

Non-fungible tokens as a framework for sustainable innovation in pharmaceutical R&D: a smart contract-based platform for data sharing and rightsholder protection

Marcelo Corrales Compagnucci, Niclas Nilsson, Paul Stankovski Wagner, Christoffer Olsson, Mark Fenwick, Timo Minssen & Kacper Szkalej

To cite this article: Marcelo Corrales Compagnucci, Niclas Nilsson, Paul Stankovski Wagner, Christoffer Olsson, Mark Fenwick, Timo Minssen & Kacper Szkalej (2024) Non-fungible tokens as a framework for sustainable innovation in pharmaceutical R&D: a smart contract-based platform for data sharing and rightsholder protection, International Review of Law, Computers & Technology, 38:1, 66-85, DOI: [10.1080/13600869.2023.2233803](https://doi.org/10.1080/13600869.2023.2233803)

To link to this article: <https://doi.org/10.1080/13600869.2023.2233803>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 20 Jul 2023.



[Submit your article to this journal](#)



Article views: 720




[View related articles](#)



[View Crossmark data](#)

Non-fungible tokens as a framework for sustainable innovation in pharmaceutical R&D: a smart contract-based platform for data sharing and rightsholder protection

Marcelo Corrales Compagnucci ^a, Niclas Nilsson^b, Paul Stankovski Wagner^c, Christoffer Olsson^d, Mark Fenwick^e, Timo Minssen^a and Kacper Szkalej^{f,g}

^aCenter for Advanced Studies in Biomedical Innovation Law (CeBIL), University of Copenhagen, Copenhagen, Denmark; ^bLU Innovation, Lund University, Lund, Sweden; ^cDepartment of Electrical and Information Technology, Faculty of Engineering (LTH), Lund University, Lund, Sweden; ^dXenit AB, Gothenburg, Sweden; ^eGraduate School of Law, Kyushu University, Fukuoka, Japan; ^fCenter for Business Law (ACLU), Faculty of Law, Lund University, Lund, Sweden; ^gLaw and Informatics Research Institute (IRI), Faculty of Law, Stockholm University, Stockholm, Sweden

ABSTRACT

Research and development (R&D) in the pharmaceutical sector traditionally operated with in closed, siloed institutional settings, driven by intellectual property rights concerns that viewed data sharing as a threat. However, the evolving scientific landscape demands a more collaborative approach involving external engagement and dynamic partnerships. To address this, a hybrid contractual framework combining smart contracts, non-fungible tokens (NFTs), and traditional licensing schemes is introduced. This framework was developed through an experimental pilot platform that adhered to FAIR data principles, allowing participants to store, find, and reuse data related to drug discovery. The platform utilizes blockchain technology to document real-world assets in an immutable digital ledger. Smart contracts and NFTs offer an open and global collaborative platform for advancing drug research assets, overcoming hurdles related to standardization, interoperability, and disclosure. This framework aims to reconcile the conflict between the demand for greater data sharing and the protection of rightsholder interests in pharmaceutical R&D. By providing mechanisms for resolving practical challenges, it facilitates further cooperation and innovation in the field.

ARTICLE HISTORY



Received 22 December 2022
Accepted 4 July 2023

KEYWORDS

Non-fungible token (NFT);
molecule; smart contract

1. Introduction

Innovation in the pharmaceutical sector has always been built on explorative collaboration. Nothing has demonstrated the need for increased collaboration to address the grand challenges of our times better than the recent COVID 2019 pandemic (Druehdahl,

CONTACT Marcelo Corrales Compagnucci  marcelo.corrales13@gmail.com  Center for Advanced Studies in Biomedical Innovation Law (CeBIL), University of Copenhagen, Karen Blixens Plads 16, DK-2300, Copenhagen, Denmark
This article has been corrected with minor changes. These changes do not impact the academic content of the article.

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

Minssen, and Price 2021). In the past, however, research and development (R&D) typically occurred in closed, siloed institutional settings. This approach was a function of a rights-based model that framed access and reuse of data (data-sharing) as an unacceptable threat to the interests of rightsholders. However, such a closed protective approach to collaboration may be ill-suited to the more complex scientific ecosystems of today, where external engagement and dynamic partnering with multiple actors and diverse information sources has become essential.

This paper focuses on intellectual property rights (IPRs) and not on privacy and data protection since there were no patients or participants involved in the project. However, it is easy to foresee how the General Data Protection Regulation (GDPR)¹ would add further issues depending on the setting if personal or sensitive data are involved.

IPRs have always been considered a crucial tool to recoup the investment of authors, inventors and researchers (Corrales Compagnucci et al. 2010).

However, an overly protectionist mindset potentially created by IPRs, risks creating tensions with open science, open source, open data, and open innovation. As such, legal interoperability creates a bottleneck in a contemporary context, and the risk of persisting with a closed, rights-oriented model is that opportunities are lost, and innovation stifled (Corrales Compagnucci 2020, 1, 2, 270, 271). The constant pressure to translate existing research into further health innovations, the exponential increase in relevant information, new platform tech and the complex network of researchers dispersed across the world add to the challenges facing the life science industry. With multiple entities scattered across the distribution chain performing different roles, traditional models of collaboration need re-visiting and improving (Kessel 2011).

In addition to thinking about more traditional IP governance tools, such as IP ordering and licensing mechanisms in the form of IP pools and clearing houses, there is a pressing need for a framework that enables a more practical, efficient, and transparent mode of sharing data and assets.

This article introduces a possible future model in the form of a hybrid contractual framework that combines the benefits of the automated functionality of smart contracts² and non-fungible tokens (NFTs)³ embedded in a blockchain with more traditional rights-based licensing schemes. The presented platform enables participants to store, find and reuse data following the FAIR data principles and Open Data.

The platform documents scientific contributions in an immutable ledger as part of a digitalization process tracking real-world physical assets in the form of the drug discovery of chemical molecules. More specifically, researchers all over the world can share and verify information about chemical molecules at different levels for commercial and non-commercial purposes. This data can then be used to document IPRs and potentially be developed towards an alternative patent process where both transparency and full recognition of the respective inventor/creator and all contributors is publicly achieved. Crucially, such a platform creates a complex, collaborative graph that precisely acknowledges all contributors and their contribution at any point in time.

The claim developed here is that blockchain technology, smart contracts, and NFTs can provide an effective and powerful alternative for digitized research assets. The decentralized database works as a digital and immutable ledger which stores information regarding who has done what and when and also who owns what (Corrales Compagnucci, Fenwick, and Haapio 2019). In other words, blockchain technology contributes to more

control over data, assets and identities while being more open and transparent, thus providing a technical platform for more open and effective collaboration (Tsarsitalidis et al. 2021; Corrales et al. 2019). As such, this model overcomes many of the practical hurdles currently obstructing efficient global collaboration in pharmaceutical R&D, as well as providing a framework to address the core conflict between the simultaneous demand for data sharing and the protection of rightsholder interests in drug discovery.

After this Introduction, Section 2 describes the limitations of current research collaboration practices and suggests that whilst open innovation models represent an improvement, they fail to institutionalize processes and mechanisms that facilitate effective collaboration over time involving multiple actors. Section 3 introduces the hybrid contractual framework which combines the benefits of the automated functionality of blockchain, smart contracts and NFTs. This model is based on the outcome of an experimental project which was developed as a blockchain-based platform (VINNOVA project 2019; Olsson and Toorani 2021). Section 4 introduces the FAIR Principles of collaborative research. Section 5 examines how licensing would operate under this scheme. Section 6 concludes with a more general discussion of likely future trends in this space.

2. The limits of traditional models of explorative collaboration and the search for alternatives

There is a growing recognition that a closed R&D model no longer functions effectively and that the world needs to develop new ways to ensure access to and development of new medicines ('Final Report – High-Level Panel on Access to Medicines' 2016). Developing new drugs addressing unmet medical needs is a key factor for pharmaceutical research, but the traditional R&D model is failing to deliver sustainable innovation, and this requires a fundamental change to the entire ecosystem (Schuhmacher, Gassmann, and Hinder 2016). Pharmaceutical R&D is increasingly global with more and more stakeholders seeking and relying on external partners for expertise, access to advanced technology and the acquisition of new research assets. For example, a recent study of cancer drugs spanning ten years found that as the price of novel oncology drugs increased, better clinical effects did not increase in a similar manner, indicating an unsustainable cost of development and a reduced level of innovation (Saluja et al. 2018).

One solution would appear to be a more open collaborative model in which assets move more freely within the ecosystem. Such an approach has already shown improved efficiency compared to traditional closed models (Deloitte US 2017).

One of the problems that needs to be overcome in this field is that the system is built on a robust framework of IPRs, and that the character protection that this framework affords creators often hinders collaboration. There is of course a well-defined need and purpose in facilitating a return on investment in R&D by patenting innovations. But when applied in a broad and unspecified manner some of the rigorous mechanisms that protect IPRs can prove detrimental to collaboration and the overall progression of science by creating barriers and transaction costs. Such barriers can impede collaboration and stifle innovation that requires joint efforts.

Joint ventures and cooperation between pharma companies in consortia have been identified as a possible strategy to increase innovation by being more open without risking business (Olk and West 2020). A scientific and innovative business environment

must keep open science and proprietary systems in balance ('Commentary: The Coronavirus Pandemic Has Shattered the Status Quo on Drug Development. We Should Build on That | Fortune' 2020).

The recent COVID-19 pandemic has also triggered calls for greater openness and sharing of both resources and knowledge. One example is the sharing of genomic data of the SARS-CoV-2 virus to speed up research, particularly about different global virus strains and variants, but some sharing platforms restrict the redistribution and use of such data as a way to control the data providers' rights (Van Noorden 2021).

Progress in pharmaceutical R&D comes from utilizing general knowledge (e.g. databases, articles, and other information) and combining that with internal specific know-how and capabilities. As more R&D is completed, it requires more data, insights, and knowledge, as well as access to specialized technical capabilities, and collaboration is needed to progress from early science to treatments.

The need, therefore, arises for systems and mechanisms to facilitate collaboration. One concrete example is the public-private consortium 'Open Target' which provides an open platform for disease and drug target exploration ('Open Targets: New Name, New Data | EMBL-EBI' 2016). Traditionally, data used to identify novel drug targets has been considered proprietary as it is business critical for drug development companies. However, it has also become exponentially more difficult to identify novel drug targets, making it more likely with increasing amounts of data both to explore opportunities as well as mitigate risk. And, as the identification of a drug target is not a patentable innovation *per se*, there is less business risk and more to gain from a scientific perspective. Instead, most patents involve the active pharmaceutical ingredients to protect the new chemical entity (NCE) or new medical entity (NME) which modified the biological function of a drug target.

The challenges, benefits, and risks of research collaboration are represented in Table 1.

There is a general consensus that working more with external partners is beneficial for R&D and innovation. We already see an increase of the proportion of externally sourced projects in the pharma pipelines and the use of contract research organizations (CRO) as extended resources and technical expertise is now part of any pharma or biotech's business (Deloitte Centre for Health Solutions 2022).

Table 1. The challenges, benefits and risks of research collaboration.

Challenge	Benefits	Risks & Problems	Example
Transparency	Better, more agile collaboration	Undermines IPR strategies and prior art	Disclosure of drug discovery molecules, structure and/or functionality
Commercial Interests	Identifying best partners & minimizing mistakes	Reduce competitive advantage	Interest in specific drug target
Sharing Data	Joint efforts in solving complex scientific problems	Unclear or no standard licensing mechanisms and reduced 'control' of open data	Facilitate early research ideas
Incentivizing & Rewarding Collaboration	Incentivizing traceability and reward and acknowledgement systems	Inadequate recognition of who has contributed what and when	Building on, or adding to third party data
Promoting Institutional Collaboration	Synergy in research infrastructure and capabilities	A lack of technical, legal, organizational, and semantic interoperability	Connecting databases, user agreements and technical interfaces

Fast-phased technical advancements, exponentially increasing amounts of data and the necessity of deeper disease understanding are examples of factors pushing towards both strategic and tactical need to improve the interface with external partners.

‘Openness’ has become a buzzword, often overused and misunderstood. Nevertheless, at the core it provides an absolute shift in how collaboration can work. By disclosing needs or capabilities, external parties can identify value creation and synergistic opportunities. Also, by allowing external participants insight and empowerment, research goals can be collectively developed and transparent (Balasegaram et al. 2017). Applied open science, open source, open access and open data in FAIR principles (Wilkinson et al. 2016) enable open innovation by focusing on the needs, data, and research instead of protectionism and traditional business restraints. This applies to early drug discovery, where scientific exploration and collaborations are paramount to generate new research and treatment initiatives. As many institutions explore more open collaboration models, the actual implementations vary to a great extent, including, in particular, the contractual details.

By reducing the transaction costs (Coase 1960) of signing an explorative research agreement and being more transparent with either, or both, what assets parties have or are seeking it is much easier for external partners to identify and assess a scientific overlap and potential collaboration. This open innovation model can be a complement to more traditional research collaborations and business development activities and provide a fast and much more simple way of engaging externally (Nilsson and Felding 2015). Several such collaboration platforms exist, and although the level of openness or transparency varies to a great extent, they all share the ambition to reach out to external partners by providing a different model than traditionally. This could be by not making any claims on commercialization rights or controlling the business process, waiving all or part of the IPRs, or disclosing scientific know-how and business-relevant information (such as a drug target of interest).

Currently such open innovation platforms are designed from their own perspective, implement different aspects of openness, and have no standard way of handling the open innovation agreements, assets, data, and resources. As such, interoperability is low and there is no overarching platform or mechanism to connect individual platforms as part of a greater ecosystem. One significant issue is the lack of common license agreements that can be readily recognized and implemented as well as incentives to share beyond the necessary and immediate scope of a collaboration (Nilsson and Minssen 2018).

The schematic in Figure 1 illustrates the three different models of collaboration. The blue and green dots represent a biotech or university whilst an arrow represents a larger pharmaceutical company. The traditional method with closed borders requires significant efforts to spark collaboration with external partners, and this method seems ill-suited to contemporary conditions. Implementing an open research platform makes it much easier for both partners to explore mutual opportunities, but bottlenecks occur around intermediaries. Finally, with a dedicated collaboration platform multiple organizations can explore, share, and collaborate more efficiently, increasing chances of match-making and joint efforts to generate value where needed. The collaborative platform is depicted in gray and spans across individual organizations. The aim of this article is to offer a workable example of the Model C-type collaborative research.

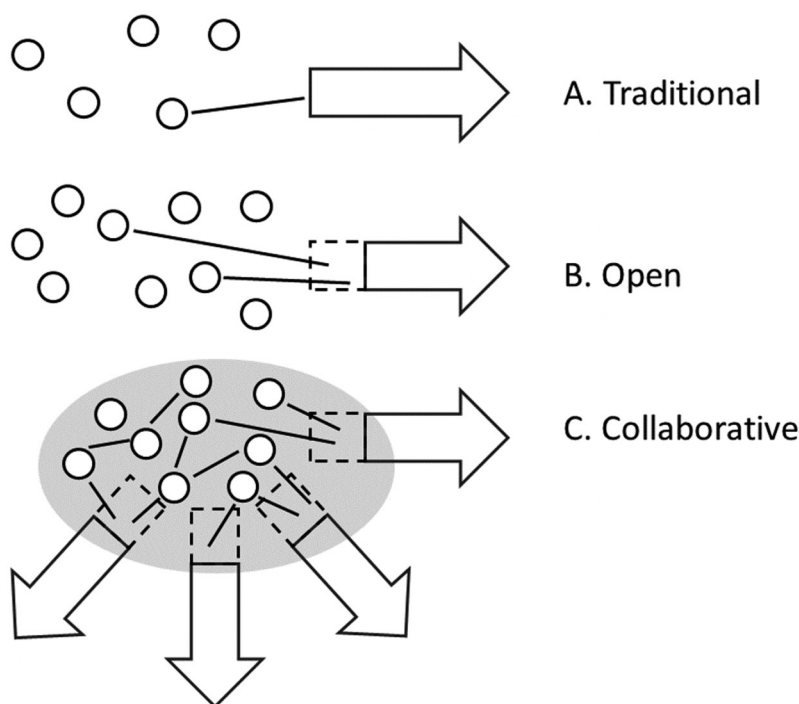


Figure 1. Three models of collaborative research.

3. A blockchain-powered alternative

3.1. New chemical entities in pharmaceutical R&D

Many existing treatments are based on chemical molecules as the active pharmaceutical ingredient, although there are many alternative modalities such as biologicals, peptides, oligonucleotides, genes and cells, and significant investments have been, and still are, made towards the development of new chemical entities (NCE). During the development of an NCE it is not uncommon that thousands or millions of molecules are utilized as part of the process and depending on the origin, nature, and opportunity, these molecules are either kept confidential or patented (Mayr and Bojanic 2009). Still, some molecules are developed simply as a tool or reference molecule to study biological and disease mechanisms mostly in academic settings, but also in open constellations where pharma contributes, e.g. the Structural Genomics Consortium (SGC) ('SGC | Structural Genomics Consortium' n.d.).

In other instances, attempts to improve collaborative innovation by sharing goals, resources and data with full openness and dedicated non-patentable outputs have been explored by e.g. Open-Source Malaria (OSM) ('Open-Source Malaria: OSM,' n.d.). Even though it can greatly boost creativity and collaboration, it is often difficult to combine with traditional business processes, making it a struggle with funding, especially as the pharma process becomes more expensive as it moves towards clinical trials. Sharing of molecules at an early stage can, therefore, significantly boost collaboration and innovation, but such information will become 'prior art' when disclosed to the

public, making it impossible to provide undisputable documentation of who has done what and when. This kind of openness is poorly compatible with commercial endeavors that rely on securing IPR traditionally relevant for biotechnological inventions (Saha and Bhattacharya 2011).

As part of drug discovery projects, internal chemistry can be initiated by testing external libraries of molecules, both publicly known as well as confidential and proprietary. External parties can provide novel chemical scaffold, direct licensing, or joint co-creation opportunities. As projects mature, lead molecules can be offered to pharmaceutical partners for their consideration, or inventors for continued development. Often a majority of the molecules are discarded, alongside their data sets, as they fail to exhibit necessary attributes or functionality. This is part of the process of building understanding and progress toward a clinical candidate. But still, as part of drug discovery, millions of molecules have been created by a multitude of different institutions, it is not trivial to share data and knowledge of such molecules, particularly when it concerns IPRs.

3.2. A blockchain-powered experimental platform

To test the practical viability of the concepts discussed above, an experimental platform was built. From a technical perspective, the experimental platform was implemented as a permissioned blockchain.⁴ More specifically, access to the blockchain was open only to participants with a known verified identity, using digital certificates (Housley et al. 1999). Digital certificates are ubiquitous components of contemporary software systems, in which every actor holds cryptographically linked keys. Namely, a private key for generating digital signatures, and a public key for verifying those signatures. The private key should only be accessible by its owner. Leaking it would allow others to impersonate the owner. The corresponding public key (for verifying signatures) can be made available globally to everyone and is published in a digital certificate. The purpose of this certificate is to have a statement from the issuer that a given public key belongs to a specific real-world actor. That is, it links the public key with an identity. In the case of the prototype, we employed a private certificate authority, issuing digital certificates to users registering on the platform. All actions taken by a user on the platform are both digitally signed with that user's private key and time stamped.

To represent molecule structures digitally, it is necessary to use a 'language'. The simplified molecular-input line-entry system (SMILES) (Weininger 1988) and International Chemical Identifier (InChI) (Heller et al. 2015) are appropriate candidates for this, and are well-established tools for this purpose, as a large class of small molecules can be represented digitally using *SMILES* and *InChI*.

These formats are known as linear string representations – they are text strings. One can trivially convert between SMILES and InChI using software. An InChIKey is a secure cryptographic hash (Gollman 1999) of an InChI text string. What this means in practice is that such an InChIKey can be used as a fingerprint of the molecular structure. Fingerprinting in this manner is a well-established and familiar application of cryptographic hash functions. Anyone that knows the molecular structure of a compound (the InChI) can derive the corresponding InChIKey. However, since a cryptographic hash function is used in this computation, it is impossible to go in the other direction and compute the molecular structure (InChI) from its fingerprint (InChIKey), as illustrated in Figure 2.

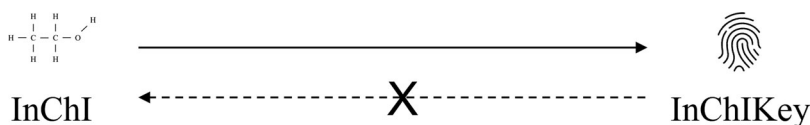


Figure 2. Molecular structure (InChi) vs. fingerprint (InChiKey).

In effect, an InChIKey can be used as proof that someone knows a certain structure without revealing the structure itself.

As a simplified but concrete example, consider the SMILES, InChI, and InChIKey for the ethanol molecule. Graphically, its structure is depicted in Figure 3.

The SMILES, InChI, and InChIKey for Ethanol are explicitly detailed in Table 2.

The SMILES consists of a sequence of characters that represents atoms. In the case of ethanol, the SMILES consists of two carbons (C) and one oxygen (O). All the hydrogen atoms are implicit. If two atoms are listed next to each other, they are connected. SMILES can contain additional information that describes more complex molecules.

An InChI starts with a fixed prefix string, 'InChI=1S', followed by a sequence of layer descriptions separated by slashes. For ethanol the first layer (C2H6O) defines which atoms the molecule consists of. Layer two (c1-2-3) defines how non-hydrogen atoms are connected. The third layer (h3H,2H2,1H3) defines how the hydrogens are connected to the non-hydrogen atoms. There exist several more layers that can be used to describe more intricate molecule features, e.g. isotopic information, but we leave these for the interested reader to study separately.

Finally, the InChIKey is derived by running an InChI text string through a cryptographic hash function in a structured manner. The result of the InChIKey calculation is a unique 27-character long text string that cannot be calculated in reverse. That is, it is practically impossible to deduce the molecular structure from the InChIKey, even though it is uniquely linked to a specific molecule.

The InChIKey makes it possible to perform an exact match search for an identical molecule without disclosing the actual structure of the molecule in the search query. This provides a practical solution to quickly identifying prior art. In fact, an InChIKey stored on a blockchain offers verifiable proof of any specific molecule's invention, without disclosing the molecule structure itself. Such a blockchain is, in effect, a ledger of inventions. With this framework, such chemical molecules can be digitized, thus creating a digital twin,

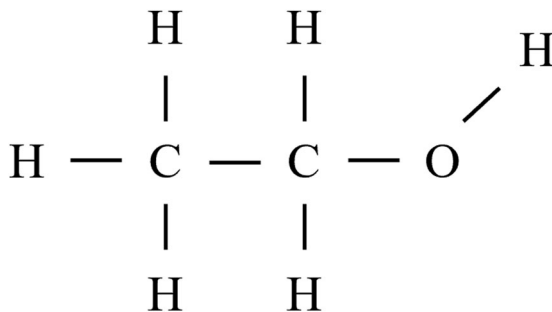


Figure 3. The ethanol molecule.

Table 2. The SMILES, InChI & InChIKey for ethanol.

Type	Value
SMILES	CCO
InChI	InChI = 1S/C2H6O/c1-2-3/h3H,2H2,1H3
InChIKey	LFQSCWFLJHTTHZ-UHFFFAOYSA-N

also known as tokenization, with a complete and immutable record of research achievements. This will improve traceability and provenance throughout the extended network of scientists dispersed across the world.

3.3. The operations of the experimental platform

The blockchain platform itself contains smart contracts that facilitate the business logic of the application. The following information and operations are defined for the purpose of using the platform (see also Table 3 below):

3.3.1. Upload molecule

Allows a user to upload a new molecule. The *Asset Administrator* will be set to the identity that uploaded the molecule. This means that she has full control of the NFT in the blockchain. The *Asset Administrator* can be changed using the *Transfer Ownership* operation. The uploader sets the *License Level* for the molecule. Depending on the license level, various structures can be set for the molecule. Under the *Fully open-source sharing*, and *Registration of a non-commercial asset* licenses the uploader can reveal the SMILES, InChI, or InChIKey for the molecule. Under the more restrictive *Collaborative Sharing*, the user either reveals nothing about the structure, or the InChIKey. The structure cannot be derived from the InChIKey as it was constructed using a secure hash function. Finally, the uploader can optionally set the other fields at will, with a couple of exceptions: *Uniqueness*, and *Upload Time* are set by the platform's services.

3.3.2. Search molecule

A user can search for molecules using any of the fields defined above. For example, the user can search for molecules that match some partial SMILES, some exact InChIKey, or biological and chemical properties (if those were configured by the uploader).

Table 3. Molecule information stored on blockchain.

Version	Version number for the data format
Asset Owner	The owner of the asset in the real world
Contributors	References to actors that have made contributions to work on the molecule
Physical & Chemical Properties	A list of physical and chemical properties for the molecule with associated values
Biological Properties	A list of biological properties for the molecule with associated values
External Data	References to data that is associated with the molecule that can be found outside of the blockchain
Molecule Structure	The SMILES, InChI, or InChIKey, or nothing depending on the license level of the molecule
License Level	The license for the molecule
Name	The name of the molecule
Asset Administrator	The identity that has the authority to administer the molecule in the blockchain
Uniqueness	Uniqueness information of the molecule. Specifies if the molecule is unique, where the uniqueness service searched for the molecule, and if applicable, where the uniqueness service found other molecules with the same InChIKey
Upload Time	A timestamp, when the molecule was uploaded

3.3.3. Transfer ownership

The administrator of an NFT can transfer her administrative capacities to another identity in the blockchain.

3.3.4. Edit molecule

All fields except *Structure*, *Name*, *Uniqueness*, and anything related to time stamps can be edited by the NFT administrator. This is used to update information on a certain molecule. This can, for example, be used to reflect that the owner of the real-world asset has changed, or that a new physical property has been discovered for the molecule.

3.3.5. List history

Allows anyone with access to the blockchain to verifiably trace the entire history of a molecule. With this level of traceability, one can track every development of a molecule and then potentially credit contributors according to their value. This functionality is enabled by having all operations in the blockchain both time-stamped and digitally signed.

Molecules that are uploaded under the *Fully open-source sharing*, or *Registration of a non-commercial asset* license can contain SMILES, InChI, or InChIKey information for the molecule. Molecules uploaded under the *Collaborative sharing* license will usually not reveal their structure due to the sensitivity of the data. They can, if the uploader desires, contain InChIKey information. If molecular structure information is present, the platform converts SMILES and InChI to the corresponding InChIKey where applicable. Recall that it is not possible to derive the structure of the molecule from the InChIKey. It then uses the InChIKey to scan the blockchain for any duplicates, followed by a search in some of the world's largest online molecular databases for the same InChIKey. Using these results, the molecule is marked in the blockchain. A web-based interface shows the user where the platform searched for duplicate InChIKeys. If no duplicates were found, the platform considers the molecule unique. If duplicates are found the user will know that she was not the first one to discover the probed molecular structure.

This unique service can be used to alter the way prior art is handled in the patent process. Currently, holders of potentially patentable intellectual property have to keep all information secret as there exists no practical way to determine who invented something first. Once the blockchain has enough molecules indexed, researchers can use the fact that InChIKeys keep molecules secret in combination with the uniqueness service to make an assessment whether their molecule is novel. If an InChIKey exists in the blockchain, one knows for sure that someone else knew about the molecule at a certain (specified) time, thanks to the time stamps. If the InChIKey does not exist in the blockchain, then a researcher can at least prove that she knew of a molecular structure at that specific time, making her claim stronger that her invention was novel. In this way, patent offices could use this service to automate the process of prior-art searches partially or fully in a new and more transparent way.

4. Technology-Enhanced resource management

Technically, we created a way of enabling more fine-grained control of the information that is valuable to the asset owner. In other words, we are enabling the asset owner to

share all information about the molecule that can possibly be shared without causing any harm to their business.

One of the greatest challenges of data-driven projects is to facilitate knowledge. The *FAIR Guiding Principles for Scientific Data management and Stewardship* published in 2016, provide guidelines to enhance the findability, accessibility, interoperability, and reusability of digital assets (Wilkinson et al. 2016). They encompass a set of principles that provide *inter alia* for a continuum of increasing reusability throughout many different implementations. They describe features and capabilities for systems and services to support the creation of valuable research outputs that could then be thoroughly evaluated and widely reused, attributing appropriate credit to the benefit of the creator and end-user (Mons et al. 2017).

To find relevant data, perform machine-analysis or employ artificial intelligence to identify patterns and correlations, data must be well-described, accessible and conform to standards. The FAIR principles are an important quality standard that has been embraced by EU policymakers and has sparked the global debate about better data stewardship in data-driven and open science projects (European Commission 2018). They have also motivated funding agencies to discuss their implementation requirements. Some of these requirements are still in their early stages, while others have matured to a set of guidelines (European Commission 2016). Generally speaking, these principles precede implementation choices and do not necessarily suggest specific technical requirements, standards, or implementation solutions. Nevertheless, they highlight ‘machine-actionability’ i.e. the ability of computational systems to find, access, interoperate, and reuse data with no or minimal human intervention. The FAIR guiding principles can be summarized in Table 4 below.

Findable, Accessible, Interoperable and Reusable data (FAIR) within the European Research area constitutes an important step towards achieving better data governance. Put simply, under the FAIR principles data are *Findable* when they are described by sufficiently rich metadata and registered or indexed in a known, accessible, and searchable user-resource. *Accessible* data objects can be obtained by humans and machines upon appropriate authorization and through a well-defined and universally implementable protocol. Interoperable data and metadata should use a formal, accessible, shared, and broadly applicable language for knowledge representation. But interoperability also comprises *technical* and *legal* interoperability. *Legal interoperability* relates to the principle that data should be ‘Reusable’. For data to be Reusable, the FAIR principles reassert the need for rich metadata and documentation that meet relevant community standards and provide information about provenance. *Reusability* requires *inter alia* that the legal conditions under which the data can be used should be transparent to both humans and machines.

There may be legitimate reasons to protect data from public access, including national security, personal privacy, competitiveness, trade secrets, patent applications, etc. The FAIR principles refer to the need to describe a process – mechanized or manual – for accessing discovered data. This requirement emphasizes the need to explicitly define the conditions under which they may be reused. Therefore, describing the context within which those data were generated, enabling evaluation of its utility, and providing clear instructions on how they should be referenced is of paramount importance (Mons et al. 2017).

Table 4. FAIR guiding principles (Wilkinson et al. 2016).

Findable:	<ul style="list-style-type: none"> Meta(data) are assigned a globally unique and persistent identifier^a Data are described with rich metadata Metadata clearly and explicitly include the identifier of the data it describes Meta(data) are registered or indexed in a searchable source
Accessible:	<ul style="list-style-type: none"> Meta(data) are retrievable by their identifier using a standardized communication protocol The protocol is open, free, and universally implementable The protocol allows for an authentication and authorization procedure, where necessary Metadata are accessible, even when the data are no longer available
Interoperable:	<ul style="list-style-type: none"> Meta(data) uses a formal, accessible, shared and broadly applicable language for knowledge representation Meta(data) use vocabularies that follow FAIR principles Meta(data) includes qualified references to other meta(data)
Reusable:	<ul style="list-style-type: none"> Meta(data) are richly described with a plurality of accurate and relevant attributes Meta(data) are released with a clear and accessible data usage license Meta(data) are associated with detailed provenance Meta(data) meets domain-relevant community standards

^aA unique and persistent identifier is a long-lasting reference to a digital resource.

In addition, FAIR data is not tantamount to OPEN data. To support uptake across the commercial sector and ‘sensitive data’ communities, FAIR does not necessarily imply complete openness. Data can be FAIR and shared under *restrictions*. When the case is made for open science, it is not argued that all research data should be open in all circumstances. Although following the idea that research data should be open, FAIR recognizes legitimate and necessary reasons for restricting access in some circumstances, and in particular in the health and life sciences. Clarity and transparency are, however, two important prerequisites governing access and reuse. While FAIR data does not need to be open in order to fulfill the conditions of reusability, FAIR data are expected to have a clear – preferably machine-readable – license (Mons et al. 2017). Recommended license schemes are described in the section below.

5. License schemes

In the field of IPRs, patents essentially extend over the use of an invention. Copyright, on the other hand, extends over the use of expressions of information. But does it cover structured datasets? The answer might not be straightforward. However, copyright law recognizes the protection of compilations of data independently of the protection of the compiled data.⁵ Hence, even if a dataset might not be protected by copyright, some of its components may be. Known in European IP law parlance as a copyright-protected database, it is intended to confer protection on the original selection and arrangement of contents,⁶ seemingly offering a rather weak, but, nonetheless, auxiliary layer of protection. Transferred into an exploitation scenario, however, in the digital environment copying an entire dataset will inevitably involve copying a copyright-protected layer and therefore involve a copyright-relevant act (Carroll 2015). However, whereas a database is broadly defined as ‘a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means’,⁷ some IP scholars maintain that mere alphabetical selection and arrangement, one that is constrained by an external function the selection or arrangement is intended to serve, or that merely follows practices or a convention standard specific to the type of expression will likely not suffice to

meet the required threshold of originality and therefore fall short of such auxiliary protection (Gervais 2019, 7).

But at the same time, European IP law also provides for a *sui generis* database right extending over the same category of subject-matter but conferring on the maker of the database a right to prevent extraction and re-utilization of the whole or substantial part of the database.⁸ Unlike the copyright-protected alternative, this IPR is not contingent on an originality standard but requires instead that the maker demonstrates a qualitative or quantitatively substantial investment in obtaining, verifying or presenting the contents of the database.⁹ This somewhat lighter 'entrepreneurial' character of protection might extend more directly over the scientific effort of compiling usable data and potentially provides a sufficient form of protection for the research community (Oliva and Corrales 2011).

A universal feature of IPR protection is that unless use is permitted, it is prohibited. Therefore, it is important to create a licensing framework where participants establish at the outset the conditions for use of the data. Needless to say, it is the prerogative of the holder of exclusive rights (or in fact the drafter of a contract, regardless of the existence of any IP) to determine the conditions for access. As discussed in the previous section, FAIR is not equivalent to open data. From a legal interoperability point of view, this means that data should be accessible under well-defined conditions (Graber-Soudry et al. 2021). The resulting blockchain platform is a repository of molecules and one essential prerequisite for this is 'data access'. Constraints on data access can cause negative effects in the upstream sector, such as universities and research institutions, but also in midstream and downstream scientific collaboration (Corrales Compagnucci 2020, 94). The other important characteristic is that data should be 'reusable'. For this reason, each level of the data-sharing model requires a different type of license scheme or a specific contract regulating access to the data in accordance with the intention of the right holder. Reliance on existing licensing schemes provides known and predictable alternatives. A licensing model could, therefore, be based on the following existing schemes.

5.1. Fully open and disclosed with no restrictions in use or reference

This level will allow participants to upload molecules to the blockchain platform with the intention to share it with the community. This approach will enable the full disclosure and sharing of molecules, including its structure and results. Any researcher could tap into it and use it without restrictions. This level is ideal for academic drug discovery such as neglected diseases and pandemic outbreaks (COVID-19, Dengue, Malaria, etc.). It follows the open data movement, where researchers may submit their data free of charge for the sake of scientific research.⁹ Full disclosure and sharing of molecules exist today, however not under a 'controlled' environment. Researchers would benefit from the blockchain-secured tracking system, which would potentially aid in fundraising.

This level requires fully open data and the most effective way to achieve this goal is by waiving copyright and sharing the data in a common pool. For practical reasons, we believe that the Creative Commons No Rights Reserved (CC0) is the recommended license for this level. The CC0 essentially enables scientists to waive their copyright and related rights to their fullest extent and place them worldwide in the public domain, so

that other participants can freely tap into it and build upon the works without restrictions. Participants who associate their data with this deed can, for instance, copy, modify, distribute and communicate the molecules without asking permission ('Creative Commons – CC0 1.0 Universal' [n.d.](#)).

The CC0 license provides a good solution for contributing work to the public domain. Similar to many open-source software licenses, CC0 is a universal instrument which is not centered on a particular jurisdiction. In other words, the CC0 empowers participants to opt-out of their exclusive rights automatically granted to creators.¹⁰ One immediate challenge, however, is the territorial character of copyright protection, related to which is a capacity of legislators to determine certain qualities of conferred rights, as set against licensing arrangements. For example, the inability in many jurisdictions to waive so-called moral rights encompassing at the least the right to be named as author and to object to derogatory treatment of the work.¹¹ Accordingly, CC0 applies 'to the greatest extent permitted by' the applicable law. But with rules on scientific misconduct in place, many R&D institutions around the world, in practice, this particular challenge may, in most situations, merely be theoretical.

Another license suitable for this level is the Open Data Commons Public Domain Dedication and License (PDDL), which is intended to allow participants to freely share, modify and use their work for any purpose and without any restrictions. The license is intended for use on databases or their content ('data'), either together or individually. Therefore, this license enables exploitation of the database and underlying data. As a consequence, the PDDL places data and databases as close as possible within the public domain. The goal is to eliminate restrictions. Thus, participants may share their molecules without restrictions or legal requirements ('Open Data Commons Public Domain Dedication and License (PDDL) v1.0 – Open Data Commons: Legal Tools for Open Data' [n.d.](#)).

5.2. Sharing and tracing of non-commercial assets

This level will allow researchers to register molecules and to search in the blockchain for other similar molecules free of charge. The blockchain in this scenario will allow the registration of specific molecules and timestamp for the event, including the possibility to search for known similarities and categorize the molecule as 'unique' if none can be found. This approach is ideal for academics with a need to document IPRs of their own creation. Therefore, the license agreements would allow to assign attribution rights.

This level also requires a fully open data model. Users are requested, however, to acknowledge the original creator (attribution rights). Creative Commons (CC) licenses are based on the 'some rights reserved' model. That means that the copyright holder when using a CC license, chooses to keep just several rights from the bundle conferred by law. Amongst the wide range of CC licenses, the CC BY 4.0 may be the most desirable for this level as it allows participants to share (copy and redistribute) the material in any medium or format. They are also entitled to make any adaptations for free or even for commercial purposes, but they must indicate if changes were made. They should also provide a link to the license and give appropriate credit to the original creator ('Creative Commons – Attribution 4.0 International – CC BY 4.0' [n.d.](#)).

When applying a CC license scheme, the downstream user will automatically receive a license to the original work as stated in the original license. However, the new derivative

work may be licensed differently from the original license. For instance, if the original data is licensed under the CC BY 4.0, a user can license derivative work under any other license, provided that credit for the original work is given to the creator of that original work.

5.3. Transparent and collaborative commercialization

This third level allows participants to establish potential collaborations and business opportunities. In such a case, the full structure of the molecule will not be disclosed, but the metadata will provide sufficient information for the researcher to find and compare molecules. As a way of illustration, this level will include an 'eBay-like' trading function which allows participants to trade their molecules in a similar fashion as to how the eBay auction system works. In essence, participants will be able to see a full list of molecules, bid on them or simply buy them as NFTs. The full structure of the molecule will not be disclosed in order to protect trade secrets and potential patent applications, but a description of the metadata will provide participants with sufficient information to sell and buy molecules. This level will operate together with a smart contract that will facilitate the execution and enforcement of the payment directly amongst the participants, if certain conditions are met.

6. Future prospects

The presented framework includes the following capabilities: evidence of creatorship/authorship; provenance to authenticate key information of a molecule; clarification of IPRs; providing evidence of uniqueness and the characteristics and properties of the molecule; establishing and enforcing license agreements or exclusive collaborative networks; and potentially making payments in real-time to IP owners.

Digitalization of physical assets can increase the transparency and speed by which drug research is conducted across both physical as well intellectual borders defined by confidentiality and ownership hurdles. The experimental blockchain solution described here provides one step towards an open collaboration platform that allows scientists and drug designers to share digital representations of early drug research molecules so that others can identify work of interest, collaborate, co-create, and handle intellectual property ownership in a much easier and more effective manner. Such a controlled sharing solution could provide a platform for laboratories, both academic and industrial, to cooperate and collaborate on a global scale. Some of the benefits from a blockchain-based collaboration platform for drug research molecules include increased speed in setting up a collaboration, reduced complexity in working with another partner, improved matchmaking in identifying new partnering opportunities, and more efficient collaboration as overlapping use of resources can be minimized. Taking each of these potential benefits in turn.

Much time is today spent on negotiating individual agreements, often because there is no standard or one-size-fits-all solution. Still, several aspects of collaborative agreements are consistent and usually addressed, including ownership of intellectual property, confidentiality, usability, and compensation. This is of course particularly true when the same kind of asset is used, i.e. chemical molecules. A reduction in complexity and a

standardized model can reduce time spent on negotiations of contractual details and terms resulting in increased speed.

There are many different ways of collaborating and even more hurdles and unknowns making it difficult to get past the initial steps of identifying, discussing, negotiating, agreeing, and executing sharing of a physical asset. Material Transfer Agreements (MTA) can be commonly defined as part of the blockchain protocol in the form of smart contracts which can be easily understood and executed. Such an improved interoperable platform makes it is much easier for different partners to 'plug-in', thus reducing much complexity and difficult hurdles required to engage with collaboration partners on a global level.

Matching a need for a drug molecule with certain properties with a supplier is both easier and more effective on a transparent platform. Whether it be for academic studies or licensing opportunities, being able to have full insights as to what is available under what licensing agreement improves the chances of both providing and seeking assets with particular properties. A platform functioning as a marketplace for drug research molecules using standardized digital licensing mechanisms will make it much easier for parties to find a match according to need or offerings globally.

With increased transparency comes an opportunity to reduce overlapping activities and repeat others' efforts making it easier to avoid aspects already addressed by others. For example, a scientist might not have to spend time and effort creating a certain chemical scaffold which someone else already completed and made available for others to further modify (similar to a creative commons license). Instead, the scientist could acquire the novel chemical scaffold as starting material and further develop particular and more advanced analogues of the molecule, all while acknowledging previous contributors' work. This is different from traditional academic dissemination as the blockchain ensures full transparency and further increases the incentive for documentation of original work to claim inventorship for recognition and also potential compensation.

This framework will allow researchers to share data and assets of chemical molecules in a secure and automated manner. This is especially important in cases where research collaboration is needed. What types of molecules are shared, for how long, under which circumstances, and with what authorization is critical in order to balance the need for internal confidentiality and data-sharing. More efficient resource utilization in general, reuse of assets, less waste, cheaper medicines, larger health problems solved, and global digital collaborations boosted.

In the context of the recent COVID-19 pandemic, it is paramount to establish research collaboration infrastructures that boost collective resources. Today, it is a common understanding amongst all stakeholders that bringing an early-stage research idea forward from inception to improving health requires joint efforts. Improving how actors work together across borders will significantly benefit drug research, both from a health, science, and business perspective. But such improved collaboration requires improved models for the sharing of both physical assets and data that relates to intellectual property in a functional manner, making sure inventorship and contributions are fully transparent, indisputable, and respected.

The prospects of blockchain technology and the gradual shift to more open innovation models highlight the importance of efficient and well-calibrated governance and use of data. From a legal perspective, it seems clear that smart contracts can fulfill many of

the traditional legal requirements. Most agreements do not need to take any specific form, such as a formal, written agreement to be legally enforceable. However, the widespread adoption of smart contracts faces several difficulties and limitations. For this reason, we adopted a hybrid approach where some functionalities can be embedded as code in the form of smart contracts while others can be depicted in the form of traditional licensing arrangements.

Against this backdrop, an experimental platform was developed as a globally and publicly accessible blockchain to share early drug research molecules. This framework has the potential to create more transparency and allow researchers to interact and contract safely. Among other benefits, what such a platform proposes is the elimination of intermediaries and a decentralization of the entire IPR management system, and a high level of security and integrity by acting as a reliable database that contains records of every transaction performed across the network.

Finally, one caveat must be added. It appears that in some strategically important technologies, not only the protection of personal data under, but also data protection through trade (and state) secrets are becoming increasingly important not only for private companies, but also for national and regional innovation and security strategies. While NFT models might work very well in some fields of biomedical innovation with less need for secrecy and privacy protection, our proposals might work less well in more sensitive areas of innovation (Kostick-Quenet et al. 2022). It appears, therefore, more important than ever to think about (a) curtailed solutions depending on the area of innovation, and (b) the broader implications of these forms of protection vis-a-vis the ideal to promote more open innovation and collaboration.

Notes

1. Regulation (EU) 2016/679 of the EP and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC, OJ 2016 L 119, 1 (General Data Protection Regulation, GDPR).
2. Smart contracts are self-executed, tamper-proof and autonomous digital agreements in the form of computer code hosted on a blockchain. For a full recount of smart contracts and different use cases see; Corrales Compagnucci, Fenwick, and Wrbka (2021) and Corrales Compagnucci, Fenwick, and Haapio (2019).
3. Non-Fungible Tokens (NFTs) are financial securities consisting of digital data representing real-world objects such as art, music, videos, etc. They are stored in a blockchain and can be transferred, sold and traded by the owner. Unlike cryptocurrencies (fungible tokens), they are characterized by having a unique identification code and metadata that distinguishes them from each other ("What Is An NFT? – Forbes Advisor" n.d.).
4. Permissioned blockchains only authorize access to certain users with permissions. Decisions are typically based on a majority vote and users may have only partial access to a selected subset of the data granted by the administrators (Corrales, Jurčys, and Kousiouris 2019, 190).
5. For example, Art. 2(5) Berne Convention, referring to 'collections'.
6. Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the Legal Protection of Databases ("Database Directive") Art. 3(1).
7. Database Directive, Art. 1(2).
8. Database Directive, Art. 7(1).
9. This level is also inspired in the Elinor Ostrom approach which focuses on the use of common-pool resources and the various routes to control such resources (Ostrom 2015; 2000).

10. Patents and trademarks are by no means affected by this license ("Creative Commons – CC0 1.0 Universal" n.d.).
11. Berne Convention, Art. 6 bis.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was partially supported by the Novo Nordisk Foundation under [grant number NNF17SA0027784] for a scientifically independent Collaborative Research Program in Biomedical Innovation Law, by the Inter-CeBIL program [grant number NNF23SA0087056], and by the VINNOVA project, 'Blockchain-powered open sharing platform for collaborative drug discovery', reference number [grant number 2019-02815].

ORCID

Marcelo Corrales Compagnucci  <http://orcid.org/0000-0001-5709-3621>

References

- Balasegaram, Manica, Peter Kolb, John McKew, Jaykumar Menon, Piero Olliari, Tomasz Sablinski, Zakir Thomas, Matthew H Todd, Els Torreele, and John Wilbanks. 2017. "An Open Source Pharma Roadmap." *PLOS Medicine* 14 (4): e1002276. <https://doi.org/10.1371/journal.pmed.1002276>.
- Carroll, Michael W. 2015. "Sharing Research Data and Intellectual Property Law: A Primer." *PLOS Biology* 13 (8): e1002235. <https://doi.org/10.1371/journal.pbio.1002235>.
- Coase, R. H. 1960. "The Problem of Social Cost." *The Journal of Law and Economics* 3 (July): 1–44. <https://doi.org/10.1086/466560>.
- "Commentary: The Coronavirus Pandemic Has Shattered the Status Quo on Drug Development. We Should Build on That | Fortune." 2020. 2020. <https://fortune.com/2020/03/26/coronavirus-vaccine-drug-development-open-science-covid-19-treatment/>.
- Corrales, Marcelo, Mark Fenwick, and Helena Haapio. 2019. "Digital Technologies, Legal Design and the Future of the Legal Profession." In *Perspectives in Law, Business and Innovation*. Singapore: Springer Nature. https://doi.org/10.1007/978-981-13-6086-2_1.
- Corrales, Marcelo, Mark Fenwick, Helena Haapio, and Erik P M Vermeulen. 2019. "Tomorrow's Lawyer Today? Platform-Driven LegalTech, Smart Contracts & the New World of Legal Design." *Journal of Internet Law* 22 (10): 3–12.
- Corrales, Marcelo, Paulius Jurčys, and George Kousiouris. 2019. "Smart Contracts and Smart Disclosure: Coding a GDPR Compliance Framework." In *Perspectives in Law, Business and Innovation*. Singapore: Springer Nature. https://doi.org/10.1007/978-981-13-6086-2_8.
- Corrales Compagnucci, Marcelo. 2020. *Big Data, Databases and "Ownership" Rights in the Cloud. Perspectives in Law, Business and Innovation*. Singapore: Springer Nature.
- Corrales Compagnucci, Marcelo, Eva Egermann, Nikolaus Forgó, and Tina Krügel. 2010. "Intellectual Property Rights in E-Health: Balancing out the Interests at Stake - A Herculean Task?" *International Journal of Private Law* 3 (3): 286–299. <https://doi.org/10.1504/IJPL.2010.033409>.
- Corrales Compagnucci, Marcelo, Mark Fenwick, and Helena Haapio. 2019. *Legal Tech, Smart Contracts and Blockchain*. Perspectives in Law, Business and Innovation. Singapore: Springer Nature. <https://doi.org/10.1007/978-981-13-6086-2>.
- Corrales Compagnucci, Marcelo, Mark Fenwick, and Stefan Wrška. 2021. *Smart Contracts: Technological, Business and Legal Perspectives*. Oxford: Hart Publishing.

- "Creative Commons — Attribution 4.0 International — CC BY 4.0." n.d. Accessed July 12, 2023. <https://creativecommons.org/licenses/by/4.0/>.
- "Creative Commons — CC0 1.0 Universal." n.d. Accessed July 12, 2023. <https://creativecommons.org/publicdomain/zero/1.0/>.
- Deloitte Centre for Health Solutions. 2022. "Nurturing Growth Measuring the Return from Pharmaceutical Innovation 2021," no. January. <https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/life-sciences-health-care/Measuring-the-return-of-pharmaceutical-innovation-2021-Deloitte.pdf>.
- Deloitte US. 2017. "Collaboration as a Key to Success in Pharmaceutical R&D." 2017. <https://www2.deloitte.com/us/en/pages/life-sciences-and-health-care/articles/health-care-current-january-17-2017.html>.
- Druedahl, Louise C., Timo Minssen, and W. Nicholson Price. 2021. "Collaboration in Times of Crisis: A Study on COVID-19 Vaccine R&D Partnerships." *Vaccine* 39 (42): 6291–6295. <https://doi.org/10.1016/j.vaccine.2021.08.101>.
- European Commission. 2016. "H2020 Programme. Guidelines on FAIR Data Management in Horizon 2020," no. 3: 12. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.
- European Commission. 2018. *Turning FAIR Into Reality: Final Report and Action Plan from the European Commission Expert Group on FAIR Data*. Publications Office. <https://doi.org/10.2777/1524>.
- "Final Report — High-Level Panel on Access to Medicines." 2016. 2016. <http://www.unsgaccessmeds.org/final-report/>.
- Gervais, Daniel. 2019. "Exploring the Interfaces Between Big Data and Intellectual Property Law." *Journal of Intellectual Property, Information Technology and E-Commerce Law*. <https://doi.org/10.2139/ssrn.3360344>.
- Gollman, Dieter. 1999. *Computer Security*. Hoboken: Wiley.
- Graber-Soudry, Ohad, Timo Minssen, Daniel Nilsson, Jakob Wested, Marcelo Corrales Compagnucci, and Benedicte Illien. 2021. "Legal Interoperability and the FAIR Data Principles: An X-Officio Study Commissioned by the EOSC FAIR Working Group." <https://doi.org/10.5281/zenodo.4471312>.
- Heller, Stephen R., Alan McNaught, Igor Pletnev, Stephen Stein, and Dmitrii Tchekhovskoi. 2015. "InChI, the IUPAC International Chemical Identifier." *Journal of Cheminformatics* 7 (1), <https://doi.org/10.1186/s13321-014-0049-z>.
- Housley, Russell, et al. 1999. "RFC 2459 - Internet X.509 Public Key Infrastructure Certificate and CRL Profile." <https://datatracker.ietf.org/doc/html/rfc2459.html>.
- Kessel, Mark. 2011. "The Problems with Today's Pharmaceutical Business—an Outsider's View." *Nature Biotechnology* 29 (1): 27–33. <https://doi.org/10.1038/nbt.1748>.
- Kostick-Quenet, Kristin, Kenneth D. Mandl, Timo Minssen, I. Glenn Cohen, Urs Gasser, Isaac Kohane, and Amy L. McGuire. 2022. "How NFTs Could Transform Health Information Exchange." *Science* 375 (6580): 500–502. <https://doi.org/10.1126/science.abm2004>.
- Mayr, Lorenz M, and Dejan Bojanic. 2009. "Novel Trends in High-Throughput Screening." *Current Opinion in Pharmacology* 9 (5): 580–588. <https://doi.org/10.1016/j.coph.2009.08.004>.
- Mons, Barend, Cameron Neylon, Jan Velterop, Michel Dumontier, Luiz Olavo Bonino da Silva Santos, and Mark Wilkinson. 2017. "Cloudy, Increasingly FAIR; Revisiting the FAIR Data Guiding Principles for the European Open Science Cloud." *Information Services & Use* 37 (February): 49–56. <https://doi.org/10.3233/ISU-170824>.
- Nilsson, Niclas, and Jakob Felding. 2015. "Open Innovation Platforms to Boost Pharmaceutical Collaborations: Evaluating External Compounds for Desired Biological Activity." *Future Medicinal Chemistry* 7 (14): 1853–1859. <https://doi.org/10.4155/fmc.15.122>.
- Nilsson, Niclas, and Timo Minssen. 2018. "Unlocking the Full Potential of Open Innovation in the Life Sciences Through a Classification System." *Drug Discovery Today* 23 (4): 771–775. <https://doi.org/10.1016/j.drudis.2018.01.002>.
- Noorden, Richard Van. 2021. "Scientists Call for Fully Open Sharing of Coronavirus Genome Data." *Nature* 590 (7845): 195–196. <https://doi.org/10.1038/d41586-021-00305-7>.

- Oliva, Michele, and Marcelo Corrales. 2011. "Law Meets Biology: Are Our Databases Eligible for Legal Protection?" *SCRIPTed* 8 (3): 226. <https://doi.org/10.2966/scrip.080311.226>.
- Olk, Paul, and Joel West. 2020. "The Relationship of Industry Structure to Open Innovation: Cooperative Value Creation in Pharmaceutical Consortia." *R&D Management* 50 (1): 116–135. <https://doi.org/10.1111/radm.12364>.
- Olsson, Christoffer, and Mohsen Toorani. 2021. "A Permissioned Blockchain-Based System for Collaborative Drug Discovery." In *ICISSP 2021 - Proceedings of the 7th International Conference on Information Systems Security and Privacy*. <https://doi.org/10.5220/0010236901210132>.
- "Open Data Commons Public Domain Dedication and License (PDDL) v1.0 — Open Data Commons: Legal Tools for Open Data." n.d. Accessed July 12, 2023. <https://opendatacommons.org/licenses/pddl/1-0/>.
- "Open-Source Malaria: OSM." n.d. <http://opensourcemalaria.org/>.
- "Open Targets: New Name, New Data | EMBL-EBI." 2016. <https://www.ebi.ac.uk/about/news/announcements/open-targets-new-name-new-data/>.
- Ostrom, Elinor. 2000. "Reformulating the Commons." *Swiss Political Science Review* 6 (1): 29–52. <https://doi.org/10.1002/j.1662-6370.2000.tb00285.x>.
- Ostrom, Elinor. 2015. *Governing the Commons*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/cbo9781316423936>.
- Saha, Chandra Nath, and Sanjib Bhattacharya. 2011. "Intellectual Property Rights: An Overview and Implications in Pharmaceutical Industry." *Journal of Advanced Pharmaceutical Technology & Research* 2 (2): 88–93. <https://doi.org/10.4103/2231-4040.82952>.
- Saluja, Ronak, Vanessa Arciero, Sierra Cheng, Erica McDonald, William Wong, Matthew Cheung, and Kelvin Chan. 2018. "Examining Trends in Cost and Clinical Benefit of Novel Anticancer Drugs Over Time." *Journal of Oncology Practice* 14 (March): e280–e294. <https://doi.org/10.1200/JOP.17.00058>.
- Schuhmacher, Alexander, Oliver Gassmann, and Markus Hinder. 2016. "Changing R&D Models in Research-Based Pharmaceutical Companies." *Journal of Translational Medicine* 14 (1): 105. <https://doi.org/10.1186/s12967-016-0838-4>.
- "SGC | Structural Genomics Consortium." n.d. Accessed July 11, 2023. <https://www.thesgc.org/>.
- Tsarsitalidis, Stylianos, Marcelo Corrales Compagnucci, George Kousiouris, and Alan Dahi. 2021. "Feeding Smart Contract Legal Requirements with Semantic and Event Detection Logic Structures from Modern Service-Oriented Supply Chains." In *Smart Contracts: Technological, Business and Legal Perspectives*, edited by Marcelo Corrales Compagnucci, Mark Fenwick, and Stefan Wrbka. Oxford: Hart Publishing. <https://doi.org/10.5040/9781509937059.ch-007>.
- Weininger, David. 1988. "SMILES, a Chemical Language and Information System: 1: Introduction to Methodology and Encoding Rules." *Journal of Chemical Information and Computer Sciences* 28 (1): 31–36. <https://doi.org/10.1021/ci00057a005>.
- "What Is An NFT? – Forbes Advisor." n.d. Accessed July 9, 2023. <https://www.forbes.com/advisor/investing/cryptocurrency/nft-non-fungible-token/>.
- Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3 (1): 160018. <https://doi.org/10.1038/sdata.2016.18>.