4biojustice.org/wp-content/uploads/2018/12/Access-to-Genetic-Resources-and-Benefit-Sharing-25-Years-On-Progress-and-Challenges.pdf.

The Natural History Museum (NHM), Royal Botanic Gardens Kew (RBG Kew), and Royal Botanic Garden Edinburgh (RGBE) (2019). 'Submission of views and information on benefit sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources' Available at: https://www.cbd.int/abs/DSI-views/2019/NHM-RBGK-RBGE-DSI.pdf.

The Natural History Museum (NHM) (2018) Natural History Museum Annual Report and Accounts 2017-2018. Presented to Parliament pursuant to the Museum and Galleries Act 1992. available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/729300/NHM_annual_report_and_accounts_2017_to_2018.pdf.

Office for Life Sciences (OLS) (2019) Life Science Sector Data. Available at: https://www.gov.uk/government/publications/life-science-sector-data-2019.

Oxford Economics (2019). The Economic Value of Royal Botanic Gardens, Kew - A Total Economic Value Approach. [online] Available at: https://www.kew.org/sites/default/files/2020-01/Read%20the%20report_0.pdf.

Perron-Welch, F., 'Blockchain Technology and Access and Benefit-Sharing' (ABS Canada, 2018) http://www.abs-canada.org/food-for-thought/blockchain-technology-and-access-and-benefit-sharing/.

Prathapan, K.D., Pethiyagoda, R., Bawa, K.S., Raven, P.H. and Rajan, P.D, (2018). 'When the Cure Kills: CBD Limits Biodiversity Research' Science 360 (6396), 1405-1406.

Price Waterhouse Coopers (PWC) (2017) The economic contribution of the UK Life Sciences industry [online] available at: https://www.abpi.org.uk/media/1371/the_economic_contribution_of_the_uk_life_sciences_industry.pdf.

Rohden, F., Huang, S., Hartman Scholz, A. and Overman, J. (2020). Combined Study on Digital Sequence Information in Public and Private Databases and Traceability. [online] Convention on Biological Diversity (CBD) - AHTEG. Available at: https://www.cbd.int/doc/c/1f8f/d793/57cb114ca40cb6468f479584/dsi-ahteg-2020-01-04-en.pdf.

Rourke, M.F., Phelan, A. & Lawson, C. (2020). Access and benefit-sharing following the synthesis of horsepox virus. Nat Biotechnol https://doi.org/10.1038/s41587-020-0518-z.

Ruiz Muller, M. (2018). Innovation and IP Access to Genetic Resources and Benefit Sharing 25 Years on: Progress and Challenges. [online] International Centre for Trade and Sustainable

Development (ICTSD). Available at: https://www.voices4biojustice.org/wp-content/uploads/2018/12/Access-to-Genetic-Resources-and-Benefit-Sharing-25-Years-On-Progress-and-Challenges.pdf.

Ruiz Muller, M. (2013) 'Protecting Shared Traditional Knowledge: Issues, Challenges and Options' ICTSD Issue Paper No. 39.

Ruiz Muller, M. (2010) 'Thinking Outside the Box Innovative Options for an Operational Regime on Access and Benefit Sharing' ICTSD Issue Paper No. 1.

Secretariat of the Convention on Biological Diversity (CBD), and World Health Organization (WHO) (2018) Implementation of the Nagoya Protocol in the context of human and animal health, and food safety: Access to pathogens and fair and equitable sharing of benefits. Geneva; 2018.

Shipman, Carissa (2012) “DNA and Its Importance in Taxonomy” California Academy of Sciences Blog. Available at: https://www.calacademy.org/blogs/project-lab/dna-and-its-importance-in-taxonomy.

Sirakaya, A. (2019). Balanced Options for Access and Benefit-Sharing: Stakeholder Insights on Provider Country Legislation. Frontiers in Plant Science, [online] 10. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6781883/.

Smith, P. and Harvey-Brown, Y. (2018). The economic, social and environmental impacts of botanic gardens (BGCI Technical Review). [online] Botanic Gardens Conservation International. Available at: https://www.bgci.org/files/IAC/IAC%202018/TechReportLowRes.pdf.

Smyth, S., Macall, D., Phillips, P., de Beer, J. (2019) Governance of Digital Sequence Information and Impacts for Access and Benefit Sharing. Centre for the Study of Science and Innovation Policy Johnson Shoyama Graduate School of Public Policy. University of Saskatchewan.

SQW (2016) “Agri-Tech Industrial Strategy: Evaluation Scoping Study and Baseline” BIS Research Paper Number 284.

Subsidiary Body on Scientific, Technical and Technological Advice (SBSTTA) (2016). Synthetic Biology - Note by Executive Secretary. CBD [online] available at: https://www.cbd.int/doc/meetings/sbstta/sbstta-20/official/sbstta-20-08-en.pdf.

The Economist (2010) ‘Chemistry Goes Green’, The Economis.com [online] available at: https://www.economist.com/business/2010/07/01/chemistry-goes-green.

Tsioumani, E., 2019. ‘Consensus Within Reach for Plant Treaty Working Group to Revise ABS System’ (ISSC) https://sdg.iisd.org/news/consensus-within-reach-for-plant-treaty-working-group-to-revise-abs-system/.

Vogel, J.H., Angerer, K., Muller, M.R. and Oduardo-Sierra, O., 2017. Bounded openness as the modality for the global multilateral benefit-sharing mechanism of the Nagoya Protocol. In Routledge Handbook of Biodiversity and the Law (pp. 377-394). Routledge.

Vogel, J.H., Angerer, K., Muller, M.R. and Oduardo-Sierra, O., (2019). '"Even best case for bilateralism supports need for a Global Multilateral Benefit-Sharing Mechanism: Common ground in ‘bounded openness over natural information’ as the modality for ABS” In response to NOTIFICATION for Submission of views and information further to decisions NP-3/13 on Article 10 of the Nagoya Protocol (SCBD/NPU/DC/VN/KG/RKi/8780.

Welch, E., Bagley, M. A., Kuiken, T., Louafi, S. (2017). Potential implications of new synthetic biology and genomic research trajectories on the International Treaty for Plant Genetic Resources for Food and Agriculture. Emory Legal Studies Research Paper. Available at: https://ssrn.com/abstract=3173781.

World Health Organisation (WHO) (2018). Genetic sequence data and databases. Factsheet. Version 2. Available at: https://www.who.int/influenza/pip/GSD_EN_V2_10Sep2018.pdf?ua=1

Annex 1 Economic significance of affected sectors

Estimation of economic subsectors that are likely to rely on DSI.

The figures in the **Error! Reference source not found.** below are based on data extracted f rom the 2018 Annual Business Survey. The election of subsectors is based on the Standard Industrial Classification (SIC) codes. This represents an approximation of the extent of the economy which relies on DSI - this approximation is likely to be underestimating the size of many of those because the industrial distinction is often blurred and business that that also treat subjects relevant to DSI may be classified under more general categories, which are then lost in the estimation.

Table A1.1 Summary of economic data for key UK sectors using DSI

Sector	Subsector	SIC code	GVA (2018) (£ million)	Turnover (2018) (£ million)	Total employees (2018)
Natural Sciences Research	Other research and experimental development on natural sciences and engineering	*72190*	3,551	20,108	120,000
Botanic	Botanicals and zoological gardens and nature reserve activities	*91040*	163	641	24,000
Biotechnology	Research and experimental development on biotechnology	*72110*	(48)	1,465	10,000
Pharmaceuticals	Manufacture of basic pharmaceutical products	*21100*	775	1,745	6,000
	Manufacture of pharmaceutical preparations	*21200*	8,074	18,638	29,000
Agriculture	Seed processing for propagation	*01640*	18	89	300
	Support activities for crop production	*01610*	555	822	9,000
	Manufacture of pesticides and other agrochemical products	*20200*	364	930	
	Manufacture of fertilisers and nitrogen compounds	*20150*	284	1,541	2,000
Horticulture	Plant propagation	*01300*	n/a	n/a	n/a
Cosmetics	Manufacture if perfumes and toilet preparations	*20420*	990	4,205	16,000
	Manufacture of soap and detergents	*20410*	919	2,768	12,000
Flavour industry	Manufacture of essential oils	*20530*	291	940	2,500
	Manufacture of basic chemicals, fertilisers and nitrogen compounds, plastics and synthetic rubber in primary forms	*20100*	4,703	17,126	n/a
	Manufacture of dyes and pigments	*20120*	396	1,278	4,000
	Manufacture of other inorganic based chemicals	*20130*	923	2,263	4,500
	Organic based chemicals	*20140*	948	3,973	9,000

Source: UK Annual Business Inquiry 2018 (ONS, 2019)

Annex 2 Excluded options and options assessment criteria

Options excluded were:

- Status quo / Nagoya-bilateral
- Open access bi-lateral
- Multilateral monetary benefit-sharing – donations

Criteria / questions excluded:

- Effectiveness
  - Questions on effects on ‘access’ were removed, anticipating that consideration of access will be indirectly considered when stakeholders respond to impact questions on costs/benefits/competitiveness.
  - Questions on the fairness and proportionality of resultant benefit sharing were removed – UK stakeholder opinion on this considered of low priority.
- Viability and impact criteria
  - The number of questions was reduced and those retained were made more general.