A blockchain system for electronic health records (EHRs), framed as a protocol through which to access and maintain health data, guarantees security and privacy through empowering the user with control of their own data. While using a blockchain architecture approaches interoperability through centralization of data, the use of Ethereum’s smart contracts enables an unprecedented ease of data sharing which transcends in simplicity of use and security. Despite this potential, these advancements depend on patients’ ability to own their health data and the establishment of a structure for identity verification. Furthermore, the establishment of these systems is contingent on the ability of patients to navigate these systems with competence. Separate even from patient use, the viability of a blockchain solution is determined by the security and standardization of the existing EHR systems. And aside from the security of a blockchain solution, there are few incentives for individual hospitals to work to make their EHRs accessible through a blockchain, and thus the government must lead this endeavor.
A FRAGMENTED DATA ECOSYSTEM

23andMe offers personal genomic-sequencing for $99.¹ Internet of Things (IoT) devices, such as Fitbits, synchronize with smartphones to enable users to capture their movements and heart rate.² Apple takes individual-level data collection a step further with ResearchKit, enabling researchers to solicit patients for large scale studies based on patient-reported data. Modern means of data collection are creating new data formats, and these new domains are bolstering clinical trial data collected by traditional means³ with considerable reliability.⁴

Despite the dearth of health data in new forms, Electronic Health Records (EHRs) created and maintained by traditional practitioners form a majority of existing medical data. Individual hospitals generally maintain their own health records,⁵ which hinders the utility of EHRs for patients, as a patient cannot access all of their data in one place, (disregarding the challenges of acquiring records from a hospital, despite HIPAA regulations).⁶ On the other hand, although storing data in a single location is advantageous for access, it also endangers large amounts of data to breaches.⁷

As a result of the deployment of new means of data collection, each with their own accompanying storage services, the lack of connectivity among traditional EHRs, has fragmented the American health record system. This fragmentation marginalizes patients, as they cannot use their existing records in conjunction and hinders research, as a majority of patient data resides in proprietary storage - inaccessible. Rather than fragmentation, imagine unification: a system of electronic health records which empowers patients through ownership of their data and enables large scale research while protecting patient privacy. What if patients could utilize a portal to access all of their EHRs across multiple hospitals, acting as a single medical history? By supplementing EHR data collected by traditional practitioners, patients could sync data from their FitBit and their 23andMe profile, and then decide to make all of their data available to researchers or their primary care provider. This portal, and this vision of American healthcare, can be achieved through blockchain technologies.

HEALTH I.T. BACKGROUND

EHRs optimize the logistical aspects of healthcare and improve patient care,⁸ but they also create new opportunities for vulnerability, such as those exploited by the WannaCry ransomware attack of May 2017,⁹ which made it impossible for afflicted hospitals to treat patients.¹⁰ A report released in June 2017 by the Health Care Industry Cybersecurity Task Force identifies many of these vulnerabilities, highlighting a lack of IT talent coupled with an abundance of outdated systems and unpatched vulnerabilities.¹¹ The report sets out a series of substantial and long term recommendations, asserting that patients should not have to choose between connectivity and security.¹²
Dissatisfaction with the fragmentation of the current EHR ecosystem is not new. The Office of the National Coordinator for Health Information Technology (ONC) envisioned nationwide interoperability in their 2015 Roadmap, highlighting the need for connectivity with EHRs. This initiative seeks to increase the value of existing EHRs by achieving patient control of data, eliminating information transfer blocking, and implementing national data standards. Through interoperability, the ONC hopes to enable large scale research to discover treatments tailored to individual patients, as part of the Precision Medicine Initiative.

The juxtaposition of these two reforms creates a challenge as increasing accessibility to data for interoperability can make data more vulnerable if security issues are not prioritized. Due to the critical nature of health data, this problem necessitates a solution that first and foremost addresses privacy and security concerns while approaching interoperability for large scale research. Furthermore, a quick to deploy solution to protect EHRs while the broader reforms detailed in the Task Force’s report are implemented would be ideal.

**Components of a Blockchain Architecture**

*This visual originally appeared in the Financial Times, “Technology: Banks Seek the Key to Blockchain,” (1 November 2015).*
• **Transactions between addresses:** A “blockchain” traditionally refers to the ongoing list of transactions arising from the exchange of a cryptocurrency like Bitcoin. In this traditional case, amounts of currency are associated with addresses, and these addresses are “stored” in wallets. Addresses do not contain identifying information, and so unless a user reveals his or her address, anonymity is preserved.

• **Encryption for trustless security:** Because of a lack of conventional means of trust, the blockchain must function as “trustless” system. In order to accomplish this, each address utilizes public key cryptography, such that the private key of the owner is required to send funds from the address, and all transactions can be verified with the corresponding public key to ensure that they come from the proper owner. This means that the user maintains control of their address and funds, and all other users can trust transactions without needing to verify the identities of the participants.

• **Creation of blocks:** The blockchain arises from the process of confirming transactions. Users pay a small fee to have their transactions processed by the network. Other users on the network, called “miners,” are incentivized to record transactions and bundle them together in “blocks,” validating the transfer of funds in return for the fees.

  In a traditional “proof of work” protocol for processing transactions, the process of creating a block involves computing a specific value from all of the non-confirmed transactions and the value of the existing block. This process requires trying many combinations of values to discover the next block, at which point the solution is broadcast to the network for verification. Once the new block is verified, all of the bundled transactions are confirmed, and the current block value is updated.

• **Chaining creates a tamper-proof record:** Because blocks are “chained” together by process of creating a new block, the blockchain is effectively immutable. Furthermore, past transactions can be verified by recalculating the current block’s value, and so the history of transactions is preserved by the current state. As blocks require a lot of computing resources to discover and a large network is invested in this process, asserting a change to the blockchain becomes increasingly difficult as time passes and more blocks have to be recalculated to catch up with the existing change. This unique design makes a blockchain ideal for creating a tamper-proof record of events which is protected by its distribution and ease of verifiability.
A look at blockchain technology

What is it?
The blockchain is a decentralized ledger of all transactions across a peer-to-peer network. Using this technology, participants can confirm transactions without the need for a central certifying authority. Potential applications include fund transfers, settling trades, voting, and many other uses.

How it works:
1. Someone requests a transaction.
2. The transaction is broadcast to a P2P network consisting of computers, known as nodes.
3. The new block is then added to the existing blockchain, in a way that is permanent and unalterable.
4. Validation: The network of nodes validates the transaction and the user’s status using known algorithms.
5. A verified transaction can involve cryptocurrency, contracts, records, or other information.
6. Once verified, the transaction is combined with other transactions to create a new block of data for the ledger.

Benefits
- Increased transparency
- Accurate tracking
- Permanent ledger
- Cost reduction

Unknowns
- Complex technology
- Regulatory implications
- Implementation challenges
- Competing platforms

Cryptocurrency
Cryptocurrency is a medium of exchange, created and stored electronically in the blockchain, using encryption techniques to control the creation of monetary units and to verify the transfer of funds. Bitcoin is the best known example.

- Has no intrinsic value in that it is not redeemable for another commodity, such as gold.
- Has no physical form and exists only in the network.
- Its supply is not determined by a central bank and the network is completely decentralized.

Potential applications
- Automotive: Consumers could use the blockchain to manage fractional ownership in autonomous cars.
- Financial services: Faster, cheaper settlements could shave billions of dollars from transaction costs while improving transparency.
- Voting: Using a blockchain code, constituents could cast votes via smartphone, tablet or computer, resulting in immediately verifiable results.
- Healthcare: Patients’ encrypted health information could be shared with multiple providers without the risk of privacy breaches.
BASIC BLOCKCHAIN FOR EHR

In the summer of 2016, the ONC organized the “Use of Blockchain in Health IT and Health-Related Research” challenge. The challenge solicited whitepapers “to address privacy, security, and scalability challenges of managing electronic health records and resources.” Of the 15 winning papers selected by the ONC, a subset proposed innovations regarding EHRs designed to empower patients through interoperability without sacrificing privacy.

For example, MIT Media Lab’s MedRec proposal succinctly demonstrates the great potential of blockchain technology for EHR storage. Utilizing an Ethereum blockchain, which enables developers to create programs that run the blockchain through a comprehensive programming language, their paper outlines a solution as follows;

• “Via smart contracts on an Ethereum blockchain, we log patient-provider relationships that associate a medical record with viewing permissions and data retrieval instructions (essentially data pointers) for execution on external databases.” This means that rather than financial transactions between users, a blockchain for EHRs could record medical interactions between patients and practitioners. Each interaction would include the anonymized addresses of the patient and the practitioner, enabling the patient to aggregate all of their records by their unique address. Due to the large size of some medical data, such as genome sequences, the specific medical record would be impossible to store directly on the blockchain, and so the digital location of the record would be stored. This location would be encrypted for patient privacy and data security, but would utilize smart contracts to be decryptable by the patient only, enabling data ownership and sharing as desired by the patient.

• “We include on the blockchain a cryptographic hash of the record to ensure against tampering, thus guaranteeing data integrity.” A hash value would be derived from the record and embedded in the transaction. This would enable the validity of the record to be recalculated and verified in the future, without needing to reveal or access its digital location.

• “Providers can add a new record associated with a particular patient, and patients can authorize sharing of records between providers. In both cases, the party receiving new information receives an automated notification and can verify the proposed record before accepting or rejecting the data. This keeps participants informed and engaged in the evolution of their records.” This platform would empower patients through access and control of their data, disrupting the current limited
options available to patients regarding their health data. At base, this could manifest as using the portal to view all of their medical history in one place, but patients could also use the portal to upload their own data, such as genomic-sequence data from a 23andMe test.

Through this portal, patients could decide to selectively share data with researchers, either for the greater scientific good or to enable studies on their unique condition. And judging from the 80% of 23andMe users who have chosen to make their genomic data available for research, enabling such simple sharing could lead to unprecedented levels of holistic data access for researchers, hopefully facilitating insights into precision medicine solutions. (This 80% comes from a population which self-selects to pay to have their genome sequenced, but the sentiment is valuable nonetheless). Furthermore, sharing access through the portal, and thus the blockchain, would be more secure than traditional sharing methods such as providing credentials to the proprietary server.

TECHNICAL CONSIDERATIONS

Implementing a blockchain solution is not the only way to create such a portal, and comes with a number of unique advantages, and some inherent challenges.

- **Ensuring the Creation of New Blocks**
  In order for a blockchain to work as an EHR ledger, (for new records to be appended, for the existing records to be immutable, for the records to be maintained by multiple parties), a number of parties must be involved in the mining process. Although an incentive-less blockchain could still be used to create a ledger, incentives encourage increased participation, bolstering the strength of the network.

  Regarding such incentives, the MedRec proposal suggests a novel solution in which medical researchers are compensated for mining blocks and maintaining the network in the form of a currency. They propose that such credit could be used to leverage the power of the network to conduct large scale analysis by querying the vast amounts of information. For such research to be useful, a significant portion of patients would have to consent to make their anonymized data available. Given the trend of data sharing among 23andMe patients, eagerness to share personal data does not seem to be an impediment, though it would be essential to ensure enough tech literacy among patients to be able to share their data if they so choose.
• **Initial Coin Offering**

Although continued mining would sustain the EHR blockchain, it would be advantageous to kick start the network to guarantee security from the onset. An Initial Coin Offering (ICO) would be one way to accomplish this. Through an ICO, the costs of creating a new blockchain protocol for EHRs could be fundraised by collecting another currency and “coining” amounts of the new token for use on the EHR blockchain.

Coining a new cryptocurrency is straightforward with Ethereum smart contracts and would enable supporters to buy in early on. As a result, ICOs have the benefit of solving the chicken-egg problem of starting a new protocol: in raising funds for the start of the blockchain, the new token is valued; creating incentives for users to mine new blocks and further establishing the network. Furthermore, the individuals who help raise the initial funds are invested in the success of the protocol created by the blockchain and are incentivized to ensure its adoption, as they benefit from the network-effect.  

Although an ICO would be useful for establishing a blockchain for EHRs, it could create ethical challenges. As the network becomes established, individuals could be incentivized to speculate on the financial value of the query token, rather than using the token only for research. Although private gains do already occur from publicly collected data and other goods, speculation on the public infrastructure itself is less ethical, as speculation ensued fluctuations could hinder the effectiveness of the protocol. If the cost of the research token became too high, research could become too expensive to conduct, stifling the potential of the network and initial incentives for mining. On the other hand, if the token crashed, there would be few incentives to mine blocks at such a low value, limiting the speed at which health records could be committed - the essential utility of the system.

In order to avoid speculation, the ICO could be restricted to researchers and miners could be vetted by the existing body of researchers to ensure intent. These measures would limit token accessibility and hopefully prohibit speculation, but it would be essential to maintain oversight to ensure that these constraints are not abused by the researchers to horde the data, as this would be unethical and stifle innovation.
Proof of Work or Proof of Stake

The Proof of Work (PoW) protocol described previously for creating blocks is advantageous for avoiding “51% attacks” in which a single party attempts to control a majority of the computing powers and thus the creation of blocks. Unfortunately, as a network grows, increasingly large amounts of computing power are required to mine competitively, and so, the amount of electricity used to maintain the network increases dramatically as well. The amount of electricity used to mine and maintain the network could be reduced by limiting the parties allowed to mine, but this could enable a monopoly to form. This would be unfortunate as they would control access to the data and thus limit the research potential of the network, rendering this solution undesirable.

An alternative approach involves implementing a “Proof of Stake” (PoS) protocol for incentivizing record processing. In this case, there are no rewards for the discovery of a new block and revenue is only derived from the fees associated with transactions. Rewards are distributed by an algorithm, and the reward a user receives depends on the portion of currency that they hold.

In the case of an established cryptocurrency, like Bitcoin, it would be very challenging to take over the network with a PoS protocol, as funds are distributed among a multitude of users and there are only a limited amount of funds available for purchase. Furthermore, PoS protocols can be designed such that users must contribute a significant amount of the currency as a wager of good intent. Contributions can be utilized as insurance against malpractice, as this amount can be revoked from the user if malpractice is detected, effectively deincentivizing misuse and ensuring decentralization.  

Designing an EHR blockchain to implement Proof of Stake from the start would avoid future energy crises as there would be no competition over discovering blocks. Furthermore, given a limited amount of parties participating in a new EHR blockchain, a PoS protocol would be safer than a PoW system - which would be susceptible to 51% attacks, as safeguards against monopoly behaviour can be built into the PoS algorithm. The main impediment would be ensuring the security of the PoS algorithm. As Ethereum is moving to transition to PoS, it is likely that their algorithm could be built off of or adapted to an EHR context in the near future.

Identity Verification

Although blockchain is well suited to preserving the declared identities of patients and practitioners due to its nature as an immutable ledger, the claimed identities must be verifiable for provenance to be of merit. To address this challenge, MedRec proposes a linking system from existing forms of ID (such as bank accounts or social security numbers) to addresses on the blockchain. This system would be
similar to that of a Domain Name Server, which associates website domain names with Internet Protocol (IP) addresses, but it would only reveal the identities of the patient and practitioner if granted through smart contacts on the blockchain.

A slightly more advanced and secure solution involves linking EHR blockchain records to a blockchain system that pairs biometrics with identity management, as recently pursued by Accenture and Microsoft.\(^\text{17}\) The development of these sorts of identity management systems is pertinent first and foremost for guaranteeing access to social services worldwide, especially in the context of the refugee crisis, but also for the success of an EHR blockchain. Furthermore, a more secure system for identity verification would eradicate healthcare fraud.\(^\text{18}\) While this system is in development, MedRec’s proposal to use existing forms of identification could work as an interim solution.

- **Smart Contracts**
  Although patients should be able to “opt-in” to select only the parts of their medical information which they desire to share, it will still be important to have some oversight of the kind of queries that researchers can submit to the database of ledgers. MIT’s Project PharmOrchard proposal suggests using “pre-fabricated” queries which have been analyzed and approved by experts in order to ensure privacy protection. Aspiring researchers would submit their queries for vetting, which would hopefully empower other researchers to “make use of them in their own context of study.”\(^\text{19}\)

  It is possible that revealing queries to researchers prior to their use could enable ethical violations, but perhaps a more open, crowdsourced approach could mitigate this, as demonstrated recently by Synlett with the peer review process.\(^\text{20}\) Releasing a pre-approved plug-and-play style formula corresponding to a large amount of pre-approved queries with the launch of the blockchain would also help to expedite research and mitigate such ethical issues with queries from the first-adopters.

- **Artificial Intelligence**
  In addition to being useful for informing traditional clinical trial research, the vast amounts of holistic data organized on the blockchain could be instrumental in training machine-learning models for healthcare, such as IBM Watson.\(^\text{21}\) A current impediment to the success of such systems is the lack of access to data with which to train the models, and data provided through the blockchain from consenting
patients would be cheaper and more holistic than what is currently used for training. As a result, such use of data could bring about the breakthroughs which would enable AI aided diagnosis, and potentially, even AI monitoring of patients outside the clinic - both of which would improve health outcomes and lower costs.

**SOLUTION EVALUATION**

- **Security**
  The individualized public key encryption scheme diminishes the vulnerability of EHRs when accessed through the blockchain. Although all of the records are available through the blockchain, every patient’s record requires a unique authentication. While the human factor is the largest threat to cybersecurity, a single user’s compromised credentials would only expose their data, rather than the larger scale vulnerabilities of existing EHR management solutions.

  Of course, this benefit is limited to attacks targeting the blockchain. A direct attack to the data server accessed by the blockchain, such as the WannaCry attack, could expose all of its records. As a result, once a blockchain solution is implemented for EHRs, it would be pragmatic to largely revoke credentials to directly access the underlying database, as such access would be accomplished more securely through the blockchain.

  An additional security benefit derives from the distributed nature of the blockchain. In the case of a Denial of Service (DoS) attack, in which access to a localized EHR system could be prohibited by overriding the data server, the records on a blockchain could still be accessed as long as part of the network remains operational. Of course, this necessitates the preservation of the network, which is also required for new records to be added, and can be ensured through the research incentives for mining as described in the MedRec proposal.

- **Interoperability**
  A functional blockchain EHR system would enable patients to organize and control all of their health data in one location, combining records stored across different locations and mediums in one “longitudinal picture of health.” Such an architecture would eliminate delays and impediments to health record access and use, as only the patient’s authentication would be required for retrieval, rather than navigating the complexity of the existing bureaucracy.
A blockchain architecture would both enable unprecedented amounts of research access to health records, and increase the utility of such access. The establishment of a permissions system for patients to consent to contribute their personal health data for research and the streamlining of the logistics of data access makes using individual-level records for research feasible, as access becomes legal and simple. The unification of health records builds upon this to make soliciting records from patients not only possible, but pragmatic, due to the holistic nature of the data.

Similar to the challenge of security, in which data is secure on the blockchain but vulnerable in their existing storage, the level of interoperability facilitated by the blockchain is contingent on the compatibility with existing data formats. Once available through the blockchain, data formats can be standardized to the specification of the Roadmap, but the blockchain will only be able to incorporate existing EHRs if they are made available in standardized formats across the nation.

**RECOMMENDATIONS**

As a result, there are a few recommendations for governance to set the foundation for innovation:

- **Work with and support the ONC Tech Lab and HHS IDEA Lab to pursue the development of prototypes of a blockchain EHR solution.** The ONC’s 2016 competition and conference demonstrated the viability of a blockchain solution and identified numerous organizations which could be integral to developing a solution. Although regulatory changes would have to occur to enable a blockchain EHR system as described to function, the development of a prototype would demonstrate the viability of the idea. The ONC Tech Lab and HHS IDEA Lab provide agile channels through which to foster the development of prototypes that build off of the solutions described in this report.

- **Pursue nationwide legislation to enable citizens to own their health data.** Patients can currently access their health data under HIPAA regulation, but this current system often involves bureaucratic delays and exorbitant prices related to the outmoded costs of paper records. A blockchain EHR system could simplify this process and empower patients by enabling them to effectively own their data through controlling access, but in order for this to occur, patients must be able to legally “own” their data. Establishing federal legislation to give patients ownership of health records would set a precedent of citizen empowerment through data rather than marginalization, and challenge the current paradigm sacrifices privacy for convenience, ushering in a new kind of patient-practitioner relationship.
• **Encourage the NSF to fund research to establish a system of identity verification and provenance.** A system of identity verification is a necessary component for the success of a blockchain EHR system and should be pursued in general as a “basic human right.”\(^{28}\) Blockchain implementation, as described by Microsoft and Accenture among others,\(^{29}\) would be a feasible way to accomplish this. Blockchain\(^{30}\) is not the only solution to this problem,\(^{31}\) so it would be pragmatic to pursue this problem objectively.

• **Work with the Interoperability Standards Advisory (ISA) of the ONC and NIST to continue developing regulatory data standards for existing EHR systems.** Current EHR system providers have few incentives to adopt similar formats, as this “locks in” customers to their product.\(^{32}\) By continuing past work\(^{33}\) to establish regulatory standards for the storage of existing EHRs, data storage security can be ensured by following NIST encryption protocols, and interoperability through the blockchain can be enabled through standardizing how healthcare data is formatted. Although this will necessitate the establishment of a shared language and structure of healthcare data, adoption will also make data more useful to large scale research by generating comparable data through shared recording protocols.

• **Establish and require educational structures and the development of infrastructure for patients and practitioners to ensure that “It just works.”** Patients must know about the patient portal, and must be able to access and utilize it, in order for their health records to be used. This would necessitate widespread announcements about the roll-out of the portal and the establishment of an educational structure. Education at the point of the practitioner could ensure widespread education but would likely have to be stipulated in order to succeed, given the demand for doctors. Furthermore, the development and adoption of technical devices to predict edge cases of use would be essential to ensuring the success of the system. For example, a device through which patients can grant authentication to their records in an emergency will be important to enable service if the patient is without a device.

• **Pursue the recommendations of the Health Care Industry Cybersecurity Task Force Report.** While a blockchain EHR system will help protect data from breaches, the recommendations of the Task Force should be still be followed. Securing data storage and access will protect a vast amount of existing data, but vulnerabilities in medical devices will still endanger new data as it is collected. Furthermore, although a blockchain system for EHR management will abstract away vulnerabilities due to direct access to all the EHRs, technical training will still be necessary to avoid vulnerabilities through human error.
ENDNOTES


Eleonore Pauwels is the Director of the AI Lab with the Science and Technology Innovation Program at the Wilson Center. She is a writer and international science policy expert, who specializes in the governance and democratization of converging technologies. Leading the AI Lab, Pauwels analyzes and compares how transformative technologies, such as artificial intelligence and genome–editing, raise new opportunities and challenges for health, security, economics and governance in different geo–political contexts.

Nathaniel Grevatt completed this work as a research assistant in the Science and Technology Innovation Program at the Wilson Center. He is a Masters student in the Science, technology and Society Program within the School of Engineering and Applied Sciences at the University of Virginia.

The Wilson Center
@TheWilsonCenter
facebook.com/WoodrowWilsonCenter
www.wilsoncenter.org

Woodrow Wilson International Center for Scholars
One Woodrow Wilson Plaza
1300 Pennsylvania Avenue NW
Washington, DC 20004-3027

The Wilson Center
@TheWilsonCenter
facebook.com/WoodrowWilsonCenter
www.wilsoncenter.org

Woodrow Wilson International Center for Scholars
One Woodrow Wilson Plaza
1300 Pennsylvania Avenue NW
Washington, DC 20004-3027