BIOCHAIN: Towards a platform for securely sharing microbiological data

Original
BIOCHAIN: Towards a platform for securely sharing microbiological data / Bonnici, V.; Arceri, V.; Diana, A.; Bertini, F.; Iotti, E.; Levante, A.; Bernini, V.; Neviani, E.; Dal Palu, A.. - (2023), pp. 59-63. (Intervento presentato al convegno 27th International Database Engineered Applications Symposium, IDEAS 2023 tenutosi a Knossos Royal Beach Resort, grc nel 2023) [10.1145/3589462.3589501].

Availability:
This version is available at: 11381/2950792 since: 2023-11-14T13:37:24Z

Publisher:
Association for Computing Machinery

Published
DOI:10.1145/3589462.3589501

BIOCHAIN: towards a platform for securely sharing microbiological data

Vincenzo Bonnici*
vincenzo.bonnici@unipr.it
Department of Mathematical,
Physical and Computer Sciences
University of Parma, Italy

Vincenzo Arceri
vincenzo.arceri@unipr.it
Department of Mathematical,
Physical and Computer Sciences
University of Parma, Italy

Alessio Diana
alessio.diana@studenti.unipr.it
Department of Mathematical,
Physical and Computer Sciences
University of Parma, Italy

Flavio Bertini
flavio.bertini@unipr.it
Department of Mathematical,
Physical and Computer Sciences
University of Parma, Italy

Eleonora Iotti
eleonora.iotti@unipr.it
Department of Mathematical,
Physical and Computer Sciences
University of Parma, Italy

Alessia Levante
alessia.levante@unipr.it
Department of Food and Drug
University of Parma, Italy

Valentina Bernini
valentina.bernini@unipr.it
Department of Food and Drug
University of Parma, Italy

Erasmo Neviani
erasmo.neviani@unipr.it
Department of Food and Drug
University of Parma, Italy

Alessandro Dal Palù
alessandro.dalpalu@unipr.it
Department of Mathematical,
Physical and Computer Sciences
University of Parma, Italy

ABSTRACT

There is a need to persuade public and private entities to share their currently unexposed bio-data banks by preserving ownership and secrecy. The reason is to make available results that can be obtained by massively exploiting the content of such data by modern machine learning approaches. Digital catalogues of data collections are being provided. However, they are not developed to protect private content that may be shared according to privileges assigned by the owners. Here, we present BIOCHAIN, a data-sharing module which will be the basis for a computational platform aimed at performing federated data analysis. The platform is intended to be used by a consortium of private and public institutions in the field of microbiology. BIOCHAIN makes use of blockchain technology to guarantee fairness among entities of the consortium by allowing them to securely share their data.

CCS CONCEPTS

• **Information systems** → **Data exchange**; *Data provenance*; Distributed database transactions.

KEYWORDS

Federated Data Sharing, Blockchain, Microbiological Data.

ACM Reference Format:
Vincenzo Bonnici, Vincenzo Arceri, Alessio Diana, Flavio Bertini, Eleonora Iotti, Alessia Levante, Valentina Bernini, Erasmo Neviani, and Alessandro Dal Palù. 2023. BIOCHAIN: towards a platform for securely sharing microbiological data. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In the current post-genomic era, artificial intelligence is becoming to be extensively used for many bioinformatics applications [7], ranging from healthcare to protein folding and microbiome analysis. Modern machine-learning techniques promise to increase the level of accuracy of computational analyses at the cost of an increased amount of processed data [14]. For this reason, there is a need to persuade public and private entities to share their currently unexposed data banks by preserving ownership and secrecy [5, 17].

An example of shareable data is the information associated with biological material deposited in bio-banks across the world, as is the case for microbial culture collections. In response to the growing interest in the exploitation of bio-resources, the Nagoya protocol, acknowledged by the European Parliament[1] and Italian Government[2], aims at defining international guidelines regarding access to genetic resources[3] and the fair and equitable sharing of benefits arising from their use [5]. The European Culture Collections' Organization has translated the protocol into a legal framework composed of models of the agreement for the deposit and the transfer of resources stored in microbial collections [18]. Digital catalogues of these collections are provided [9]. However, they are not developed to protect private content that may be shared unwittingly in violation of copyrights and agreements.

---

[*] Corresponding author.

[1] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32015R1866

[2] https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:1994-02-14;124

[3] https://www.cbd.int/abs

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

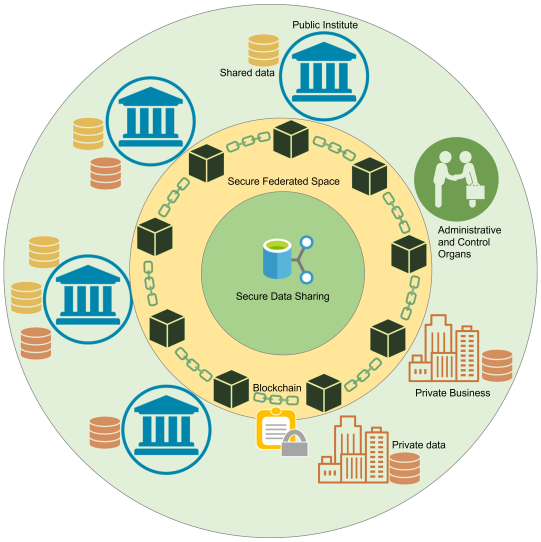


Figure 1: Overview of the BIOCHAIN platform paradigm.

The BIOCHAIN project aims at providing a prototype module for managing biological data regarding microbial collections. The overall platform is intended to be used by a consortium of public and private entities that have the need to share their data and integrate their knowledge for gaining advantages by federated computational analyses [21]. Members of the consortium share data under specific law agreements that must be traced in the digital world. For this reason, the proposed system relies on a security layer based on blockchain technology . By means of such technology, the platform is able to immutably store the transaction history of all the members of the consortium. The blockchain provides a decentralised solution where any member can verify the transaction history, and it is protected against malicious or negligent members.
Figure 1 shows an overview of the paradigm from the BIOCHAIN platform. Each member stores its own data off-chain and decides to whom and for which purpose the data can be shared. The blockchain exploits *smart contracts* to track request transactions and their responses, ensuring compliance with agreements by every party.

This study reports the ongoing activity regarding the development of BIOCHAIN, by exposing the currently available results. BIOCHAIN is the basis for a larger computational platform called BIOCHAIN-AI[4]. The complete platform will provide an environment for performing federated machine learning: a decentralised machine learning approach with the peculiarity that a partial execution of the algorithm is run on the servers holding local data, without the need to exchange them. Partial executions are then aggregated into a merged result. The request for a specific task is performed by a given member, then the other members can approve or reject it.

It is expected that the prototype, as well as future developments of it, will have a high impact on the national and European network of institutions that host microbial collections. In particular, to those Nations and institutions who are in agreement with the Nagoya protocol, given the importance of these collections as a reservoir of biodiversity, having not only an ecological interest but also the opportunity to contribute to innovation in the field of bio-based industry. In particular, MIRRI is the pan-European distributed Research Infrastructure for the preservation, systematic investigation, provision and valorisation of microbial resources and biodiversity[5]. Recently, the University of Parma, hosting the University of Parma Culture Collection (UPCC), has been associated with the Joint Research Unit (JRU) of MIRRI-IT, aimed at the development of the Italian network of collections of microbial resources. These are of great interest to many research fields such as agrifood (bio-fertilisers and bio-pesticides for sustainable agriculture, starter, protective or aromatic strains to produce fermented foods), human health (nutraceuticals, probiotics, human microbiomes for personalised medicine), environment (bioremediation strategies and development of bio-processes aimed at obtaining compounds of technological interest, like nano-materials).

2 RELATED WORKS

Recent approaches, especially in the field of machine learning, exploit the huge amount of data that is made available via recent data acquisition technologies [14]. However, a great portion of biological data is still not available to be shared [8]. This aspect regards both non-standard information and resources protected by copyright. In particular, copyrighted data is owned by private businesses, or its access is restricted by specific agreements. A way to reach this goal is to apply federated learning techniques by preserving differential privacy [22].

Federated learning techniques are recently changing the paradigm of model training, moving from centralising the data to decentralising the model [13]. The new paradigm copes with constraints posed by privacy and intellectual property regulations. Collaborations among pharmaceutical institutions can be also useful for drug discovery preserving intellectual property and related financial interests [3]. Two other interesting examples are presented in [2] and [19]. In the former, the authors presented an automatically single-cell type identification tool. While in the latter, the authors proposed a federated approach to evaluate gene expression for privacy-preserving disease diagnosis. Another advantage of the federated approach lies in the ability to process a huge amount of data as demonstrated in [10], where the authors used a cross-pharma dataset of more than 2.6 billion confidential experimental activity data points. Unfortunately, even if there are plenty of machine learning applications for the analysis of microbiome data and bacterial population, there is no current approach which made use of the federated paradigm. The reason is possibly due to a lack of a sharing and computational platform that easily allow such analyses for this specific field of application.

Blockchain technology has recently disrupted the field of digital security. It is a form of distributed ledger which aims at providing a consensus of replicated, shared, and synchronised digital data in a decentralised environment with no central authority. It is now applied to a wide variety of applications that also include the biomedical field [11]. The advantages of blockchain technology for data-sharing have been widely investigated in the field of

---

[4] https://ahead-lab.unipr.it/biochain-ai

[5] http://www.mirri-it.it

healthcare [4], IoT [15] and industrial applications [20]. Traceability, security, and consensus (acceptance as well as direct revocation) are some of the main gains of using blockchains for such types of data-sharing systems. The introduction of blockchain in federated learning models is very recent [12], however, no models specialised in microbiology are currently available.

3 METHODOLOGY

The BIOCHAIN module traces requests to access private data, and their responses. In particular, what has been requested, who requested the data, when, and for what purpose. It implements the security aspects of the platform and it ensures fairness among the members of the consortium. In fact, by means of a distributed ledger system (a blockchain), data-sharing transactions are registered such that every member actively stores and verifies them. The data itself is privately stored, and it will be identified by a unique identifier. Access to data is guaranteed by smart contracts running over the blockchain. The smart contracts interact with the second layer of the platform that is responsible for defining the protocol for exposing the data and requesting it.

Unlike more popular blockchains that are public and permissionless, such as Ethereum, BIOCHAIN should be *private*, i.e., each blockchain node must be registered to the blockchain, and *permissioned*, i.e., just certain nodes are allowed to perform some operations. A trusted administration node manages nodes' registrations and their policy accesses. In the proposed prototype, the administration node allows any node to insert and request biological data units, but more strict policies will be implemented based on specific requirements verified by the administration node (e.g., the reputation of an organisation, and potential data shareable by an organisation). The biological data-sharing layer over the proposed blockchain is orchestrated by smart contracts. A smart contract is a program immutably stored in a blockchain used to automatically execute agreements between two or more nodes. BIOCHAIN implements a single smart contract usable by organisations registered to the blockchain for (i) making available a biological data unit (ii) requiring a specific biological data unit, and (iii) accepting or denying a data request. Any interaction between blockchain participants is immutably tracked on the blockchain, allowing any organisation to access the history of the data requests and who owns each specific biological data. Being a smart contract immutable, before deploying it in the blockchain we should prove the absence of notable vulnerabilities, such as non-determinism and numerical overflow, that may cause the whole blockchain to halt or stop [16]. Hence, we rely on sound automatic program analysis to formally prove that the smart contract does not contains such vulnerabilities. Given the desired features previously explained, BIOCHAIN is implemented in Hyperledger Fabric [1], allowing one to develop a fully customizable permissioned private blockchain and allowing the development of smart contracts.

4 THE BIOCHAIN ARCHITECTURE

BIOCHAIN[6] has been developed using the framework Hyperledger Fabric (HLF for short). HFL is a blockchain framework designed to build private, permissioned, and highly customizable blockchains. Moreover, it allows the writing of custom smart contracts (called *chaincodes*), namely computer programs immutably stored in a blockchain that define a set of rules to regulate asset exchanges among users/nodes of the blockchain. A graphical representation of the interactions between the blockchain global state and a chaincode is reported in Figure 2. Two methods are represented: `InsertData` and `GetData`. The first method allows to store any kind of data into the global state of the blockchain using a unique identifier to index it. The second one allows to retrive a specific data from the global state using the corresponding key. In order to communicate, nodes must be part of the network, which is the infrastructure that allows nodes to exchange assets and messages. In HFL, there exist two classes of nodes: *peer nodes*, managing the ledger and chaincodes, and *orderer nodes*, managing the order with which the transactions must be executed in the network. BIOCHAIN's network contains at least one ordered node and a peer node associated with each participating organisation. Every peer node maintains a copy of the blockchain global state (i.e., the ledger) that is kept synchronised among all the other nodes as new transactions occur. Similarly to other blockchains, HFL distributed ledger is a set of key-value pairs. BIOCHAIN's ledger tracks, for each biological data, the following information: the owner (an organization participating in the blockchain), a description of the biological data, and a unique identifier of the biological data. In this way, any blockchain participant can inform other participants that certain biological data is available and can query the blockchain to know which biological data are offered by the other participants. These operations are managed via the chaincode that we will explain later.

Clearly, the actual biological data cannot be stored in the ledger, being otherwise publicly exposed. In order to store biological data in the blockchain without publicly exposing it to the other blockchain participants, we rely on HFL private data collections[7]. Private data collections are managed by their owners and kept inside the organisations' peer nodes. The access policy of a private data collection can be customised on the owner's needs (e.g., sending the data to another participant) but only the owner of the collection is able to read data inside it. Hence, descriptions of which biological data are available are publicly stored in the ledger, while the actual data are stored in the private data collection of the owner. The biological data

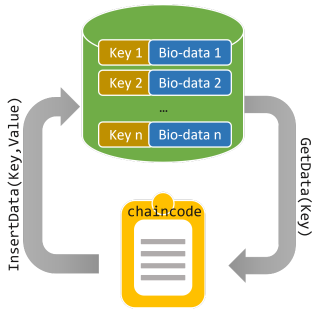
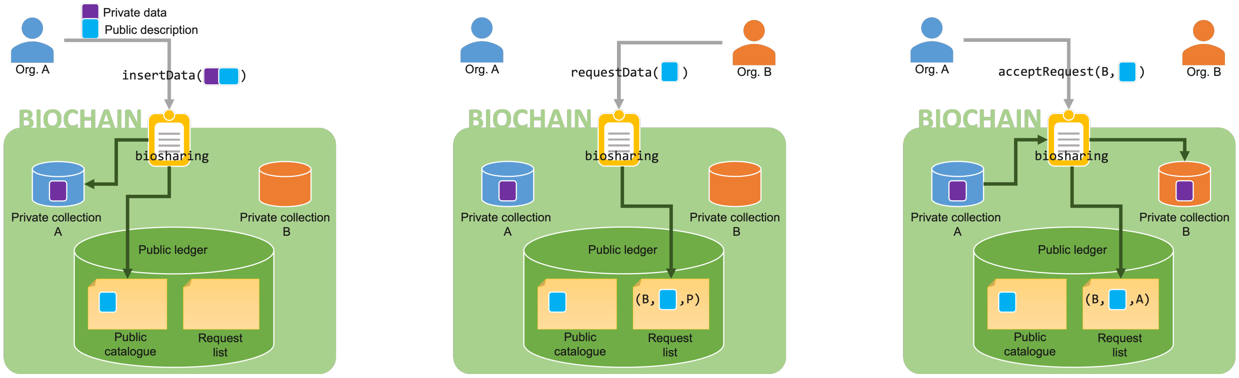


Figure 2: Interaction with the public distributed ledger.

---

[6] Available at https://github.com/biochain-ai/biochain-ai.

[7] https://hyperledger-fabric.readthedocs.io/en/latest/private-data/private-data.html



(a) (b) (c)

Figure 3: (a) `insertData`, (b) `requestData`, (c) `acceptRequest` methods graphical representation.

transfer from an owner to another organisation is automatically managed by the chaincode that we will present shortly.

Once an organisation makes available certain biological data in the blockchain, another organisation can request it. Hence, the ledger stores also *sharing requests*, storing which data is desired, which organisation requires the data, and from which organisation the data is required. Then, the owner may accept or deny a sharing request. Given the immutability of the blockchain, such an architecture allows tracking the history of any sharing request.

The chaincode. All the transactions, including sharing requests, accepting or denying responses, and data exchange over BIOCHAIN are managed by the chaincode `biosharing`[8], implemented in Go. In order to verify the absence of non-determinism and numerical overflow vulnerabilities, we have analysed the chaincode with GoLiSA [16], a sound static analyser for HLF chaincodes based on LiSA [6], proving the absence of such vulnerabilities.

The chaincode exposes a range of methods corresponding to the actions that an organisation can perform over BIOCHAIN: store and share biological data, ask for resources from other organisations, and accept or deny requests for data sharing. In the following, we describe in detail all the methods offered by the chaincode `biosharing`.

`insertData`. It allows an organisation to insert a biological data unit into the blockchain. Calling this method, an organisation simultaneously stores in its own private data collection a given data unit, and notify in the public ledger that the data unit is made available to the consortium, and in turn, it can be requested by interested organisations. While calling this method, in order to hide the actual biological data from other organisations, we rely on HLF transient data, which is a way of passing data to methods without being visible during the transaction. Figure 3a graphically represents the effects of this method on the ledger. Note that, the actual biological data unit (purple block in Figure 3a) is not stored in the public ledger, thus it is not visible to other organisations.

`viewCatalogue`. It allows an organisation to query the public ledger to get all the biological data units made available from the other organisations, i.e., that can be requested.

`viewPersonalData`. It allows an organisation to query the public ledger to retrieve which biological data units owned by the caller are available to the other organisations.

`getPrivateData`. Similar to the previous method, an organisation can call this method for querying the biological data unit in its own private data collection.

`requestData`. Once an organisation A has made available a biological data unit (using `insertData`), another organisation B can require it by calling the method `requestData`. This will put in the public ledger a new sharing request (initially in the pending state) from the organisation B. Figure 3b graphically represents the effects of this method on the ledger.

`viewRequests`. It allows an organisation to query the history of the sharing requests from other users. The method queries all the sharing requests associated with biological data units owned by the caller in any state: accepted, denied, and pending. Each sharing request is associated with a unique serial number that must be specified when an organisation wants to accept or deny a sharing request.

`accept/denyRequest`. An organisation can call these methods for accepting or denying a certain sharing request previously saved in the public ledger. The fact that a sharing request is accepted or denied is tracked in the public ledger. In case of denial, the previous sharing request passes from the pending state to the denied state. In case of approval, the biological data unit stored in the private data collection of the owner accepting the request is automatically copied to the private data collection of the organisation that requested the data unit. Then, the previous sharing request passes from the pending state to the accepted state. Figure 3c graphically represents the effects of the method `acceptRequest` on the ledger. It is worth noting that, after the request is accepted, the actual biological data unit owned by organisation A is copied into the private data collection of the organisation B, without passing through the

---

[8] Available at https://github.com/biochain-ai/biochain-ai/blob/main/biosharing/go/main.go.

public ledger. Thus, it will be accessible just by organisations A and B (through their private data collections).

5 CONCLUSIONS AND FUTURE WORKS

In this paper we have presented BIOCHAIN, a prototype of a blockchain-based platform for securely sharing biological data, allowing members of a consortium to securely share and exchange microbiological data. The project is publicly available at https://github.com/biochain-ai/biochain-ai. BIOCHAIN is still under development and what is presented in this paper should be intended as a prototype. Nevertheless, the current state of BIOCHAIN already allows to (i) securely store biological data in the blockchain, (ii) create sharing requests to make biological data available to a consortium, (iii) exchange biological data between blockchain participants, (iv) approve or deny sharing requests from other blockchain participants, and (v) keep track of the history of the exchanged biological data and requests.

Currently, our efforts are focusing on extending BIOCHAIN to offer more functionalities, in order to bring it from a prototype to a usable product. For instance, we are working on implementing the possibility for organisations to have multiple users in order to have accountability over everyone's actions. Over the presented blockchain, we are developing a decentralised application (DApp) usable mainly by biology scientists and employees of the involved institutions, namely, end-users that do not have expertise in blockchain technology. We are working on defining and implementing RESTfull (REpresentational State Transfer) APIs over BIOCHAIN, on which the DApp will be built. Finally, one feature that is missing in this prototype is the possibility for organisations to update or delete biological data units. We left this for future work since, once a biological data unit is deleted or updated, the other BIOCHAIN participants having a pending request should be notified, in order to guarantee data integrity and availability. A proper protocol for managing this scenario deserves further investigation.

ACKNOWLEDGMENTS

This project has been founded by the University of Parma (Italy), project number MUR_DM737_B_MAFI_BONNICI. V. Bonnici and A. Dal Palù are partially supported by INdAM-GNCS, project number CUP_E55F22000270001. V. Bonnici is also supported by the CINI InfoLife laboratory.

REFERENCES

[1] Elli Androulaki, Artem Barger, Vita Bortnikov, Christian Cachin, Konstantinos Christidis, Angelo De Caro, David Enyeart, Christopher Ferris, Gennady Laventman, Yacov Manevich, Srinivasan Muralidharan, Chet Murthy, Binh Nguyen, Manish Sethi, Gari Singh, Keith Smith, Alessandro Sorniotti, Chrysoula Stathakopoulou, Marko Vukolic, Sharon Weed Cocco, and Jason Yellick. 2018. Hyperledger fabric: a distributed operating system for permissioned blockchains. In *Proceedings of the Thirteenth EuroSys Conference, EuroSys 2018, Porto, Portugal, April 23-26, 2018*, Rui Oliveira, Pascal Felber, and Y. Charlie Hu (Eds.). ACM, 30:1–30:15. https://doi.org/10.1145/3190508.3190538

[2] Shaoqi Chen, Bin Duan, Chenyu Zhu, Chen Tang, Shuguang Wang, Yicheng Gao, Shaliu Fu, Lixin Fan, Qiang Yang, and Qi Liu. 2022. Privacy-preserving integration of multiple institutional data for single-cell type identification with scPrivacy. *Science China Life Sciences* (2022), 1–13.

[3] Shaoqi Chen, Dongyu Xue, Guohui Chuai, Qiang Yang, and Qi Liu. 2020. FL-QSAR: a federated learning-based QSAR prototype for collaborative drug discovery. *Bioinformatics* 36, 22-23 (2020), 5492–5498.

[4] Xu Cheng, Fulong Chen, Dong Xie, Hui Sun, and Cheng Huang. 2020. Design of a secure medical data sharing scheme based on blockchain. *Journal of medical systems* 44, 2 (2020), 52.

[5] Tom Dedeurwaerdere, Paolo Melindi-Ghidi, and Arianna Broggiato. 2016. Global scientific research commons under the Nagoya Protocol: Towards a collaborative economy model for the sharing of basic research assets. *Environmental science & policy* 55 (2016), 1–10.

[6] Pietro Ferrara, Luca Negrini, Vincenzo Arceri, and Agostino Cortesi. 2021. Static analysis for dummies: experiencing LiSA. In *SOAP@PLDI 2021: Proceedings of the 10th ACM SIGPLAN International Workshop on the State Of the Art in Program Analysis, Virtual Event, Canada, 22 June, 2021*, Lisa Nguyen Quang Do and Caterina Urban (Eds.). ACM, 1–6. https://doi.org/10.1145/3460946.3464316

[7] Loveleen Gaur, Arun Solanki, Samuel Fosso Wamba, and Noor Zaman Jhanjhi. 2021. *Advanced AI techniques and applications in bioinformatics.* CRC Press.

[8] Justin Guinney and Julio Saez-Rodriguez. 2018. Alternative models for sharing confidential biomedical data. *Nature biotechnology* 36, 5 (2018), 391–392.

[9] Manzour Hernando Hazbón, Leen Rigouts, Marco Schito, Matthew Ezewudo, Takuji Kudo, Takashi Itoh, Moriya Ohkuma, Katalin Kiss, Linhuan Wu, Juncai Ma, et al. 2018. Mycobacterial biomaterials and resources for researchers. *Pathogens and disease* 76, 4 (2018), fty042.

[10] Wouter Heyndrickx, Lewis Mervin, Tobias Morawietz, Noé Sturm, Lukas Friedrich, Adam Zalewski, Anastasia Pentina, Lina Humbeck, Martijn Oldenhof, Ritsuya Niwayama, et al. 2022. MELLODDY: cross pharma federated learning at unprecedented scale unlocks benefits in QSAR without compromising proprietary information. (2022).

[11] Tsung-Ting Kuo, Hyeon-Eui Kim, and Lucila Ohno-Machado. 2017. Blockchain distributed ledger technologies for biomedical and health care applications. *Journal of the American Medical Informatics Association* 24, 6 (2017), 1211–1220.

[12] Qinbin Li, Zeyi Wen, Zhaomin Wu, Sixu Hu, Naibo Wang, Yuan Li, Xu Liu, and Bingsheng He. 2021. A survey on federated learning systems: vision, hype and reality for data privacy and protection. *IEEE Transactions on Knowledge and Data Engineering* (2021).

[13] Tian Li, Anit Kumar Sahu, Ameet Talwalkar, and Virginia Smith. 2020. Federated learning: Challenges, methods, and future directions. *IEEE signal processing magazine* 37, 3 (2020), 50–60.

[14] Yu Li, Chao Huang, Lizhong Ding, Zhongxiao Li, Yijie Pan, and Xin Gao. 2019. Deep learning in bioinformatics: Introduction, application, and perspective in the big data era. *Methods* 166 (2019), 4–21.

[15] Samuel D Okegbile, Jun Cai, and Attahiru S Alfa. 2022. Performance Analysis of Blockchain-Enabled Data-Sharing Scheme in Cloud-Edge Computing-Based IoT Networks. *IEEE Internet of Things Journal* 9, 21 (2022), 21520–21536.

[16] Luca Olivieri, Fabio Tagliaferro, Vincenzo Arceri, Marco Ruaro, Luca Negrini, Agostino Cortesi, Pietro Ferrara, Fausto Spoto, and Enrico Talin. 2022. Ensuring determinism in blockchain software with GoLiSA: an industrial experience report. In *SOAP '22: 11th ACM SIGPLAN International Workshop on the State Of the Art in Program Analysis, San Diego, CA, USA, 14 June 2022*, Laure Gonnord and Laura Titolo (Eds.). ACM, 23–29. https://doi.org/10.1145/3520313.3534658

[17] Henning Perl, Yassene Mohammed, Michael Brenner, and Matthew Smith. 2014. Privacy/performance trade-off in private search on bio-medical data. *Future Generation Computer Systems* 36 (2014), 441–452.

[18] Gerard Verkley, Giancarlo Perrone, Mery Piña, Amber Hartman Scholz, Jörg Overmann, Aurora Zuzuarregui, Iolanda Perugini, Benedetta Turchetti, Marijke Hendrickx, Glyn Stacey, et al. 2020. New ECCO model documents for Material Deposit and Transfer Agreements in compliance with the Nagoya Protocol. *FEMS microbiology letters* 367, 5 (2020), fnaa044.

[19] Qingyong Wang and Yun Zhou. 2022. FedSPL: federated self-paced learning for privacy-preserving disease diagnosis. *Briefings in Bioinformatics* 23, 1 (2022), bbab498.

[20] Zhiyuan Wang, Zhiqiang Zheng, Wei Jiang, and Shaojie Tang. 2021. Blockchain-enabled data sharing in supply chains: Model, operationalization, and tutorial. *Production and Operations Management* 30, 7 (2021), 1965–1985.

[21] Qiang Yang, Yang Liu, Tianjian Chen, and Yongxin Tong. 2019. Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)* 10, 2 (2019), 1–19.

[22] Huadi Zheng, Haibo Hu, and Ziyang Han. 2020. Preserving user privacy for machine learning: local differential privacy or federated machine learning? *IEEE Intelligent Systems* 35, 4 (2020), 5–14.