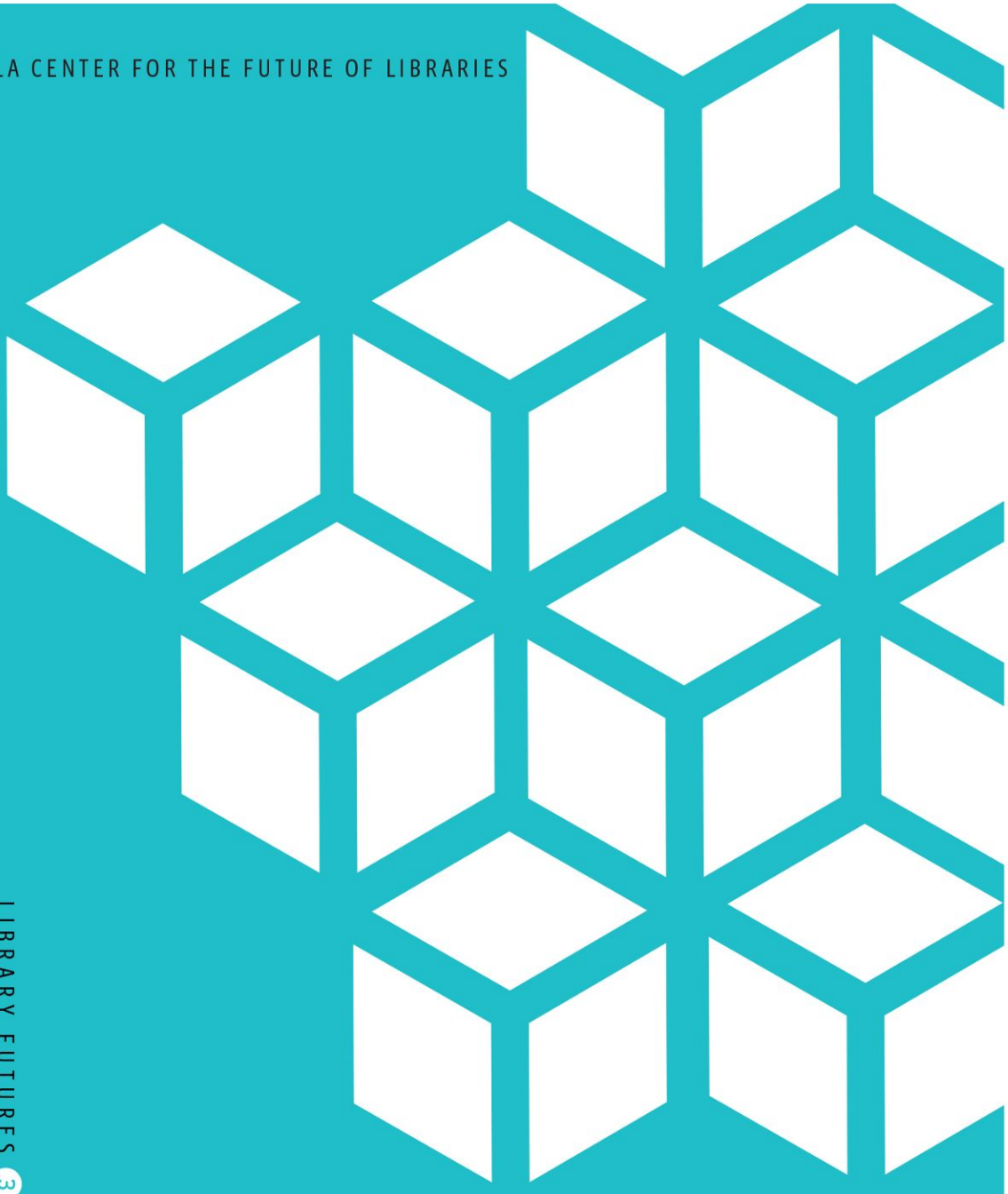
ALA CENTER FOR THE FUTURE OF LIBRARIES

LIBRARY FUTURES 3

BLOCKCHAIN

Edited by SANDRA HIRSH *and* SUSAN ALMAN

ALA CENTER FOR THE FUTURE OF LIBRARIES



LIBRARY FUTURES 3

BLOCKCHAIN

Edited by SANDRA HIRSH *and* SUSAN ALMAN

BOOKS IN THE LIBRARY FUTURES SERIES

Resilience by Rebekkah Smith Aldrich
Anonymity by Alison Macrina and Talya Cooper
Blockchain edited by Sandra Hirsh and Susan Alman
Design Thinking (forthcoming) by Rachel Ivy Clarke

ALA Neal-Schuman
CHICAGO | 2020

ALA CENTER FOR THE FUTURE OF LIBRARIES

BLOCKCHAIN

BLOCKCHAIN

EDITED BY SANDRA HIRSH AND SUSAN ALMAN

LIBRARY FUTURES 3

© 2020 by the American Library Association

Extensive effort has gone into ensuring the reliability of the information in this book; however, the publisher makes no warranty, express or implied, with respect to the material contained herein.

ISBNs
978-0-8389-1743-5 (paper)
978-0-8389-4682-4 (PDF)
978-0-8389-4681-7 (ePub)
978-0-8389-4683-1 (Kindle)

Library of Congress Control Number: 2019040312

Cover design by Kimberly Thornton. Image © lunar_cat/Adobe Stock.

ALA Neal-Schuman purchases fund advocacy, awareness, and accreditation programs for library professionals worldwide.

CONTENTS

Acknowledgments

Foreword by Miguel A. Figueroa

Introduction: An Investigation of Blockchain Applications: Beginnings and Implications by Sandra Hirsh and Susan Alman

UNDERSTANDING BLOCKCHAIN

From Medieval Origins to Modern Applications
CHRISTINA CORNEJO and STACEY JOHNSON

Blockchain and Decentralization: Big-Picture Opportunities and Risks
JASON GRIFFEY

Blockchain: Merits, Issues, and Challenges
BOHYUN KIM

BEFORE THE HYPE

Standards
TODD A. CARPENTER

Legal Considerations
DAN BLACKABY

Security
TONIA SAN NICOLAS-ROCCA

Healthy Skepticism
TOBY GREENWALT

FOR THE FUTURE: SPECULATIVE APPLICATIONS

Support for Scholarship and Scholarly Communications
MacKENZIE SMITH

Credentialing and Continuing Education
HEATHER McMORROW and AMY JIANG

Distributed Access to Library Metadata
TIMOTHY A. THOMPSON

Data Collection and Assessment
ANNIE NORMAN

Lessons from Health Information Management
VICTORIA LEMIEUX

Electronic Health Records
FRANK CERVONE

Born-Digital and Digital-First Content
JOHN BRACKEN and MICHAEL DELLA BITTA

Community-Based Collections
M. RYAN HESS

FOR THE PRESENT

Blockchain Education for Communities
LINK SWANSON

Conclusion
Selected Resources
About the Editors

ACKNOWLEDGMENTS

WE WANT TO THANK THE CONTRIBUTORS TO THIS PUBLICATION for their pioneering spirit, and for engaging in a national dialogue focused on the possibilities that blockchain technology can bring to the information professions. Collectively, these contributors have identified key issues and use cases that need further investigation, and their work has made our crystal ball a little clearer as we gaze into the emerging future.

FOREWORD

BY MIGUEL A. FIGUEROA
Center for the Future of Libraries
American Library Association

FOR ITS 2019 GLOBAL BLOCKCHAIN REPORT,[1] **DELOITTE SURVEYED** 1,386 senior executives in a dozen countries, finding that 53 percent of respondents view blockchain as a critical priority for their organizations and 83 percent see compelling use cases for blockchain in their sectors. Across news articles and reports, 2019 appears to have been the year when blockchain shifted from novel technology to something more viable that could provide solutions to problems across industries. And while this enthusiasm could be dismissed as part of a normal competitive business enthusiasm (the executives surveyed for Deloitte's report represented US companies with US$500 million in revenue or more and international companies with US$100 million or more), there is growing curiosity and interest in blockchain across healthcare, government, the art world, and telecommunications.

Blockchain grows out of Drs. Sandra Hirsh's and Susan Alman's experience with a project funded by the Institute of Museum and Library Services to investigate the possible uses of blockchain technology in libraries. What I appreciate about that project's process and eventual National Forum—and what I appreciate about this volume—is the mix of perspectives, enthusiasm, and skepticism that have been brought to this discussion. In so many ways, it represents a spirit of librarianship that recognizes the emergence of a new need or interest within the community and brings to bear a process of inquiry, reflection, and discernment to help understand libraries' roles in relation to that need. For those readers who bring an enthusiasm for technology generally or blockchain specifically, there will be much to find in this collection. For those who bring a healthy skepticism, so too, will there be contributions that reinforce and some that challenge our assumptions. And for those who hold an enthusiasm for libraries' ongoing development and adaptation, this collection provides an

accessible introduction to a topic that many people are discussing and that we would all do well to be more informed about.

In line with the previous volumes in this series, editors and contributors have done important work to not only frame the emerging trend and issue, but to also root their discussion in fundamental and timeless library roles, purposes, and values.

“Understanding Blockchain” helps readers establish the fundamentals of blockchain technology, the risks and opportunities, and how it fits into broader movements in technology. Jason Griffey contextualizes the innovative importance of blockchain, focusing on the potential user-facing services and systems that this decentralized technology might provide. Bohyun Kim promotes the positive potential for the technology while also tempering enthusiasm with considerations for interoperability, scalability, security, energy consumption, and the very basic concern that libraries may not yet have compelling use cases to justify investments of time and money.

“Before the Hype” presses pause on ambitions to consider some of the issues still being sorted out for the technology. In addition to Todd Carpenter’s consideration of standards, Dan Blackaby’s legal considerations, and Tonia San Nicolas-Rocca’s security concerns, Toby Greenwalt provides a very practical perspective on the value trade-off in time, training, and infrastructure for operations that may improve existing library services and practices or challenge issues of trust, storage, bandwidth, and cost.

“For the Future” brings together eight speculative use cases that might improve operations and experiences in libraries. The range of uses spans academic libraries, instruction, and programming (MacKenzie Smith on scholarly communication; Heather McMorrow and Amy Jiang on credentialing); content, collections, and the organization of information (Timothy A. Thompson on metadata; John Bracken and Michael Della Bitta on born-digital content); administration and promotion of libraries (Annie Norman on assessment; M. Ryan Hess on community-based collections); and medical libraries and health information (Victoria Lemieux, health information; Frank Cervone on health records). While these contributions outline very specific use cases, they also help prepare a framework for thinking more broadly about how blockchain could change operations and practices in libraries and in the sectors around us.

Finally, a brief “For the Present” offers Link Swanson’s thoughts for libraries’ roles in educating communities about blockchain’s possibilities.

Is a path forward with blockchain assured? Absolutely not. The conversation is still very much a debate. With this in mind, *Blockchain* is not meant to be a guide or manual, but a conversation starter. These contributions from experts and innovative thinkers will help strengthen our discussion and refine our prioritization for this technology.

1. https://www2.deloitte.com/content/dam/Deloitte/se/Documents/risk/DI_2019-global-blockchain-survey.pdf.

INTRODUCTION

An Investigation of Blockchain Applications: Beginnings and Implications

BY SANDRA HIRSH AND SUSAN ALMAN

BLOCKCHAIN HAS GENERATED A LOT OF BUZZ DUE TO THE potentially disruptive and transformative possibilities it offers. It is clear from the expanding literature that blockchain technology is on the brink of revolutionizing the public and private sectors. Conferences, books, white papers, start-ups, and numerous back-channel discussions have explored the ways blockchain technology can address seemingly endless processes. The extensive media coverage of the growing popularity of blockchain technology, and their predictions of its being a game changer, caught our attention, and we began to think of possible blockchain use for library applications and how this could be investigated. As we looked into it more, we discovered that blockchain technology was being discussed widely, but without clear directions for library applications.

As we started exploring blockchain possibilities, we were frequently asked: can you please provide a *simple* explanation of what blockchain is? The problem is that blockchain, as a new and emerging technology, still feels like an abstract concept. Blockchain still has very few actual implementations, which makes it hard for many of us to get our minds wrapped around what it is, how it can be used, and why people are so excited about it.

Simply put, blockchain is a decentralized and distributed ledger technology (DLT) that has no central authority. It is used to record (and store) the same information in a block across many computers. These blocks are linked together in a chain. A record cannot be altered after it is created unless the majority of users agree to make the change in the chain or blocks. As Darra Hofman, one of the experts who participated in our blockchain project and a PhD candidate at the University of British

Columbia, explains: “A blockchain is basically a distributed database where lots of different parties can read and write transactions to the database. Instead of a third party checking those transactions, the blockchain has a built-in consensus mechanism that checks transactions to make sure they’re good (that I have the $20 I just tried to spend, for example).”[1]

Each record that is put into the chain is given a unique digital signature created by a cryptographic hash function or algorithm; this acts like a digital fingerprint. In addition to a hash, each block in the blockchain also includes the hash of the previous block, a timestamp, and data about the proof-of-work algorithm (this is how new blocks are mined and verified). This creates the “chain” part of the blockchain. This is also what makes blockchains hard to tamper with, because if anything changes in any of the blocks (even if it is just the deletion of a period), this changes the hash (the unique identifier) of the block, which then breaks the blockchain.

The information in a blockchain cannot be changed unless the majority of users in this decentralized group agree to it. If there is no agreement, then the record reverts to its original value. This feature has enabled blockchain to be a trusted and reliable technology. There are instances from cryptocurrency blockchains where more than fifty percent of the network was controlled by members who could make changes or block new transactions from taking place or being confirmed. A 51 percent take-over is less likely to occur in a private or permissioned blockchain where users are not able to participate individually. Centralized organizations like libraries are most likely to use permissioned or private blockchains.

As we learned more about blockchain, we wondered if the buzz over it was all hype, or whether the technology could be used effectively in libraries and information centers. To answer this question, in 2017, San José State University’s (SJSU) School of Information secured a grant from the Institute of Museum and Library Services to launch an eighteen-month examination to explore the ways that libraries can utilize blockchain for practical library applications while the implementation of the technology is still in development. An advisory committee helped to develop the scope of the project, which included five main components: (1) develop recommendations on the future uses of blockchain technology within the information professions; (2) create a *Blockchain* blog[2] and website; (3) offer a free virtual conference, called Blockchain Applied: Impact on the

Information Profession;[3] (4) engage in deep discussions with leading experts at a National Forum,[4] an invitational gathering to discuss the feasibility of blockchain applications in libraries; and (5) develop and offer a blockchain MOOC (massive open online course).

Through this process, we have discovered that blockchain technology has many potential applications in libraries. However, the path forward includes a number of issues and tasks that must be resolved. This book includes a compilation of what librarians and information professionals need to understand about blockchain technology, and offers several speculative visions for how blockchain could support the core work of libraries. Librarians need to understand new technologies in order to determine which ones will improve operations and services, while also avoiding new and unproven technologies that would not be useful.

As a next step in this project, we plan to develop models that will indicate if blockchain applications in library settings are feasible and scalable. These pilot projects will allow librarians to experiment with real-life examples. Blockchain is currently an unproven technology for the information profession, with many advocates, skeptics, and opponents. We have a professional obligation to examine this technology's potential applications in order to make informed decisions about it. The essays in this book are written by information professionals who have interest and experience in blockchain technologies, and who offer their ideas for exploration and consideration. We invite you to explore their ideas in this book, and also contribute to the Blockchain blog at https://ischoolblogs.sjsu.edu/blockchains/.

1. https://americanlibrariesmagazine.org/2019/03/01/library-blockchain-reaction/.
2. https://ischoolblogs.sjsu.edu/blockchains/blog/.
3. http://www.library20.com/page/blockchain.
4. https://ischoolblogs.sjsu.edu/blockchains/national-forum/.

UNDERSTANDING BLOCKCHAIN

WHAT IS DLT AND HOW DID IT BEGIN? WHAT IS blockchain technology? What are the opportunities that blockchain technology may offer to libraries, and what obstacles will libraries need to overcome in adopting it? These are critical questions for understanding blockchain in the information professions.

FROM MEDIEVAL ORIGINS TO MODERN APPLICATIONS

Christina Cornejo, Librarian, Stockton-San Joaquin County Public Library; and Stacey Johnson, Library Manager and Technical Services Librarian, Chino Valley Public Library

For many people, the first thing that comes to mind at the mention of blockchain is Bitcoin and other cryptocurrencies. The term *blockchain* first appeared in a 2008 white paper written by an anonymous developer or group of developers known as Satoshi Nakamoto, who developed a peer-to-peer method of exchanging electronic currency.[1] Their research became the basis for Bitcoin, the best-known of the cryptocurrencies.[2] Blockchain provides the foundation for how transactions in this cryptocurrency are managed, stored, and verified, so without it, there really is no Bitcoin. But the core of blockchain, the idea of a distributed ledger that is shared by more than one entity to secure data for record-keeping, is not new. Indeed, there is reason to suggest that a version of this concept existed in medieval times.

In the transition period between oral forms of record-keeping and written ones, information may have been stored as objects or tokens. “Today, what once had a material form can be essentially dematerialized. Paper currency can be transformed into cryptocurrency. Land, fine wine, artwork, diamonds, food, and other material objects—though still physically in

existence—can be transformed into virtual representations called ‘tokens.’ In this way, in a tokenized, blockchain record-keeping system, literally everything potentially becomes a record.”[3]

The reliance on physical objects or tokens to prove authenticity has been replaced by the ability to keep digital records. Unlike physical objects, digital records can be changed and must be secured through the use of encryption to protect the transfer of data on the Web from hackers.

Blockchain Explained

As already noted, *blockchain* is the term used to describe a new technology that began as the underlying system behind Bitcoin. Bitcoin is a system of digital currency that provides a secure, and sometimes nearly anonymous, way to earn and use electronic cash without enlisting the services of a bank or other financial entity as the broker. Blockchain technology ensures that Bitcoins cannot be copied and re-used in multiple locations (e.g., no one is able to purchase an item using a Bitcoin at one store and then reuse a *copy* of the same Bitcoin at another store). Blockchain was developed to keep track of each Bitcoin, enable the transfer of ownership of Bitcoin if needed, and prevent the tampering with or copying of Bitcoins.

Blockchain is comprised of *blocks* of information or records that store data using sophisticated algorithms. The *blocks* are connected or *chained* together and distributed to a peer network of multiple trusted sources. Each member of the network has a key or an individualized digital signature that is assigned to the transactions made by that member. It is difficult, but not impossible, to hack a blockchain, since there are a number of inherent characteristics built into the technology that ensure the integrity of each record—making the network difficult to breach.

Several design features make blockchain a secure and ideally unalterable ledger or collection of informational transactions. Each block can contain unrelated data or a distinct group of related information. To make a blockchain, the distinct blocks are *chained* together by computer code—enabling the network to track whether each block has been altered in any way. A good way to visualize this is to consider the blockchain as a book with each page being a different document or set of information, like a contract or an academic transcript. The pages are numbered in order, and if someone tries to add a page or change an existing page within the book, it

will be obvious that a change has been made. As a book, the blockchain would have an unlimited number of pages, but you can only add to the end of the book, without impacting the previous pages and pagination.

Continuing with the analogy of a book, this means that copies of the blockchain or the book are sent to multiple locations: locally, regionally, or internationally. The content in the book is *chained* together, and identical copies are distributed throughout the peer-to-peer network. The blockchain is immutable and cannot be changed without agreement from more than 50 percent of the network; however, in the simplest explanation, the blockchain can be hacked or altered if it is attacked by a group that controls more than 51 percent of the network (a 51 percent attack). Fifty-one percent attacks are more likely to target the more valuable cryptocurrency networks rather than other types of blockchain usage. There are different levels of security for blockchains depending on how they are made and how far they are distributed.

The ideal way to use a blockchain is to have a long list of information that needs to be recorded, visible to many people, and never change in the future; for example, sale contracts which can be used to prove ownership of an item or property, art provenance, or public records. Even personal records, such as academic transcripts, could be a good use of blockchain. The transparent blockchain is not something that will necessarily be used on its own, but as a way to provide access to documents and information with the security and assurance that they cannot be altered.

Combining Old Tech with New Ideas

Blockchain applications are being tested by governments and organizations in a number of inventive ways, from applications in smart cities such as the Blockland Cleveland initiative[4] to providing international aid to refugees. Librarians are beginning to think of possible blockchain applications for a system that is decentralized, tamperproof, and independent without the need for a broker to assist in holdings or transactions.

Professional information organizations such as the American Library Association’s Library Information and Technology Association, the National Information Standards Organization (NISO), and the Digital Public Library of America (DPLA) have placed blockchain technology on their radar for further exploration.[5]

BLOCKCHAIN AND DECENTRALIZATION: BIG-PICTURE OPPORTUNITIES AND RISKS

Jason Griffey, Director of Strategic Initiatives, National Information Standards Organization

While technological change is rapid and unceasing, it is not often that something truly novel emerges, something that is fundamentally new and different in the world of the digital. Most new digital services or techniques are enhancements of previous work, and are evolutionary and predictable at least in scope, if not in exact form. But in 2008, something completely new and different did enter the world, and its anonymous creator called it *blockchain*. While it is true that blockchain leverages well-understood digital technology such as hashing[6] and encryption, and that philosophically it shares a lineage with Merkle trees, it is safe to say that nothing like it had taken root and really worked prior to the Bitcoin white paper and initial mining event that began the era of blockchain.[7]

One of the reasons why it has taken so long to see the potential of a distributed system like blockchain is that it was not until around 2014 that people started seeing the decentralized database nature of a public blockchain system as a platform for the creation of user-facing services. It seems obvious in retrospect, but it took several years for programmers to see blockchain systems as possible answers to the current centralized state of web properties. This centralization, driven by market capture and the network effect, has caused the existing World Wide Web infrastructure to be far more centralized in actual use than it was originally planned to be in theory. Several huge multinational companies control the vast majority of services that people consume online, and those services are just as fragile as they are critical to most people's experience of the Web. When something like Facebook or Amazon goes down, huge swathes of web content temporarily die with it.

The promise of blockchain is not the current state of its use. A few hundred initial coin offerings and a half-dozen obscure services do not a revolution make. However, the new architectures that systems inspired by blockchain are using, and the concepts behind blockchain that have been

enabled, are likely to make a huge difference to the world over the next decade. A Web that is increasingly difficult to censor, one that defies local control and enables a truly robust set of distribution and archival tools for digital content, is a Web that is better for both producers and consumers.

For librarians, these systems will be significantly better than the existing highly centralized infrastructure currently in use. The services that are emerging from this initial round of decentralization, like Dat Project and Matrix, are extremely exciting since they point in the direction of a Web that is far less vulnerable than the one we have now.[8]

Opportunities and Risk for Libraries

So where do libraries intersect with this potential new world of decentralized services and protocols? There are at least three areas that libraries and librarians should be watching:

1. When new decentralized services launch, librarians should watch them carefully and use them when they are beneficial to our specific needs. Decentralized services that provide significant advantages technologically, economically, and perhaps even socially/politically should be adopted where possible.
2. At the protocol level, it is possible that the development of new tools by librarians could be assisted by using new decentralized systems. The decision to build using a particular set of tools is one that should be made judiciously and carefully, but library developers should be watching services like InterPlanetary File System (IPFS), Dat Project, and even Ethereum as platforms upon which new and different sorts of information systems could be constructed.[9]
3. At the social level, libraries should support these new decentralized tools as a possible foundational platform of an informed democracy. Because libraries are a cornerstone for creating an informed citizenry, and an informed citizenry is necessary for a functional democracy, libraries should assist in providing access to those tools that also promote those ideals. When libraries can provide open, community-controlled

> informational systems and services which resist censorship, that is a good thing.[10]

This is not to say that decentralization is entirely positive. By emphasizing distribution and independence, decentralization has the capacity to further exacerbate filter-bubbles and self-directed informational limitations that have had deleterious effects on our society and culture over the last several years. It is possible that even with robust federation, informational silos could be re-created through accidental infrastructure, and rather than providing robustness and freedom, these services could devolve into pockets of self-organized biases. Especially in the area of social networking, there is the risk of existing social structures simply being replicated and resisting cross-pollination in the way that utopianists might describe. These risks should not be a reason to decry the development of these new services entirely; rather, these concerns should be seen as warning signs to heed in the development and deployment process. Having a plan for how to proactively avoid these risks will be far more effective than reacting to them when things begin to break.

Blockchain and other decentralized systems may afford libraries easier, more trustworthy, and more secure ways to track digital provenance, store metadata, access secure authentication tokens, verify transactions, run decentralized data stores, and more. Over the next five years, the multiplicity of services that use blockchain or a similar decentralized technology will be evident. Librarians who are not paying attention to this trend risk not only being late adopters, but possibly missing the opportunity to be a vital and driving part of this new decentralized Web. Librarians ceded their opportunity to be significant actors in the first incarnation of the World Wide Web to commercial entities. In this next-generation decentralized Web, librarians have the opportunity to be partners with the new technology leaders, working in cooperation with them to build the necessary infrastructure to allow for fully decentralized services to emerge.

Make no mistake, whether librarians act as partners or as users, these blockchain-derived technology infrastructures will enable new and interesting services and systems to emerge. The decentralized-yet-federated nature of these new systems will present intriguing possibilities to share content and media, and produce discovery layers that enable searching rich and deep sets of collections; most significantly, the content and collections

shared across a decentralized network will then be accessible to all. The next generation of the Web will not be like the current generation, and libraries and librarians must watch these developments in order to ensure that their organizations emerge from the resulting rebuilding as resilient and remarkable as ever.

BLOCKCHAIN: MERITS, ISSUES, AND CHALLENGES[11]

Bohyun Kim, Chief Technology Officer, University of Rhode Island Libraries

As discussed previously, blockchain secures data and prevents tampering by its technical implementation alone, and it changes the current role of a third-party authority. Traditionally, a third-party authority is brought into many types of transactions, such as fund transfers, real estate purchases and sales, insurance, and any type of credentialing ranging from school graduation to marriage certification. Its role is to guarantee the authenticity of the transaction and the integrity of the recordation process. However, blockchain renders this authorizing and recording role of a third party unnecessary.

Without the need of the mediating third party, a transaction becomes immediate, and its cost becomes much lower, while remaining secure. For example, when blockchain is adopted in real estate transactions, people will no longer need to go through a laborious process to officially record the purchase or sale of a real estate property, which involves the recorder's office of the city or county office and private title insurance. Blockchain can make the process of recording a property-related transaction more straightforward and efficient, thereby saving time and money.[12] Similarly, with blockchain implemented in banking, people will be able to quickly and safely transfer funds across different countries without going through banks. The World Food Programme is already using blockchain to increase their humanitarian aid to refugees.[13] Such practices can drastically lower the fees involved, and the transaction will take effect much more quickly, if not immediately.

In this respect, blockchain can serve as an alternative trust protocol that decentralizes traditional authorities. As such, it can be used for a wide range of purposes beyond real estate or financial transactions. From this, one can imagine the scale of the impact the adoption of blockchain can bring should it be adopted widely. Blockchain can lead to efficiency, convenience, and cost savings in areas where the authenticity and security of records are of paramount importance. These areas include electronic health records, digital identity authentication/authorization, digital rights management (DRM), digital provenance, and historic materials that may be contested or challenged due to the vested interests of certain groups and digital provenance.

However, there are some issues with blockchain technology. Since blockchain achieves its security by public key cryptography, if one loses one's own personal key to the blockchain ledger holding one's financial or real estate asset, then this will result in the permanent loss of such assets.[14] With the third-party authority gone, there will be no institution to step in and remedy the situation. Other issues with blockchain include:

- the lack of interoperability between different blockchain systems,
- the scalability of blockchain at a global scale with large amounts of data,
- potential security issues such as a 51 percent attack, and
- the huge energy consumption that a blockchain requires to add a block to a ledger.[15]

The last issue, energy consumption, has both environmental and economic ramifications because it can cancel out the cost savings gained from eliminating a third-party authority and related processes and fees.

There is growing interest in blockchain among information professionals, but there are also some obstacles to the technology gaining momentum and moving further towards wider trial and adoption. One obstacle is the lack of general understanding about blockchain in the larger audience of information professionals. Due to its original association with Bitcoin, many mistake blockchain for cryptocurrency. Another obstacle is technical. The use of blockchain requires setting up and running a node in a blockchain network, such as Ethereum, which may be daunting to those who are not tech-savvy.[16] This creates a high entry barrier for those who are

not familiar with command line scripting and yet still want to try out and test how a blockchain functions.

The last and most important obstacle, however, is the lack of compelling use cases for blockchain technology in libraries, archives, and museums. To many, blockchain is an interesting new technology, but even many blockchain enthusiasts are skeptical of its practical benefits at this point, when all the associated costs are considered.[17] Of course, this is not an insurmountable obstacle. The more familiar people get with blockchain, the more ways people will discover to use blockchain in the information profession that are uniquely beneficial for specific purposes. In order to determine what may make a compelling use case of blockchain, the information profession would benefit from considering the following:

1. What kind of data/records (or series thereof) must be stored and preserved exactly the way they were created?
2. What kind of information is at great risk to be altered or compromised by changing circumstances?
3. What types of interactions may need to take place between such data/records and their users?[18]
4. How much would be a reasonable cost for implementation?

These questions will help connect the potential benefits of blockchain with real-world use cases and take the information profession one step closer to the wider testing and adoption of the technology. Despite all of the issues discussed above, blockchain is expected to spread to many industries due to the unique benefits it offers, and it eventually will have an impact on libraries and the information profession.

1. https://en.wikipedia.org/wiki/Satoshi_Nakamoto.
2. https://www.bitcoin.com/.
3. https://phys.org/news/2017-11-blockchain-technology-medieval-roots.html.
4. https://www.blocklandcleveland.com/.
5. http://www.ala.org/lita/; https://www.niso.org; and https://dp.la/.
6. A hash is a one-way mathematical function that derives a unique value from a given input.
7. https://blockonomi.com/merkle-tree/ and https://bitcoin.org/bitcoin.pdf.
8. https://datproject.org/ and https://matrix.org/blog/home/.
9. https://ipfs.io/ and https://www.ethereum.org/.
10. https://boingboing.net/2016/03/28/how-libraries-can-save-the-int.html.

11. This section is a revised version of an article originally posted at *the ACRL TechConnect* blog. See https://acrl.ala.org/techconnect/post/blockchain-merits-issues-and-suggestions-for-compelling-use-cases/.
12. http://www.chicagotribune.com/classified/realestate/ct-re-0715-blockchain-homebuying-20180628-story.html.
13. Blockchain may be used not only for financial transactions, but also for the identity verification of refugees in the future. See https://www.technologyreview.com/s/610806/inside-the-jordan-refugee-camp-that-runs-on-blockchain/.
14. https://en.wikipedia.org/w/index.php?title=Public-key_cryptography&oldid=851218348.
15. https://www.investopedia.com/terms/1/51-attack.asp. And https://www.forbes.com/sites/shermanlee/2018/04/19/bitcoins-energy-consumption-can-power-an-entire-country-but-eos-is-trying-to-fix-that.
16. https://dev.to/legobox/how-to-setup-an-ethereum-node-41a7.
17. See https://go-to-hellman.blogspot.com/2018/06/the-vast-potential-for-blockchain-in.html; https://eprint.iacr.org/2017/375, and http://doyouneedablockchain.com/.
18. The interaction can also be a self-executing program when certain conditions are met in a blockchain ledger. This is called a "smart contract." See https://www.technologyreview.com/s/610718/states-that-are-passing-laws-to-govern-smart-contracts-have-no-idea-what-theyre-doing/.

BEFORE THE HYPE

BEFORE BELIEVING THE HYPE AROUND BLOCKCHAIN technology, librarians should weigh the benefits of adopting the technology against the risks and costs associated with doing so. Moreover, a strong foundation needs to be built for blockchain technology that includes establishing standards, addressing legal concerns, and ensuring security.

STANDARDS

Todd A. Carpenter, Executive Director, National Information Standards Organization

Looking beyond the hype surrounding blockchain technology, at their core, each of the applications of an open ledger system is essentially a solution for metadata management problems. These metadata might describe the ownership, provenance, transaction status, ownership rights, or status changes of things, be they virtual or physical. Regardless of the application or the object described, blockchain systems are simply tools for managing metadata. If there is one thing that librarians and library systems have mastered, it is managing metadata, and as such, adapting to an open ledger technology in libraries should be something that librarians are well-positioned to accomplish. There are many metadata management issues within the library context, so the potential applications that could be built using blockchain systems are abundant. Whether blockchain-based systems are the best approach to solving a specific problem, or whether they can supplant existing metadata management systems already in wide use, is an open question.

The Need for Standards

One of the first requirements when developing a metadata management system is to establish standards. Standards come in many forms, from the

informal online community practice of agreement and adoption, to the more formal world of standards developed by organizations and governing bodies, such as NISO, the Library of Congress (LOC), the World Wide Web Consortium (W3C), and the International Standards Organization (ISO).[1] How we set out to describe collections, the items in those collections, the transactions we process, and the rights governing those items all need to be standardized for metadata management systems to function. With blockchain systems, the same problems exist as with more traditional systems such as identifier registries, authority control lists, and integrated library systems (ILS).

Therefore, standards are needed to clarify what is, or is not, included in a metadata system. Standards are needed to explicate how we establish the characteristics we want to capture. Standards provide protocols for the basic elements of every metadata system. Just like metadata systems from the past, a new system built using blockchain will also require these basic elements of standardization and an agreement about how to structure metadata elements.

Beyond describing the metadata managed by an open ledger, there are also a host of standards questions around blockchain technology itself. In a 2017 white paper, *Reinforcing the Links of the Blockchain*, the IEEE identified three core elements necessary to advance blockchain technology.[2] The second among these elements is "a strategy for adopting technical standards."[3] Like every digital system, the baseline standards we adopt at the outset impact many other options further down the chain. Some of these standards address technical and interoperability questions surrounding the system itself, whereas others are more practical business-function questions related to how a system might operate and the business rules that govern its operation. The technical issues begin with the data to capture and in what form, but also delve into process management and data retention. Policy-related questions might include status change issues, determinations of provenance, and authenticity. Other issues might also interrelate to particular legal or ethical guidelines that govern how information is managed or shared in the library community, such as privacy.

Standards for Blockchain Adoption

There are at least five areas of focus where standards development can speed the adoption of blockchain technology. These areas are process management, data retention, status changes, the provenance of information, and the authenticity of information.

Process Management and Data Retention

The topic of process management encapsulates many of the current questions and issues of how blockchain functions. One of the initial issues to address is the metadata in the block elements of blockchain. There are a variety of hashing algorithms that can be used to capture information about a record, each with specific strengths and implications. Because blockchain systems can be distributed across a network, how those network nodes are assigned and managed must be set forth. Some blockchain systems allow for limited public data availability, whereas others hold information in private, so there needs to be agreement on what those data availability rules are. Inevitably, there will be errors introduced into the system through dual posting, inaccurate information, or other reasons, and rules need to be established on how to correct those errors.

Data retention, similarly, is functional and has to do with how to address questions about maintenance of the data in the blockchain. As any system grows, there are issues with managing the size and scale of the ledger. Although the distribution of nodes helps address this issue, the robustness of the network could still be hindered by the management issues of an ever-growing repository of data included in the system. This begs the question: what are the data management rules for the network, such as how long to retain data? Rules will need to be established around what data is critical to keep versus what data can be discarded. For instance, within the context of the European Union (EU) General Data Protection Rule (GDPR), there is a new "right to be forgotten," in which EU citizens can request to have their private data deleted from an online system.[4] This could have implications for how the blockchain would be managed.

Status Changes, Provenance of Information, and Authenticity of the Information

The remaining three areas for libraries to address in establishing blockchain use standards relate to business processes and data quality issues impacting the data that are captured in an open ledger. The first of these areas deals with systems that track status changes, and requires agreement on which

states should be captured within the system and how those states are defined. Additionally, what constitutes a change in status and which status changes are relevant to track need to be established. Similarly, if status is being tracked over time, who within the system determines when an object has passed from one state to another? Finally, at a system level, there should be broad agreement about how timely the tracking of state changes needs to be. Some applications might require data to be real-time, whereas other applications might be sufficient if updated daily or weekly.

The second area has to do with the provenance of the data in the system and the rules for establishing a chain of control. The core of the question of provenance is the agreement and establishment of the current state of the community. In other words, should librarians simply accept the current status quo, can they go back to some original state of being, or should they collectively accept the potential inequalities that might go along with the establishment of the current state? Once the system is implemented, more practical issues arise regarding the rules for transfer from one entity to another, and how the system can ensure that every transfer takes place within the system.

Finally, there are standards focused on the authenticity of the data in the system. If the data is important enough, someone might have a motivation to cause problems with the system. Implementers should plan for how the system ensures against fraudulent access and record creation. If problems are noted, fallback plans are required to return the object to its previous or original state, prior to when the problems or erroneous data were posted.

As with all metadata systems, it is important to ask with blockchain if the cost of managing the system over time and at scale is worth the investment. Can a blockchain system be used to make the process of managing library metadata better, faster, more accurate, more secure, or cheaper overall? Librarians should use these questions to analyze every new technology when considering implementation. If the new system is not a significant improvement over existing systems, then there is concern about collectively buying into the hype of the new system. Standards can help bring these costs down and speed adoption, but standards cannot address the core question of whether the new system is, in fact, a better solution.

LEGAL CONSIDERATIONS

Dan Blackaby, Head of Technology Initiatives, Cornell University Law Library

Blockchain by its very nature presents a number of legal issues. The supposedly *immutable* nature of its contents, combined with the everywhere and nowhere nature of DLT, make blockchain a very odd and uncomfortable fit into the existing legal schema of data ownership, control, and use.

The concept of data ownership is in some ways synonymous with the control of said data. When a court decides to subpoena data records, often the subpoena will not only be served on the owner of the data, but also on the *controller*—an IT data storage provider, such as Google, or on the person who owns the physical server where the data is stored. Many countries have taken steps to make sure that access to the data owned by countries that do business within their borders is available to them by instituting what are called *data localization laws*—laws which mandate that the data be stored within the physical and legal jurisdictions they are operating in.[5] When centralized databases are the prime storage medium, this approach makes a lot of practical sense, since the database has to *live* somewhere.

The blockchain distributed ledger system either makes this much more difficult or dramatically easier, depending on how it is viewed. If a company uses a DLT, then the company either has

1. obviated its need to conform to data localization laws by having a copy available on any computer that has accessed and used the database,
2. made its data open to subpoena by every country where a user resides (and thereby has a copy of the ledger), or
3. both.

Also, given the varying requirements of the different laws around the world, the use of a blockchain may dictate data curation techniques that are either

onerous or unclear, given the possible conflicting data standards in the various jurisdictions.

This type of control is at the heart of the EU's GDPR.[6] The reach of the GDPR is long. It applies not only to organizations within the EU, but to any organization that offers goods or services to, or monitors the behavior of, EU data subjects. This law applies both to the controllers of data (those who determine the purposes, conditions, and means of the processing of personal data) and to processors, which in this case means among other things, cloud storage providers. Per the GDPR, personal data is "any information related to an identifiable person who can directly or indirectly be identified in particular by reference to an identifier."[7]

One of the most problematic sections of the EU's GDPR in the context of blockchain is the incorporation in Article 17 of the "right to be forgotten."[8] This requires the controller of the data to erase personal data under certain circumstances, which conflicts directly with blockchain's assertions of immutability. Article 4 of the GDPR defines personal data in a very broad sense, and this, when combined with the lack of a definition of what it means to *erase* data in Article 17, left blockchain users with a seemingly all-or-nothing proposition.[9] As things are now, either blockchain users conform to the GDPR and make all the data physically erasable (which means that they just don't use a blockchain) in order to maintain their ability to erase data, or they use a blockchain in defiance of the GDPR.[10] Although it is technically possible to comply with the GDPR, this is still a problem that needs to be addressed both logistically and legally.

Another possible issue with the immutability of blockchains involves a particular application, the integration of so-called *smart contracts* within a blockchain. First proposed by Nick Szabo in 1994, smart contracts are "computerized transaction protocols that execute [the] terms of a contract."[11] It is important to note, as Szabo did, that these protocols are not the contract itself. If asked, a contracts lawyer will probably focus on the subordinate nature of that definition, in that smart contracts are not contracts per se. There are many things that go into a *contract* (e.g., offer, acceptance, capacity, a meeting of the minds, etc.) that a lawyer using a strict construction would not see in a *smart contract* in a blockchain. Computer code in this case would facilitate the *executable* part of the contract by sending money to an account, ordering the shipping of materials, or

authorizing some other action as part of the understanding of a contract that actually would likely exist outside of the code. Essentially, it is code as action, an *if-then* statement, which implements the larger contractual agreement.

In action, smart contracts will have internal callouts to other databases and benchmarks, for example, referencing live data on the Web for things such as commodity prices, interest rates, and other terms of the action at the time of performance. This aspect of smart contracts is probably the most appealing one to the users of blockchain since it makes the actions self-executing and, in many ways, self-constructed. But what happens when someone wants to change or cancel the contract's terms or delete the code that serves as essentially a non-cancellable executable?

This need necessitates a kind of *living* nature to the code that can be altered. So, how does a *living* smart contract function within a database that is immutable? If new terms are written, does that cancel the previously ordered execution (as it may have, given standard contract construction)? These are questions that have really not been answered yet, but this has not stopped legislators from going forward with legislation to recognize the use of smart contracts and blockchain.[12]

Both Tennessee and Arizona have legislation with regard to smart contracts, with Arizona additionally joining Wyoming in passing legislation on blockchain and cryptocurrencies. Many critics in the legal world feel that these states are jumping the gun by passing laws with regard to technologies that are still incipient in their formation. Andrew Hinkes of New York University's Leonard N. Stern School of Business remarks: "Laws should not attempt to define technologies that do not have a widely held definition in their relevant technical communities."[13]

Perhaps most concerning from a sheer evidentiary perspective are the laws passed by Tennessee and Arizona with respect to blockchain that call the contents of the blockchain an *uncensored truth*.[14] From a data curation standpoint, this is highly problematic. One of the oldest tenets in computer science is *garbage in, garbage out*. The contents of the blockchain are true representations only of what the data provider intended to include—not of the truth or validity of the nature of those contents. Putting a false statement into a blockchain would not suddenly make that false statement true, so the

truth would lie in its being a *true record* of what was input, not of the content of the record. That, however, is not the language of the statutes.

Blockchain does have great promise for a multitude of legal functions, and there are times when using a blockchain would not only be appropriate, but of great advantage.[15] Some of the best uses are when an extended *chain* record of events is required, such as in payment systems, an evidentiary chain of custody records, and perhaps the area where blockchain has received its widest use outside the realm of cryptocurrencies—land registries.[16]

Land registries in the past have taken the form of massive books, large file cabinets, and, famously, huge piles of paper bound by ribbons. Doing research into the history of a particular parcel of land with all of the possible transactions, changes in ownership, retained future interests, easements, and various other entanglements is one of the most onerous chores that many legal researchers ever face. A blockchain seeks to streamline this process. With a system built on individual transactions and changes in contract terms, land registries seem tailor-made to be converted to a blockchain. This would allow for both findability and a clear record of current status. Many countries have bought into this idea, with varying levels of success.[17] The most notable experiments have had their problems, however.

For instance, Estonia has positioned itself at the forefront of blockchain technology with its e-citizen program. However, given its relatively small scale, Estonia's e-citizen program has yet to prove that such a program can scale up to the needs of a world power, or even a global-sized corporation with a large history of records to be retrofitted into the system, as any land registry would normally require.[18] Estonia, as a former Soviet vassal state, had the rare advantage of starting over from scratch with much of its land registry and technology infrastructure. Trying to implement the same kind of system in a country such as France or the United States would prove vastly more complicated.

Brazil is testing a blockchain-based land registry in order to counter its history and reputation of governmental corruption and graft. Blockchain can go a long way toward establishing immutable records. However, as one article notes, "registration involves a legal analysis . . . with a lawyer checking the authenticity of identity documents, title deeds and government

allocations, as well as cross-checking to avoid double titling—tasks that blockchain cannot perform."[19] Blockchain cannot in the end convey legal value or validity because it cannot alter the underlying *garbage in, garbage out* nature of its contents. That validity comes from the people and the legal process and is not solved by changing the nature of the database. The New America think tank has published a white paper establishing criteria for how to implement a blockchain-enabled land registry and examining the complexities involved in doing so.[20]

There are numerous other possible applications for blockchain in a legal context, ranging from voting systems and digital public record systems in the United Kingdom, to the United Nations' World Food Programme administering refugee aid.[21] Banks across the world are creating trade finance platforms with blockchain features, and the United States government is pursuing blockchain's use by agencies ranging from the U.S. Postal Service to the Centers for Disease Control.[22] These applications all come with various legal definitional hurdles similar to those faced by the land registries and by any company dealing with the GDPR. The real concern comes with the idea that the applications and, implicitly, the technology are outstripping the ability of the legal infrastructure to keep pace.

So what does this mean for libraries? Libraries have many functions and processes that might seem well-suited for the promise of blockchain. One of the most frequently suggested applications for blockchain within a library context is interlibrary loan. As a pure exercise in shipment and supply chain tracking, using blockchain seems like a no-brainer. However, when paired with the library culture's ethical stance in regard to user privacy, along with libraries' functioning as a data store of personal data under the GDPR, the number of questions that need to be answered before implementation can go forward multiply, and it may prove difficult to know when all the bases have been covered. The efficiency that a blockchain may enable needs to be balanced against the ethical interests of protecting user data.

These and other kinds of questions call for a patience that the speed of technology uptake seems ready to overtake. Patience, discernment, and judgment will be necessary to determine not only when the use of blockchain is or is not appropriate, but perhaps more realistically to

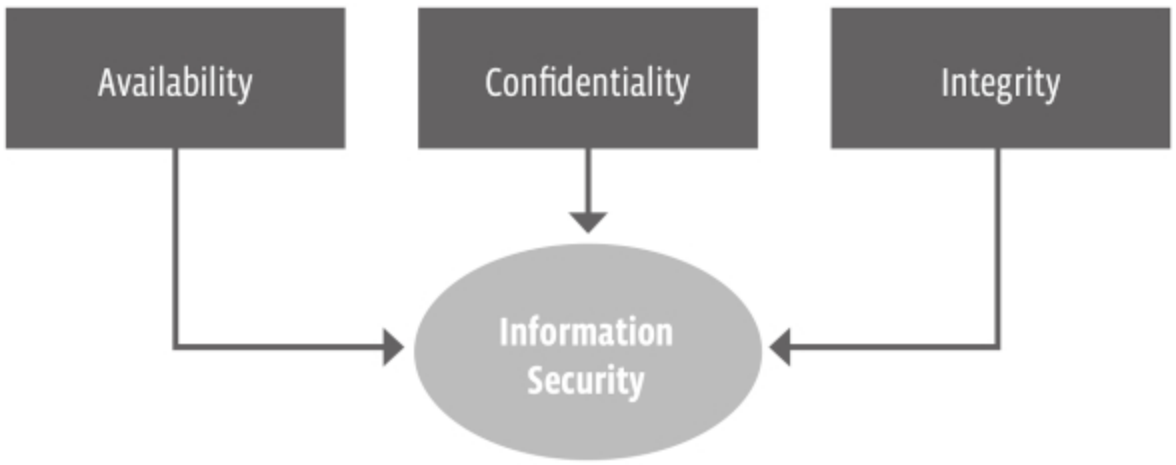
determine what definitions and values need to be changed in the face of what can be accomplished with this new technology.

SECURITY

Tonia San Nicolas-Rocca, Associate Professor, San José State University School of Information

Blockchain is a young technology that is evolving into a solution that could support efforts to protect and safeguard information and information system resources. To keep information and information resources secure and safeguarded, three fundamental principles must be met: availability, confidentiality, and integrity (see figure 1). Efforts to safeguard information and information system resources must provide one or more of these security principles.

Figure 1
Principles of Security



Availability

Information and information system resources must be available to authorized users when they need it. In the event of a disruption, the system

should be able to recover quickly in order to minimize any downtime.

The goal is to provide authorized users with timely and reliable access to information and information system resources. By having the same information stored on multiple computers in a distributed network, a blockchain has no single point of failure, and is therefore less vulnerable to system downtime or complete lack of access.

Confidentiality

Confidentiality ensures that sensitive content is not disclosed to unauthorized users, programs, or processes. Libraries collect a wide range of information, some more sensitive than others, requiring a higher level of confidentiality. Access controls should be used for identification and access management (I/AM), ensuring that only the right people access the appropriate resources. Four elements are required for adequate protection and trust: identification, authentication, authorization, and accountability. Authenticated users must only be granted authority to perform authorized actions and be held accountable for their actions taken within a system.

Blockchain supports the principle of confidentiality with the use of public key cryptography to strengthen I/AM. This makes it nearly impossible for unauthorized users to access information that is stored or transmitted in an encrypted format.[23]

Integrity

Information must be complete, dependable, and trustworthy. Integrity is maintained when the accuracy and reliability of information and information system resources are neither tampered with nor altered and are guarded against unauthorized modification or destruction.

Information and information resources are susceptible to unauthorized modification or destruction. Users may affect the integrity of information by accidentally inputting data into a system incorrectly. For example, a librarian may enter a patron’s personally identifiable information (PII) incorrectly into a database, making it difficult to retrieve that person’s account at a later time. The integrity of information or an information system may be compromised by an active attack in which an attempt is made to alter or modify information stored on a system or while it is in

transit. Integrity may also be compromised through malware leading to the unauthorized modification or destruction of information. For example, a user may open an e-mail attachment, download a file off the Internet, or visit an infected website, thereby allowing a virus, logic bomb, or backdoor into an information system. This could lead to the unauthorized modification or destruction of information.

The integrity principle ensures that information and information system resources are guarded against unauthorized modification or destruction. Multi-signature is a digital signature scheme, which allows a group of users to sign a single document rather than including individual digital signatures.[24] Blockchain's multi-signature protection is resistant to the modification or destruction of information. Blocks store transactions across many interlocking computers that cannot be tampered with or altered without altering other subsequent blocks. Since blockchain is a secure ledger that is updated regularly, stored across multiple systems, and requires public key cryptography to authorize a transaction process, it mitigates active attacks, and ensures that all transactions are legitimate and properly recorded.

Blockchain technology provides built-in measures that uphold the three principles of security and preserve the immutability of the data; it is a technology to be considered for the security that it provides.

HEALTHY SKEPTICISM

Toby Greenwalt, Director of Digital Strategy and Technology Integration, Carnegie Library of Pittsburgh

As libraries begin to consider the potential applications of DTL, it will be important to consider the wider implications of library operations running on *everyone* else's computer.

At their core, blockchain mechanics hold a great deal of promise for reshaping library services. Distributed database operations create opportunities for many foundational operations of library technology. Records management for both patron and bibliographic information could

be addressed in this way, and blockchain mechanics demonstrate even more potential for digitization, archives, and applied digital humanities. In many ways, blockchain as a conceptual framework could represent the next evolutionary step for data-centric institutions that are moving their operations to the cloud.

That said, blockchain-based operations are just close enough to many existing library practices that one has to ask if the gains are big enough to warrant the investment in time, training, and infrastructure. There is an "uncanny valley" element to some of the proposed elements of blockchain technology in libraries—particularly public libraries. Many applications for blockchain appear very close to what libraries are seeking to offer, but the subtle differences act to highlight just how much ground remains to be covered. For example, blockchain could serve as the backbone of a municipal ID program, providing one form of identification that works for schools, libraries, public transportation, and other city services. Similarly, blockchain could serve to address some of the issues with the provenance of documents, offering a universal record for citations and publication records. Issues with privacy and transparency require a healthy skepticism as the profession evaluates blockchain's merits. Creating the right value-adds for blockchain-based library services will be an important test for any early adopters. Ensuring that the bar to entry is low enough to develop a critical mass among libraries of all types is the next test. Making this work will require a great deal of effort. What will it take to turn blockchain-based library technology into something scalable, sustainable, and accessible to most? Here are a few of the areas requiring greater exploration.

Trust: A distributed database requires a common data framework in order to be effective across institutions. In many libraries, it is difficult to create common data standards across departments. How can the framework be flexible enough to allow for institutional variation, while still maintaining the level of consistency necessary for blockchain to be effective? How can an institution avoid *lock-in* to a particular framework or platform, while ensuring that its users' data stays exclusive to them, portable to other platforms and, most of all, private?

Storage: In the past five years, the overall size of a blockchain (i.e., the total size of the ledger, including all headers and transactions) has grown from 4GB to over 190GB as of this writing.[25] This exponential growth in

size will require data storage tools that can keep pace with it. If growth continues at this pace (a rate which will only increase as new applications are built on the blockchain), then it will likely be a struggle for current hardware or cloud storage tools to grow at a corresponding rate.

Even if data storage technology can keep pace, it is another thing entirely to be able to absorb the costs of the upgrade cycle. Speaking as a member of a consortium with a pretty broad spectrum of library size and technical ability, it is difficult to imagine every institution investing the resources necessary to properly leverage blockchain-based operations. Even for the institutions that possess the necessary infrastructure, how does one plan for smart growth, even with an entirely cloud-based environment?

Bandwidth: These same constant rates of growth will require an increase in available bandwidth. How will this fit in with existing demands for robust, steady, high-speed Internet access in public library spaces—especially as the rate of consumption continues to grow? With broadband infrastructure still lacking in remote, rural, and tribal parts of the country, basic connectivity is enough of a struggle—let alone more complex, resource-intensive applications.[26]

Environmental Impact: See Kim.

Cost: Each of these factors will come at a financial cost. Whether developed in-house or provided through a vendor, blockchain-based library services will impact the library's budget structure in ways we have not yet articulated. Will these be reasonable? Will the gains provided by the applied uses of blockchain provide an effective return on investment? Without a stronger set of use cases (and their accompanying business models), these questions are yet to be resolved.

Some of the cost factors may be addressed through government prioritization. Just as thousands of libraries rely on the Federal Communications Commission's E-rate program to facilitate access to high-speed broadband, there is a strong chance that many of these same libraries will be unable to leverage blockchain technology without a similar type of federal assistance.[27]

Where We Stand Now

At present, there are more questions than answers regarding blockchain as a practical tool for libraries. Given that libraries exist on such a wide spectrum of funding, infrastructure, and technical ability, it is entirely possible that blockchain could serve as an equalizer. But we should not abandon a healthy skepticism; as with all presumptively disruptive technologies, there is a good reason to maintain a critical eye while moving forward.

Moreover, offsetting internal library operations to a decentralized ledger comes with risks beyond the ones listed above. What does it mean to place so much trust in a system which is still likely to rely on external gatekeepers, commercial vendors, and other entities that may not share our professional values? If things like account services are moved to the blockchain, what does it mean for libraries to compel our users to adopt a system of data aggregation that may be used against them? The Wild West nature of the current blockchain ecosystem does not offer much clarity. The responsibility is on us as a profession to look beyond the raw technological opportunities presented by blockchain and consider the larger cultural and market forces at work.[28]

Because all of these rules are being written (and rewritten) in the moment, it is important to maintain a healthy level of skepticism regarding blockchain's merits. Moving specific library operations and services to a distributed ledger-based system requires a massive leap of faith, and would be another attempt to build on the spirit of radical trust that has driven the Internet since its early days. But as public libraries evaluate the applications of the blockchain, they must consider their functions as responsible stewards of community funds, advocates for thoughtful digital fluency, and protectors of patron privacy. In the rush to advance, libraries must bring some of their own core values to the table.

1. https://www.niso.org; https://www.loc.gov/standards/standard.html; https://www.w3c.org; and https://www.iso.org.
2. https://blockchain.ieee.org/images/files/pdf/ieee-future-directions-blockchain-white-paper.pdf.
3. Ibid.
4. EUR-Lex, "Right to erasure ('right to be forgotten')" in *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016,* Section 3, Article 17, https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1532348683434&uri=CELEX:02016R0679-20160504.
5. http://stlr.org/2018/03/01/the-emerging-trend-of-data-localization/.

6. https://eugdpr.org/.
7. https://eugdpr.org/the-regulation/gdpr-faqs/.
8. https://gdpr-info.eu/art-17-gdpr/.
9. https://gdpr-info.eu/art-4-gdpr/.
10. A possible alternative is a "referential" blockchain in which the blockchain contains only links to the personal data, which is stored in another, editable database elsewhere. See https://www.blockchainandthelaw.com/2018/04/blockchain-personal-data-and-the-gdpr-right-to-be-forgotten/.
11. https://www.investopedia.com/terms/s/smart-contracts.asp.
12. https://www.technologyreview.com/s/610718/states-that-are-passing-laws-to-govern-smart-contracts-have-no-idea-what-theyre-doing/.
13. Ibid.
14. https://www.ethnews.com/tennessee-establishes-legal-authority-of-blockchain-technology-smart-contracts; and "Signatures and records secured through blockchain technology; smart contracts; ownership of information; definitions," 44-7061, Arizona State Legislature, https://www.azleg.gov/ars/44/07061.htm.
15. https://www.coindesk.com/information/why-use-a-blockchain/.
16. https://e-resident.gov.ee/.
17. https://www.gov.uk/government/news/hm-land-registry-to-explore-the-benefits-of-blockchain; and https://cointelegraph.com/news/netherlands-land-registry-to-test-blockchain-solution-for-real-estate.
18. https://www.investinblockchain.com/estonia-blockchain-model/.
19. https://www.reuters.com/article/us-brazil-property-blockchain/can-blockchain-save-the-amazon-in-corruption-mired-brazil-idUSKBN1FE113.
20. https://www.newamerica.org/documents/2121/Graglia_Mellon_blockchain.pdf.
21. http://blockchain.surrey.ac.uk/projects/archangel.html.
22. "Major Banks Target $1.5tn Coverage Gap with Trade Finance Platform," *Financial Times*, https://www.ft.com/content/a886eb72-d274-11e8-a9f2-7574db66bcd5; Selva Ozelli, "US Government Implements Blockchain Programs to Improve Transparency and Efficiency: Expert Blog," Cointelegraph: The Future of Money, 2018, https://cointelegraph.com/news/us-government-implements-blockchain-programs-to-improve-transparency-and-efficiency-expert-blog.
23. Not all data on the blockchain are encrypted; for example, sensitive information in plain text in Bitcoin graffiti.
24. https://medium.com/@safehavenio/transferring-ownership-of-multi-signature-private-keys-in-the-event-of-death-or-disability-a116167aadc8.
25. https://www.blockchain.com/charts/blocks-size.
26. https://www.fcc.gov/reports-research/reports/broadband-progress-reports/2018-broadband-deployment-report.
27. "Search Commitments - Schools and Libraries Program - USAC.org," numbers from 2015 dataset used for reference, USAC.org., 2018, https://www.usac.org/sl/tools/commitments-search/AdvancedNotification.aspx.
28. http://martin.wolske.site/wp-content/uploads/2017/07/A-Radical-Reconsideration-of-Digital-Literacy-ISC-Issue-36.pdf.

FOR THE FUTURE

Speculative Applications

WHAT WAYS MIGHT LIBRARIES USE BLOCKCHAIN technology in the future? There are many possible applications, ranging from supporting scholarly communication to credentialing, to community-based collections, to health information management.

SUPPORT FOR SCHOLARSHIP AND SCHOLARLY COMMUNICATIONS

MacKenzie Smith, University Librarian and Vice Provost of Digital Scholarship, University of California at Davis

Blockchain technology has significant implications for the future support of scholarship and scholarly communication—both in terms of the operations and the research activities supported by libraries.

Operations

The obvious applications of blockchain technology for libraries include scholarly resources, particularly *metadata* and *digital objects*. The fundamental concepts of provenance and authenticity in special collections and archives, including records management, allow the authoritative tracking of ownership and other properties of the collection. Blockchain technology could support broad access to provenance and authenticity metadata about library collections, offering a superior solution to the current fragile, labor-intensive record-keeping workflows. For example, recording sale transactions on a blockchain throughout the lifetime of an item leaves no question about its history and provenance, assuming that the

transaction is either native to the blockchain or, if originating off the blockchain, recorded correctly. Recording changes made to an item on a blockchain (e.g., reformatting a digital asset for preservation or amending a database) could make its authenticity simple to verify.

A related application of blockchain is in *research data curation*. Current digital asset management systems have various customized methods of tracking digital asset sources and integrity, such as digital hash values to track unintended changes to digital objects. Blockchain distributed ledgers might be ideal for tracking digital objects on a large scale, as well as tracking locations, owners, stewards, and other metadata that should be reliable and traceable over time.

Similarly, applying blockchain technology to *metadata about information resources*, from news items to research results, might improve public trust in that information by providing new ways to evaluate sources and changes over time. For example, a website such as Climate Feedback, which is now annotation-based, could use blockchain to sign notations or criticisms by scientists using a ledger-based comment system; the signed notations and criticisms could then easily be inspected by readers to establish the credibility of the annotators.[1]

Blockchain platforms could support *new distributed, large-scale metadata systems*, obviating the need for centralized systems like OCLC's WorldCat, Crossref, or ORCID.[2] While technically possible, the advantages of reforming large-scale metadata systems are unclear, given the inefficiency of blockchain technology and the fact that improved trust and greater decentralization are not top priorities for library and research-related metadata systems. Other new technologies, such as linked data, offer equally interesting alternatives for improving efficiency and greater decentralization.

Blockchain-based financial systems could be used to purchase scholarly resources. Finally, research libraries buy scholarly resources from all over the world, in every currency, and currency fluctuations can wreak havoc on library budgets. Blockchain-based financial systems, Bitcoin for example, offer intriguing possibilities for using cryptocurrency to regain control over international financial transactions, between libraries and publishers or among libraries, potentially eliminating exchange rate uncertainty while streamlining acquisition procedures. Achieving this potential would require

large-scale change by many stakeholders, but smaller scale experiments could test the idea, for example, among library consortia that support fee-based interlibrary lending.

Research and Publishing Activities

Today, many libraries manage research data of all kinds. Research data management involves not only storing and curating digital data, but also organizing that data for discovery, providing data governance, and supporting open scientific workflows across the research life cycle. Outside of libraries, there are many efforts underway to apply blockchain technology to *research workflows for improved accountability and reproducibility*, such as Orvium, artifacts.ai, and protocols.io, to name just a few.[3] These offer the potential for better policy compliance-monitoring by institutions and government funding agencies, and for helping to restore public trust in research. If blockchain catches on with researchers, libraries will need to adapt their data services to the technology. Ideally, libraries would be involved in the design and deployment of blockchain platforms and applications for research, and would add their expertise in archiving, digital preservation, and metadata management.

In the area of *scholarly publishing and communications*, there are similar efforts underway to apply blockchain to aspects of those activities, such as version tracking, peer review, and content management. If libraries continue to acquire and manage the outcomes of those activities—books, journals, websites, databases, media, and so on—and make these available to scholars over many years, then libraries need to get involved in defining how blockchain technology will be applied to the scholarly record.

Considerations for the Future

While blockchains are distributed and decentralized, the existence of these new systems could lead to even stronger centralized control of information resources. A particular issue for libraries with regard to the adoption of blockchain technology in publishing and scholarly communication is the potential of blockchain to significantly tighten intellectual property controls and DRM. For example, content distribution using smart contracts on the Ethereum blockchain platform could cripple legal tools like fair use and

eliminate *digital first sale* by creating verifiable transaction records that use licenses and transfer-tracking to limit owner rights and eliminate the possibility of rights expiration.

An area of great interest for blockchain advocates is identity management, such as giving individuals control over their personal information, rather than allowing companies like Facebook, Amazon, and Google to own this information. Identity information is stored on a blockchain using public key and private key cybersecurity protocols and is extremely secure. For researchers who currently have to give information about themselves to a plethora of digital platforms to have an online presence (ResearchGate, Google Scholar, Microsoft Academic, etc.), this notion of regaining control over which platforms get what information is very appealing. But research libraries are all too familiar with professors and students who frequently misplace things.

Another consideration is the nature of research and knowledge production. As explained earlier, blockchain technology is best suited to transactions where immutability is important; data on the blockchain cannot change. While certain aspects of research and scholarly communication are transactional and immutable, like the records of experimental research, scholarship in general is neither transactional nor immutable. Scholarly research is an evolutionary discovery process that is marked by healthy disagreement and sudden paradigm shifts. Unlike finance and purchasing, science and the humanities are messy, and the blockchain is ill-suited to complex and chaotic data.

CREDENTIALING AND CONTINUING EDUCATION

Heather McMorrow, Level Pre-College Program Director, Northeastern University; and Amy Jiang, Library Technology Coordinator, University of La Verne

Universities and governments have been the arbiters of academic accreditation and professional certification for centuries. This has been the most functional and expeditious way to operate. However, this closed

system has also meant ceding authority over our identities and narrowing what is validated as authentic knowledge at a time when the knowledge economy is only expanding.[4] Decentralized applications, such as blockchain, can dramatically improve the issuance and data management of academic credentials, as well as shift the societal view of knowledge and skills verification, return identity, and academic sovereignty; improve accountability; impede corruption; and effectively knock down barriers to economic and social mobility. Decentralized systems place libraries in a unique position to innovate in support of these goals.

Applications

Three particular types of blockchain implementations (cryptocurrency, smart contracts, and systems of record) could play a role in education, impacting a range of services and experiences from university credentials to migration and mobility.[5]

One of the most commonly discussed applications of blockchain in higher education is for *traditional university credentials*. For example, MIT has implemented a Blockcert system to issue digital certificates, as well as dual paper and blockchain diplomas to students for its finance degree and its master's degree in media arts.[6] This Blockcert system gives students the option of receiving a digital diploma in addition to a paper diploma, and because the mobile app gives students unique private keys, a student can therefore prove ownership of his or her diploma. This system does not host student artifacts, but rather simply provides a unique university artifact, in this case a digital diploma, to each student. In Europe, the Open University and the University of Nicosia have also begun experimenting with Blockcerts.

Blockchain technology can also be used to recognize *lifelong learning activities* and to provide *credentialing* for continuing professional education. Governmental and nongovernmental bodies, individual institutions, and funding organizations have explicit directives to recognize alternative paths to credentialing.[7]

Access to Personal Records

Blockchain technology has possibilities for developing economies that are disproportionately impacted by extreme poverty, institutional corruption, climate change, natural and man-made disasters, and a lack of infrastructure.[8] Even in more developed countries, many people can still encounter similar challenges, depending upon their economic status and access to resources. Blockchain technology provides ways for people to retain access to important information during disruptive migration and mobility experiences.

Pros and Cons

Looking specifically at one blockchain application in higher education, credentialing, there are several pros and cons to consider. These pros and cons may vary based on the system of blockchain implementation (e.g., Ethereum, Quorum, Bitcoin), the type of consensus mechanism (e.g., proof of work, proof of stake), the level of permissions included in the credential data, and whether the system is used in a public or private sector. A partial list of pros (cost effectiveness, accuracy, immutability, and identity sovereignty) and cons (access and interoperability) need to be considered.

Some of the benefits are that the distributed nature and advanced cryptography of blockchain make it next to impossible to tamper with. The lack of a centrally held database can eliminate the cost of accessing transcripts and professional credentials. Moreover, blockchain can create automated processes that remove the potential for human error or falsification. This could solve the challenge of exaggerated qualifications and fake credentials and contribute to sovereignty over one's identity. It could allow those who are globally mobile by choice, or by necessity (e.g., people who are unemployed or underemployed, poverty-stricken, employees working or studying internationally, displaced persons, and refugees), to have agency and access over the distribution of their personal documentation. The technology has the potential to reduce dependence on international aid and provide more accurate and authentic data collection, transparency, and accountability for organizations, governments, and regulators.

Blockchain, when used for personal identification, can also provide greater national security by ensuring the swift and accurate identification of individuals—without a central government controlling or accessing such

information for nefarious purposes. This will provide a powerful counterargument to protectionist governments and growing global nationalist movements whose citizenry are fearful of accepting immigrants. It would also help ensure that children who are separated from their parents are not lost in a poorly run bureaucracy or exploited by individuals posing as parents or guardians.

However, there are some limitations to blockchain, particularly in terms of access. From reliable electricity to Internet, technology access is still economically disparate. Hardware is difficult to come by and electricity can be spotty. The issuance and acceptance of blockchain credentials as part of an application packet to college or for a job is limited.

Another issue relates to interoperability and standards. Universities run multiple platforms for information storage and sharing. It is not yet known how the distributed nature of blockchain will affect or cooperate with existing systems. There is also no global consortium, as there is with the Internet's W3C, that establishes standards and practices or manages updates for blockchain technology.[9] Most blockchain implementations have their own consortium.

Considerations for the Future

Libraries need to consider how blockchain implementations would fit into the other IT systems provided to users. It is also important to coordinate blockchain efforts across libraries and institutions; one-off solutions are unlikely to succeed. One promising possibility is for libraries to use blockchain technology in credentialing in the mastery of information literacy units. Libraries could also educate community members about how blockchain technology can help them achieve self-sovereign identity.

DISTRIBUTED ACCESS TO LIBRARY METADATA

Timothy A. Thompson, Discovery Metadata Librarian, Yale University Library

Authority control, controlled vocabularies, and the intellectual and physical control of collections—the idea of control is part of a professional ethos that tends to privilege centralization and uniformity in the organization of information. Centralization on the level of organization and representation maps to centralization on the level of maintenance and exchange: libraries rely on gatekeepers and centralized services to manage knowledge organization systems such as the LOC Subject Headings and related authority files. Ultimately, questions of control are questions of trust: centralized systems and workflows allow libraries to define boundaries within which community norms and standards can be enforced. At the same time, they are exclusionary and limit participation to a set of authorized contributors. Blockchain technology has the potential to reconfigure the relations of information exchange among libraries and to shift the locus of trust from centralized services to distributed systems built on the premise of peer-to-peer interactions.

As an illustration, consider the example of the NACO-México project, which began in 2003 as a cooperative effort among academic libraries in Mexico to contribute records to the Name Authority Cooperative Program (NACO), maintained by the Program for Cooperative Cataloging and the LOC.[10] Currently, an institution that wishes to contribute to the Name Authority File must do so through one of two authorized gateways, OCLC or SkyRiver, both of which charge a membership fee.[11] In 2011, participants in NACO-México were compelled to move from OCLC to SkyRiver because they could no longer afford the cost of OCLC membership. Subsequently, the project dealt with discrepancies in its contribution statistics as recorded by the LOC, which has reported contribution totals that were lower than the project's internal numbers. If NACO itself were to run on a blockchain network rather than as a centralized service, membership fees might be eliminated or replaced by marginal transaction fees, making participation affordable for a broader range of international contributors. Discrepancies in contribution statistics could be eliminated through reference to the blockchain as a permanent, time-stamped record of transactions. Individual libraries might sell or purchase data on demand from peer contributors, thereby decreasing local workloads and creating new opportunities for international collaboration.

For several years, academic libraries have been discussing and undertaking efforts to migrate from legacy metadata formats to linked open

data. This transition has been a difficult one, however. Although work developing semantic ontologies such as BIBFRAME has allowed the library community to examine and reconsider some of its fundamental data models, the implementation of linked open data in academic libraries has been impeded by the absence of an underlying computational architecture that can support new models for data-sharing and production.[12] A major selling point of linked open data is its support for internationalization and integration with the wider web of data. However, without an open, distributed market for the exchange of data, centralized bottlenecks will continue to undermine attempts at systemic change.

Many information professionals have been skeptical of blockchain applications for the cultural heritage domain, echoing the standard critique of blockchains in general: that blockchain applications have been peddled as a panacea when in reality these applications are little more than inefficient databases that offer limited functionality in exchange for troubling amounts of energy (in Proof-of-Work systems) or wealth (in Proof-of-Stake systems).[13] Even leading Bitcoin advocates such as Jimmy Song have argued that the constraints imposed by blockchain technology make it appropriate for only a very limited set of applications (in his view, currency and the exchange of value).[14] For individuals and organizations that are investigating blockchains as a technical solution, it is important from the outset to establish a framework for evaluating their applicability and appropriateness.[15] On its own terms, a blockchain is simply a means to an end, one possible approach to achieving consensus among nodes in a distributed network—and doing so even in the presence of arbitrary system failure or malicious behavior (the so-called Byzantine Generals' Problem in computer science). The question for libraries is whether participation in decentralized networks is desirable for a given use case or can help further their core mission in ways that may not have been possible before.

For example, the problem of entity resolution, also known as record linkage or data-matching, is one that has a direct impact on the work of information professionals in academic and research libraries. In library units responsible for catalog management, many workflows center on the procedure known as copy cataloging, which aims to expedite the processing of new acquisitions. Copy cataloging involves searching a shared database for records created by another cataloging agency, but which describe identical publications that have been acquired by one's local institution. In

the current environment, the global exchange of library catalog data is controlled in large part by OCLC. Although OCLC provides data aggregation and storage services that allow libraries to share their data, this vendor-driven paradigm entails the acceptance of a business model that, in effect, charges libraries for serving their own data back to them, albeit enhanced with different forms of added value.

Libraries have a tradition of experience with data-matching and automation, but now stand to benefit from the increasingly mainstream availability of algorithms and routines developed within the context of data science and machine learning. Sophisticated algorithms for string comparison and probabilistic record linkage have long been available, but are not widely used by libraries, with notable exceptions such as the Virtual International Authority File,[16] which is itself a project of OCLC. As machine learning tools and methods have become more accessible, large-scale, real-time access to library metadata has not necessarily followed suit. The catalog of a large academic library may contain several million records. By comparison, as of August 2018, the OCLC catalog database, WorldCat, contained 427,501,671 bibliographic records in 491 languages.[17] As long as central hubs or service providers maintain control over the aggregated metadata of research libraries, the large-scale computational analysis and utilization of this data will remain out of reach for most.

Of course, when discussing decentralization, there are a range of new technologies that should be considered. Blockchain may or may not be the most appropriate one for a particular use case—or it may need to be used in conjunction with other technologies in order to enable decentralized exchange. Several efforts are underway to develop systems for decentralized file storage using distributed hash tables, one of the most prominent being the IPFS. In a way similar to the software versioning protocol Git, IPFS uses hash values to capture the state of a file at a particular point in time and then serves it on a peer-to-peer network. IPFS hashes might be referenced as links in blockchain transactions in order to decouple the storage layer from the accounting layer.[18]

Technologies and protocols for distributed systems, such as blockchains and distributed hash tables, could allow research libraries to form robust peer-to-peer networks that would enable data-sharing on both macro and micro levels. Public blockchains such as Ethereum and Bitcoin are severely

limited in the amount of data that can feasibly be stored on the chain, but alternative blockchain platforms that address this limitation have recently been developed. For example, the blockchain-based database service BigchainDB offers a robust data-storage solution while ensuring Byzantine fault tolerance and providing blockchain features such as immutability and an asset-based transactional model.[19] By running a "consortium blockchain" network using a system like BigchainDB, academic libraries could be empowered to move away from centralized models and begin managing their data collectively.[20] Instead of paying a centralized metadata hub to distribute their catalog records, libraries could use blockchain technology to share their metadata—whether in batch or as discrete bits—on a peer-to-peer exchange. Many blockchain systems support the creation of so-called smart assets, a term used prominently by the New Economy Movement project.[21] Smart assets can be modeled as nonfungible tokens that represent an object on a blockchain and allow it to be exchanged. Metadata professionals are familiar with the concept of the record or descriptive unit as a surrogate for a real object. As smart assets, metadata objects themselves could be tokenized and represented at the appropriate level of granularity, whether as linked data statements (triples) or as record-like objects. By providing an international peer-to-peer marketplace, blockchain networks could facilitate the free flow of library metadata, potentially creating new revenue streams for individual libraries and replacing some of the costly subscription services that currently predominate.

DATA COLLECTION AND ASSESSMENT

Annie Norman, State Librarian of Delaware and Director, Division of Libraries

Blockchain may provide the salvation for libraries' data challenges. Effectively quantifying and showcasing the value of libraries, especially public libraries, has been a perpetual problem that impacts funding. Blockchain may help public libraries measure their performance and outcomes at scale. Libraries are an open book and share willingly among

themselves, but the sharing is not as efficient and effective as it could be, since library services and data are segregated into silos owned by various vendors and governances. Meanwhile, businesses such as Google and Amazon are taking more aggressively innovative steps relative to the strategies currently employed by most libraries. These businesses have a single-entity advantage that enables monitoring of the entire value chain for customers and all the processes of their businesses, at scale. To leap beyond traditional data management, libraries and library vendors need to collaborate in a much more seamless and transparent way.

Currently, the Chief Officers of State Library Agencies' "Measures That Matter" project is marshalling efforts with IMLS and other stakeholders to assess and resolve public libraries' data challenges. New technological solutions such as a distributed ledger could provide the necessary authentication and privacy capabilities for public libraries to support services such as a universal library card, a union catalog with provenance and ownership verification, data-sharing, and more.[22]

Applications

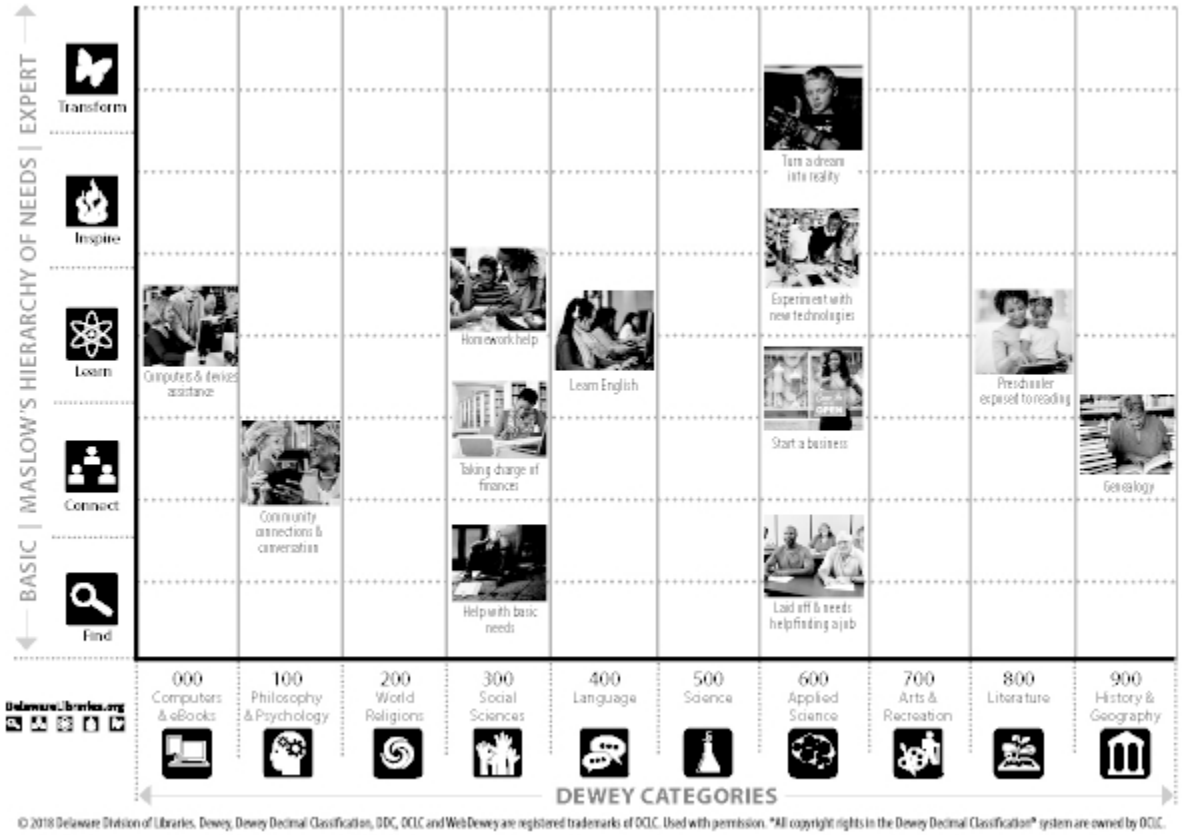
Public libraries typically collect and analyze data using existing library technologies. Blockchain's distributed nature and privacy capabilities may provide libraries with the potential for faster experimentation and more effective and seamless data collection solutions. Here are some ways that Delaware libraries are thinking about using blockchain to address their data collection and analysis needs.

> Could a user-facing blockchain-based system help libraries better collect and understand data about patron needs, about what patrons are trying to do, and, ultimately, how libraries might provide services to meet those needs?
>
> Traditional library statistics, such as totals of circulation, program attendance, and reference questions, yield data about the user in the life of the library. Libraries also need new ways of looking at data from the patrons' viewpoint and showing how patrons are using library services while still respecting patron privacy. Delaware libraries use two macro-organizers to encompass all potential services and community needs, a Dewey/Maslow

framework (see figure 2). The Dewey Decimal Classification System (x axis) is the installed base for Delaware libraries' collections, and the same taxonomy is used to align circulation with program attendance and reference questions *by subject*. A modified version of Maslow's hierarchy of needs (y axis) organizes *all needs* that libraries support, from basic needs to transformational ones. The Dewey/Maslow framework addresses all disciplines and subjects, includes all library and Delaware partner services, and can be used in a variety of ways for planning and assessment.[23]

Figure 2

Dewey/Maslow Framework



Data alignment may be easier to manage using blockchain solutions. By strategically measuring library performance across services, it would be possible to understand how libraries

contribute to improving a variety of community indicators. Consequently, libraries could become even more effective and influential in helping communities to evolve and transform.

What are the cumulative benefits of reading books, attending one-hour programs, and so on, over a year, and over a lifetime?

In Delaware, a method was developed for individuals to capture the benefits that accrue from books read, programs attended, and so on. Tips, tools, and techniques are provided, and the public is encouraged to track their reading and learning over time.[24] Tracking enables members of the public to enhance their learning and quantify the benefits they have received from libraries. A blockchain alternative may help ensure secured and sustainable patron access over the long term.

What proportion of the population do libraries actually serve?

In order to truly understand library performance, the population's use of a library should be analyzed in relation to community development. The stretch target, or capacity measure, of a library is in place to serve the entire population (or close to it), and the library landscape should be seamless across school, public, and academic libraries. Data gathered for this measure will be easier to manage with the blockchain technology noted above.

Considerations for the Future

A holistic view of performance is needed to better understand the cause and effect of library services, which is what funders want to know. Transparency of inputs, outputs, outcomes, and community impact at scale across all libraries would foster a clearer understanding of library performance and ways to improve it. Library leaders are responsible for creating a transparent library infrastructure, and would benefit from working together to implement blockchain more seamlessly to provide evidence of the public library's contributions and value to society.

LESSONS FROM HEALTH INFORMATION MANAGEMENT

Victoria Lemieux, Associate Professor of Archival Science,
University of British Columbia

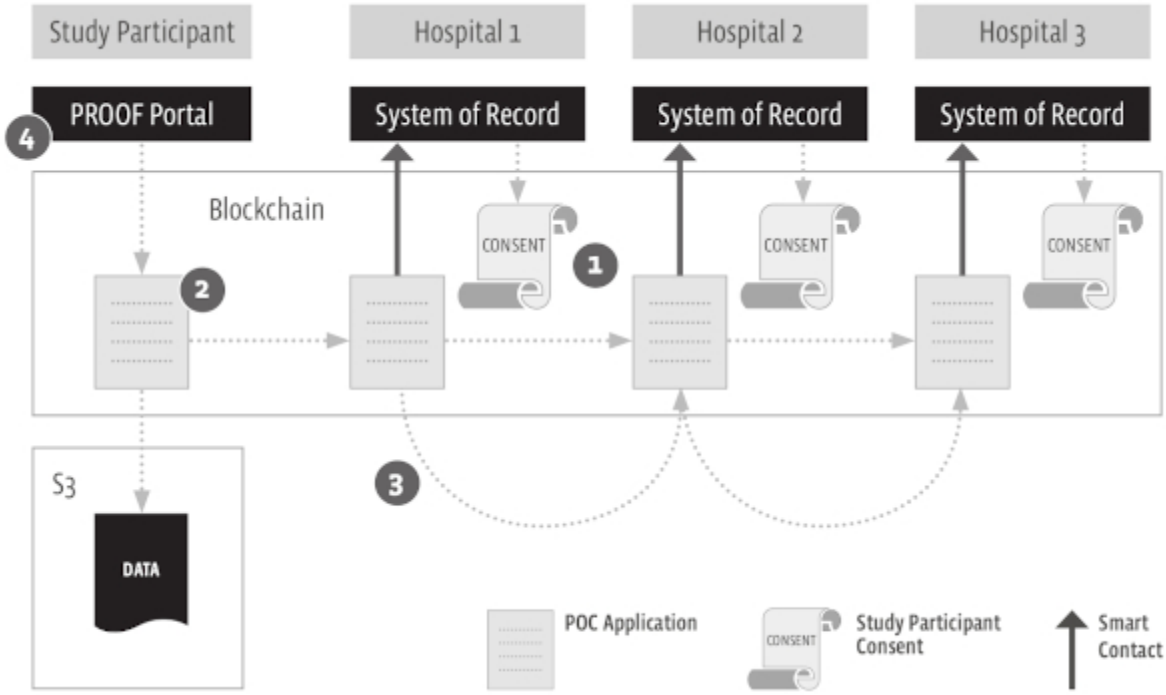
Among the growing interest for applying blockchain technology to the health sector is its potential application as a means of giving an individual direct control over access to her medical records and consenting to secondary use of her health data for research purposes.[25]

Applications

A case study, undertaken as part of the University of British Columbia's "Records in the Chain" Project, focused on a blockchain prototype developed collaboratively by the Prevention of Organ Failure Centre of Excellence and the Deloitte accounting firm in which a member of the Record in the Chain Project was embedded as an observer.[26] The purpose of the blockchain prototype was to manage users' consent to the use of their clinical data for precision health research. The prototype was built for the purpose of making the enrollment of study participants more efficient, eliminating the need for the researcher as an intermediary between study participants and clinical sites in the consent management process, and providing study participants with greater transparency about the usage of their personal data.

The Deloitte proof of concept (PoC) used the blockchain to build a single decentralized, disintermediated system to serve as an interface between participants, researchers, and hospitals. The system allows participants to enroll and consent to the use of their personal data through a web portal, and access time-stamped audit trails of their interactions with the system. This system allows researchers to create studies and invite participants to contribute personal data to them, and also allows researchers to request patient data from other institutions within the system. The data-sharing user journey (see figure 3) shows how the system integrates and coordinates the steps and participants in the previously manual process of researchers requesting data from other institutions.

Figure 3
PoC User Journey



The solution used a Nuco Ethereum private blockchain, which is an extension of Ethereum.[27] Ethereum is an open-source blockchain protocol suite that was originally designed as an alternative to the Bitcoin blockchain platform. The solution stack incorporated a custom front-end graphical user interface and an Amazon S3 file server, with Amazon Web Services. Smart contracts developed in Solidity controlled the workflow.[28]

Pros and Cons

Blockchain-based user consent management of health data has the potential to solve inefficiencies in the health data consent management process, and this could enable more rapid advances in precision medicine to the benefit of all. It can also support more transparency for end users about the consent process and how their personal data are being used.

On the other hand, reliance on blockchain technology for this purpose also raises several critical issues. The first of these is privacy and the protection of personally identifiable information (PII). A number of

jurisdictions have passed legislation regulating the processing of personal data, based on the principles of lawfulness, fairness, transparency, purpose limitation, data minimization, accuracy, storage limitation, integrity, and confidentiality.[29] Beyond these laws, there are a number of principles of privacy protection and the handling of PII that should be respected, such as those established by the International Association of Privacy Professionals' Principles of Fair Practice.[30]

How these laws and principles and their specific implementations can be addressed in blockchain solutions is an open question, as yet. It is unclear, for example, how PII that has been stored on a blockchain would be removed if an individual were to make the request under the "right to be forgotten" provisions of the GDPR. With data breaches on the rise globally, and many of these targeting health records, security is also a major concern. Blockchains, like all digital record-keeping systems, rely upon software and computing technology that can have vulnerabilities due to poorly written code, poor system design, backdoors, and so on, which means that, despite the use of encryption in blockchain systems, information could be at risk.[31]

Another area of concern is usability. While it may be true that users want more control over the custody of their health data and more transparency when it comes to consent to the use of their data, the added complexity of using an unfamiliar and emerging technology may be asking too much of end users. Some blockchain systems rely on users to remember their private keys and keep them secure, and have no key recovery service in the event that a private key is lost. There have been several reported cases of even very sophisticated users of the Bitcoin blockchain losing their private keys and, consequently, access to thousands of dollars' worth of cryptocurrency.[32] What might happen if users lose the private keys that control access to their health data?

Another issue is the legal status of smart contracts. See Blackaby.

These are but a few of the potential issues that may arise in blockchain technology in health information management. Though challenging, there are potential solutions or ways to mitigate risks, as discussed in the following subsection.

Considerations for the Future

Steps that might be taken to mitigate the downside of using blockchains to manage health information include:

- Adopt the principles of privacy by design.[33] As an example, in the PROOF use case, Amazon S3 was chosen for data storage because it avoids placing personal health data on the blockchain and complies with data localization laws in British Columbia.
- Develop a security model that encompasses the entire blockchain solution stack, for example, including decentralized applications (DApps), such as wallets and smart contracts, and not just the core blockchain processing layer, since the former tend to be the most vulnerable components of a blockchain system, based on experience with the use of blockchain technology to date.
- Employ user-centered design principles and conduct usability testing to ensure that end users can interact with the system effectively. This may indicate the need for a key recovery or escrow service, thus giving up some decentralization in order to create a better user experience.
- Understand the legal status of smart contracts and blockchain-based records and ensure that systems are designed and implemented to conform to principles of good records management, such as ARMA International's "Generally Accepted Recordkeeping Principles" and the "International Records Management Standard," ISO 15489.[34]

ELECTRONIC HEALTH RECORDS

Frank Cervone, Executive Director for Information Services, School of Public Health, University of Illinois at Chicago

Within the realm of health information management, both proposed and currently operational blockchain projects provide clues for useful applications in libraries and information organizations. This should not be too surprising, given the close historical relationship between medical information management and librarianship. Within the health care field,

blockchain continues to rapidly move from theoretical discussions to specific applications impacting areas ranging from pharmacology and medical device supply chains to the recruitment of patients for clinical trials, and to improving the security and interoperability of medical devices. In the realm of health care, blockchain is now a reality.

Applications

Applying blockchain technology to health information management is now possible because of the standards for medical information interoperability, which are conceptually equivalent to those in the library and information science fields. For example, Fast Healthcare Interoperability Resources (FHIR) is a developing standard that defines data formats and elements, along with providing publicly accessible application programming interfaces for the purpose of exchanging electronic health records (EHR). This standard is the health care equivalent of a completely normalized machine-readable cataloging system with predefined access interfaces. FHIR offers the potential to extend EHRs outside the constructs of traditional electronic health care systems to mobile and cloud-based applications, medical device integration, and flexible or customized health care workflows.[35]

Several current, real-world applications of blockchain technology in health information management are providing baseline functionality. *Provider directory services* is a joint project between Optum, Quest, and Humana to provide common, distributed health plan provider directories. This is a mega-directory of health care professionals within specific health care systems and a good example of an application that is trying to address the problem of managing multiple sources of truth through reconciliation between those sources.[36] *Validation of patient identities* is a project by the government of Estonia to create a blockchain-based framework to validate patient identities.[37] All citizens are issued a smartcard, which links each individual's EHR data with his or her blockchain-based identity. In light of recent concerns about scheduling fraud at the U.S. Department of Veterans Affairs and the risk for data manipulation of implantable medical devices, such as pacemakers, this type of system has several potential benefits to guarantee that any modifications to an individual's health care record are secure and auditable.

MedRec is a project between the MIT Media Lab and the Beth Israel Deaconess Medical Center that provides a decentralized platform for managing permissions, authorization, and data-sharing between health care systems.[38] A *Data Provenance Toolkit* is being developed by RAIN Live Oak technology to support the creation and validation of provenance records.[39] While focused on health technology, this application could be useful in a number of other settings, including library and information science.

A major benefit of blockchain for EHRs is that an immutable audit trail guarantees the integrity and provenance of data. Once a transaction is committed, it cannot be changed, which guarantees the integrity of the transaction data as well as the provenance of the data. Records are *signed* by the source, which allows the legitimacy of records to be verified and false records to be plausibly denied. Furthermore, security and privacy are increased because data is encrypted in the blockchain and can only be decrypted using the patient's private decryption key. Therefore, even if the network is infiltrated by a malicious party, there is no feasible way for them to read patient data.

Pros and Cons

Blockchain presents both benefits and challenges in health care. For example, blockchain technology will allow health data to be collected over the life span of an individual. The benefit to the patient is a complete medical record which accurately details that individual's entire health history. A patient could then make this data available to any health care provider, and individual health outcomes should be improved. Furthermore, the agglomerated data could also better support research as a rich base of information to be analyzed in order to help predict future health concerns at the population level.

A critical factor in this approach is that the patient would have control over who has access to his data. The patient could then provide access at varying levels, ranging from full access to all of his health records, to just a limited subset based on the application.[40] For example, a patient visiting his dentist could then restrict his health data to just those aspects that would be relevant for dental practice. Similarly, patients participating in research, such as drug trials, can limit the amount and type of their health data that is

made available to researchers, as well as what personal information the researchers can subsequently share with others. This same type of disclosure scheme could be applied within bibliographic systems to allow the patron to control his personal information as appropriate to the context, such as opting out of sharing circulation or electronic journal access information from recommender systems.

Currently, much work is being done on creating an environment where a common database of patient information could be built using blockchain technology. This same idea could be applied to libraries and information organizations through the use of a common database of user information, allowing for a common patron database that would enable universal access across library systems.

Considerations for the Future

Clearly, the basic architectural features of blockchain, such as the immutability of transaction logs, are of benefit to EHRs. This same functionality of blockchain applies to libraries as well, such as the ability to demonstrate provenance; however, this also poses issues where data needs to be forgotten, such as with circulation and access information. Nonetheless, current applications like health data exchanges can be examples of creative initiatives to share common data, such as using health data reporting models as a basis for a common library reporting system to gather data for advocacy and efficacy initiatives.

BORN-DIGITAL AND DIGITAL-FIRST CONTENT

John Bracken, Executive Director, Digital Public Library of America; and Michael Della Bitta, Director of Technology, Digital Public Library of America

As various entities across sectors begin to explore the uses and implications of blockchain technology, libraries have a critical role to play. First, the public trust in libraries, which is sadly diminishing among so many other

civic institutions, advances the public good by building and advocating for new tools, services, and approaches. Second, the inherent networked and cryptographically authenticated nature of the blockchain, and its reliance on open source, plays to the library's strength as a nurturer of community. Third, the skills of the library's information technologists who build systems enable libraries to build and contribute to blockchain technology, placing the library in a position that guides the crafting of blockchains for the public good.

Libraries are still in the early days of distributed digital media, and thus in the first moments of being able to imagine the possibilities of blockchain technology. In this period of exploration and experimentation, one of the library's challenges is to look beyond the blockchain's function as a medium of data storage.

Instead, information professionals need to interrogate how the distributed network of disparate parties with common interests that powers a blockchain system can also be employed for the common good. By creating networks in which libraries and members of their communities—or civic society more broadly—can work together to record and verify knowledge, there is the potential to build new, dynamic, trusted sources for public information.

Born-Digital and Digital-First Content

The move to digital-first media and platforms makes verifiable archives difficult because it is very easy to introduce alterations into the archived data. As more and more essential communications and publications are born-digital, it is incumbent that archival practices scale to match the new problems these media and platforms introduce.

It would be easy to assume that the practices that grew out of the movement to archive digitally captured physical media are applicable to born-digital media. However, the landscape of born-digital media is changing rapidly, and with these changes come new forms of attack. There are established practices for using cryptographic signatures to ensure against the authoritarian tampering of archival materials, but there are problems when this approach happens under a single business domain.

In contrast to the beginning days of the Internet, storage and bandwidth are now extraordinarily inexpensive. The cost to store a gigabyte has fallen from the near-million-dollar range in the early 1980s to single-digit cents in this decade.[41] And the costs to transmit a gigabyte of data have also declined proportionally. Along with these declining costs, the variety of channels used to store files has increased.

The fixity of digitally captured physical objects was originally conceived of as a way to protect against file corruption, and not one to counter deliberate forgery or attack by a hostile party. Since digital surrogates are commonly backstopped by a physical copy of the object, it would be easy to expose a campaign to forge digital copies of an artifact by referring back to the original. However, in the case of a born-digital object, it is hard to identify an original using intrinsic characteristics.

Given these realities, one attack on born-digital content archives would flood channels with subtly changed copies of the archival content. Since it is hard to authenticate which copy of the file is “original,” and traditional fixity measures are maintained by the same administration that houses the content, it would be impossible to determine which copy is real beyond relying on the reputation of the archival institution.

In contrast, a digital archival system that records auditable, independent observations of the source material to prove that they are copies of the source eliminates the problems of objects being manipulated during or after ingest. Blockchains are an ideal means for recording these observations because they are decentralized and cryptographically secure, which protects against tampering by a single institution and provides an audit trail. There is no need for trust in a third-party authority.

Pros and Cons

Blockchains help establish trust by sidestepping organizational reputation issues through the use of overlapping data structures, decentralization, and cryptography. However, a cornerstone of this approach is that the information being independently verified must be observable by multiple disparate parties. This is possible in cases where the archival material is public or at least accessible by different types of institutions.

Unfortunately, not all archives work in this manner. Records are routinely archived and only made available for viewing at a later date. Using a blockchain solution in this situation would require that those records at least be shared with independent third parties. Distributed trust would be weakened in this scenario, since the authority to select which of these third parties are allowed to verify the archival material could be viewed as self-serving. An overarching system of selecting parties for independent verification could be established, but that creates other trust issues. Also, storing the archival media on the blockchain may be infeasible, given bandwidth and storage considerations. However, a system that stores the archival content in a separate storage medium, and merely records the fixity information on the blockchain, will scale more easily.

Considerations for the Future

Ideally, a proof-of-concept implementation of this idea would be undertaken by a small number of institutions that have the means to build shared library technological infrastructure and could create a sponsoring organization to author standards, evangelize the new technology, and shepherd its growth among institutions that are interested in decentralized archiving. The Digital Public Library of America (DPLA), which was founded as an open, distributed network, can help to facilitate these connections and conversations.[42] In fact, as information professionals explore this new topic together, the process that led to the founding of the DPLA could provide lessons. The DPLA brought together diverse stakeholders inspired by an idea, and through an intentional, national collaborative process, they iterated on that concept to build a plan of action.

COMMUNITY-BASED COLLECTIONS

M. Ryan Hess, Digital Initiatives Manager, Palo Alto City Library

The blockchain opens up the possibility of extending library-like services beyond the library walls. In leveraging blockchain’s facility for establishing trust between total strangers, libraries could deploy a community sharing

platform, wherein community members can exchange resources. These resources could be anything from books, objects, and tools to services and know-how. The blockchain would serve as a ledger of transactions between individuals, with smart contracts utilized to govern loan periods and other borrowing policies. This community-based collection solution could empower individuals to participate in a peer-to-peer lending service.

Community-Based Collections Platform

The community-based collections platform could be composed of five major elements: the blockchain ledger, smart contracts, application, participants, and collections.

The Blockchain Ledger

The blockchain ledger can be used for a number of critical services, including providing an inventory of resources that participants have made available, a record of transfers of resources between participants, and a balance sheet of participant accounts.

Smart Contracts

A programmatic layer, or smart contract, will be required to govern policies. Policies might include setting borrowing periods, defining borrowing privileges and limitations, fines, due dates, hold limits, and so on. The smart contract functions like an integrated library system, establishing rules for using collections. But it differs in that participants would be able to create contracts with rules that meet their needs.

The Application

Participants will require a user interface to interact with the community-based collections. There are two ways to provide this interface: a distributed application (DApp) or a standard mobile app. This app would query the blockchain ledger to enable any number of activities. It would also interact with the smart contract to restrict and guide usage of the collections via the app. Features of the app might include:

- a catalog search of shared resources
- collection management for adding or removing shared resources

- network management for creating, configuring, joining, and leaving sharing networks
- account management
- fine payment

Participants

Participants can be both individuals and institutions with resources to share. Participants will form networks on the platform as defined by policies in the smart contract. For example,

- members of a municipal area may restrict access to their network to only those members living within a defined geographic area
- a community organization might restrict access to its membership
- a group of friends might restrict access to their resources to only those in their social circle

Both individuals and institutions can operate within multiple networks (e.g., a student may simultaneously participate in a sharing network with his school, local library, and social circle).

Collections

Resources within the community-based collections can include anything that could be assigned a unique identifier, or what is known as a hash. Resources can include objects like tools, books, art, musical instruments, and technology. Resources might also include services and skills such as expertise, training, or advice, as long as the service in question can be tracked and authenticated. For this reason, community-based collections may require auxiliary credentialing blockchains that could vouch for the quality of the service before they can be reliably incorporated.

Technical Components

There are a number of ways to build the community-based collections platform. Some potential components are already available. Others are emerging.

Ethereum

Ethereum is a versatile blockchain technology that could provide the blockchain ledger for tracking the various transactions and provenance of items. Even more importantly, Ethereum includes smart contract functionality, allowing for programmable rules to govern accounts, networks, and collection usage.

IPFS

IPFS is an emerging peer-to-peer Internet protocol that allows computers to exchange data without requiring centralized web servers to host that data. IPFS makes a truly distributed web possible, including the deployment of DApps made available by peers on the IPFS network. Like the blockchain, DApps behave more like a shared operating system, encoded into the Internet itself and available to anyone.

The front-end code of community-based collections could be made available as a distributed app via IPFS. This is not a requirement, but it makes sense in many ways. For example, using IPFS negates the need for any one organization to maintain expensive servers to keep the code online. Instead, the app would be made available by numerous peers. These peers could be libraries, but also supporters of the library in the community.

Identity Management Blockchains

Emerging blockchain-based identity management platforms are in the works. Examples include uPort, Hyperledger, and Civic.[43] The community-based collections might be able to leverage these blockchains to define who has access to a sharing network. Such a platform might also be a good candidate for validating skills or credentials, as well as providing proof of service quality.

Building a Prototype

Phase I: Investigation

A working group of blockchain experts, programmers, and librarians should convene to consider the feasibility of community-based collections. The focus of these groups would be to answer the outstanding questions, resolve technical issues, and determine whether or not to proceed.

Phase II: A Prototype

Assuming the working group green-lights further work, a new project team should be assembled to develop a prototype. The first stage for this group would be to develop a project plan for building a prototype. Included in the plan would be technical and functional requirements. The plan would also need to include a budget for developing the prototype. The second stage would include obtaining a new grant to build the prototype and test it with users. The testing will ensure the security and usability of the app.

Phase III: Public Beta

Once testing is complete, a new iteration of the app can be built using another round of grant funding. This app would then be refined until it was ready for deployment as a public beta.

Pros and Cons

The blockchain suffers from some notable limitations which pose challenges to a community-based collections platform. Two major concerns are the energy costs of maintaining the blockchain, and the slowness of validating transactions. While future innovations will likely reduce these obstacles, we must ask ourselves if such a system warrants the effort. Indeed, centralized applications might be able to provide a similar sharing service, without blockchain's limitations.

Conceptually, centralized versions of a community-based collection platform are possible using existing, non-blockchain technology. Integrated library systems are one example in which a centralized server governs the usage of library-controlled resources. It is conceivable that additional features could be developed to incorporate community-based resources into an ILS. Centralized sharing services like Uber, Airbnb, and Bird Scooters already provide the basis for peer-to-peer sharing networks.

However, the *killer app* of a blockchain-based sharing network is its ubiquity. Such a DApp would essentially embed the library service into the fabric of the Internet. Like viral code, once released onto the Internet, such a platform could be utilized freely by anyone to begin sharing with anyone else. Unlike a library which must maintain funding and support to operate, the community-based collections platform would require only a network of participants to ensure its existence.

Moreover, the flexibility of smart contracts means that the platform can be deployed for endless use cases. Unlike a centralized application where an institution or company writes the rules, on a community-based collection platform, users could write rules that meet their specific needs.

Assuming that the technical issues related to the blockchain can be overcome, the community-based collection platform would be a *killer app* for resource-sharing. Put another way, the blockchain platform has the potential of infecting the Internet with library values.

Considerations for the Future

Much work remains to realize this vision, and there are a number of key questions that must be answered before implementing a blockchain project:

- What is the optimal incentive system for such a platform?
- How do we ensure robust security, particularly around the smart contracts and front-end code?
- What is the library's role in such a platform? Will the platform be completely autonomous?
- What would the energy costs be to process transactions?
- Who would build such a system, and with what resources?

Therefore, this project requires a long-term, funded enterprise that can turn vision into reality.

1. https://climatefeedback.org/.
2. https://www.worldcat.org/, https://www.crossref.org, and https://orcid.org/.
3. https://orvium.io/; "Building the Ledger of Record for Research," ARTiFACTS, artifacts.ai; and "Make Your Science More Reproducible," protocols.io, https://www.protocols.io/.
4. Ibid.
5. https://www.mercycorps.org/research/block-block.
6. https://medium.com/mit-media-lab/what-we-learned-from-designing-an-academic-certificates-system-on-the-blockchain-34ba5874f196.
7. https://ec.europa.eu/ploteus/en/content/descriptors-page; http://education.ohio.gov/Topics/Quality-School-Choice/Credit-Flexibility-Plan; and https://www.nzqa.govt.nz/assets/About-us/Publications/Brochures/introducing-nzqa.pdf.
8. "Global Warming of 1.5°C: Summary for Policymakers," Intergovernmental Panel on Climate Change, 2018, http://report.ipcc.ch/sr15/pdf/sr15_spm_final.pdf.
9. https://www.w3.org/.

10. Julia Margarita Martínez Saldaña, “Informe sobre las actividades del proyecto NACO-México,” PowerPoint presentation, Sites/Cites, Texts, and Voices in Critical Librarianship: Decolonizing Libraries and Archives, Seminar on the Acquisition of Latin American Library Materials (SALALM) 63, Mexico City, Mexico, July 2, 2018.
11. https://www.oclc.org; https://theskyriver.com.
12. https://www.loc.gov/bibframehttps.
13. https://web.archive.org/web/20180812122624/https://go-to-hellman.blogspot.com/2018/06/the-vast-potential-for-blockchain-in.html.
14. https://web.archive.org/web/20181031220327/; https://medium.com/@jimmysong/why-blockchain-is-hard-60416ea4c5c.
15. Brian A. Scriber, “A Framework for Determining Blockchain Applicability,” *IEEE Software* 35 (July/August 2018): 70–77.
16. https://viaf.org/.
17. https://web.archive.org/web/20181029213224/https://www.oclc.org/en/worldcat/inside-worldcat.html.
18. https://web.archive.org/web/20181010171040/https://medium.com/@mycoralhealth/learn-to-securely-share-files-on-the-blockchain-with-ipfs-219ee47df54c.
19. https://www.bigchaindb.com.
20. https://web.archive.org/web/20181027023321/https://blog.ethereum.org/2015/08/07/on-public-and-private-blockchains/.
21. https://web.archive.org/web/20190420165029/https://nem.io/technology/.
22. https://measuresthatmatter.net/summary-of-the-measures-that-matter-implementation-group-meeting-november-28-29-2018/.
23. https://guides.lib.de.us/departners.
24. http://guides.lib.de.us/unleashinnergenius.
25. http://dci.mit.edu/assets/papers/eckblaw.pdf.
26. D. Hofman, C. Shannon, B. McManus, K. Lam, S. Assadian, R. Ng, V. Lemieux, “Building Trust & Protecting Privacy: Analyzing Evidentiary Quality in a Blockchain Proof-of-Concept for Health Research Data Consent Management,” in 2018 *IEEE Conference on Internet of Things, Green Computing and Communications, Cyber, Physical and Social Computing, Smart Data, Blockchain, Computer and Information Technology, Congress on Cybermatics.* Presented at the iThings / GreenCom / CPSCom / SmartData / Blockchain / CIT / Cybermatics (2018): 1650–56.
27. https://www.ethereum.org/.
28. https://medium.com/coinmonks/introduction-to-solidity-programming-and-smart-contracts-for-complete-beginners-eb46472058cf.
29. L.-D. Ibáñez, K. O’Hara, and E. Simperl, “On Blockchains and the General Data Protection Regulation,” presented at the EU Blockchain Observatory and Forum: Workshop on GDPR, Data Policy and Compliance, University of Amsterdam, 2018.
30. https://ischoolblogs.sjsu.edu/blockchains/blockchain-chock-full-of-problems-for-medical-data-privacy-by-jessica-berger-mlis-cipm/.
31. I.-C. Lin and T.-C. Liao, “A Survey of Blockchain Security Issues and Challenges,” *IJ Network Security* 19, no. 5 (2017): 653–59.
32. S. Eskandari, J. Clark, D. Barrera, and E. Stobert, “A First Look at the Usability of Bitcoin Key Management,” 2018, arXiv preprint, https://arxiv.org/abs/1802.04351.

33. http://dataprotection.industries/wp-content/uploads/2017/10/privacy-by-design.pdf.
34. ARMA International, “Generally Accepted Recordkeeping Principles,” https://www.arma.org/docs/bookstore/theprinciplesmaturitymodel.pdf?sfvrsn=2; ISO 15489-1:2016 – Information and Documentation – Records Management–Part I: General, ISO, Geneva.
35. https://www.healthit.gov/sites/default/files/12-55-blockchain-based-approach-final.pdf.
36. https://www.forbes.com/sites/brucejapsen/2018/04/02/unitedhealths-optum-and-humana-in-blockchain-deal-to-improve-doctor-directories/#528b59d13998.
37. https://e-estonia.com/solutions/healthcare/e-health-record/.
38. https://medrec.media.mit.edu/.
39. http://www.liveoaktech.net/about.html.
40. http://bulletin.facs.org/2017/12/blockchain-technology-in-health-care-a-primer-for-surgeons/#.WxWFgyAh3mF.
41. http://www.mkomo.com/cost-per-gigabyte-update.
42. https://dp.la/.
43. https://www.uport.me/; https://www.hyperledger.org/; and https://www.civic.com/.

FOR THE PRESENT

WHAT CAN LIBRARIANS DO TODAY TO PREPARE FOR blockchain possibilities in libraries? Education is the key ingredient in getting ready to adopt new technologies in libraries—both for librarians and for the community.

BLOCKCHAIN EDUCATION FOR COMMUNITIES

Link Swanson, Systems Architect, Minitex

A recent survey by the Pew Research Center revealed that a large majority of the American public expects libraries to provide technology education programs: "78 percent of those 16 and older say libraries should *definitely* offer programs to teach people how to use digital tools such as computers, smartphones and apps."[1] This expectation extends to blockchain technologies. As the attention and hype surrounding blockchain increase, more individuals are looking to libraries to help them understand the fundamentals of this technology. This phenomenon was cited at a recent Senate Banking, Housing and Urban Affairs Committee hearing on virtual currencies, where Christopher Giancarlo, chairman of the U.S. Commodity Futures Trading Commission, stated:

> America's libraries are a place where a lot of people go and research Bitcoin . . . One of the most frequently searched items from a library computer is "Bitcoin." So, we're teaming up with the U.S. Consumer Finance Protection Bureau to go out to America's libraries and educate librarians to direct patrons to use our Bitcoin website and other resources.[2]

Libraries can meet this growing demand for blockchain education by understanding some of the core challenges that learners encounter in their attempts to acquire a basic understanding of blockchain fundamentals, and by adopting practical, targeted strategies to integrate into their own blockchain education curricula.

Library-Based Blockchain Education

Library-based blockchain education offerings make sense for a number of reasons. One reason, mentioned above, is that the American public expects libraries to offer education programs that can teach people about new technologies, and blockchain is no exception to this expectation. Because private blockchain education and training services can be very expensive, libraries can help underprivileged and low-income communities become familiar with these emerging technologies and help prevent income-based gaps in access and knowledge, which is one of the core values of librarianship. Another reason is that blockchain technologies can be applied to use cases that are highly consistent with the traditional mission of libraries, such as accurate record-keeping, information availability, censorship resistance, data transparency, and privacy. Finally, if libraries and library-focused organizations hope to launch successful blockchain projects of their own, the success of these projects is highly dependent upon the level of understanding held by project leaders, designers, stakeholders, and—most importantly—users. This final point is worth stressing: the success of a blockchain is directly proportional to the number of engaged participants on that blockchain, and the key to increasing the number of active participants is *education*.

What do library-based blockchain education programs look like? There are three distinct types of blockchain education. The first is research-based, where learners are provided with a list of educational materials from which they can acquire an understanding of the history of blockchain, the technical fundamentals, and common use cases, such as cryptocurrencies. The second type consists of traditional instructor-led training and hands-on learning, where learners complete courses designed for beginner, intermediate, and advanced levels of blockchain understanding. The third type is the concept of a blockchain “starter kit” or “empowerment package,” which would contain all of the steps needed to install a wallet

application for personal use of a cryptocurrency in order to facilitate self-guided, hands-on experience using a blockchain application.

Blockchain Education: The Charge and the Challenges

The task of developing effective blockchain curricula poses a unique set of challenges. At the root of these challenges are *cognitive barriers* to understanding blockchain. Blockchain-based software systems have novel properties—decentralization, peer-to-peer architecture, and trustless interactions—that make them quite different from traditional systems, which are typically centralized, built on a client-server architecture, and reliant on a trusted authority. While these novel properties make blockchain a uniquely promising technology, they also create cognitive tension in learners who are accustomed to using traditional software systems. For example, years of experience using traditional database applications leads an individual to expect that when an account password is forgotten, it can be recovered—or if a transaction is made by mistake it can be reversed—with the help of the system administrator's privileged access to the system. However, this expectation does not translate to blockchain systems because there is no administrator who can recover the account or reverse a transaction with privileged access; there is no privileged access to the system, period.

Examples like this (of which there are many) illustrate how expectations and concepts inherited from years of interacting with traditional software systems can create cognitive barriers to understanding blockchain technologies. Blockchain education offerings should target these kinds of cognitive barriers in order to maximize educational effectiveness.

Addressing the Challenges with Targeted Implementation

There are some practical approaches that can serve as starting points for the development of blockchain education offerings, and are designed to target the cognitive barriers mentioned above. Broadly, these strategies aim to impart a combination of three distinct types of learning:

1. Historical understanding of the evolution of blockchain
2. Understanding of the component technologies that blockchain is built from

3. Hands-on experience interaction with and participation in a live blockchain

Historical Understanding of the Evolution of Blockchains

Knowing the history of Bitcoin and later cryptocurrencies is an effective first step in understanding blockchains. Bitcoin was created to address a need and solve a problem; it combined existing technologies in novel ways to accomplish these goals. Books and documentaries about the birth of blockchain can be integrated into educational offerings to serve as a foundation for explaining how blockchain is different from other software systems.[3]

Understanding the Component Technologies of a Blockchain

Blockchain technologies are novel and innovative in part because of the ways in which they combine *existing* technologies, such as public key cryptography, decentralized peer-to-peer networks, distributed systems, and open-source software communities. The acquisition of a (basic) understanding of each of these component technologies is a powerful way to understand blockchain fundamentals. Moreover, educational materials *already exist* for these topics and can be integrated directly into blockchain education offerings.

Hands-on Experience: Learners Can Interact with a Live Blockchain

Hands-on learning is incredibly effective. The most straightforward way to gain hands-on experience interacting with a blockchain is simply by using cryptocurrencies. For example, learners can acquire a small amount of Bitcoin and use it to complete learning exercises, such as:

1. Install a desktop wallet on your local computer and create a new Bitcoin address.
2. Buy five dollars' worth of Bitcoin from an exchange.
3. Send your funds from the exchange to your local desktop wallet.
4. Install a mobile wallet on your smartphone.
5. Transfer funds from your desktop wallet to your mobile wallet.
6. Create a paper wallet.
7. Send some funds to your paper wallet.
8. View the balances of your wallets' public addresses using a block explorer.

9. Back up your private keys from each wallet.
10. Buy goods and services from merchants who accept cryptocurrencies.[4]

It is important to keep in mind that cryptocurrencies are the first *successful live production implementation* of blockchain technology. Experience using cryptocurrencies is a unique and powerful tool for learners to begin to understand the fundamental concepts of blockchain via hands-on learning.

Hands-on Experience: Learners Can Run a Full Node on Personal Computers

Many of the unique advantages of blockchain systems come from the participating *nodes* or *peers* in the peer-to-peer network. A learner can gain an understanding of the function of blockchain nodes by *becoming one*: she becomes a "peer" by running a "full node" on her own computer. During this process, learners download the entire blockchain to their hard drive, which facilitates further understanding of the decentralized nature of the blockchain. Becoming a peer can be an effective educational exercise which can demonstrate how a blockchain functions as a collection of decentralized nodes.

Hands-on Experience: Learners Can Code an Application That Interacts with a Live Blockchain

At a more advanced level, learners can build a simple *exercise* software application that interacts with a live public blockchain. This hands-on process will present learners with many opportunities to deepen their understanding of blockchain technologies. Coding exercises of various levels of difficulty could be integrated into blockchain education offerings.

There is an increasing demand for blockchain education—and there is an expectation that libraries will offer these educational programs. Libraries can help ensure that underprivileged and low-income communities have opportunities to become competent and equal participants in public blockchains. Furthermore, education is central to the success of any blockchain project, since it facilitates a clear understanding of the capabilities of the technology among project teams, prevents misapplications of the technology, and promotes greater numbers of users, which are fundamental to the health and success of any blockchain.

1. www.pewinternet.org/2015/09/15/libraries-at-the-crossroads.

2. Senate Banking, Housing and Urban Affairs Committee Hearing on Virtual Currencies, 2018, https://www.c-span.org/video/?440770-1.
3. For a book-format introduction, see A. M. Antonopoulos, *Mastering Bitcoin: Programming the Open Blockchain* (Sebastopol, CA: O'Reilly, 2017), available for free at https://github.com/bitcoinbook/bitcoinbook. For a documentary film, see C. Cannucciari, *Banking on Bitcoin* (documentary), IMDB.com, https://www.imdb.com/title/tt5033790/.
4. *Banking on Bitcoin* (documentary), IMDB.com, https://www.imdb.com/title/tt5033790/.

CONCLUSION

IS BLOCKCHAIN SNAKE OIL OR IS IT THE FUTURE? DESPITE the hype around blockchain technology, there are also many realistic possibilities for blockchain in libraries in the future. Librarians should learn more about blockchain technology in order to gain a solid understanding of it and take a leadership role in envisioning how this new technology can enhance and improve library operations and access.

Throughout our eighteen-month investigation of blockchain technology and its implications for libraries, we identified the following as some of the most promising ideas for testing out blockchain technology in libraries. International ILLs could use blockchain to help libraries easily pay for international ILL requests using cryptocurrencies and track ILL transactions. Universal library cards could be created using blockchain technology and potentially improve access to materials in public libraries across library systems. Libraries could get involved in credentialing, which could use blockchain technology to verify and hold some of the credentials and identity documentation for stateless and other individuals in the blockchain, and they could also use blockchain to issue credentials to users who achieve certain skills (such as information literacy skills). Blockchain could also be used for archival records, particularly to determine the accuracy and consistency (validity) of data over its life cycle.

While many benefits of blockchain technology are touted, such as its immutability, its elimination of the middleman in transactions, and its near tamper-proof security, there are also many concerns about this technology: its costs, its environmental impact, what happens if someone loses or forgets his private key, and how to address privacy concerns such as the right to be forgotten. As blockchain technology matures, librarians should pay close attention to the development of standards, privacy protections, security, and legal implications to ensure that these align with library goals and support the needs of libraries.

As we look to the future, the next step that is needed is to create testbeds for some of these ideas so that libraries can get a full sense of where and

how blockchain could help them and where other technologies might offer better solutions. Creating testbeds presents a great opportunity for libraries to work collaboratively together to develop blockchain pilots and lead the way into the future.

SELECTED RESOURCES

Blockchains for the Information Profession (website and blog): https://ischoolblogs.sjsu.edu/blockchains/

Blockchain and Decentralization for the Information Industries (MOOC): https://www.canvas.net/browse/sjsu/courses/blockchain-decentralization

Blockchain Glossary: https://blockchainhub.net/blockchain-glossary/

ABOUT THE EDITORS

DR. SANDRA HIRSH is director and professor of the School of Information at San José State University. She is a past-president of ASIS&T and is the president-elect of ALISE. Her research interests include information seeking behavior, online education, and LIS education, and she was the co-PI of the IMLS-funded blockchain investigation. Hirsh cofounded and cochairs the global virtual Library 2.0 conference series.

DR. SUSAN ALMAN is a member of the San José State University School of Information faculty. Prior to this appointment, she taught at the University of Pittsburgh's School of Information Sciences and the University of Michigan. She was the co-PI in the IMLS-funded project, Investigation of Possible Uses of Blockchain Technology. Alman's areas of interest are in online teaching and learning, foresight studies and emerging technologies, marketing/PR, and interpersonal communication.