

# Digital Transformation and Governance Innovation for Public Biobanks and Free/Libre Open Source Software Using a Blockchain Technology

Citation for published version (APA):

Evangelatos, N., Upadya, S. P., Venne, J., Satyamoorthy, K., Brand, H., Ramashesha, C. S., & Brand, A. (2020). Digital Transformation and Governance Innovation for Public Biobanks and Free/Libre Open Source Software Using a Blockchain Technology. *OMICS-a journal of Integrative Biology*, 24(5), 278-285. <https://doi.org/10.1089/omi.2019.0178>

## Document status and date:

Published: 01/05/2020

## DOI:

[10.1089/omi.2019.0178](https://doi.org/10.1089/omi.2019.0178)

## Document Version:

Publisher's PDF, also known as Version of record

## Document license:

Taverne

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

Download date: 24 Apr. 2024

# Digital Transformation and Governance Innovation for Public Biobanks and Free/Libre Open Source Software Using a Blockchain Technology

Nikolaos Evangelatos,<sup>1-3,\*</sup> Sudhakara P. Upadya,<sup>4,\*</sup> Julien Venne,<sup>5</sup> Kapaettu Satyamoorthy,<sup>6</sup> Helmut Brand,<sup>5,7</sup> C.S. Ramashesha,<sup>4</sup> and Angela Brand<sup>3,6,8</sup>

## Abstract

Digitalization and digital health are transforming research practices, while economic growth is increasingly driven by the information commons. In the case of biological sciences, information commons, such as public biobanks and free/libre open source software (FLOSS), are of paramount importance for both research and the bioeconomy. In a time of digitalization, however, information commons are vulnerable to violations, such as the free-rider problem, that render the commons unsustainable. Consequently, it has been argued that the enclosure of the informational common resources is the only means to effectively exploit them. Given the social and economic importance of the information commons, the new digital environment in biology and health requires governance innovation that will regulate the social embedding of the commons and their relationship to the free market, that is, a new political economy is needed. In this context, the need for a core common infrastructure, stretching from the physical to the logical and content layer of the information environment, that will guarantee the protection of the commons from both violations and enclosures, has been highlighted. Focusing on the interaction between two biological/bioinformatics commons, namely public biobanks and the FLOSS, we have set up an ecosystem relying on a blockchain technology. The proposed governance mechanism protects the information commons from the free-rider problem and guarantees their sustainability without hampering their operational framework. Our model demonstrates the interdependence and protection of the information commons not as an abstract theoretical exercise, but rather as a physical reality on the digital ontological matrix.

**Keywords:** blockchain, digitalization, free/libre open source software, public biobanks, political economy, the commons

## Introduction

**D**IGITALIZATION TRANSFORMS HUMAN ENVIRONMENTS, business endeavors, and practices (Brennen and Kreiss, 2016). In the specific case of the bioeconomy, the significance of digitalization largely relies on (and is exemplified

by) its function as the enabler for the emergence of Big Data, mainly in the form of “-omics” data that can be separately stored, analyzed, shared, and exchanged over public or private research infrastructures known as biobanks.

Public biobanks are platforms where researchers from around the world can freely upload and/or study and analyze

<sup>1</sup>Intensive Care Medicine Unit, Department of Respiratory Medicine, Allergology, and Sleep Medicine, Paracelsus Medical University (PMU), Nuremberg, Germany.

<sup>2</sup>Dr. TMA Pai Endowment Chair in Research Policy in Biomedical Sciences and Public Health, Prasanna School of Public Health (PSPH), Manipal Academy of Higher Education, Manipal, India.

<sup>3</sup>UNU-MERIT (Maastricht Economic and Social Research Institute on Innovation and Technology), Maastricht University, Maastricht, The Netherlands.

<sup>4</sup>School of Information Sciences, Manipal Academy of Higher Education, Manipal, India.

<sup>5</sup>Department of International Health, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands.

<sup>6</sup>Department of Biotechnology, School of Life Sciences, Manipal University, Manipal, India.

<sup>7</sup>Prasanna School of Public Health (PSPH), Manipal Academy of Higher Education, Manipal, India.

<sup>8</sup>Public Health Genomics, School of Life Sciences, Manipal Academy of Higher Education, Manipal, India.

\*These two authors have equally contributed.

information in the form of meaningfully aggregated “-omics” data. Of importance, the operational framework of public biobanks, that allows free access to informational shared resources, renders them, in essence, information commons arrangements (Boggio, 2017).

Typically, information commons consists of a shared informational resource (i.e., a form of knowledge), the actors that manage it (i.e., the commoners), and the rules and social norms that determine its function, that is, information commons is the institutional arrangement of all these elements. For instance, in the case of the International Nucleotide Sequence Database, the biological information in the form of sequences is the shared resource and the researchers who upload/use the sequences are the commoners. The operating framework is the one providing free and unrestricted access to all sequence data records without fees, restrictions, or licensing requirements placed on the redistribution or use of the data (INSDC, 2019).

The generation of “-omics” data and their further analysis (e.g., over public biobanks) relies extensively on the utilization of free/libre open source software (FLOSS). Of note, researchers using and/or modifying a FLOSS are also forming (even if unconscious of it) an information commons. In this case, the researchers using and/or modifying the FLOSS are the commoners, the source code of the FLOSS is the shared resource, and the operating framework is the one determined by the corresponding license (Evangelatos et al., 2018). For the scope of this article, we will abstractly consider licenses, such as the General Public License, that provide free access to the source code along with the freedom to use, copy, study, and voluntarily modify it (GNU General Public License, 2007).

Of importance, each time a researcher (a commoner) modifies the source code of a FLOSS, she/he adds knowledge in the form of the modifications. This new knowledge builds on the existing knowledge “deposited” in the previous versions of the FLOSS. Theoretically, every time a FLOSS is modified, the researcher who has taken advantage of the previous knowledge and added her/his own is morally obliged to make the modified version publicly known and freely available to every member of the community that constitutes this certain commons, that is, to the other commoners.

Indeed, in accordance with the terms of the corresponding general public license, the right to modify the source code is predicated on the moral obligation to make the new version publicly and freely available. However, abiding on this obligation depends exclusively on the ethos of the people who use and modify the FLOSS, and violations in the form of free riding may occur. More specifically, in cases where a modification has taken place but not been released to the public, the party who has modified the FLOSS has capitalized on the previous efforts of the other commoners (who have developed it up to this point) without, however, letting the community benefit from her/his contribution. Although informational resources, as opposed to natural ones, cannot be depleted, free riding has clearly occurred, threatening the sustainability of information commons (Hardin, 1968).

In the specific case of FLOSS, that would mean that the commoners might refrain from making their modifications publicly and freely available because they would not be willing to allow free riding on the shared resource (that reflects the collective effort) by third parties. Such developments would undermine the commons and threaten their sustainability. As a result, many have called for the enclosure

of the informational common resources as the only means to effectively exploit them.

In the case of natural commons, as Elinor Ostrom has brilliantly demonstrated, the regulation of the actions of the commoners based on horizontal, nonhierarchical, self-regulation models may deter free riding, thus resulting in sustainable development and the preservation of the commons (Hess and Ostrom, 2007; Ostrom, 1990).

Correspondingly, in the case of information commons, the challenge is to construct a regulating mechanism that will prevent free riding without changing the core operational framework of the commons, that is, leaving the informational common resource free for everyone to use and modify. In the case of FLOSS, this would mean to create a mechanism that prevents third parties from taking advantage of the collective effort, without however restricting access to the source code for use and/or modification.

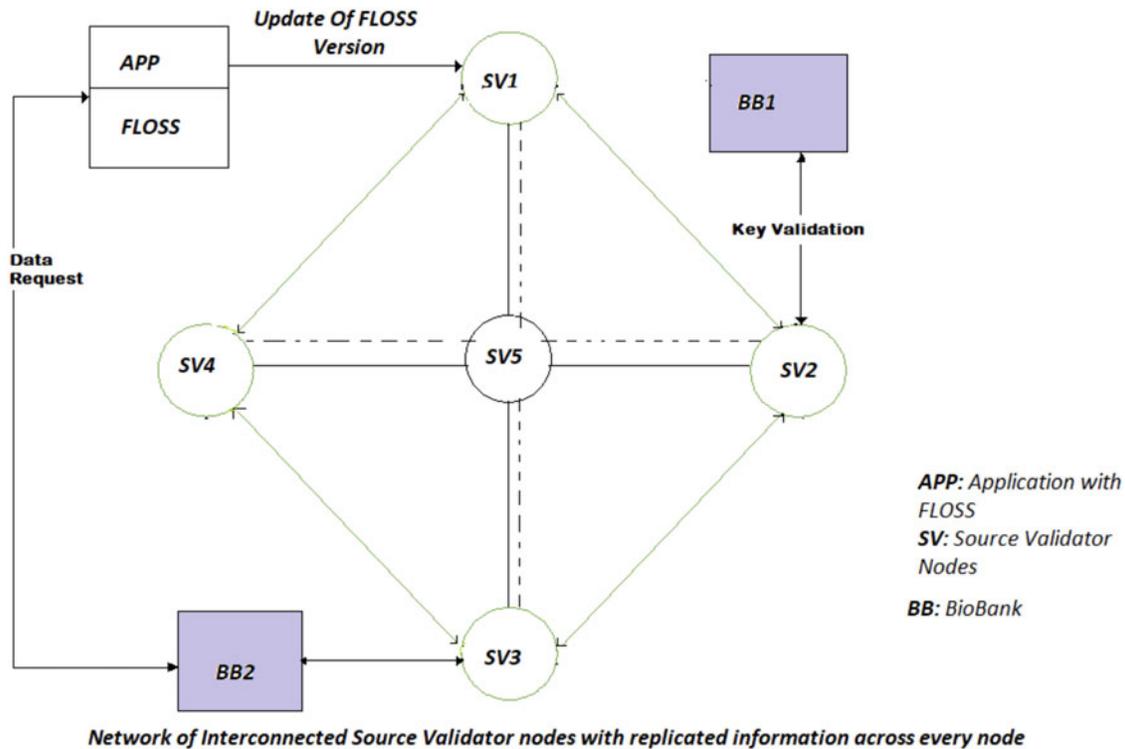
In essence, this is a data governance problem. In the past, many efforts have been made toward securing the optimal distribution of data among the members of the scientific community. A noteworthy initiative from the International Committee of Medical Journal Editors has led to the sharing of the de-identified individual patient data of clinical trials within 6 months after publication as a condition for publication in the member journals (Taichman et al., 2016). However, in the case of FLOSS, tying the disclosure of the (modified) source code to the publication of the results would not prevent a violator from using her/his new source code offline and disclosing an old version of the source code.

To circumvent such problems of trust one should employ a technology that can create networks and secure trust between parties without relying on their good will or altruism. Being in essence a combination of established technologies, a blockchain is a technology that registers the exchange of data (transactions) between users keeping a chronological order. It does so by utilizing a distributed ledger technology and cryptography to verify and store information among members of a network often in the absence of a central regulatory authority. There is not just one type of blockchain and the various types differ on the level of openness (public or private), the level of the required permissions in order to be able to add information to the chain, and a number of other variable features. Blockchain has already started reshaping the world of transactions, changing the landscape in many sectors including health care (Berryhill et al., 2018).

Focusing our empirical lens on the interaction between two information biological/bioinformatic commons, that is, public biobanks and FLOSS, we show how the interaction of information commons arrangements with each other can protect them from the free-riding problem, thus securing their sustainability. To this end, we have modeled an ecosystem using a decentralized blockchain technology that protects the information commons from the free-riding problem, guarantees their sustainability, and regulates their social embedding within and along free market models.

## Materials and Methods

The proposed ecosystem (Fig. 1) consists of four components: the FLOSS (library), the applications that use FLOSS, the biobanks that have the data requested from the applications, and the source validators. The latter act as the



**FIG. 1.** Regulatory system for usage of FLOSS and biobank data built using a network of source validator nodes. FLOSS, free/libre open source software.

computing nodes of the blockchain: they keep track of the different versions of the FLOSS maintained by the various applications and provide the validation check on whether the version of the FLOSS used by an application is registered with the ecosystem. In essence, the source validators maintain the ledger of the different versions of FLOSS. When a random node is updated with a new version, the node distributes the information to all other nodes in the network. This ensures that the ledger of FLOSS versions is in synch across every node in the regulatory network.

To identify the version of the FLOSS, the system generates a unique key based on the compiled code of the library. This unique key is considered the identification key of the FLOSS used by a certain application. Any change to the source code (even addition of a single comment) will result in the generation of a different key for the library. Each time an application that utilizes the FLOSS services requires access to data from a biobank that belongs to the ecosystem, it is required to provide the identification key of the employed FLOSS to the biobank.

The application passes the compiled version of the FLOSS to the biobank along with its request for data. The biobank, in turn, forwards the FLOSS details to the most readily available source validator for verification. The latter computes the key based on the incoming FLOSS version, compares the key with its ledger of FLOSS, and determines whether the current FLOSS version is already registered or not.

If the source validator identifies the key, then the biobank allows access to the requested data.

In case the FLOSS key is not identified as already registered by the source validator, the biobank requests from the

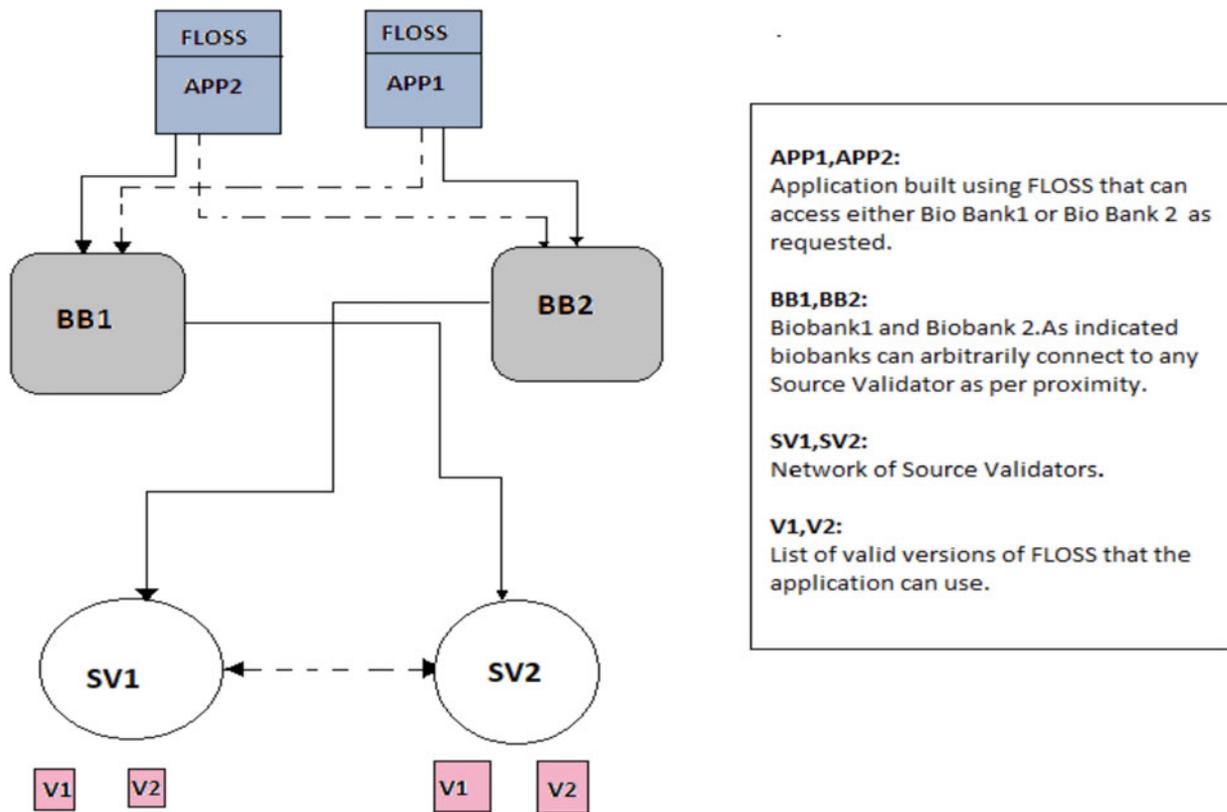
application to register the new version (i.e., to provide the modified source code of the updated FLOSS) with the source validator node. The latter compiles the newly uploaded source code, computes a new key, updates its FLOSS ledger, and distributes the updated information to every node in the network. The application now requests access to the biobank data again, and, because the current FLOSS version is already updated at the source validator ledger, access to data is granted.

## Results

### Experimental set up

To illustrate the regulatory concept between the two commons (biobank and FLOSS) using the network of source validator nodes, we have implemented a prototype as a proof of concept (Fig. 2). The prototype consists of the following components:

1. A library (represents a FLOSS) that has a method `computeData()`.
2. A data provider (represents a biobank) that has the method `getDatavalue` (encoded FLOSS source file) to allow data access to the calling application.
3. The data provider calls the Source Validator Node that has the keys of all registered versions of the Library and also provides the lookup mechanism to check whether an identification key corresponds to a registered version of FLOSS. The Source Validator further provides the end user with the option to upload a new version of the library.



**V1** - Updated at Source validator1, replicated across other nodes.

**V2** - Updated at Source validator2, replicated across other nodes.

**FIG. 2.** Experimental set up of the prototype.

4. An application 1 and an application 2 that utilize the library for online work with data provided by the provider (biobank).
5. In the experimental setup, we show Biobank 1 connecting to Source Validator 2 to illustrate that each biobank may freely contact any source validator available in the network. Correspondingly, the applications are also freely configured to contact either Biobank 1 or Biobank 2 based on the application's data needs.

#### *Design of the prototype*

Application 1 and Application 2 use a library for their functionality implementation. The prototype uses the library's `computeData()` method to find the sum of the incoming data sequence. Two data sources are utilized for illustration purposes, representing two different biobanks. In the prototype, Data Provider 1 returns the sequence of "5.52, 6.19" and Data Provider 2 returns the sequence "105.52, 800.19." The application can contact any of the data providers and request access to their data. In the prototype, the application will contact Data Provider 1 if type entered is B1 and Data Provider 2 if type entered is B2.

Before gaining access to the data, the application needs to get the version identifier key of the library it is using. In this

prototype, we are using an MD5 digest of the executable file of the library as the version identifier key of the library. The MD5 is a message-digest algorithm initially designed to be used as a cryptographic hash function. The MD5 function works on the principles of cryptography (Rivest, 1992) and produces the 128-bit message digest or fingerprint of the input supplied to it. The MD5 algorithm is intended for the digital signature of the applications. In the prototype, the MD5 digest is used to identify a particular version of the library. Any changes to the library will result in a different MD5 digest.

The application gets the source version key value by contacting the data provider. The class file of the library is sent to the data provider in the form of a Base64 Encoded String. The string is decoded and an MD5 Hash code is generated for the .class file. The generated key of the library is passed along with the request to get access to the data of the provider. The latter provides access to the application if the library's source code version is recognized by the source validator. The applications need to register any changes to the source code of the library to the repository maintained at the source validator node(s) in order to get access to the biobank data. The data provider checks if the library used by the application is registered with the source validator by passing the key to the latter. Once the source validator confirms the library version (by comparing the library's key to the existing ledger), the data provider allows access to the requested data.

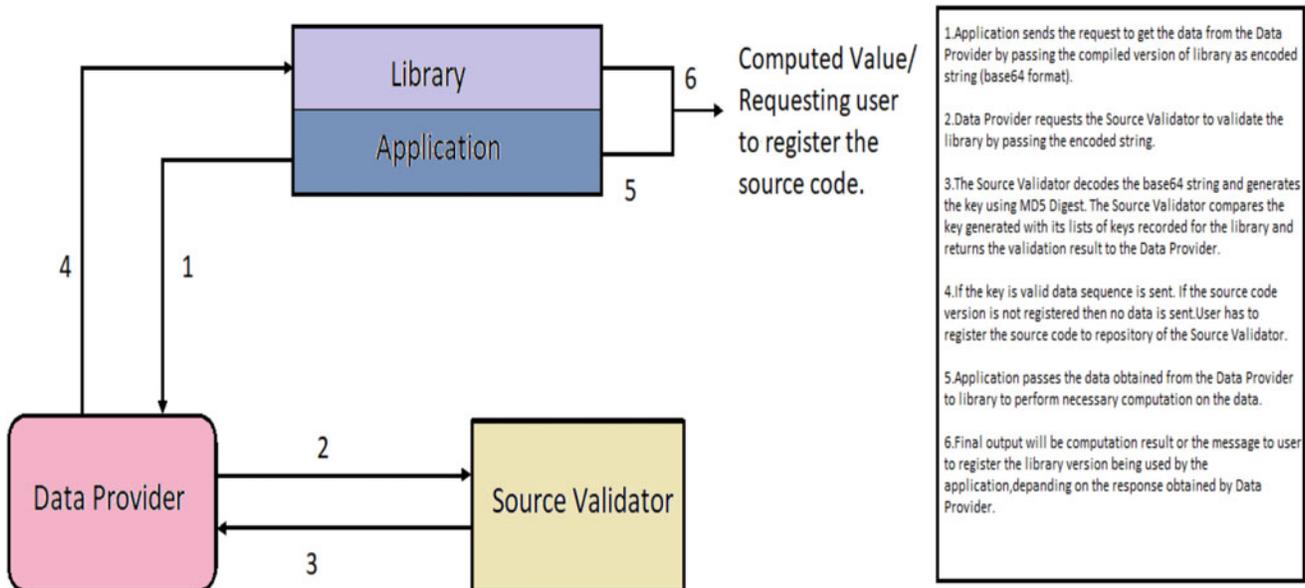


FIG. 3. Workflow of the prototype.

There are two versions of the library used for the illustration in the prototype. Each of them computes the sum of the data obtained by the data provider. The calculation of the sum represents the computation that can be applied to the incoming data sequence by FLOSS in a real scenario. The library version 1 computes the data obtained by the data provider to bind it to a two-digit precision output value. The second version of the library computes the data obtained and binds it to a single-digit precision output value. The sequence of activities of the prototype is given in Figure 3 (for more information, see Supplementary Data).

## Discussion

It has been argued that, in the current Age of Information and Communication, “technoscience” reconstructs itself based on an informational, rather than a material ontology (Forman, 2007). Consequently, it is informational and not natural resources that have allowed for a new view of the natural and artificial phenomena, ranging from “...black holes to DNA, from the organization of cells to the processes of human thought, and from the management of corporations to the allocation of global resources” (Mahoney, 1988). This new “Weltanschauung,” based on digitalization, has monumental economic and social implications and has practically paved the way to the current knowledge economy (Powell and Snellman, 2004).

In this article, we have shown how a blockchain technology can be used in cases where information commons interact with each other over digital networks, protecting the former from the free-riding problem and securing their sustainability.

More specifically, in the case of the biological sciences, digitalization and informational resources have been changing the economic landscape toward a bioeconomy “[...] based on the use of research and innovation in the biological sciences to create economic activity and public benefit” (National Bioeconomy Blueprint, 2012). Importantly, the

bioeconomy differs substantially from the traditional free market system, as it relies to a great extent on commons arrangements of intangible, informational resources (Evangelatos, 2018).

For instance, precision medicine and the subsequent economic growth are to a great extent based on research platforms consisting of information commons, such as public biobanks and FLOSS, and is a prominent example of the relevance of such commons arrangements for the bioeconomy (Huang et al., 2017). The increasingly emerging health data cooperatives, built around biodata owned and governed by the citizens, are also vivid examples of the role of information commons arrangements both on the economic and social level (van Roessel et al., 2017).

In this context, and although the concept of the commons initially referred to the exploitation and management of common natural resources, the similarities between the natural and information commons have been highlighted and enthusiastically (although sometimes uncritically) embraced (Hess and Ostrom, 2003).

Not surprisingly then, in an analogy to the first enclosure and the privatization of common natural resources that took place in (pre)industrial England (Yelling, 1977), appropriation and commodification of informational common resources nowadays have led to what is known as the second enclosure movement (Boyle, 2003).

The history of the first enclosure has fueled a fierce intellectual debate between those who consider the enclosure to have propelled innovation and those who argue that this progress has occurred at the cost of social inequity (Bradley, 2001).

Nowadays, numerous scholars in the fields of intellectual property rights and economic history have argued for and against the sustainability of the information commons and many have advocated their enclosure because of the free-riding problem and the so-called tragedy of the commons (Travis, 2000). In other words, the free-riding problem is used as evidence by the advocates of privatization that the

commons are intrinsically unsustainable because some of the commoners free ride on the shared resource leading the commons to their demise.

Premised on the assumptions of economic rationality and the tenets of neoclassical liberalism, the free-riding problem emphasizes the risk of common resources depletion by individuals who act independently and rationally according to each one's self-interest, and who behave contrary to the whole group's long-term interests. Despite being primarily concerned with natural common resources, the argument has also set the basis for the contemporary repudiation of the information commons as "... too horrifying to contemplate" (Hardin, 1968).

The argument is not without merit. In sharp contrast to natural ones, informational resources cannot be over-exploited and depleted. However, free riding produces social inequality and may, therefore, discourage commoners from contributing to the commons. Perhaps even more importantly, such developments would undermine the economy as a whole, because commons arrangements and classical economic systems coexist in a rather symbiotic relationship. This is exemplarily demonstrated by the case of the Linux operating system, which in the 1990s was the sole alternative to the digital monopoly of Microsoft.

Although the Linux kernel was licensed under the General Public Licence (GPL), already in the early 1990s nearly 85% of the Linux code was written by employees of major private concerns. These private companies clearly benefited from the effort put from other private companies and from individual contributors to the improvement of the shared resource (Vercellone et al., 2015). Moreover, a multitude of private companies has adopted open-source business models, where the generation of commercial value shifts away from the actual products and derives from the product "halo." For instance, by providing a range of paid services (e.g., installation, customization, and training of personnel) around the freely available Linux system, private companies present impressive total revenues (RedHat, 2018).

It may thus be argued that in the current economic environment, where value derives increasingly from informational resources, the privatization of the information commons would have negative consequences for the economy. Despite these alarming concerns, we are currently witnessing a reconfiguration of the social embedding of the new informational environment in the direction of increasingly enclosed informational resources (Travis, 2000). This may have profound consequences for the bioeconomy because of its high reliance on informational resources. Indeed, enclosing informational resources is anticipated to hamper innovation and growth in the bioeconomy through an anti-commons effect (Gold et al., 2007).

In this context, and based on the similarities between the governance of common natural resources and the governance of information commons, relevant scholarship has pointed out the need for a politics of the information age (Frischmann et al., 2014). In other words, a new political economy that will regulate the social embedding of the commons and their relationship to the free market system is needed. Specific policy measures in that sense would include a core common infrastructure, that would "... stretch from the very physical layer of the information environment to its logical and content layers" (Benkler, 2003).

Elsewhere, we have demonstrated how information commons arrangements can be used for the promotion of innovation through the facilitation of clinical trial data sharing among private biopharmaceutical companies. By implementing a single excludability criterion (*quid pro quo*), we have treated knowledge (in the form of meaningfully aggregated "-omics" data) as a club good and developed a commons arrangement symbiotic to the free market operational models (Evangelatos et al., 2016).

Here, we sought to address the free-riding problem by providing a regulatory mechanism that protects the information commons without, however, interfering with their core operational framework, that is, without imposing operational restrictions on the interaction between the commoners and the shared resource. Rather, we have focused on the interaction between independent commons as wholes and we have built an ecosystem consisting of the two interacting commons, namely the public biobanks and the applications using FLOSS, and the regulatory mechanism. The latter comprises a network of interconnected source validators that provide notary and authentication services. To this end, we have utilized a blockchain technology that serves as a "physical" intermediate layer at the interface between interacting commons.

In the context of informational common resources, the problem of free riding is in principle a data governance problem. Blockchain-based platforms have already been used as decentralized approaches to the governance problems of data management and sharing (Shabani, 2019). Of interest, these technologies may be the answer to the problem of the so-called digital "toxicity" (Özdemir, 2018).

Recently, blockchain has been used to ensure the integrity and nonrepudiation of transactions between data consumers (humans or programs) and biomedical databases. By using small digital contracts, the blockchain application guarantees that the retrieved data cannot be modified by either the data consumer or the database (Kleinaki et al., 2018). More specifically, the application ensures both the integrity and nonrepudiation of the data that is acquired from a biobank. The notary service records the query, and the resultant data are being added by the biobank to the existing blockchain as a new block so that the related information can be recalled in the future for any query on the transaction.

In our case, the blockchain supports the functionality of the ecosystem through the application of a controlling unit, namely the source validator. The prototype is developed on a network module using the web services for the regulation of the interactions between the various components. The source validator and the data provider services are deployed on an internet accessible system and tested for proper functioning. In our model, the source validators are the nodes that provide the required processing power and generate the new keys while at the same time, they maintain the ledger of the transactions (i.e., the updated versions of FLOSS). This way we secure that violations of the operational premises of the information commons cannot occur.

The technology resembles a blockchain in that it consists of interconnected, synchronized nodes that generate keys (allowing the interaction between biobanks and FLOSS) and maintain a ledger of values (in the form of the updated versions of FLOSS) using a peer-to-peer network.

More specifically, the regulating mechanism in our system is implemented as a network of individual validating units that perform the same tasks and have common information regarding the FLOSS versions. The individual unit of the source validator is similar to a node of the traditional blockchain and the information regarding the different versions of FLOSS corresponds to the information blocks. Furthermore, each validating node in the system is connected to every other node as in the classical blockchain. This interconnection ensures that any new version of the FLOSS (new block of information in the blockchain) updated at any node is immediately synchronized across all the other nodes maintaining the shared ledger concept of the FLOSS versions.

Implementing the source validator in the lines of blockchain principles provides the system with the required redundancy, thus avoiding single points of failure. Furthermore, it offers a performance benefit by allowing the biobanks to contact any available validating unit closer to its location. Of importance, this setup further gives us the flexibility of implementing regulatory operations using publicly available blockchain implementation models such as Ethereum.

Although many different tools for software versioning exist (such as Git), our model has far-reaching consequences as it provides a generic scaffold for the functional regulation of information commons. However, it also has limitations. First, it is functional only in cases where FLOSS is employed for exclusively online use of data stored in biobanks. In case the data are downloaded to private servers, then it would be possible for the violator to use an old version of the FLOSS to request and download the data, and then use the modified version offline. This way the violator would have evaded the controlling mechanism.

To avoid this elusion, a mechanism may be introduced, whereby the data from the biobank are sent in an encrypted form. In this case, the source validator would send the key to the biobank based on the FLOSS version that was passed in; the biobank would then encrypt the data using the encoded key and then forward it to the application. The computing function of the FLOSS would now have to first decrypt the data before the application can use it. The decryption service would be built into the source validator. The computed method of FLOSS would then pass the encrypted data along with the current class version to the decryption service.

Finally, the source validator would determine the key from the class code that was passed in, apply the encoding algorithm on the key, and then try to decrypt the data that would have been passed in. Decryption thus would only be possible if the two FLOSS versions (the one used to request the data and the other performing the computation) are the same.

As far as the design of the prototype is concerned, we have used MD5 although it has been cryptographically broken for more than a decade. Indeed, there are other approaches, such as for example, SHA-2/3, Keccak, or Blake2b, that could be used. However, every new technique will probably also be broken in the future rendering the technical part not the most important aspect of our work, which primarily focuses on the reconceptualization of the information commons and their protection.

A second limitation is that the model presupposes a readiness of the public biobanks to become members of the

ecosystem and allow for the source validators to check their “clients”. Finally, public or private research organizations should provide the necessary infrastructure (in the form of servers with adequate processing power and storing capacity) that would constitute the hardware scaffold (in the form of the source validators) of the ecosystem.

Despite these limitations, we believe that our model offers a horizontal, nonhierarchical, self-regulatory mechanism utilizing a blockchain technology for the protection of the commons. This technology is not limited to the biological sciences and the bioeconomy, but it could be used in cases where information commons interact with each other over digital networks, serving as a core common infrastructure. In the case of health data cooperatives and similar social arrangements/contracts, where citizens own and govern their biodata, this technology could be implemented to “re-design” our societal democratic models and the very role of citizens. In this respect, it has far-reaching implications as it could serve as a generic law in the new political economy of the digital era.

## Conclusions

A commons is the economically relevant institutional arrangement of human agents around a shared resource and, as such, an integrated socioeconomic paradigm. Information commons arrangements enabled by digitalization are of paramount importance for the bioeconomy. However, a major critique has been that information commons are vulnerable to violations that threaten their sustainability, and that the only way to effectively exploit informational resources is their enclosure. By employing a blockchain technology at the interface between two interacting commons, that is, public biobanks and FLOSS, we have developed an ecosystem that protects the information commons from the free-riding problem and guarantees their sustainability without hampering their operational framework.

Our model demonstrates the interdependence of the commons not as an abstracted theoretical exercise, but rather as a physical reality that leads to the transformation of the research, economic and social landscape in the biological sciences. This technology could be used in cases where information commons interact with each other over digital networks, such as in the case of health data cooperatives and similar social arrangements, where citizens own and govern their biodata. In this respect, it has far-reaching implications as it could serve as a generic law in the new political economy of the digital era.

## Acknowledgments

The authors thank Andria Pazarloglou, MSc and Efthymios Thanasoulis for their valuable insights.

## Author Disclosure Statement

The authors declare they have no conflicting financial interests.

## Funding Information

No funding was received for this article.

## Supplementary Material

Supplementary Data  
 Supplementary Figure S1  
 Supplementary Figure S2  
 Supplementary Figure S3

## References

- Benkler Y. (2003). The political economy of commons. *Eur J Inform Prof* 4, 6–9.
- Berryhill J, Bourgerly T, and Hanson A. (2018). Blockchains unchained: Blockchain technology and its use in the public sector. *OECD Working Papers on Public Governance*, No. 28. Paris, France: OECD Publishing, <https://doi.org/10.1787/3c32c429-en> Accessed January 31, 2020.
- Boggio A. (2017). Population biobanks' governance: A case study of knowledge commons. In: *Governing Medical Knowledge Commons. Cambridge Studies on Governing Knowledge Commons*. Strandburg K, Frischmann B, and Madison M, eds. Cambridge, United Kingdom: Cambridge University Press, 102–120.
- Boyle J. (2003). The second enclosure movement and the construction of the public domain. *Law Contemp Probl* 66, 33–74.
- Bradley H. (2001). *The Enclosures in England. An Economic Reconstruction*. Kitchener: Batoche Books Limited.
- Brennen JS, and Kreiss D. (2016). Digitalization. In: *The International Encyclopedia of Communication Theory and Philosophy*. Jensen KB, Rothenbuhler EW, Pooley JD, and Craig RT, eds. Wiley-Blackwell, United Kingdom.
- Evangelatos N. (2018). *Digitalization, Big Data and the Commons: How '-Omics' Technologies Transform the Research, Economic, and Social Landscape in Biomedicine*. Maastricht, The Netherlands: Gildeprint.
- Evangelatos N, Reumann M, Lehrach H, and Brand A. (2016). Clinical trial data as public goods: Fair trade and the virtual knowledge bank as a solution to the free rider problem—a framework for the promotion of innovation by facilitation of clinical trial data sharing among biopharmaceutical companies in the era of omics and big data. *Public Health Genomics* 19, 211–219.
- Evangelatos N, Satyamourthy K, Levidou G, et al. (2018). Use of free/libre open source software in sepsis “-omics” research: A bibliometric, comparative analysis among the United States, EU-28 Member States, and China. *OMICS* 22, 365–372.
- Forman P. (2007). The primacy of science in modernity, of technology in postmodernity, and of ideology in the history of technology. *History Technol* 23, 1–152.
- Frischmann BM, Madison MJ, and Strandburg KJ. (2014). *Governing Knowledge Commons*. Oxford University Press, New York.
- GNU General Public License. (2007). <http://webcitation.org/76n6fNbMW> Accessed January 31, 2020.
- Gold ER, Herder M, and Trommetter R. (2007). The role of biotechnology intellectual property rights in the bioeconomy of 2030. Report prepared for OECD International Futures Program. OECD, Paris, France.
- Hardin G. (1968). The tragedy of the commons. The population problem has no technical solution; it requires a fundamental extension in morality. *Science* 162, 1243–1248.
- Hess C, and Ostrom E. (2003). Ideas, artifacts and facilities: Information as a common-pool resource. *Law Contemp Probl* 66, 111–145.
- Hess C, and Ostrom E. (2007). *Understanding Knowledge as a Commons: From Theory to Practice*. MIT Press, Cambridge, MA.
- Huang C, Mezencev R, McDonald JF, and Vannberg F. (2017). Open source machine-learning algorithms for the prediction of optimal cancer drug therapies. *PLoS One* 12, e0186906.
- INSDC. (2019). <http://webcitation.org/76n8BSev7> Accessed January 31, 2020.
- Kleinaki AS, Mytis-Gkometh P, Drosatos G, Efraimidis PS, and Kaldoudi E. (2018). A blockchain-based notarization service for biomedical knowledge retrieval. *Comput Struct Biotechnol J* 16, 288–297.
- Mahoney MS. (1988). The history of computing in the history of technology. *IEEE Ann Hist Comput* 10, 113–125.
- National Bioeconomy Blueprint. (2012). [https://whitehouse.gov/sites/default/files/microsites/ostp/national\\_bioeconomy\\_blueprint\\_april\\_2012.pdf](https://whitehouse.gov/sites/default/files/microsites/ostp/national_bioeconomy_blueprint_april_2012.pdf) Accessed June 9, 2016.
- Ostrom E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press, Cambridge, United Kingdom.
- Özdemir V. (2018). The dark side of the moon: The internet of things, industry 4.0, and the quantified planet. *OMICS* 22, 637–641.
- Powell Walter W, and Snellman K. (2004). The knowledge economy. *Annu Rev Sociol* 30, 199–220
- RedHat. (2018). <https://investors.redhat.com/financial-information/financial-statements> Accessed October 31, 2018.
- Rivest R. (1992). *RFC 1321—The MD5 Message-Digest Algorithm*. <https://www.ietf.org/rfc/rfc1321.txt> Accessed March 6, 2020.
- Shabani M. (2019). Blockchain-based platforms for genomic data sharing: A de-centralized approach in response to the governance problems? *J Am Med Inform Assoc* 26, 76–80.
- Taichman DB, Backus J, Baethge C, et al. (2016). Sharing clinical trial data: A proposal from the international committee of medical journal editors. *PLoS Med* 13, e1001950.
- Travis H. (2000). Pirates of the information infrastructure: Blackstonian copyright and the first amendment, 15 *Berkeley Tech. L.J.* 777. <https://doi.org/10.15779/Z382D6R> Accessed January 31, 2020.
- van Roessel I, Reumann M, and Brand A. (2017). Potentials and challenges of the health data cooperative model. *Public Health Genomics* 20, 321–331.
- Vercellone C, Bria F, Fumagalli A, et al. (2015). Managing the commons in the knowledge economy. <http://webcitation.org/76nCrjcD9> Accessed January 31, 2020.
- Yelling JA. (1977). *Common Field and Enclosure in England 1450–1850*. London, United Kingdom: Macmillan. <https://doi.org/10.1007/978-1-349-15797-6> Accessed January 31, 2010.

Address correspondence to:  
 Nikolaos Evangelatos, MD, PhD  
 Intensive Care Medicine Unit  
 Department of Respiratory Medicine,  
 Allergology, and Sleep Medicine  
 Paracelsus Medical University (PMU)  
 Prof. Ernst-Nathan-Strasse 1  
 Nuremberg 90419  
 Germany

E-mail: nikos.evangelatos@gmail.com

### Abbreviations Used

FLOSS = free/libre open source software