Scientific Industry Input on Potential Open-Access Policy Options for DSI

Monday, July 13, 2020 | 12:30 -15:30 Central European Daylight Time

Online workshop

Hosted by the Leibniz Institute DSMZ and the Leibniz Institute IPK Gatersleben.

Workshop Report

Table of Contents

Background and Brief Overview

The 15th Conference of the Parties (COP) to the Convention on Biological Diversity (CBD) will meet next year in Kunming, People's Republic of China to negotiate and define the Post-2020 Global Biodiversity Framework.[1] An important topic in the upcoming negotiations will be the issue of access and benefit-sharing (ABS) from "digital sequence information on genetic resources" (DSI) under the Nagoya Protocol (NP) and CBD. At stake in these negotiations for biologists worldwide is open access to sequence data with respect to the large public databases. Science-based solutions for benefit sharing that do not endanger open access are urgently needed.

To this end, the German Federal Ministry of Education and Research (BMBF) has funded an interdisciplinary research project (WiLDSI) led by the Leibniz Institute DSMZ and the Leibniz Institute IPK Gatersleben to research DSI policy options. The project, "**Wissenschaftsbasierte Lösungsansätze für Digitale Sequenzinformation (WiLDSI)"**, aims to research viable open-access DSI policy options and to actively involve the scientific (academic and industry) stakeholder community in Germany and Europe.[2]

To proactively involve the private sector scientific stakeholders in Europe, an online workshop on "Scientific Industry Input on Potential Open-Access Policy Options for DSI" was held on 13th July, 2020 (12:30-15:30 CEST). This online workshop was third in a series of three stakeholder workshops[3] organised by the WiLDSI project and the goal of this workshop was to receive feedback on the priorities and options that were developed and researched by the WiLDSI project members over the period between September 2019 and July 2020. These open access policy options for DSI will be presented in a final WiLDSI White paper that will be launched in the coming months.

The online workshop was attended by 33 participants (with 17 private sector representatives across Europe). Industry representatives from Plantum, Roche, Novozymes, BASF, IFPMA (International Federation of Pharmaceutical Manufacturers & Associations), Bayer Crop Science, DSM Nutritional Products Ltd, KWS SAAT, EuroSeeds, Astra Zeneca and BASF Vegetable Seeds attended the online workshop.

[1] This meeting has now been postponed from October 2020 to Autumn 2021 due to the COVID-19 pandemic.
[2] For more information on the WiLDSI project: https://www.dsmz.de/collection/nagoya-protocol/digital-sequence-information
[3] For further information about the two prior WiLDSI workshops please see: https://www.dsmz.de/collection/nagoya-protocol/digital-sequence-information/dsi-workshop-january-2020 and https://www.dsmz.de/collection/nagoya-protocol/digital-sequence-information/dsi-workshop-march-2020.

Presentations

The online workshop was opened by the chair, Prof. Esther van Zimmeren (University of Antwerp), who welcomed the participants and gave a brief introduction explaining the aim and objectives of the workshop. She explained the workshop would be conducted under Chatham House Rules and any discussion during the course of the workshop would not be attributed to individuals. This was followed by two presentations from Alicja Kozlowska and Dr. Amber.H.Scholz respectively.

The political landscape around DSI (Alicja Kozlowska)

Alicja Kozlowska (Policy Officer ABS, European Commission) gave an overview on the CBD processes surrounding DSI. Her presentation covered briefly the following topics: a) the legal framework of CBD and the Nagoya Protocol (NP) b) the reasons why DSI is an issue, c) discussions at the two previous Conferences of the Parties (COP13 and COP14) concerning DSI so far, and d) other international fora that are also discussing DSI. She explained that the CBD in 1993 put in place new principles whereby genetic resources were no longer considered 'heritage of mankind' and that States have sovereign rights over their genetic resources. The CBD introduced the concept of access and benefit- sharing (ABS) and the discussions on DSI are taking place in this context. She went on to explain how the Nagoya Protocol (NP) was agreed in 2010 in order to 'balance' the Aichi targets and to further advance the implementation of the 3rd objective of the CBD, namely ABS. Important elements were introduced in the legal landscape such as Prior Informed Consent (PIC) and Mutually Agreed Terms (MAT) on Genetic Resources (GR) and associated Traditional Knowledge (TK). The basic bargain envisioned was that developed countries get transparent access legislation and developing countries receive fair and equitable sharing of benefits. Based on her personal reflections and experience, many amongst the developing countries seem to be frustrated with the narrow scope of the Protocol and dissatisfied with the perceived insufficient (monetary) benefits generated by the Protocol. Many in the developed world are dissatisfied with intransparent processes and significant delays in research (much of which is in the interest of the public good). In recent years, DSI has begun to be perceived by some as a loophole/leak in the system – a way to circumvent the NP obligations and reason why few monetary benefits were being realised. She described how the DSI issue was addressed at the last two COPs (COP13 and COP14) and explained the inter-sessional processes that were established and the tasks of the Ad Hoc Technical Expert Group (AHTEG) in March 2020 and the third meeting of the Open-ended Working Group on the Post-2020 Global Biodiversity Framework (OWEG 3) in 2021. DSI is an integral part of the discussion in the post-2020 considerations and knowledge from different studies and submissions is being gathered during this inter sessional period (2018-2020). Also, to be noted is that DSI is being discussed by other fora such as the ITPGRFA, FAO CGRFA, WHO, BBNJ-UNCLOS and WIPO.

The WiLDSI Project: Scientific priorities & 5 open-access DSI Policy Options (Dr. Amber H.Scholz)

Dr. Amber H.Scholz (Deputy to the Director, Leibniz Institute DSMZ) provided a brief overview of the WiLDSI project timeline and stated the two guiding principles of the WiLDSI project namely; 1) quantify and describe what open access DSI enables and 2) assess whether monetary benefit-sharing is compatible with open access. She mentioned research performed by the project members as well as input from the two prior WiLDSI workshops (Bonn and Brussels) which culminated into the development of some open-access DSI options. The idea for this workshop was to explore the perspective of private sector representatives. Ideas and input from this workshop could be integrated into the WiLDSI white paper expected to be released in the fall of 2020. She described Option 0: the extensive and expensive non-

monetary benefit sharing from DSI. Based on the CBD study on traceability and databases,[4] most DSI is not sourced from Low and Middle Income countries (LMICs), in fact 52% of nucleotide sequence data (NSD) comes from 4 countries (China, United States, Canada and Japan). There are 10-15 million users worldwide and interestingly, 50% of users live in countries that do not contribute to the NSD infrastructure and operating costs (approx. 50 million US dollars annually). Moreover, there are countries that have more users when compared to the sequence data that they provide most of which are middle- or low-income countries. These data show that benefit-sharing around DSI is already happening.

The WiLDSI project has also been assessing the scientific priorities for ABS and DSI prior to the upcoming negotiations such as:

1. Open access
2. Minimize administrative overhead & „future-proof“
3. Legal certainty
4. Time until value delivery
5. Genetic resources

Dr. Scholz briefly described 5 policy options that were being examined that could potentially be compatible with open-access to DSI [5]:

Option 1: Blockchain for DSI Metadata

Option 2: Commons Licenses for DSI

Option 3: Cloud-based Fees

Option 4: Membership Fee

Option 5: Micro-levy

She concluded her talk by explaining the purpose for the workshop to receive inputs that can inform the drafting of the white paper prior.

Break-out session to discuss the 5 DSI policy options

Three parallel breakout sessions were set-up in the online platform and participants were asked to discuss the following questions:

- Do you agree with the scientific priorities outlined by the project? Are there any missing?
- What ideas are interesting to you in the options presented? Are there any elements that generate a strong reaction for you (negative or positive)?
- Are there ideas around DSI not represented here that should be?
- Are there policy opportunities through DSI to re-think/re-do/re-invent the ABS system?

After approx. 50 minutes of intense discussion, each group summarized the main points of their discussion.

[4] Rohden, F., Huang, S., Dröge, G., and Scholz, A., (2020) Combined Study on Digital Sequence Information (DSI) in Public and Private Databases and Traceability. Secretariat of the Convention on Biological Diversity. Montreal, Canada. https://www.cbd.int/meetings/DSI-AHTEG-2020-01.

[5] See Appendix for background paper briefly describing the policy options for DSI. NOTE: These options were ultimately re-ordered in the October 2020 white paper.

Synthesis and discussion

Outcome of discussion for Group 1

- Do you agree with the scientific priorities outlined by the project? Are there any missing?
    - Legal certainty
        - Scope needs to be further explored
        - Is retroactivity part of this?
        - From a scientist's perspective – They want to know if they have complied with what's required.
        - Excluding this could narrow the options, e.g. the Blockchain option.
        - Inclusion of costs and conditions.
        - Term needs more consideration.

- What ideas are interesting to you in the options presented? Are there any elements that generate a strong reaction for you (negative or positive)?
  Option 1: Blockchain
    - Open access is potentially compromised.
    - Administrative burden – This might pose a problem
    - Difference between sequence data and metadata in usage (normal scientific practice)?
    - Appears complex and could take years to establish.
    - Metadata and sequence data belong together.
    - Time until value materialisation? For options 1 &2, this could delay research and benefits would be slow to accrue.
    - Different forms of "value". Intrinsic value of openness and completeness of the system.

- **Option 2: Commons Licenses**
    - Less investment in infrastructure as compared to the Blockchain option.
    - Licenses may be complex to establish, but once established it may be easier.
    - A major issue for option 1 & 2 is that of traceability.
    - For both the above bilateral options there is the likelihood of jurisdiction shopping. Also, where there is legal uncertainty (as to national law), there will be avoidance.

- **Option 3: Cloud-based Fees**
    - How much to pay?
    - For larger companies, this option could promote working through academia and SMEs who have used the information. Depends on what the cost would be. Could be significant amounts where millions of sequences are used.
    - Options 3 & 4 do not differentiate between NP and non-NP sequences. This has significant consequences for pharma sector (e.g. human sequences).

- **Option 4: Membership fee**
    - Difference in fees to option 3 - important for different sectors. Same concerns as option 3.
    - Unlimited access is appealing.
    - International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) - need to learn lessons from this, these discussions failed. Concerns about giving up sovereign rights.
    - What is the impact of a MLS on the operation of the NP? What is the relationship with GR that has been sequenced?

- **Option 5: Micro-levy**
    - Revolutionary approach.
    - Could raise more money than other options.
    - Could be broadly applicable, broader discussion on financing biodiversity conservation.

Outcome of discussion for Group 2

- Do you agree with the scientific priorities? Are there any missing?
    - Issue of retroactivity and how to deal with it.
    - Time until value materializes shouldn't be a priority or an issue as it takes time to perform research and development. Not an issue in this debate.
    - GR in DSI ABS framework: one system for everything is needed. We should not have overlapping ABS systems.
    - Some missing elements: 1. public-private partnerships can be hampered by the DSI discussion – important to ensure that they can continue to work, 2. look into sector-specific needs e.g. plant breeding sector – common approach use materials without payment obligations, 3. Legal systems should not be retroactive, 4. minimize complexity if you want it to work, 5. consider the non-monetary aspect.
    - With respect to COVID-19 and the public health sector, any model that imposes financial burden needs to be considered. We need to take into account interests of scientists in developing countries. They do not have the budget to pay for access to DSI.
    - In general, everyone in this group was happy with the scientific priorities.
    - Retroactivity concerns could be connected to legal certainty and made explicit.

- What ideas are interesting to you in the options presented? Are there any elements that generate a strong reaction for you (negative or positive)?
    - All the options have their flaws; the main flaw being that they will probably not generate significant money – expect for possibly the last one (Option 5: Micro-levy). There is also a barrier to open-access in all 4 options except the micro-levy.
    - Micro levy could generate more money, in particular if it could be applied more broadly, not just to sequencing.
    - The options are still quite complex but, at the same time, much more detail would be needed to understand them fully.
    - Cost effectiveness of the Blockchain option and cloud-based fees option is very questionable.
        - Blockchain option: However, if you compare the costs of this system – database with administration system (PIP framework 6 million dollars per year but this is only a single organism. Can this be an example of what such a system would cost? What can be extrapolated for a system that would cover the diversity of CBD?)
        - Technical perspective – If such a system is developed, it would be useful to build it phase by phase and not try to have one major global system. Firstly, make clear choices what you want to achieve and to protect .To not try to do too much at once as there will be a risk of failure.
        - Not feasible from a developing countriy perspective.
    - In case of the Micro-levy option, we could learn lessons from UNITAID (Airline levy initiated by France and taken up later by other countries). Would be interesting to simulate countries through a multilateral system and through the CBD. However, in the end, it would be critical to have the commitment of several countries in order to work.

- Are there ideas around DSI not represented here that should be?
    - Capacity development initiatives.
    - Option 3-5 show a multilateral system and giving up the sovereign rights. This however doesn't look too hopeful and probable.
    - Micro-levies is the only option that keeps open access and fosters innovation and research. Other options seem to impose barriers for this.
    - Sector-specific treatment. Needs of different sectors should be taken into consideration.
    - Open access is very critical.
    - Some of the options appear that they could add delay that could be detrimental.
    - COVID-19 sharing of sequence data was very important. Speed is very important for the public health domain.
    - Also important for food and agriculture-easy access to pathogens for food security – especially in the tropics
    - Cost benefit ratio not discussed. Expectation management. Start with pilot projects and then proceed gradually.
    - Emphasis on non-monetary BS

Outcome of discussion for Group 3

- Do you agree with the scientific priorities? Are there any missing?
    - There needs to be a clear definition of what is open access in order to ensure a common understanding.
    - Include GR in a DSI framework and to link it to rethink ABS framework (for DSI and GR). This would avoid additional layer of complexity.
    - Optimise value creation.
    - Priorities should be for people engaging in science, no discrimination between commercial and academic science. A better term- 'research priorities'. To ensure inclusion of scientists from different countries.

- What ideas are interesting to you in the options presented? Are there any elements that generate a strong reaction for you (negative or positive)?

  Option 1: Blockchain

    - Blockchain option seems an inflexible system.
    - How would this option work? Seems very clunky.
    - Tracking and tracing may be an issue.
    - Very complex. This will make the current system even more complicated.

  Option 2: Commons Licenses for DSI

    - Also very complex.
    - In this option, the license needs to travel with data (seems undesirable for science).
    - This option could also create friction

For both the above bilateral options, there seems to be diminished freedom to use data and this may have a negative impact on science. Fluidity of data use and exchange enables science, and seems affected/compromised in the two bilateral options.

For the remaining multilateral options, these options will only work when parties agree to delegate sovereign rights. How realistic is this? This is not to be underestimated. How would these options impact scientists in provider countries? This is not addressed. Bilateral systems could have very negative impacts for them, possibly some multilateral as well.

Option 3: Cloud-based fee

- If one downloads more, you end up paying more. Is this workable or feasible?
- There could be the issue of confidentiality.

Option 5: Micro-levy

- Is there a possibility to combine different options? Combine Micro-levy option with other options?

Are there ideas around DSI not represented here that should be?

- Reinforce capacity building – Option 0
- To capture the value of non-monetary sharing
- Value creation is key
- The idea that open access creates value could be stronger. This can be linked with including more scientists from developing countries (capacity building) but also better documentation of existing collaborations and inclusive efforts that are not curenly calculated into the equation.
- Any option that disincentives research is not acceptable

Summary

There was a general preference for the micro-levy option (option 5) although attendees noted this was very early stages of discussions. The microlevy maximizes open access to DSI and at the same time seems best-positioned to generate sufficient monetary benefit sharing. There was also an emphasis on addressing in further detail the non-monetary benefit-sharing aspect as well as focus on capacity development initiatives.

Glossary of acronyms

ABS	Access and Benefit-Sharing
AHTEG	Ad Hoc Technical Expert Group
CBD	Convention on Biological Diversity
COP	Conference of Parties
DSI	Digital Sequence Information
GR	Genetic Resources
ITPGRFA	International Treaty on Plant Genetic Resources for Food and Agriculture
LMICs	Low and Middle Income countries
MAT	Mutually Agreed Terms
MLS	Multilateral System of Access and Benefit-sharing
NP	Nagoya Protocol
NSD	Nucleotide Sequence Data
OECD	Organisation for Economic Co-operation and Development
OEWG	Open-ended Working Group on the Post-2020 Global Biodiversity Framework
PIC	Prior Informed Consent
UNCLOS	United Nations Convention on the Law of the Sea
WHO	World Health Organization
WiLDSI	Wissensbasierte Lösungsansätze für Digitale Sequenzinformation (Scientific approaches for DSI)
WIPO	World Intellectual Property Organization

Appendix

Background document for the scientific industry stakeholder DSI workshop

Organized by the BMBF project "Science-based solutions for DSI" (WiLDSI)

July 13, 2020 | Online Workshop

Background & the WiLDSI Project

The 15th Conference of the Parties (COP) to the Convention on Biological Diversity (CBD) will meet next year in Kunming, China to negotiate and define the Post-2020 Global Biodiversity Framework. An important topic in these negotiations will be the issue of access and benefit sharing (ABS) from "digital sequence information on genetic resources" (DSI) under the Nagoya Protocol (NP) and CBD. The contents of public databases for DSI are growing exponentially and some countries of origin fear that uncontrolled access to freely available DSI undermines the rightful sharing of benefits from the utilization of genetic resources. At the same time, DSI and its free accessibility are essential for life science, including biodiversity research, food security, and human/plant/animal health and beyond.

The goal of the WiLDSI (Wissenschaftsbasierte Lösungsansätze für DSI) project is to provide scientific input prior to the negotiations (and expected political compromises) that could have long-reaching consequences for the global scientific research community. This workshop is the third and final in a series of scientific stakeholder workshops. The first two workshops took place in January 2020 (Bonn) and March 2020 (Brussels) and, together, complement research by the WiLDSI project steering committee members and the study for the CBD Secretariat on DSI databases and traceability.[6]

The DSI options paper

The workshops and research are the basis for the final project deliverable: a white paper on the scientific perspective around DSI policy, which is expected to be released in early September 2020. The paper presents the following arguments:

- The DSI issue is politically urgent but should be discussed in the context of shared responsibilities and connected to broader resource mobilization discussions. DSI and access and benefit sharing, more generally, ought to be more tightly connected to the biodiversity targets that they support and the incentives and mechanisms of any future policy framework should more closely align these goals.

[6] Rohden, F., Huang, S., Dröge, G., and Scholz, A., (2020) Combined Study on Digital Sequence Information (DSI) in Public and Private Databases and Traceability. Secretariat of the Convention on Biological Diversity. Montreal, Canada. https://www.cbd.int/meetings/DSI-AHTEG-2020-01.

- The DSI challenge arises from the fact that the open-access system that has developed over more than four decades is fundamentally incompatible with the bilateral ABS system developed by the CBD and NP. The tension arises between a desire for control over genetic resources (GR) (and the sequence data resulting from GR), and the scientific assessment that the value of sequence data can actually only be realized *IF* the system is open and as comprehensive as possible. The collision of these two perspectives has brought high levels of political tension in the past.
- The open-access system for DSI enables a tremendous amount of non-monetary benefit sharing and represents significant monetary investment. Key take-home messages on non-monetary benefit sharing include:
    - The core infrastructure for DSI is the International Nucleotide Sequence Data Collaboration (INSDC) which has annual running costs of ca. $50 million USD paid for almost exclusively by governments in the USA, EU, and Japan.
    - There are users of the INSDC in every country in the world. Half of users live in countries whose governments do not fund the infrastructure.
    - There are four countries of origin (geographical source) for over half (52%) of DSI in the INSDC: USA, China, Canada, and Japan.
    - The "provider-user" dichotomy often portrayed in political discussions is not borne out by the data.
- Based on the stakeholder workshops and extensive discussions, five scientific priorities on DSI are described: Open access; Minimize administrative overhead & "future-proof"; Legal certainty; Time until value materializes; An option for the inclusion of GR in a new DSI ABS framework
- Nevertheless, given the urgency of the biodiversity crisis and the political acrimony surrounding ABS, the WiLDSI project recognizes that now is the time for proactive engagement. Therefore, five open-access policy options (two bilateral, three multilateral) for DSI with monetary benefit sharing components were developed and are presented briefly with diagrams below. Options 3-5 offer an opportunity for a harmonized DSI system that could be used by other international fora.

This paper is complemented by a technical report and annex as well as two workshop reports that provide complementary analysis.[7]

Workshop Goals

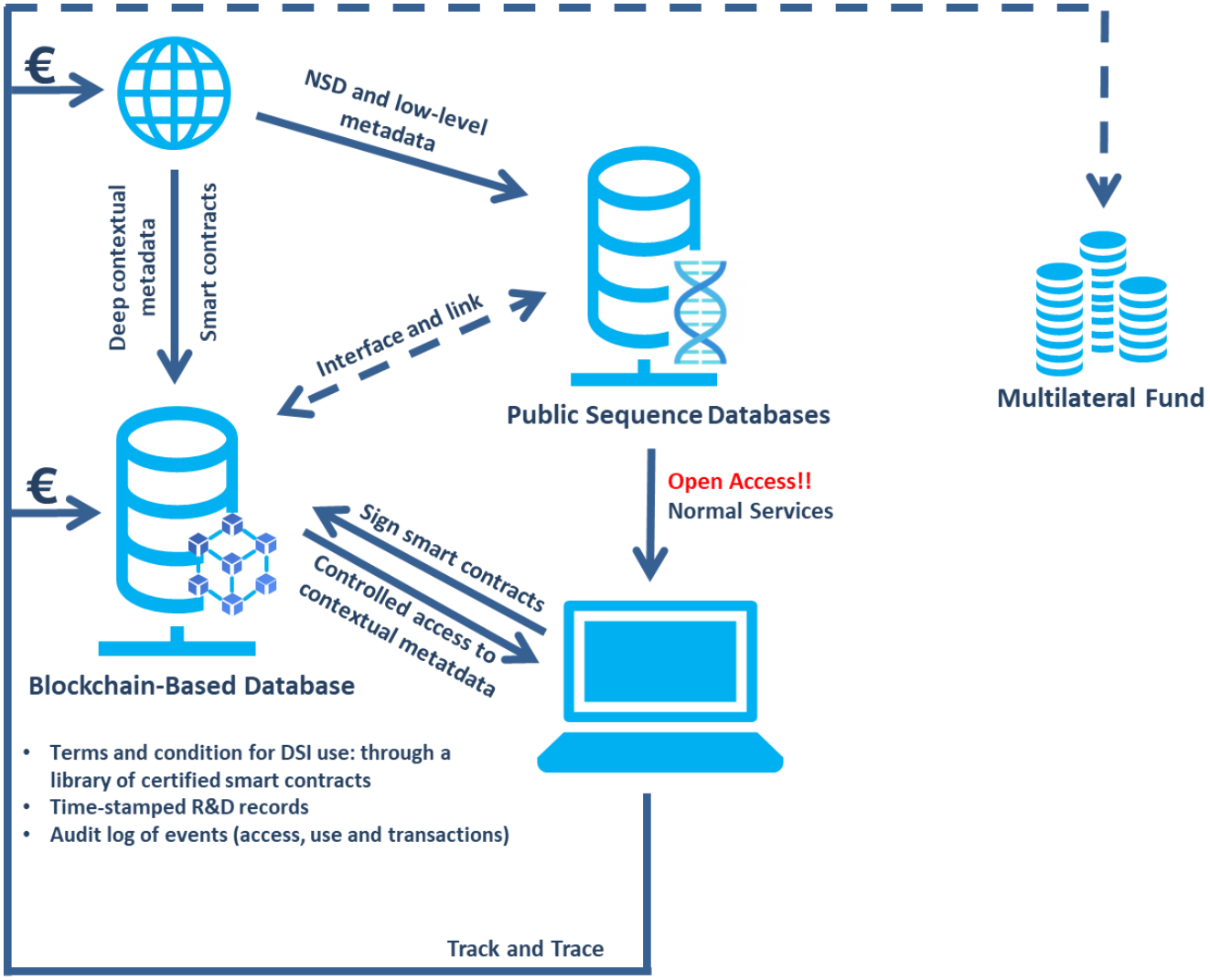
The goal of this workshop is to engage scientific colleagues from the private sector. Since the project is now at the stage of putting forth options with monetary components, for which the private sector would theoretically likely have a higher contribution, we would like to discuss our ideas, identify pros and cons, or general impracticalities. In short, this short online workshop is an opportunity for open discussion, analysis, fresh thinking, and early socialization of the ideas the project has developed to-date.

[7] https://www.dsmz.de/collection/nagoya-protocol/digital-sequence-information

Option 1: Blockchain on DSI metadata

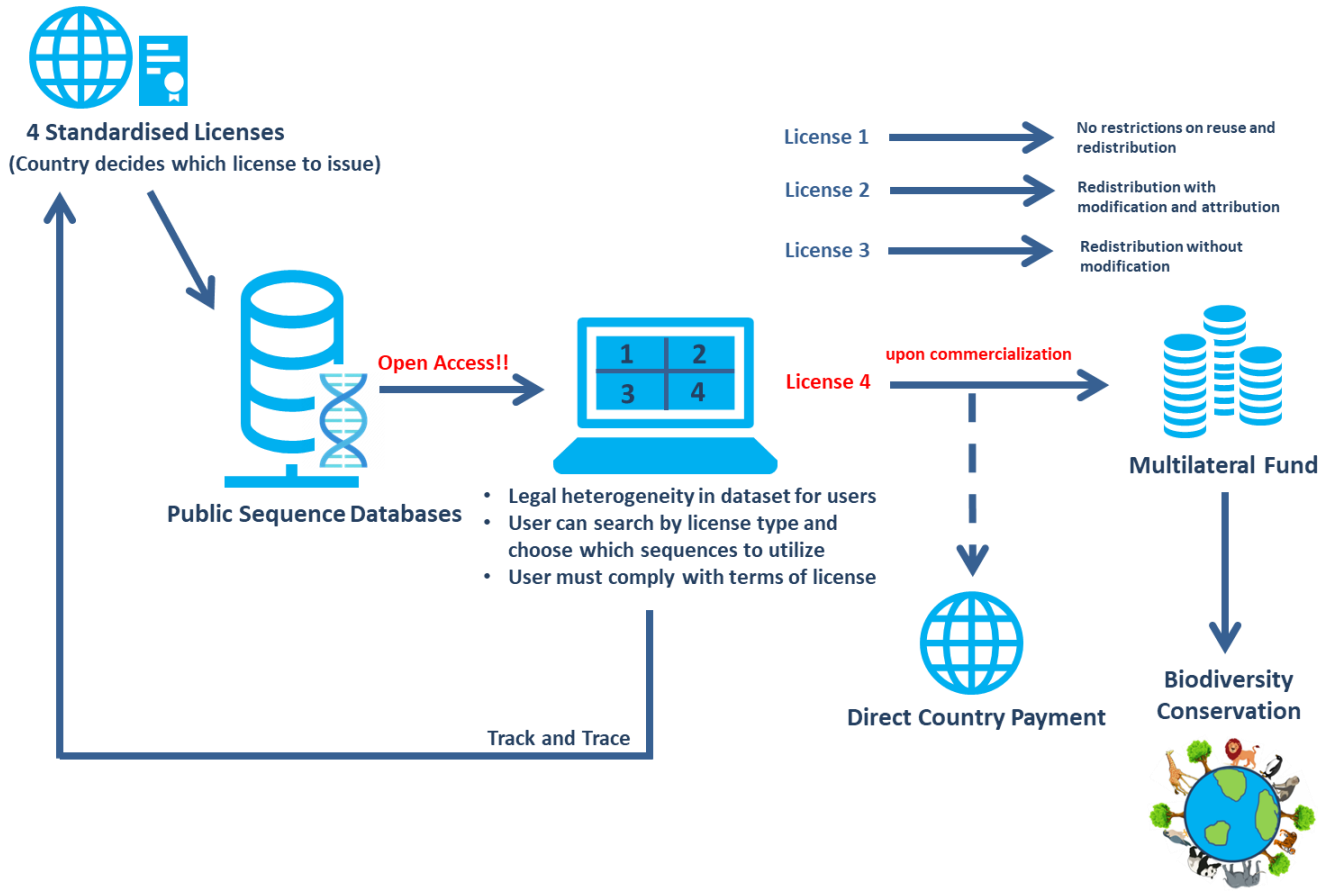
The blockchain technology has been repeatedly discussed in CBD circles as it could provide the "security" that provider countries seek. However, there are large and unresolved issues around the technological costs, infrastructure, and feasibility as well as the assumed incompatibility with the open-access system if sequence data itself were to be put into the blockchain. One compromise implementation could be to decouple the sequence data from some associated metadata and place only certain types of scientific and legal metadata into the blockchain for which access would be controlled and tracked, while continuing to use the open-access infrastructure for the sequences (DSI). This approach would require tracking and tracing by the user during the R&D process, which is a considerable challenge since sequence datasets commonly exceed millions of sequences analyzed. Monetary benefits could be triggered either at the point of access to the metadata or at the point of commercialization.

Technical note: for all diagrams, dotted lines indicate possible connections whereas solid lines indicate necessary connections/interactions.



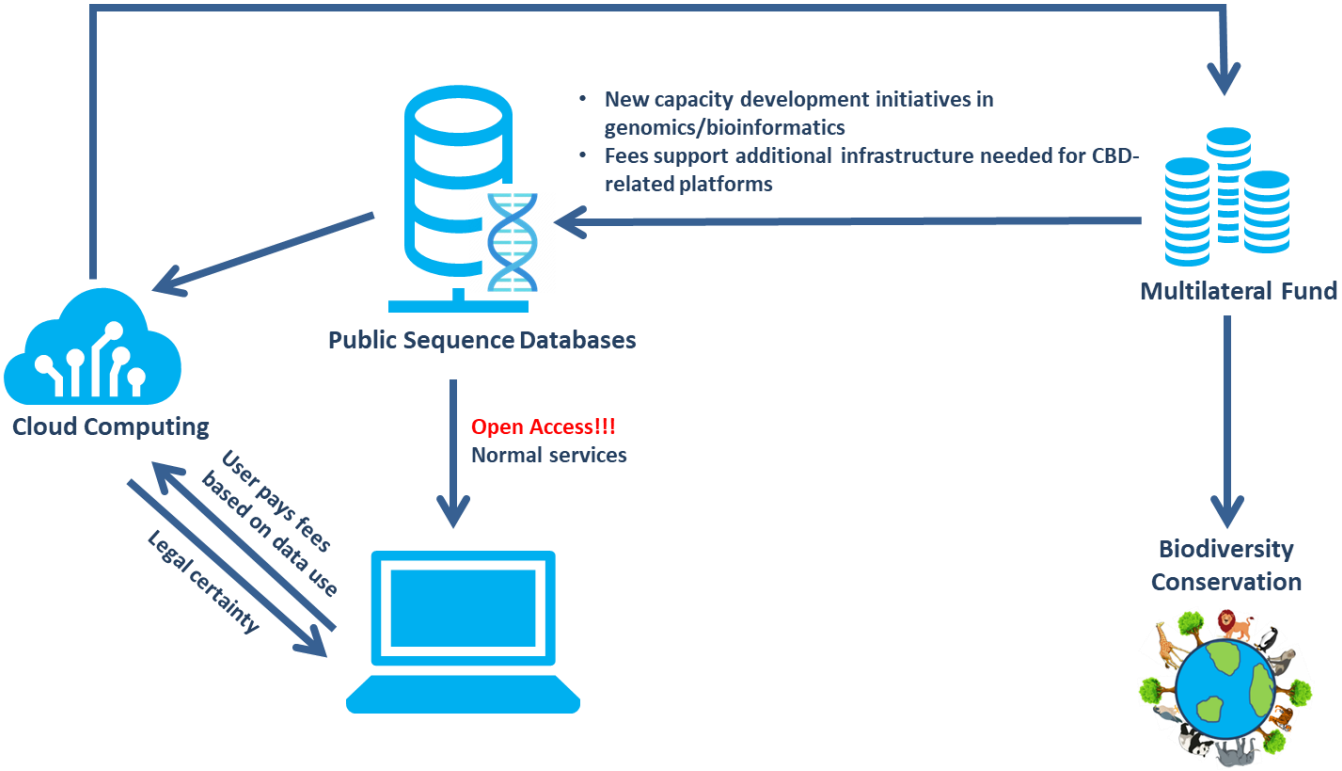
Option 2: Commons Licenses for DSI

The widespread development of open-source software provides an instructive model for how a system can provide both legal clarity and stimulate innovation through openness and reusability. In this scenario, Parties would have the option to require users to use and associate a standardized license to all publicly-available DSI. Databases would need to allow licenses to be associated with DSI and users would need to track and trace DSI used during R&D. (For example, ISNDC does not allow licenses to be associated with sequence data at present, but the Global Biodiversity Information Facility (GBIF) does.) Monetary benefits could be triggered at the point of access or at the time of commercialization. Alternatively, a license could require users to upload sequences to the cloud-based infrastructures (option 3). Enforcement and compliance mechanisms are not well-defined.



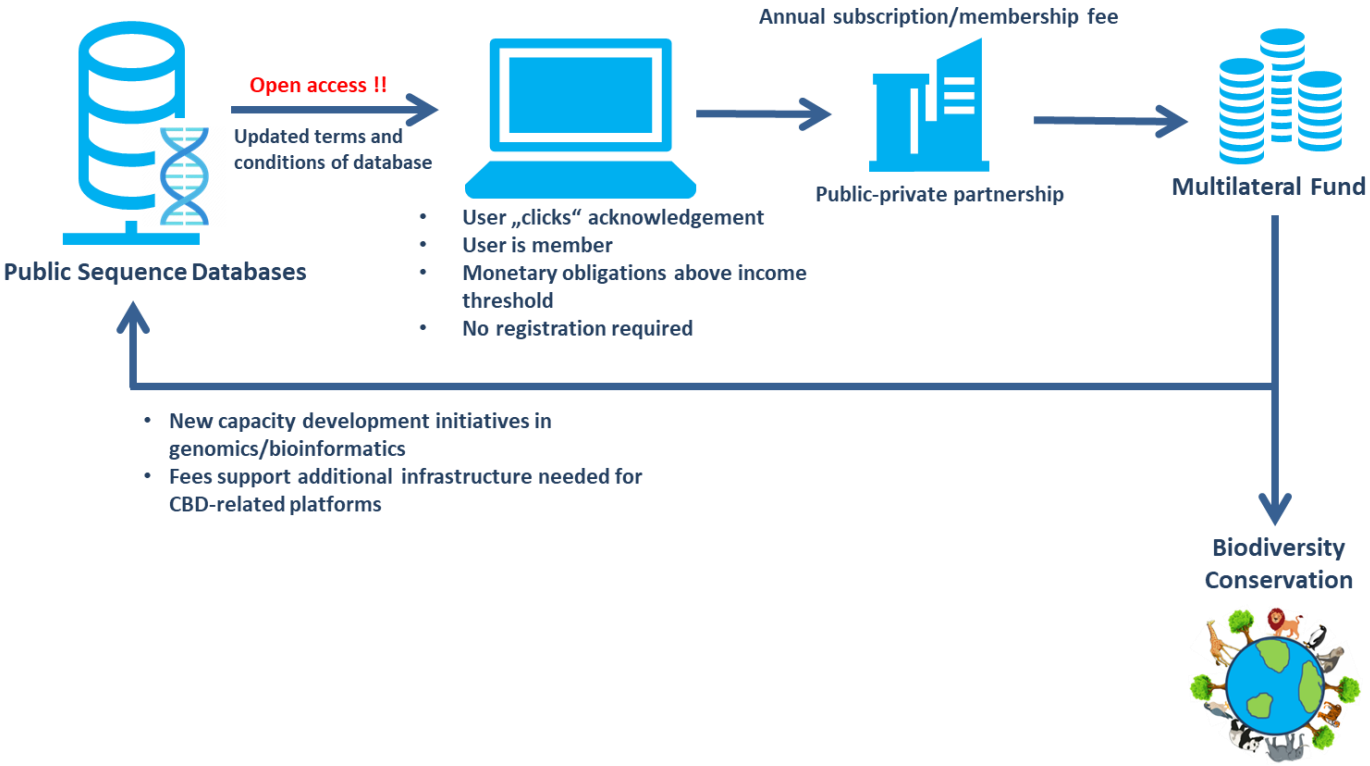
Option 3: Cloud-Based Fees

The explosion of data of all types in recent years has led to the centralization and consolidation of data storage and related services – use of "the cloud" rather than local storage and server maintenance. In this scenario, DSI would be offered through a "pay as you go" cloud-based model as an additional access mechanism to the standard non-cloud portals. The cloud portal could potentially offer additional services (e.g., storage or even sector-specific workbenches). A natural starting point for a first cloud-based platform would be to provide an additional entry point to the INSDC dataset but in later phases could be extended to other core database infrastructures. Fees generated by a cloud-based system would cover (additional) infrastructure costs, cloud costs, and then return a portion to biodiversity. Users of cloud-based biodiversity infrastructure would receive legal certainty and would not need to track and trace.



Option 4: Membership fee

Within discussion around the International Plant Treaty for Genetic Resources for Food and Agriculture (IPTGRFA), a new subscription system was proposed and was discussed by Parties (although not agreed to) in November 2019. This model very closely mirrors the concepts and ideas put forth by the IPTGRFA working group. The model simply requires annual "membership" payments for any users that are above a (yet unspecified, conceivably sector-tailored) income threshold. One concern within the WiLDSI project was that the word "subscription" triggers the idea of a paywall – a financial barrier the precedes access to DSI. A paywall system is NOT a model that would maintain, support, or enable open access and, as such, we have opted for the word "membership" instead to emphasize the separation between access and payment.



Option 5: Micro-levy

One opportunity for completely disconnecting monetary benefit sharing and access to DSI is to generate funds at the beginning of the value chain. Micro-levies are very small charges on high-volume purchases that are "barely felt" by the purchasing customer. The micro-levy concept presented here closely follows the airline micro-levies that fund UNITAID, a HIV/AIDS, malaria, and tuberculosis purchasing facility. The target of the micro-levy could be as broad or narrow as is politically feasible and could be applied to DNA sequencing/synthesis services, laboratory reagents or equipment or, more radically, mass- market consumer electronics. The micro-levy model requires that Parties give up "control" of the sequence data arising from their GR. In order to make this politically palatable, it is likely that a strong case will need to be made for the simplicity of this model and, in particular, its ability to generate much more significant monetary benefit streams than the other models.