Access and benefit sharing biological materials for machines: Artificial intelligence, machine learning and deep learning

Author

Lawson, C, Englezos, E, Rourke, M

Published

2025

Journal Title

Plants People Planet

Version

Version of Record (VoR)

DOI

10.1002/ppp3.70007

Rights statement

Downloaded from

Griffith Research Online

Received: 13 November 2024 | Revised: 17 January 2025 | Accepted: 6 February 2025

DOI: 10.1002/ppp3.70007

RESEARCH ARTICLE

Plants People Planet PPP Open Access

Access and benefit sharing biological materials for machines: Artificial intelligence, machine learning and deep learning

Charles Lawson[1] | Elizabeth Englezos[1] | Michelle Rourke[2]

[1]Griffith Law School, Griffith University, Southport, Queensland, Australia

[2]Griffith Law School, Griffith University, Nathan, Queensland, Australia

Correspondence
Charles Lawson, Griffith Law School, Griffith University, Parklands Dr, Southport, Queensland 4222, Australia
Email: c.lawson@griffith.edu.au

Societal Impact Statement

Future research and development of biological materials for foods, feeds, fibres, materials and medicines will increasingly rely on information and knowledge using Artificial Intelligence (AI) applications for detecting patterns to make useful decisions. The access and use of this information and knowledge is increasingly being regulated under international laws according to an ideal that delivers money and other benefits from the uses of the information and knowledge. We conclude these issues will require specific attention to ensure the ideals of fair and equitable benefit sharing are sustainable and can deliver real benefits.

Summary

- The global regulation of access and benefit sharing (ABS) biological materials is starting to impose complex rules for governing information and knowledge about those materials. Artificial Intelligence (AI) poses existential challenges for the research and development of those materials for foods, feeds, fibres, materials and medicines under these ABS schemes.
- We speculate these challenges are in three distinct scenarios: (1) data used to train the AI models; (2) data used to test (and verify) the AI models; and (3) data used in applying the models to reveal useful patterns. Building in ‘explainability’ to the AI algorithms may be a solution, at least in part, to delivering on fair and equitable ABS.
- We then posit two issues for ABS: (1) negotiating the ownership status to use the input data with each data owner for training, testing (and verifying) and using the models (although a further complication here is that most data is without an owner because it is already open and free as a public domain without intellectual property restrictions); and (2) following the data per se through the AI models (explainability).
- We conclude that how ABS will be addressed in developing and applying AI models will require careful consideration to avoid the apparent dulling effects of current ABS regulation and the potentially significant consequences this may have for biology-based research and the commercialisation of biology-based products and services.

KEYWORDS
access and benefit sharing, artificial intelligence, convention on biological diversity, explainability, genetic resources, information, Nagoya Protocol

1 | INTRODUCTION

Artificial Intelligence (AI) poses existential challenges for the regulatory complex in international law forums around the access and benefit sharing (ABS) of biological materials and their associated information and knowledge[1] used for foods, feeds, fibres, materials and medicines. These forums are the United Nations *Convention on Biological Diversity* (CBD) (CBD, 1992) and its *Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity* (Nagoya Protocol) (CBD, 2014), the Food and Agriculture Organisation of the United Nations *International Treaty for Plant Genetic Resources for Food and Agriculture* (Plant Treaty) (FAO, 2001; see also CBD, 2011, [103] and Annex (Decision X/1, Preamble)), the World Health Organisation of the United Nations *Pandemic Influenza Preparedness Framework* (PIP Framework) (WHO, 2011) and the United Nations *Agreement under the United Nations Convention on the Law of the Sea on the Conservation and Sustainable Use of Marine Biological Diversity of Areas Beyond National Jurisdiction* (BBNJ Agreement) (GA, 2023).[2]

Each of these regulatory ABS schemes accommodate the existing ideal of all information and knowledge being open and accessible with only: (1) contract terms and conditions limiting the uses of information and knowledge between the parties to the agreement about shared biological materials but not third parties; and (2) time limited restrictive intellectual property claims (such as copyrights, confidences and trade secrets) mandated by the World Trade Organisation's *Agreement on Trade Related Aspects of Intellectual Property Rights* (TRIPS) (WTO, 1994) that apply to all including third parties. These ABS schemes are moving to imposing binding and onerous legal obligations on the users (including third parties) of some information and knowledge about the covered biological materials (see Lawson et al., 2023, 2024). In addition to these ABS schemes, there are claims for 'cultural property' under the United Nations Educational, Scientific and Cultural Organization *Convention on the Means of Prohibiting and Preventing the Illicit Import, Export and Transfer of Ownership of Cultural Property* (Article 1) (UNESCO, 1970) and for 'cultural heritage, traditional knowledge and traditional cultural expressions' of Indigenous Peoples under the United Nations *Declaration on the Rights of Indigenous Peoples* (Article 31.1) (UNDRIP, 2007) (Figure 1).

The essence of ABS under the CBD, Nagoya Protocol, Plant Treaty, PIP Framework and BBNJ Agreement[3] is that users are required to account for their uses of the covered biological materials and share any monetary and non-monetary benefits that flow from utilising those materials (Lawson, 2012, pp. 1–22). The foundational example that inspired the original CBD imagination and reflected in the later forms of regulation was the agreement for access to Costa Rica's rainforests to discover active chemicals in exchange for money payments, and the transfer of technology between the National Biodiversity Institute of Costa Rica and the United States pharmaceutical company Merck (see Svarstad, 1994, pp. 53–54). Unfortunately, these CBD, Nagoya Protocol, Plant Treaty and PIP Framework ABS schemes have not delivered the expected monetary benefits (see, e.g., Laird et al., 2020, p. 1202; Prathapan et al., 2018, p. 1405). And as a direct response, the recent *Kunming-Montreal Global Biodiversity Framework* has mandated that 'by 2030, facilitating a significant increase of the benefits shared, in accordance with applicable international access and benefit-sharing instruments' requires action to deliver substantially more monetary benefits (CBD, 2023, [114] and Decision 15/4, [1] and Annex (Target 13)). How this will be achieved is currently being intensely negotiated in the various ABS forums with a focus on applying the same kinds of ABS ideas about tangible physical materials to some information and knowledge about those tangible physical materials (see, e.g., Scholz et al., 2023; Ebert et al., 2023; Lawson et al., 2020, 2019; Humphries et al., 2021). The role and place of similar claims under the UNESCO Convention and UNDRIP are also being addressed although they are still poorly characterised (Fredriksson, 2021; Lawson, 2017; Lawson et al., 2023).

The purpose of this article is to engage with some of the challenges posed by AI to the ABS ideal that information and knowledge can be followed, and benefits be fairly and equitably shared. Recall that ABS is founded on the proposition of benefits flowing from utilisation being the action of making practical and effective use of the accessed resource or information and knowledge (Humphries et al., 2021). In this formulation track and trace means an ability to follow the accessed materials and identify each step of their journeys into the hands of those required to share benefits, like a parcel being tracked and traced from the warehouse to the consumer's home (Humphries et al., 2020, p. 4). Meanwhile, traceability means the ability to find the origin of the accessed materials in the hands of those required to share benefits by following back to the origin, like a car manufacturer recall where the car can be traced to the manufacturer

[1] One match is data (unprocessed facts), one match in a matchbox is information (structured data) and that matches go in matchboxes in knowledge (information with experience, context and interpretation).

[2] The World Intellectual Property Organisation has recently concluded the concluded the *WIPO Treaty on Intellectual Property, Genetic Resources and Associated Traditional Knowledge* that requires certain disclosures in patent applications concerning some biological materials and associate Traditional Knowledge (WIPO, 2024) and there are current negotiations about a pandemic prevention, preparedness and response treaty (see Hampton et al., 2023).

[3] Being 'genetic resources' (CBD, 1992, Articles 2 and 15; CBD, 2014, Articles 2 and 5), 'plant genetic resources for food and agriculture' (FAO, 2001, Articles 2 and 13), 'genetic sequences' and 'genetic sequence data' (WHO, 2011, Articles 4.2 and 5.2 respectively) and 'marine genetic resources' and 'digital sequence information on marine genetic resources' (GA, 2023, Articles 1.8 and 14).

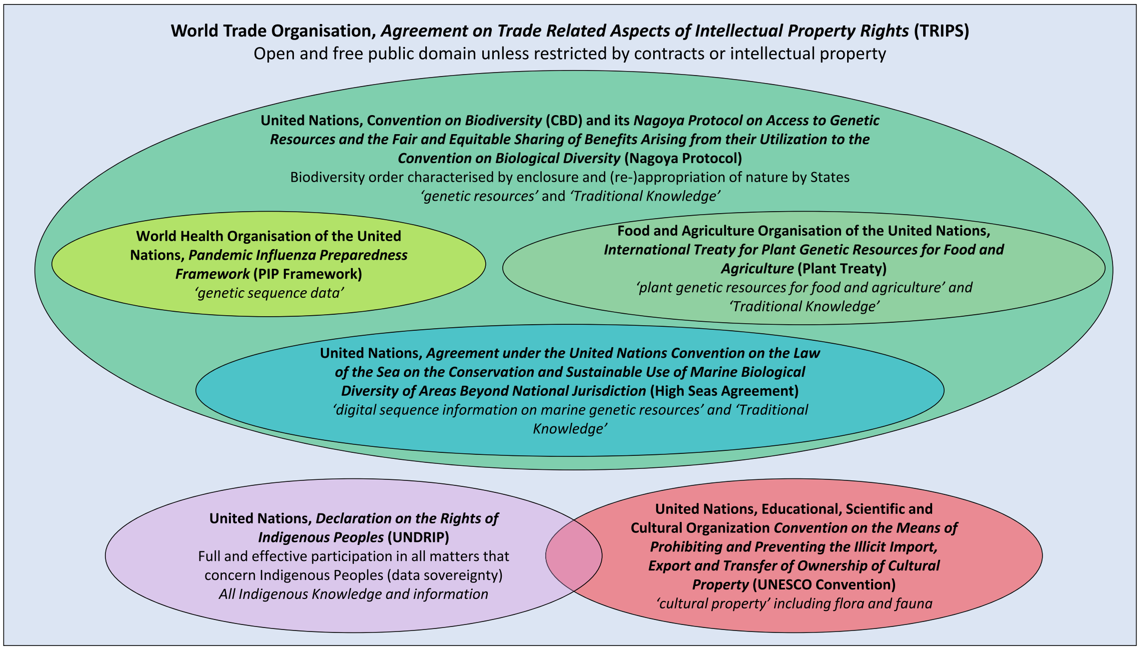

FIGURE 1 The regulatory complex governing information and knowledge about the access and uses of covered biological materials. The existing ideal under TRIPS is that all information and knowledge is open and accessible unless restricted by intellectual property. Under the CBD and Nagoya protocol (and various related kinds of information and knowledge under the high seas agreement, the plant treaty and the PIP framework) there are binding and onerous legal obligations for the users of information and knowledge for 'digital sequence information' and 'traditional knowledge'. Distinct and separate schemes for 'cultural property' under the UNESCO convention and indigenous Peoples' 'cultural heritage, traditional knowledge and traditional cultural expressions' (often termed data sovereignty) under UNDRIP with an uncertain overlap with the CBD, Nagoya protocol, high seas agreement, plant treaty and PIP framework schemes although all (presently) consistent with TRIPS (Lawson et al., 2023, 2024).

when there is a problem (Humphries et al., 2020, pp. 4–5). This is a challenge with some history. CBD negotiators, and especially the G77 plus China (Brand et al., 2008),[4] have tried to introduce monitoring and compliance mechanisms on users with some success including the Nagoya Protocol imposing some obligations on Parties to ensure compliance with ABS obligations in their jurisdiction (CBD, 2014, Articles 15–16), checkpoints (CBD, 2014, Article 17.1), compliance certificates (CBD, 2014, Article 17.2) and dispute resolution mechanisms (CBD, 2014, Article 18). Other forums have now added to this monitoring and compliance through a system of identifiers under the BBNJ Agreement to trace the results of research and development back to the originally collected materials and some of the associated information (see Humphries, 2024), a patent disclosure requirement for some inventions based on biological materials and associated traditional knowledge (see WIPO, 2024, Article 3; Nolff, 2024) and recognising the rights of Indigenous Peoples to associated Traditional Knowledge (see CBD, 2014, Articles 7 and 12; Lawson et al., 2023). None of these has settled on a mechanism for track and trace or traceability (see Humphries, 2024; Richerzhagen, 2014, pp. 141–142; Bavikatte & Robinson, 2011, pp. 47–48), and notably, the CBD and Nagoya Protocol forums accept that 'tracking and tracing of all digital sequence information on genetic resources is not practical' (CBD, 2023, [143] and Decision 15/9, [5]). The ABS ideal remains that any uses of the accessed biological materials and any associated information that are used in a way that generates benefits *should* be shared back to the origin sources of the biological materials – a simple ideal of accounting for and sharing the benefits.

This article speculates that AI applications pose challenges for data input in three distinct scenarios: (1) data used to train the AI models; (2) data used to test (and verify) the AI models; and (3) data used in applying the models to reveal useful patterns. Building in 'explainability' to the AI algorithms to account for the uses of information and knowledge may be a solution, at least in part, delivering on fair and equitable ABS, although this seems limited given the unexplainability of most AI models (see Lipton, 2018). The article first clarifies the ABS AI issues setting out key context and concepts and establishing the problem as 'explainability'. After that, the analysis engages with the ideal of explainability before setting out a discussion about the current proposals to regulate information in the CBD and other forums. The article concludes that AI presents an existential challenge for ABS because of: (1) the need for many contracts in different forms to deal with the different sources and uses of information and knowledge likely resulting in the 'tragedy of the anticommons'; and (2) how to deal with information and knowledge that has no identifiable originator or owner resulting in 'orphaned information and knowledge'. As such, AI is yet another lacuna for ABS regulation that will require specific attention to ensure the ideals of ABS are

[4]Notably the CBD draws a distinction between low- and high-income countries (using the terminology 'developing' and 'developed' respectively): see CBD, 1992, Preamble, Articles 8 (m), 9(e), 12, 16(4), 17(1), 18(2), 19(1) and 20.

sustainable and can deliver real benefits without the dulling burden of inefficient and ineffective regulation.

2 | THE DATA PROVENANCE PROBLEM?

The essence of AI is the goal of machines replicating human intelligence rather than a definition per se (McCarthy et al., 2006, p. 12; see also Turing, 1950; Buchanan, 2005; McCorduck & Cfe, 2004). In recent times a working definition of AI is 'the creation of intelligent systems that can perform tasks that would normally require human intelligence, such as learning, problem-solving, and decision-making' (Holzinger et al., 2023, p. 17).[5] This broad scope is usefully distinguished from the subfield of machine learning (ML) which is 'training digital computers to perform tasks without explicit instructions, using patterns and insights from data', and the subset of deep learning (DL, often termed Artificial Neural Networks) that 'uses artificial neural networks with many layers to learn and make decisions … particularly useful for tasks that involve analysing large amounts of data' (Holzinger et al., 2023, p. 17; see also Sah, 2020; LeCun et al., 2015). Figure 2 illustrates the useful conceptual distinctions (see also Haupt et al., 2022; Segessenmann et al., 2023) and Figure 3 illustrates the process (technique) distinctions between ML and DL. The key point for our purposes is that ML and DL bring together large and often disparate data sets and then apply an AI model to an analysis of these data sets atomising those data sets using some data and not others in making the ultimate predictions.

The foundation of these AI, ML and DL processes is mathematics using the language of numbers and formulas. These are applied through linear algebra (vector spaces and linear equations), statistics (collecting, analysing and interpreting data) and calculus (derivatives and integrals for rates of change and optimisation) using learned adjustments (by creating, classifying, linking and weighting data) through the network (layers) so that internal hidden units based on the learned adjustments affect the outputs (see, e.g., Widrow & Lehr, 1990). The key point is that these are machines based on mathematics that detect patterns to make useful decisions without any understanding of the underlying meaning or significance of the data in the way a human might understand meaning or significance (see Plebe & Grasso, 2019, pp. 526–528; Marcus, 2018, pp. 5–14).

The distinction between ML and DL reflects the techniques applied to data: ML uses computers with algorithms applying rules to analyse data and draw inferences applied to generally structured data sets and makes simple correlations between data points – in other words, ML reveals underlying patterns within data sets (Sah, 2020, p. 1); while, DL layers computer units and algorithms into an (artificial) neural network that uses connections and connection weights learned from training data to make complex, non-linear correlations between data points – in other words, DL reveals through the processing of multiple layers, representations of abstracted data (LeCun et al., 2015, p. 436; see also Segessenmann et al., 2023, pp. 3–9). 'Deep learning

[5]For an overview of definitions see Collins et al., 2021, pp. 6–7.

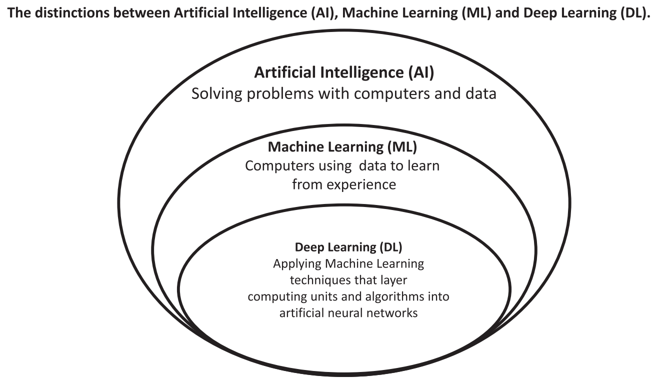

FIGURE 2 The kinds of artificial intelligence. The useful conceptual distinctions between artificial intelligence (AI), machine learning (ML) and deep learning (DL).

… is essentially a statistical technique for classifying patterns, based on sample data, using neural networks with multiple layers' (Marcus, 2018, p. 3). The 'deep' in DL refers to the increasing number of processing layers applying connections and connection weights to the input data, most often in hidden layers, mapping the inputs to desired outputs. More training data improves the connections and connection weights to set the structure applied to the input data, and theoretically, with enough training data any input data can be mapped with sufficient generalisation to be intelligent – 'deep neural networks can approximate every conceivable input-output mapping, and in principle a huge amount of cognitive tasks can potentially be formulated as a classification problem' (Tsimenidis, 2020, p. 3). The training data can be structured data (e.g., numerical data in tabulated forms) or unstructured data (e.g., texts and images) depending on the ML or DL model and the learning can be supervised, unsupervised and reinforced depending on the internal parameters of the algorithm being adjusted according to the known input and associated output data, adjusted according to patterns and similarities determined by the algorithm itself and adjusted according to reward and incentive feedback imposed on the algorithm respectively (see Hesami et al., 2022, pp. 3509–3511). The learnings from one ML or DL model can also be used as further inputs to other ML and DL models as transfer learnings (see Agarwal et al., 2021).

AI has already proven its application in drug discovery (see, e.g., Dara et al., 2022; Sellwood et al., 2018), functional and structural genomics (see, e.g., Caudai et al., 2021; Lin & Ngiam, 2023), proteomics (see, e.g., Xiao et al., 2021; Mund et al., 2021), metabolomics (see, e.g., Petrick & Shomron, 2022), phenotyping (see, e.g., Fiorani & Schurr, 2013; Tester & Langridge, 2010), plant breeding (see, e.g., Yoosefzadeh Najafabadi et al., 2023; Hesami et al., 2022), pharmacology (see, e.g., van der Lee & Swen, 2023), pharmacogenetics and pharmacogenomics (see, e.g., Roche-Lima et al., 2020; Lin et al., 2020), systems biology (see, e.g., Reel et al., 2021) and so on (see also Bhardwaj et al., 2022; Dixon et al., 2022; Kim, 2019; Oliveira, 2019). Both ML and DL, however, pose problems for ABS because ABS requires the end user to be able to follow data back to an originating source, either proving that information and knowledge

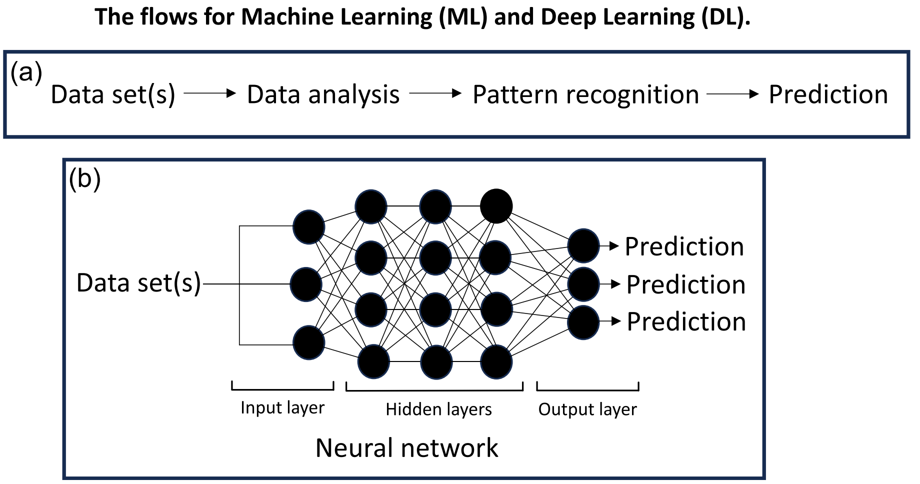

FIGURE 3 The flows of data. (a) Machine learning (ML); (b) deep learning (DL).

about the genetic resource (or Traditional Knowledge associated with the genetic resource) was the source or disproving it as a source (including that no source is known), for the purposes of benefit sharing. This arises in three distinct ABS scenarios:

1. *Training data* – Linking the training data in the dataset about the original genetic resources to the change in internal parameters that affect the later application of the ML or DL model. This is because the data in the dataset is not linked to the end user product or service although it affects how the data in the dataset about the original genetic resources links the information in the end user product or service after the ML or DL model has been trained or applied. All the (input) training data will be used because that is the data that establishes the patterns (for ML) and weightings (for DL).
2. *Testing (and verification) data* – Applying the testing (and verification) data to the trained ML or DL models to test and verify the already trained internal parameters that affect the later application of the ML or DL model to evaluate the qualities of the ML or DL model (usually their predictive performance or validate their predictions). This is because the data in the dataset while not necessarily linked to the end user product or service tests (or verifies) the original genetic resources linked to the information in the end user product or service after the application of trained ML or DL model to a dataset. Again, all the (input) testing (and verification) data will be used because that is the data that confirms or adjusts the patterns (for ML) and weightings (for DL).
3. *Using the models* – Linking the data in the dataset about the original genetic resources to the information in the end user product or service after the application of trained and tested (or verified) ML or DL to the dataset. This is because the data in the dataset is linked to the end user product or service either as the prediction revealed by the ML and DL application or the pattern or weighting imposed by the training or testing (and verification) data on the ML or DL application. Importantly, this is both the data set to which the ML or DL is applied (so, using the model) and the effects of the training or testing (and verification) data on the ML or DL application.

The ABS ideal anticipates evidence of provenance and an explanation between the original biological materials and uses/users. Benefit sharing then also anticipates some quantification of the uses of the original biological material in the product or service delivering benefits including the quantity or quality of the information and knowledge associated with those physical materials embodied in the product or service. ML and DL all have problems with proving a satisfactory provenance or an explanation for the sources of information (inputs) that contribute to the outputs. Generally, ML models are easier to explain than DL because they are easier to comprehend and self-interpret (using, for example, rule sets or decision trees), although some questions remain unanswerable (Doshi-Velez & Kim, 2017), and they are commonly termed glass (or white) box or grey box problems depending on their degree of explainability (Ali et al., 2023, pp. 2–3). Generally, DL models cannot provide explanations (or reasons) for the output, and this is commonly called the black box problem (see, e.g., Castelvecchi, 2016; see also Rudin, 2019). This is because DL models (and also many ML models) process numbers that have been applied to extracted features and vectorisation that represent the input data objects and these are then processed to identify statistical regularities among the numbers rather than qualitative relationships like causation, hierarchy, and so on (see Marcus, 2018, pp. 5–14). These statistical regularities between the quantitative relationships among the numbers cannot be explained in any meaningful way for human comprehension, hence the black box problem (Tsimenidis, 2020, p. 7). While some ML models may be readily explainable, more complex models can transcend human cognition making them also unexplainable (see Lipton, 2018, p. 3). While the DL models (and many ML models) may never satisfy human meaningful explanations (or reasons) for the outputs, some measures of explanation (or reasons) may be possible.

Explainability (often termed 'eXplainable AI' or 'XAI') is understood in contrast to interpretability, although these terms are often used interchangeably (see Linardatos et al., 2020, pp. 2–3). There is a difference between 'interpretation' as 'the mapping of an abstract concept (e.g. a predicted class) into a domain that the human can make sense of' and 'explanation' as 'the collection of features of the interpretable domain, that have contributed for a given example to produce a decision (e.g., classification or regression)' (Montavon et al., 2018, p. 2; see also Ali et al., 2023, p. 8; Guidotti et al., 2018, p. 5). Put slightly differently, 'interpretation' is opening the black box (or grey box) to understand how the model works (a technical question[6] – generally an *ante hoc* interpretation) and 'explanation' is understanding what the model does (a representation question – generally a post hoc explanation) (Berry, 2024, p. 137; Lisboa et al., 2023, pp. 28–30; see also Ali et al., 2023, pp. 7–8; Guidotti et al., 2018, p. 5).[7] This treats these terms differently, and for our purposes, explainability means the capacity of black-box AI models to be explained using any external means such as visualisations that are self-contained and do not need further explanations (Guidotti et al., 2018, p. 5).

A further refinement is 'transparency' as the 'model is … by itself … understandable' and 'understandable' as 'the characteristic of a model to make a human understand its function – how the model works – without any need for explaining its internal structure or the algorithmic means' (Arrieta et al., 2020, pp. 84–85). In other words, '*transparency* can make *incomprehensible* systems *interpretable*, and *explainability* can make *interpretable* systems *understandable*' (emphasis in original) (Verhagen et al., 2021, p. 122; see also Bellucci et al., 2021, p. 241–250; Vilone & Longo, 2020). A further problem is that the ML and DL outputs also have to be understandable to humans, and if the outputs are not, then those outputs need to be made human understandable so they can be useable and useful (Holzinger et al., 2019, pp. 3–4). This has been addressed through a range of other terminologies that remain under development: causality, completeness, informativeness, justifiability, reliability, responsibility, reversibility, satisfaction, simplicity, transferability, trustworthiness and so on (Ali et al., 2023, p. 8 and the references therein). The conundrum is that generally unexplainable DL models are more accurate (deliver better predictions, recommendations or classifications) than more generally explainable ML models with a trade-off between accuracy and explainability (Ali et al., 2023, p. 8).[8] So, what actually is explainability or XAI?

3 | EXPLAINABLE AI (OR XAI)

While explainability has attracted a considerable research effort, there is no universally accepted definition of explainability.[9] The terminology and conception of the black box problem is generally tied to the kinds of application with strands of research focussing on medical diagnosis (see, e.g., Göndöcs & Dörfler, 2024; Al Kuwaiti et al., 2023), insurance and financial services (see, e.g., Nallakaruppan et al., 2024), autonomous vehicles and transportation (see, e.g., Kuznietsov et al., 2024; Tekkesinoglu et al., 2024), law and law enforcement (see, e.g., McGregor Richmond et al., 2024; Vale et al., 2022), bioinformatics (see, e.g., Biswas et al., 2023), military applications (see, e.g., Szabadföldi, 2021), employment decisions (see, e.g., Chan, 2024) and so on. Generally, more critical decisions require more explanation and often there are specific requirements for a particular application (Burkart & Huber, 2021, p. 250). More broadly for our purposes, explainability has a legal dimension (rules and requirements that account for the rationale of the decision) and a technical dimension (the criteria or thresholds). For example, a legal decision requires a justification through reasons for the particular output decision rather than a necessarily causal understanding of how the algorithm produced an output, so the focus of explainability is on reasons (McGregor Richmond et al., 2024, p. 7). In contrast, a medical diagnosis merely requires a technical assessment of the diagnostic criteria as a scientific understanding rather than how a specific outcome was arrived at (Göndöcs & Dörfler, 2024). In other words, the requirements for explainability also depend on the context for requiring an explanation.

The main forms of the XAI research effort have been on strategies for explanation. These include comprehensible text explanations (Lipton, 2018, pp. 15–17), mathematics accounts (Ribeiro et al., 2016, pp. 1137–1,140), visualisations (Lipton, 2018, pp. 17–18) and so on (see Ali et al., 2023, p. 6 [Table 1]), with various classifications such as application-grounded, human-grounded and functionally-grounded (Doshi-Velez & Kim, 2017, p. 2), taxonomies of methods (Gilpin et al., 2018, p. 86), taxonomies of types of problems (Guidotti et al., 2018, p. 11) and so on (see Molnar, 2020). In a practical sense, and as a generalisation, XAI are a set of procedures and techniques (sometimes called methods) that can be either or both applied to and/or integrated into ML and DL models to make them explainable, although the suitability of explanations will always depend on the kinds of user needing the explanations, and what it is they want to know (see Ali et al., 2023; Arrieta et al., 2020; Ribera & Lapedriza, 2019). This is, of course, only a generalisation and ABS is a specific context for XAI.

Recalling the ABS problem addressed by this article is about the end user being able to follow back information and knowledge to an originating source to either prove that information or knowledge about the genetic resource (or Traditional Knowledge associated with the genetic resource) was the source or disprove it as a source (including that no source is known) (section 2). Put another way, explanations need to facilitate the exercise of the data originator's ABS interests in the original biological materials embodied in the input and output data, being the data used to train, test and then use the ML and DL models. Characterised this way the explainability question then is about accounting for what information and knowledge is actually used, distinguishing the information and knowledge subject to ABS from other information and knowledge so that benefit sharing can be appropriately addressed.

[6]So, interpretation is understanding the mechanics of a car, but explainability is understanding that a car can take you from location A to location B

[7]*Ante hoc* interpretation: Someone can explain to you how your car works using a scale model of a combustion engine (therefore, you know how it works). Post hoc explanation: You can see that the car got you from location A to location B (therefore, you now know what it does).

[8]But also see Rudin et al., 2022, p. 6 ('there is no scientific evidence for a general trade-off between accuracy and interpretability').

[9]See, for examples, Al-Ansari et al., 2024; Hassija et al., 2024; Ali et al., 2023; Saranya & Subhashini, 2023; Lisboa et al., 2023; Yu, 2023; Li et al., 2022; Samek et al., 2021; Burkart & Huber, 2021; Confalonieri et al., 2021; Vassiliades et al., 2021; Alicioglu & Sun, 2022; Das & Rad, 2020; Arrieta et al., 2020; Fernandez et al., 2019; Mueller et al., 2019; Guidotti et al., 2018; and so on.

A useful categorisation of the procedures and techniques (or methods) for these assessments is to consider the dimensions of the ML and DL models[10]: data explainability about the kinds of data, data distributions, data inferences and so on, that are used in the training of ML and DL models; model explainability about the parameters, objectives and actions of the ML and DL model including what is taken into account and what is rejected; post hoc explainability about the various text, visual explanations and features accounts for how and why a model generated a particular output; and, an assessment of explainability about evaluating the quality of explainability (see Ali et al., 2023, pp. 10–34; Burkart & Huber, 2021, pp. 253–298). The key point being that ML and DL models are explainable either because they intrinsically include a level of explainability through constraints imposed on the complexity of the model (called intrinsic explainability), or procedures and techniques (or methods) are applied after the models are trained to explain the outputs (called post hoc explainability) (Carvalho et al., 2019, p. 12).[11] For ABS to be successful and the benefits from using the information and knowledge to be fairly and equitably shared, the ML and DL models need to account for their uses of data (from the information and knowledge).

There are many approaches to intrinsic and post hoc explainability. The particular approach will depend on the specific requirements of the end users. Data scientists and developers are probably more interested in intrinsic explainability about model performance while domain experts and end users are probably more interested in post hoc explainability about how and why a model generated particular results and conclusions (Ali et al., 2023, p. 9). For post hoc explainability there are generally two approaches: local explanations that focus on the data and provide individual explanations for specific outputs, and global explanations that focus on the model and provide some understanding of the mechanisms by which the model works. The point is that explainability is possible either driven by the type of ML and DL model (intrinsic explainability) or the procedures and techniques (or methods) that are applied after the models are trained (post hoc explainability). And importantly for ABS, XAI will need to engage with the specific context so that the data uses can be explained (including quantified) to 'determine the extent to which a particular input was determinative or influential on the output' (see Doshi-Velez et al., 2017, p. 4). Presumably, where input data is not used (there are no connections) then it would be unfair to treat this the same way as treating data that is used. Further, when combining data from multiple sources that is then used with different weightings then different claims to the shares of benefits might be expected with a greater share going to those whose data is more determinative or influential. Importantly, however, this also means that the context of ABS will need to be factored into making the ML and DL models explainable.

[10]These are not settled and there are others: see, e.g., Sheh & Monteath, 2018; Preece, 2018.

[11]Again, there is variation in this terminology, some making the distinction between, e.g., *ex ante* transparency (intrinsic explainability) and *ex post* transparency (post hoc explainability): see House of Lords, 2018, p. 38.

4 | THE REGULATORY CHALLENGE

There is some evidence that ABS schemes in domestic laws have reduced the collection of physical biological materials and the uses of information and knowledge, and predominantly information in publicly available databases (see, e.g., Engels et al., 2024; Michiels et al., 2021; Prathapan et al., 2018). This reflects the increasing complexity and costs in complying with the ABS obligations (Ebert et al., 2023; Hampton, 2023; Michiels et al., 2021) and the apparent regulatory gap in the current ABS schemes for information and knowledge (see, e.g., Mulatinho Simoes & Birchfield, 2024; Bond & Scott, 2020; Lawson et al., 2020). The response in the various CBD, Nagoya Protocol, Plant Treaty, PIP Framework and BBNJ Agreement forums has been to extend ABS to information and knowledge (and associated Traditional Knowledge) (see Lawson et al., 2024). The recent *Kunming-Montreal Global Biodiversity Framework* in the CBD and Nagoya Protocol forums mandated finding ABS solutions to significantly increasing the benefits shared (CBD, 2023, [114] and Decision 15/4, [1] and Annex (Target 13)) and more recently decided on a multilateral fund (called the 'Cali Fund') from taxing certain users of 'digital sequence information on genetic resources' (CBD, 2024, Decision 16/2, [1] and [2]). The Plant Treaty forum is also considering a package of measures including monetary benefit-sharing from the use of 'digital sequence information/genetic sequence data' (Working Group on Multilateral System, 2024a, 2024b) and the BBNJ Agreement explicitly includes benefit-sharing for 'digital sequence information on marine genetic resources' (GA, 2023, Article 10.1).[12] Each forum is seeking a solution for only *some* categories of information and knowledge essentially creating two forms of regulation and four ways to limit information and knowledge that would otherwise be open and accessible unless limited by intellectual property claims[13]: (1) Bespoke terms and conditions for information and knowledge in contracts for accessing biological materials (including associated Traditional Knowledge) under the CBD and Nagoya Protocol that is not 'digital sequence information on genetic resources'; and (2) a multilateral mechanism for some information and knowledge (and some associated Traditional Knowledge): (i) under the CBD and Nagoya Protocol that is 'digital sequence information on genetic resources'; (ii) under the Plant Treaty that is 'digital sequence information/genetic sequence data'; and (iii) under the BBNJ Agreement that is 'digital sequence information on marine genetic resources'.[14] A common theme among all these regulatory forms and forums is that users of information and knowledge of covered biological materials may be subject to limits on using data that are bespoke or multilateral terms and conditions depending on the origins (within or beyond national jurisdictions: CBD, Nagoya Protocol or BBNJ Agreement) or forms of

[12]In another forum the World Health Organisation of the United Nations is negotiating an agreement dealing with ABS for 'digital sequence information' and pathogens: see Rourke, 2024.

[13]Noting that the PIP Framework provides that 'genetic sequence data' of influenza viruses with human pandemic potential should be made available in public-domain or public-access databases: WHO, 2011, Article 5.2.2.

[14]Although this might change for other forms of information and knowledge if the Conference of the Parties determines different modalities: WHO, 2011, Article 14.7.

the material (influenza viruses with human pandemic potential, plant genetic resources for food and agriculture, or other: CBD, Nagoya Protocol, PIP Framework or Plant Treaty). The following discussion focusses on the CBD and Nagoya Protocol using bespoke contracts for information and knowledge (and associated Traditional Knowledge) that is not 'digital sequence information on genetic resources'. This is because, presumably, the CBD and Nagoya Protocol, Plant Treaty and BBNJ Agreement multilateral systems will develop their own modalities including for dealings with the 'digital sequence information' clusters they have identified. The main point here is that these clusters will need to clarify the scope of information and knowledge as there is no consensus about what these 'digital sequence information' terminologies mean. For example, the CBD forum has no definition of 'digital sequence information on genetic resources' and the agreed processes for negotiating the modalities includes taxing profits of revenues of '[u]sers of digital sequence information on genetic resources in sectors that directly or indirectly benefit from its use in their commercial activities' (CBD, 2024, [1] and Annex, [3]) with those taxes going to the multilateral Cali Fund (CBD, 2024, Decision 16/2, [1]). This might be a narrow scheme covering just DNA sequences in some publicly accessible databases, although it could also be very broad with the scope of 'digital sequence information on genetic resources' encompassing DNA sequences, RNA sequences, amino acid sequences and perhaps even associated phenotypic and environmental data (see CBD, 2020, [11]).

Returning to the CBD and Nagoya Protocol using bespoke contracts for information and knowledge (and associated Traditional Knowledge) that is not 'digital sequence information on genetic resources'. The increasing scope and scale of datasets available make ML and DL exciting prospects for biology (see, e.g., Holzinger et al., 2023). The analysis in this article so far shows that data made available as training data, testing (and verification) data and data used in particularly opaque black box ML and DL models are problematic for ABS because the causal link (the evidence of provenance and an explanation) between the original biological materials and uses/users is not clear. And as the scale of data sets with disparate origins increases this is likely an even bigger problem in the future. There are then two parts to this problem in developing regulation: (1) negotiating the ownership status to use the input data with each data owner for training, testing (and verifying) and using the models (although a further complication here is that most data is without an owner because it is already open and free as a public domain without intellectual property restrictions – so-called 'orphaned information and knowledge'); and (2) following the data per se through the ML and DL models (explainability).

These are overlapping problems that would be resolved by comprehensive explainable DL and ML models facilitating track and trace or traceability. In those circumstances, each of the owners of the input data would reach an agreement of the particular quantities and qualities of the data in the distinct ABS scenarios with each of the users. This is then simply an efficiency problem of stacking the input data contracts and reconciling the various terms and conditions of use in the outputs, including benefits to be shared with each input owner. As the analysis shows, explainability may be possible so that, depending on the particular DL and ML models, the uses of information and knowledge for training, testing (and verification) and model application can be determined, contracts agreed and the benefits shared. The analysis also revealed that this is complicated, such as intrinsic and post hoc explainability and the specific context of the explainability user and may require careful consideration in developing the DL and ML models. Whether this might be specific programming code requirements, visualisation standards, text explanations, classifications and so on, will need to be considered and may not necessarily be resolved by a single or uniform standard applicable to all ML and DL models. Put simply, the forms of explainability depend on the kinds of explanations required by particular users and these explanations about the sources and specific uses of data will also have different forms depending on the training, testing and uses in ML and DL models. A further consideration will be reach-through terms and conditions applying to future uses of the ML and DL model outputs (predictions) and the requirements to further deal with those outputs subject to those reach throughs and seeking to avoid the 'tragedy of the anti-commons' – the underuse of scarce resources because too many owners can block each other for future uses (Heller & Eisenberg, 1998).

Where comprehensive explainable DL and ML models are not possible, and the causal link between the original biological materials and uses/users is not apparent, then contracts between the owners of the input data (that is not in the public domain) and each of the users are likely to be particularly difficult. Such contracts can no longer focus on the quantities and qualities of the data in the distinct ABS scenarios but instead make assumptions about uses and the contributions to the ML and DL model outputs (predictions). The result will be that *all* data owners with input data for the training, testing and uses in ML and DL models will likely make broad claims to the outputs and the benefits are unlikely to address actual causal contributions from original biological materials. Again, reach through terms and conditions applying to future uses are then likely to be even more restrictive aggravating the 'tragedy of the anti-commons'. As alluded to above, a further complication for the large datasets is that most data are likely without an identifiable owner because the information and knowledge are already open and free in the public domain without intellectual property restrictions – so-called 'orphaned information and knowledge'. All this information and knowledge from biological materials would have had an originating owner at some stage. Without a clear provenance, there will always be potential claims from possible owners and these pose considerable reputational risks and informal regulation as users seek to comply with soft law measures like journal requirements, institutional research frameworks, database submission requirements and so on (see MacDonald, 2025; Sherman & Henry, 2020). The same problems are apparent from copyright orphaned works where a work cannot be used without permission and the copyright owner cannot be located (see, e.g., Sarid & Ben-Zvi, 2023). The same kinds of concerns will apply to information and knowledge that is UNESCO Convention 'cultural property' and UNDRIP 'cultural heritage, traditional knowledge and traditional cultural expressions' (see Lawson et al., 2023). These are potentially significant consequences for biology-based research and the commercialisation of biology-based products and services.

While the multitude of contracts and problems of explainability might be characterised as another instance that will dull the quest to collect physical biological materials and use of the resulting information and knowledge, it might also be an opportunity to reconsider ABS. This can take into account a broader range of perspectives with a renewed focus on conservation, equity and a regulatory form more relevant to the kinds of benefits that suit the political and economic realities of communities (Halewood et al., 2023; Leskien, 2023; Scholz et al., 2023; Wynberg, 2023). More broadly, this is also an opportunity to (re)organise the biosciences in ways that the wealth and value of biological materials and the associated information and knowledge is distributed.

5 | CONCLUSIONS

This article shows that AI, ML and DL processes, specifically black box ML and DL, are problematic for ABS. A careful resolution will be necessary to avoid the dulling effect on research and development from imposing overly onerous obligations on using data. As always, the challenge is finding common ground between facilitating seamless research and development with free and open access to the physical material, information and knowledge inputs and delivering on conservation and equity. This is also an opportunity to revisit ABS with a refreshed enthusiasm to address research and development in the context of the kinds of benefits that suit the political and economic realities of communities – *real* fair and equitable benefit sharing.

AUTHOR CONTRIBUTIONS

C.L., E.E. and M.R. planned and designed the research and wrote the manuscript.

ACKNOWLEDGEMENTS

We are grateful for the useful discussions about this work with Fran Humphries and Todd Berry and the comments provided by Fran Humphries and the referees on earlier versions of this paper.

CONFLICTS OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

ORCID

Charles Lawson https://orcid.org/0000-0003-0253-026X
Elizabeth Englezos https://orcid.org/0000-0003-2743-248X
Michelle Rourke https://orcid.org/0000-0003-1560-8943

REFERENCES

Agarwal, N., Sondhi, A., Chopra, K., & Singh, G. (2021). Transfer learning: survey and classification. In S. Tiwari, M. Trivedi, K. Mishra, A. Misra, K. Kumar, & E. Suryani (Eds.), *Smart innovations in communication and computational sciences* (pp. 145–155). Springer. https://doi.org/10.1007/978-981-15-5345-5_13

Al Kuwaiti, A., Nazer, K., Al-Reedy, A., Al-Shehri, S., Al-Muhanna, A., Subbarayalu, A., Al Muhanna, D., & Al-Muhanna, F. (2023). A review of the role of artificial intelligence in healthcare. *Journal of Personalized Medicine*, *13*(6), 951. https://doi.org/10.3390/jpm13060951

Al-Ansari, N., Al-Thani, D., & Al-Mansoori, R. S. (2024). User-centered evaluation of explainable artificial intelligence (XAI): A systematic literature review. *Human Behavior and Emerging Technologies*, *2024*(1), 4628855. https://doi.org/10.1155/2024/4628855

Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., Guidotti, R., Del Ser, J., Díaz-Rodríguez, N., & Herrera, F. (2023). Explainable artificial intelligence (XAI): What we know and what is left to attain trustworthy artificial intelligence. *Information Fusion*, *99*, 101805. https://doi.org/10.1016/j.inffus.2023.101805

Alicioglu, G., & Sun, B. (2022). A survey of visual analytics for explainable artificial intelligence methods. *Computers & Graphics*, *102*, 502–520. https://doi.org/10.1016/j.cag.2021.09.002

Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., & Chatila, R. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Bavikatte, K., & Robinson, D. (2011). Towards a people's history of the law: Biocultural jurisprudence and the Nagoya Protocol on access and benefit sharing. *Law, Environment and Development Journal*, *7*(1), 35–52. http://wwwlead-journal.org/content/11033.pdf

Bellucci, M., Delestre, N., Malandain, N., & Zanni-Merk, C. (2021). Towards a terminology for a fully contextualized XAI. *Procedia Computer Science*, *192*, 241–250. https://doi.org/10.1016/j.procs.2021.08.025

Berry, D. (2024). Algorithm and code: explainability, interpretability and policy. In R. Regine Paul, E. Carmel, & J. Cobbe (Eds.), *Handbook on public policy and artificial intelligence* (pp. 133–145). Edward Elgar. https://doi.org/10.4337/9781803922171.00017

Bhardwaj, A., Kishore, S., & Pandey, D. (2022). Artificial intelligence in biological sciences. *Lifestyles*, *12*(9), 1430. https://doi.org/10.3390/life12091430

Biswas, A., Kumari, A., Gaikwad, D., & Pandey, D. (2023). Revolutionizing biological science: The synergy of genomics in health, bioinformatics, agriculture, and artificial intelligence. *OMICS: A Journal of Integrative Biology*, *27*(12), 550–569. https://doi.org/10.1089/omi.2023.0197

Bond, M., & Scott, D. (2020). Digital biopiracy and the (dis)assembling of the Nagoya Protocol. *Geoforum*, *117*, 24–32. https://doi.org/10.1016/j.geoforum.2020.09.001

Brand, U., Görg, C., Hirsch, J., & Wissen, M. (2008). *Conflicts in environmental regulation and the internationalisation of the state: Contested terrains*. Routledge. https://doi.org/10.4324/9780203928509

Buchanan, B. (2005). A (very) brief history of artificial intelligence. *AI Magazine*, *26*(4), 53–60.

Burkart, N., & Huber, M. (2021). A survey on the explainability of supervised machine learning. *Journal of Artificial Intelligence Research*, *70*, 245–317. https://doi.org/10.1613/jair.1.12228

Carvalho, D., Pereira, E., & Cardoso, J. (2019). Machine learning interpretability: A survey on methods and metrics. *Electronics*, *8*(8), 832. https://doi.org/10.3390/electronics8080832

Castelvecchi, D. (2016). Can we open the black box of AI? *Nature*, *538*(7623), 20–23. https://doi.org/10.1038/538020a

Caudai, C., Galizia, A., Geraci, F., Le Pera, L., Morea, V., Salerno, E., Via, A., & Colombo, T. (2021). AI applications in functional genomics. *Computational and Structural Biotechnology Journal*, *19*, 5762–5790. https://doi.org/10.1016/j.csbj.2021.10.009

CBD. (1992). Convention on biological diversity, opened for signature 5 June 1992, 1760 U.N.T.S. 79 (entered into force 29 December 1993) (CBD)

CBD. (2011). Report of the tenth meeting of the conference of the parties to the Convention on Biological Diversity. UN Doc. UNEP/CBD/COP/10/27. Montreal: Secretariat of the Conference of the Parties to the Convention on Biological Diversity.

CBD. (2014). Nagoya protocol on access to genetic resources and the fair and equitable sharing of benefits arising from their utilization to the convention on biological diversity, opened for signature 29 october 2010, 3008 U.N.T.S. 3 (entered into force 12 october 2014) (Nagoya Protocol)

CBD. (2020). Digital sequence information on genetic resources: Concept, scope and current use. UN Doc. CBD/DSI/AHTEG/2020/1/3. Montreal: Secretariat of the Conference of the Parties to the Convention on Biological Diversity.

CBD. (2023). Report of the conference of the parties to the Convention on Biological Diversity on the second part of its fifteenth meeting. UN Doc. CBD. /COP/15/17. Montreal: Secretariat of the Conference of the Parties to the Convention on Biological Diversity.

CBD. (2024). Digital sequence information on genetic resources: Decision 16/2. UN doc. CBD/COP/DEC/16/2. Montreal: Secretariat of the Conference of the Parties to the Convention on Biological Diversity.

Chan, G. (2024). Ai employment decision-making: Integrating the equal opportunity merit principle and explainable AI. *AI & Society*, *39*(3), 1027–1038. https://doi.org/10.1007/s00146-022-01532-w

Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management*, *60*, 102383. https://doi.org/10.1016/j.ijinfomgt.2021.102383

Confalonieri, R., Coba, L., Wagner, B., & Besold, T. (2021). A historical perspective of explainable artificial intelligence. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *11*(1), e1391. https://doi.org/10.1002/widm.1391

Dara, S., Dhamercherla, S., Jadav, S., Babu, C., & Ahsan, M. (2022). Machine learning in drug discovery: A review. *Artificial Intelligence Review*, *55*, 1947–1999. https://doi.org/10.1007/s10462-021-10058-4

Das, A. and Rad, P. (2020). Opportunities and challenges in explainable artificial intelligence (XAI): A survey. https://arxiv.org/pdf/2006.11371

Dixon, T., Freemont, P., Johnson, R., & Pretorius, I. (2022). A global forum on synthetic biology: The need for international engagement. *Nature Communications*, *13*(1), 3516. https://doi.org/10.1038/s41467-022-31265-9

Doshi-Velez, F. and Kim, B. (2017). Towards a rigorous science of interpretable machine learning. https://arxiv.org/pdf/1702.08608

Doshi-Velez, F., Kortz, M., Budish, R., Bavitz, C., Gershman, S., O'Brien, D., Scott, K., Schieber, S., Waldo, J., Weinberger, D., and Weller, A. (2017). Accountability of AI under the law: The role of explanation. https://arxiv.org/pdf/1711.01134

Ebert, A., Engels, J., Schafleitner, R., Hintum, T., & Mwila, G. (2023). Critical review of the increasing complexity of access and benefit-sharing policies of genetic resources for genebank curators and plant breeders – A public and private sector perspective. *Plants*, *12*(16), 2992. https://doi.org/10.3390/plants12162992

Engels, J., Ebert, A., & van Hintum, T. (2024). Collaboration between private and public genebanks in conserving and using plant genetic resources. *Plants*, *13*(2), 247. https://doi.org/10.3390/plants13020247

FAO. (2001). International treaty on plant genetic resources for food and agriculture, opened for signature 3 November 2001, 2400 U.N.T.S. 303 (entered into force 29 June 2004) (Plant Treaty).

Fernandez, A., Herrera, F., Cordon, O., del Jesus, M., & Marcelloni, F. (2019). Evolutionary fuzzy systems for explainable artificial intelligence: Why, when, what for, and where to? *IEEE Computational Intelligence Magazine*, *14*(1), 69–81. https://doi.org/10.1109/MCI.2018.2881645

Fiorani, F., & Schurr, U. (2013). Future scenarios for plant phenotyping. *Annual Review of Plant Biology*, *64*(1), 267–291. https://doi.org/10.1146/annurev-arplant-050312–120137

Fredriksson, M. (2021). Dilemmas of protection: Decolonising the regulation of genetic resources as cultural heritage. *International Journal of Heritage Studies*, *27*(7), 720–733. https://doi.org/10.1080/13527258.2020.1852295

GA. (2023). agreement under the United Nations Convention on the Law of the Sea on the conservation and sustainable use of marine biological diversity of areas beyond national jurisdiction, UN Doc. A/CONF.232/2023/4 (19 June 2023, open for signature 20 September 2023). Geneva: Secretariat of the General Assembly of the United Nations (BBNJ Agreement).

Gilpin, L., Bau, D., Yuan, B., Bajwa, A., Specter, M., & Kagal, L. (2018). Explaining explanations: An overview of interpretability of machine learning. In *Proceedings of the 2018 IEEE 5th international conference on data science and advanced analytics* (pp. 80–89). IEEE. https://doi.org/10.1109/DSAA.2018.000

Göndöcs, D., & Dörfler, V. (2024). AI in medical diagnosis: AI prediction and human judgment. *Artificial Intelligence in Medicine*, *149*, 102769. https://doi.org/10.1016/j.artmed.2024.102769

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys*, *51*(5), 1–42. https://doi.org/10.1145/3236009

Halewood, M., Bagley, M., Wyss, M., & Scholz, A. (2023). New benefit-sharing principles for digital sequence information. *Science*, *382*(6670), 520–522. https://doi.org/10.1126/science.adj1331

Hampton, A. (2023). Pathogen dematerialization and the ABS loophole. *Journal of Law and the Biosciences*, *10*(1), lsad002. https://doi.org/10.1093/jlb/lsad002

Hampton, A., Eccleston-Turner, M., Rourke, M., & Switzer, S. (2023). 'Equity' in the pandemic treaty: The false hope of 'access and benefit-sharing'. *International and Comparative Law Quarterly*, *72*(4), 909–943. https://doi.org/10.1017/S0020589323000350

Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M., & Hussain, A. (2024). Interpreting black-box models: A review on explainable artificial intelligence. *Cognitive Computation*, *16*(1), 45–74. https://doi.org/10.1007/s12559-023-10179-8

Haupt, S., Gagne, D., Hsieh, W., Krasnopolsky, V., McGovern, A., Marzban, C., Moninger, W., Lakshmanan, V., Tissot, P., & Williams, J. (2022). The history and practice of AI in the environmental sciences. *Bulletin of the American Meteorological Society*, *103*(5), E1351–E1370. https://doi.org/10.1175/BAMS-D-20–0234.1

Heller, M., & Eisenberg, R. (1998). Can patents deter innovation? The anticommons in biomedical research. *Science*, *280*(5364), 698–701. https://doi.org/10.1126/science.280.5364.698

Hesami, M., Alizadeh, M., Jones, A., & Torkamaneh, D. (2022). Machine learning: Its challenges and opportunities in plant system biology. *Applied Microbiology and Biotechnology*, *106*(9), 3507–3530. https://doi.org/10.1007/s00253-022-11963-6

Holzinger, A., Keiblinger, K., Holub, P., & Heimo Müller, H. (2023). AI for life: Trends in artificial intelligence for biotechnology. *Nature Biotechnology*, *74*, 16–24. https://doi.org/10.1016/j.nbt.2023.02.001

Holzinger, A., Langs, G., Denk, H., Zatloukal, K., & Müller, H. (2019). Causability and explainability of artificial intelligence in medicine. *Wiley*

Interdisciplinary Reviews: Data Mining and Knowledge Discovery, *9*(4), e1312. https://doi.org/10.1002/widm.1312

House of Lords. (2018) Ai in the UK: ready, willing and able? Report of Session 2017–19. Select Lindon: Committee on Artificial Intelligence https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf

Humphries, F. (2024). *Decoding marine genetic resource governance under the BBNJ agreement*. Springer. https://doi.org/10.1007/978-3-031-72100-7

Humphries, F., Gottlieb, H., Laird, S., Wynberg, R., Lawson, C., Rourke, M., Tvedt, M., Oliva, M., & Jaspars, M. (2020). A tiered approach to the marine genetic resource governance framework under the proposed UNCLOS agreement for biodiversity beyond national jurisdiction (BBNJ). *Marine Policy*, *122*, 103910. https://doi.org/10.1016/j.marpol.2020.103910

Humphries, F., Rabone, M., & Jaspars, M. (2021). Traceability approaches for marine genetic resources under the proposed ocean (BBNJ) treaty. *Frontiers in Marine Science*, *8*, 661313. https://doi.org/10.3389/fmars.2021.661313

Kim, H. (2019). AI, big data, and robots for the evolution of biotechnology. *Genomics and Informatics*, *17*(4), e44. https://doi.org/10.5808/GI.2019.17.4.e44

Kuznietsov, A., Gyevnar, B., Wang, C., Peters, S., and Albrecht, S. (2024). Explainable AI for safe and trustworthy autonomous driving: A systematic review. arXiv preprint arXiv:2402.10086. https://arxiv.org/pdf/2402.10086

Laird, S., Wynberg, R., Rourke, M., Humphries, F., Muller, M., & Lawson, C. (2020). Rethink the expansion of access and benefit sharing. *Science*, *367*(6483), 1200–1202. https://doi.org/10.1126/science.aba9609

Lawson, C. (2012). *Regulating genetic resources: Access and benefit-sharing in international law*. Edward Elgar. https://doi.org/10.4337/9781781001660

Lawson, C. (2017). Relationship between the work of the WIPO IGC and other forums. In D. Robinson, A. Abdel-Latif, & P. Roffe (Eds.), *Protecting traditional knowledge: The WIPO intergovernmental committee on intellectual property and genetic resources, traditional knowledge and folklore* (pp. 56–82). Routledge.

Lawson, C., Humphries, F., & Rourke, M. (2019). The future of information under the CBD, Nagoya Protocol, Plant Treaty, and PIP Framework. *Journal of World Intellectual Property*, *22*(3–4), 103–119. https://doi.org/10.1111/jwip.12118

Lawson, C., Humphries, F., & Rourke, M. (2023). Genetic resources as culture and heritage: Repatriation and benefit sharing. *Melbourne Journal of International Law*, *24*(1), 27–53.

Lawson, C., Humphries, F., & Rourke, M. (2024). Challenging the existing order of knowledge sharing governance with digital sequence information on genetic resources. *Journal of Intellectual Property Law and Practice*, *19*(4), 337–357. https://doi.org/10.1093/jiplp/jpad129

Lawson, C., Rourke, M., & Humphries, F. (2020). Information as the latest site of conflict in the ongoing contests about access to and sharing the benefits from exploiting genetic resources. *Queen Mary Journal of Intellectual Property*, *10*(1), 7–33. https://doi.org/10.4337/qmjip.2020.01.01

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Leskien, D. (2023). Accessing biological control genetic resources and sharing benefits resulting from their utilization. *BioControl*, *68*(3), 225–233. https://doi.org/10.1007/s10526-023-10193-7

Li, X., Xiong, H., Li, X., Wu, X., Zhang, X., Liu, J., Bian, J., & Dou, D. (2022). Interpretable deep learning: Interpretation, interpretability, trustworthiness, and beyond. *Knowledge and Information Systems*, *64*(12), 3197–3234. https://doi.org/10.1007/s10115-022-01756-8

Lin, E., Lin, C.-H., & Lane, H.-Y. (2020). Precision psychiatry applications with pharmacogenomics: Artificial intelligence and machine learning approaches. *International Journal of Molecular Sciences*, *21*(3), 969. https://doi.org/10.3390/ijms21030969

Lin, J., & Ngiam, K. (2023). How data science and AI-based technologies impact genomics. *Singapore Medical Journal*, *64*(1), 59–66. https://doi.org/10.4103/singaporemedj.SMJ-2021-438

Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2020). Explainable AI: A review of machine learning interpretability methods. *Entropy*, *23*(1), 18. https://doi.org/10.3390/e23010018

Lipton, Z. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, *16*(3), 31–57. https://doi.org/10.1145/3236386.3241340

Lisboa, P. J., Saralajew, S., Vellido, A., Fernández-Domenech, R., & Villmann, T. (2023). The coming of age of interpretable and explainable machine learning models. *Neurocomputing*, *535*, 25–39. https://doi.org/10.1016/j.neucom.2023.02.040

MacDonald, H. (2025). Conceptual uncertainties and practical challenges in voluntary Nagoya Protocol compliance: The Australian situation. *Biopreservation and Biobanking*, in press. https://doi.org/10.1089/bio.2024.0090

Marcus, G. (2018). Deep learning: A critical appraisal. arXiv preprint arXiv:1801.00631. https://arxiv.org/pdf/1801.00631

McCarthy, J., Minsky, M., Rochester, N., & Shannon, C. (2006). A proposal for the Dartmouth summer research project on artificial intelligence. *AI Magazine*, *27*(4), 12. https://doi.org/10.1609/aimag.v27i4.1904

McCorduck, P., & Cfe, C. (2004). *Machines who think: A personal inquiry into the history and prospects of artificial intelligence*. AK Peters/CRC Press. https://doi.org/10.1201/9780429258985

McGregor Richmond, K., Muddamsetty, S., Gammeltoft-Hansen, T., Olsen, H., & Moeslund, T. (2024). Explainable AI and law: An evidential survey. *Digital Society*, *3*(1), 1. https://doi.org/10.1007/s44206-023-00081-z

Michiels, F., Feiter, U., Paquin-Jaloux, S., Jungmann, D., Braun, A., Sayoc, M., Armengol, R., Wyss, M., & David, B. (2021). Facing the harsh reality of access and benefit sharing (ABS) legislation: An industry perspective. *Sustainability*, *14*(1), 277. https://doi.org/10.3390/su14010277

Molnar, C. (2020). *Interpretable machine learning: A guide for making black box models explainable*. Lean Publishing.

Montavon, G., Samek, W., & Müller, K. (2018). Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, *73*, 1–15. https://doi.org/10.1016/j.dsp.2017.10.011

Mueller, S., Hoffman, R., Clancey, W., Emrey, A., and Klein, G. (2019). Explanation in human-AI systems: A literature meta-review, synopsis of key ideas and publications, and bibliography for explainable AI. arXiv preprint arXiv:1902.01876. https://arxiv.org/pdf/1902.01876

Mulatinho Simoes, R., & Birchfield, V. (2024). Biodiversity and the digital transformation: Rethinking private property and global governance in the twenty-first century. *Environmental Ethics*, *46*(1), 47–69. https://doi.org/10.5840/enviroethics202422270

Mund, A., Coscia, F., Hollandi, R., Kovács, F., Kriston, A., Brunner, A., Bzorek, M., Naimy, S., Rahbek Gjerdrum, L., Dyring-Andersen, B., Bulkescher, J. (2021). Ai-driven deep visual proteomics defines cell identity and heterogeneity. arXiv preprint arXiv:2021–01. https://doi.org/10.1101/2021.01.25.427969

Nallakaruppan, M., Balusamy, B., Shri, M., Malathi, V., & Bhattacharyya, S. (2024). An explainable AI framework for credit evaluation and analysis. *Applied Soft Computing*, *153*, 111307. https://doi.org/10.1016/j.asoc.2024.111307

Nolff, M. (2024). Finally! WIPO treaty on disclosure of origin or source of genetic resources and associated traditional knowledge in patent applications: The end or just the beginning? *Journal of the Patent and Trademark Office Society*, *104*(3), 397–414.

Oliveira, A. (2019). Biotechnology, big data and artificial intelligence. *Biotechnology Journal*, *14*(8), 1800613. https://doi.org/10.1002/biot.201800613

Petrick, L., & Shomron, N. (2022). Ai/ml-driven advances in untargeted metabolomics and exposomics for biomedical applications. *Cell Reports Physical Science*, *3*(7), 1–16. https://doi.org/10.1016/j.xcrp.2022.100978

Plebe, A., & Grasso, G. (2019). The unbearable shallow understanding of deep learning. *Minds and Machines*, *29*(4), 515–553. https://doi.org/10.1007/s11023-019-09512-8

Prathapan, K., Pethiyagoda, R., Bawa, K., Raven, P., Rajan, P., & 172 Co-Signatories. (2018). When the cure kills – CBD limits biodiversity research. *Science*, *360*(6396), 1405–1406. https://doi.org/10.1126/science.aat9844

Preece, A. (2018). Asking 'why' in AI: Explainability of intelligent systems – Perspectives and challenges. *Intelligent Systems in Accounting, Finance and Management*, *25*(2), 63–72. https://doi.org/10.1002/isaf.1422

Reel, P., Reel, S., Pearson, E., Trucco, E., & Jefferson, E. (2021). Using machine learning approaches for multi-omics data analysis: A review. *Biotechnology Advances*, *49*, 107739. https://doi.org/10.1016/j.biotechadv.2021.107739

Ribeiro, M., Singh, S. and Guestrin, C. (2016). Why should I trust you? Explaining the predictions of any classifier. in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016, 1135–1144. https://doi.org/10.1145/2939672.2939778

Ribera, M. and Lapedriza, A. (2019). Can we do better explanations? A proposal of user-centered explainable AI. In Joint Proceedings of the ACM IUI 2019 Workshops, Los Angeles, USA, 20 March 2019, 2327, 38. https://openaccess.uoc.edu/bitstream/10609/99643/1/explainable_AI.pdf

Richerzhagen, C. (2014). The Nagoya Protocol: Fragmentation or consolidation? *Resources*, *3*(1), 135–151. https://doi.org/10.3390/resources3010135

Roche-Lima, A., Roman-Santiago, A., Feliu-Maldonado, R., Rodriguez-Maldonado, J., Nieves-Rodriguez, B. G., Carrasquillo-Carrion, K., Ramos, C., da Luz Sant'Ana, I., Massey, S., & Duconge, J. (2020). Machine learning algorithm for predicting warfarin dose in Caribbean Hispanics using pharmacogenetic data. *Frontiers in Pharmacology*, *10*, 1550. https://doi.org/10.3389/fphar.2019.01550

Rourke, M. (2024). One health and pathogen sharing: What's missing in the pandemic treaty's proposed pathogen access and benefit-sharing (PABS) system? Available at SSRN: https://doi.org/10.2139/ssrn.4965622

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, *1*(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Rudin, C., Chen, C., Chen, Z., Huang, H., Semenova, L., and Zhong, C. (2022). Interpretable machine learning: fundamental principles and 10 grand challenges. arXiv preprint arXiv: 2103.11251. https://doi.org/10.48550/arXiv.2103.11251

Sah, S. (2020). Machine learning: A review of learning types. Preprint 202007.0230. https://doi.org/10.20944/preprints202007.0230.v1

Samek, W., Montavon, G., Lapuschkin, S., Anders, C., & Müller, K. (2021). Explaining deep neural networks and beyond: A review of methods and applications. *Proceedings of the IEEE*, *109*(3), 247–278. https://doi.org/10.1109/JPROC.2021.3060483

Saranya, A., & Subhashini, R. (2023). A systematic review of explainable artificial intelligence models and applications: Recent developments and future trends. *Decision Analytics Journal*, *7*, 100230. https://doi.org/10.1016/j.dajour.2023.100230

Sarid, E., & Ben-Zvi, O. (2023). A theoretical analysis of orphan works. *Cardozo Arts & Entertainment Law Journal*, *40*(3), 585–619.

Scholz, A., Nunez-Vega, G., Weissgold, L., & Wussmann, K. (2023). The future of access and benefit-sharing: What next after the adoption of the global biodiversity framework and decision on digital sequence information? *Diversity*, *16*(1), 27. https://doi.org/10.3390/d16010027

Segessenmann, J., Stadelmann, T., Davison, A., and Dürr, O. (2023). Assessing deep learning: A work program for the humanities in the age of artificial intelligence. AI and Ethics, Epub ahead of print 21 December 2023. https://doi.org/10.1007/s43681-023-00408-z

Sellwood, M., Ahmed, M., Segler, M., & Brown, N. (2018). Artificial intelligence in drug discovery. *Future Medicinal Chemistry*, *10*(17), 2025–2028. https://doi.org/10.4155/fmc-2018–0212

Sheh, R., & Monteath, I. (2018). Defining explainable AI for requirements analysis. *KI-Künstliche Intelligenz*, *32*(4), 261–266. https://doi.org/10.1007/s13218-018-0559-3

Sherman, B., & Henry, R. (2020). The Nagoya Protocol and historical collections of plants. *Nature Plants*, *6*(5), 430–432. https://doi.org/10.1038/s41477-020-0657-8

Svarstad, H. (1994). National sovereignty and genetic resources. In V. Sánchez & C. Juma (Eds.), *Biodiplomacy: Genetic resources and international relations* (pp. 45–65). ACTS Press. https://idl-bnc-idrc.dspacedirect.org/server/api/core/bitstreams/2817a950-e0ae-47b1-a715-394071a4d358/content

Szabadföldi, I. (2021). Artificial intelligence in military application: Opportunities and challenges. *Land Forces Academy Review*, *26*(2), 157–165. https://doi.org/10.2478/raft-2021–0022

Tekkesinoglu, S., Habibovic, A. and Kunze, L. (2024). Advancing explainable autonomous vehicle systems: A comprehensive review and research roadmap. arXiv preprint arXiv:2404.00019. https://arxiv.org/pdf/2404.00019

Tester, M., & Langridge, P. (2010). Breeding technologies to increase crop production in a changing world. *Science*, *327*(5967), 818–822. https://doi.org/10.1126/science.1183700

Tsimenidis, S. (2020). Limitations of deep neural networks: A discussion of G. Marcus' critical appraisal of deep learning. arXiv preprint arXiv: 2012.15754. https://arxiv.org/pdf/2012.15754

Turing, A. (1950). Computing machinery and intelligence. *Mind*, *59*(236), 433–460. https://doi.org/10.1093/mind/LIX.236.433

UNDRIP. (2007). United Nations declaration on the rights of indigenous peoples, 13 September 2007 GA res 61/295, 61st sess, 107th Plen MTG, agenda item 68, un doc, a/res/61/295 (2 October 2007, adopted 13 September 2007) (undrip)

UNESCO. (1970). Convention on the means of prohibiting and preventing the illicit import, export and transfer of ownership of cultural property, opened for signature 14 November 1970, 823 U.N.T.S. 321 (entered into force 24 April 1972) (UNESCO Convention).

Vale, D., El-Sharif, A., & Ali, M. (2022). Explainable artificial intelligence (XAI) post-hoc explainability methods: Risks and limitations in non-discrimination law. *AI and Ethics*, *2*(4), 815–826. https://doi.org/10.1007/s43681-022-00142-y

van der Lee, M., & Swen, J. (2023). Artificial intelligence in pharmacology research and practice. *Clinical and Translational Science*, *16*(1), 31–36. https://doi.org/10.1111/cts.13431

Vassiliades, A., Bassiliades, N., & Patkos, T. (2021). Argumentation and explainable artificial intelligence: A survey. *Knowledge Engineering Review*, *36*, e5. https://doi.org/10.1017/S0269888921000011

Verhagen, R., Neerincx, M., & Tielman, M. (2021). A two-dimensional explanation framework to classify AI as incomprehensible, interpretable, or understandable. In D. Calvaresi, A. Najjar, M. Winikoff, & K. Främling (Eds.), *Explainable and transparent AI and multi-agent systems: Third international workshop, EXTRAAMAS 2021, virtual event, 3–7 May 2021, revised selected papers* (pp. 119–138). Springer International Publishing. https://doi.org/10.1007/978-3-030-82017-6_8

Vilone, G. and Longo, L. (2020). Explainable artificial intelligence: A systematic review. arXiv preprint arXiv:2006.00093. https://arxiv.org/pdf/2006.00093

WHO. (2011). World health organisation, pandemic influenza preparedness: Sharing of influenza viruses and access to vaccines and other benefits, sixty-fourth world health assembly. UN Doc. wha64.5. Geneva: World Health Organisation (pip framework)

Widrow, B., & Lehr, M. (1990). 30 years of adaptive neural networks: Perceptron, madaline, and backpropagation. *Proceedings of the IEEE*, *78*(9), 1415–1442. https://doi.org/10.1109/5.58323

WIPO. (2024). Diplomatic conference to conclude an international legal instrument relating to intellectual property, genetic resources and traditional knowledge associated with genetic resources, WIPO treaty on intellectual property, genetic resources and associated traditional knowledge. UN Doc. GRATK/DC/7. Geneva: World Intellectual Property Organisation.

Working Group on Multilateral System. (2024a). Eleventh meeting of the ad hoc open-ended working group to enhance the functioning of the multilateral system. UN Doc. IT/OWG-EFMLS-11/24/Report. Rome: Secretariat of the International Treaty of Plant Genetic Resources for Food and Agriculture.

Working Group on Multilateral System. (2024b). Twelfth meeting of the ad hoc open-ended working group to enhance the functioning of the multilateral system. UN Doc. IT/OWG-EFMLS-12/24/Report. Rome: Secretariat of the International Treaty of Plant Genetic Resources for Food and Agriculture.

WTO. (1994). Agreement on trade related aspects of intellectual property rights, opened for signature 15 April 1994, 1867 U.N.T.S. 154 (entered into force 1 January 1995) (TRIPS)

Wynberg, R. (2023). Biopiracy: Crying wolf or a lever for equity and conservation? *Research Policy*, *52*(2), 104674. https://doi.org/10.1016/j.respol.2022.104674

Xiao, Q., Zhang, F., Xu, L., Yue, L., Kon, O. L., Zhu, Y., & Guo, T. (2021). High-throughput proteomics and AI for cancer biomarker discovery. *Advanced Drug Delivery Reviews*, *176*, 113844. https://doi.org/10.1016/j.addr.2021.113844

Yoosefzadeh Najafabadi, M., Hesami, M., & Eskandari, M. (2023). Machine learning-assisted approaches in modernized plant breeding programs. *Genes*, *14*(4), 777. https://doi.org/10.3390/genes14040777

Yu, S. (2023). Towards trustworthy and understandable AI: Unraveling explainability strategies on simplifying algorithms, appropriate information disclosure, and high-level collaboration. In 26th International Academic Mindtrek Conference (Mindtrek'23), 3–6 October 2023, Tampere, Finland, 133–143. https://doi.org/10.1145/3616961.3616965

How to cite this article: Lawson, C., Englezos, E., & Rourke, M. (2025). Access and benefit sharing biological materials for machines: Artificial intelligence, machine learning and deep learning. *Plants, People, Planet*, 1–13. https://doi.org/10.1002/ppp3.70007