RDA Health Data IG

Edwin Morley-Fletcher (co-chair), Lynkeus

RDA meets Nordic researchers, 14 June 2017, Göteborg, Sweden
Brief History

• BoFs @RDA P6 in Paris and @RDA P7 in Tokyo
• Health Data Interest Group established @RDA P8 in Denver (September 2016) # over 100 members
• Second meeting at @RDA P9 with focus on **Data Protection** (GDPR) and **Health Data Mapping**
  https://docs.google.com/spreadsheets/d/1WUGU1McZH3jo5x7g2FCMsvPdF5DKUHKL6AqJWBUmB0/edit#gid=1612672872
• BOF on **Blockchain** for Health @RDA P9 in Barcelona (April 2017), to be held again @RDA P10, in view of establishing a dedicated Working Group.
• Quarterly confcalls
User scenarios/focus areas

• Data access and protection
  • sharing best practice on pseudonymisation and anonymisation
  • developing models for consent that protect patients while enabling research
  • providing a forum for discussing, explaining and responding to data protection regulation
  • secure opening up of data to facilitate research

• Data-based healthcare for personalised medicine
  • disease signatures identification
  • stratification of patient groups
  • patient-specific simulation and prediction

• Data literacy in Health care
  • providing materials for education of healthcare professionals on use and misuse of data

• Patient data repositories/patient-centric data gathering systems

• In-silico drug development and clinical trials
  • representing interests of the data-based healthcare community to policy makers
  • identifying and discussing related challenges, interdisciplinary research needs and potential roadmaps.

• Blockchain applications to health data
Target participation areas

- Clinicians (to use data technology to improve practice)
- Biomedical researchers (using analytical techniques)
- Healthcare Data Analytics (with data mining, machine learning, physiological modelling and image processing expertise)
- HPC and distributed computing experts
- Policy-makers for Healthcare
- Health bioinformatics legal experts
- Healthcare administrators - Health Maintenance Orgs
- Pharmaceutical industry researchers and manufacturers
- Medical equipment researchers and manufacturers
- In silico modelling, testing and clinical trial experts
- Blockchain Applications in Health
Looking forward to an RDA WG focusing on Blockchain in Health?

• The HD-IG is sponsoring the idea of establishing a WG focusing on Blockchain in health data with the aim of debating in depth the potential of such a system and whether the blockchain can ensure compliance with advanced data protection requirements (such as those defined, for instance, by the EU General Data Protection Regulation – GDPR), yet making it happen seamlessly and efficiently, at scale.

• Due to its scope, the next (second) preliminary BoF meeting ideally fosters relations to a number of RDA WGs and IGs that may be able to contribute with their results to, or benefit from, the proposed future WG’s activities.
Blockchain Hype?

• The Economist went so far as to state that, at first sight, “the notion of shared public ledgers may not sound revolutionary or sexy. Neither did double-entry book-keeping or joint-stock companies. Yet, like them, the blockchain is an apparently mundane process that has the potential to transform how people and businesses cooperate. [...] A realisation that systems without centralised record-keeping can be just as trustworthy as those that have them may bring radical change. [...] A world with record-keeping mathematically immune to manipulation would have many benefits.”
Blockchain Hype ? IBM

• “Blockchain promises to put privacy and control of data back in the hands of citizens. Trust and integrity will be established without reliance on third-party intermediaries. IBM believes blockchain is an extraordinarily important phenomenon with the potential to transform industries and upend business models”.

• “In healthcare, new research is seeking to apply blockchain’s distributed ledger and decentralized database solutions to the critical issues of interoperability, security, record universality, and more. Intriguing uses in other industries are being extended to healthcare, such as extending blockchain’s smart contracts to provider network management or connecting myriad medical devices through common, blockchain-enabled systems of information relationships. While technical consensus on a distributed ledger for healthcare has yet to emerge, with debate ongoing regarding scalability, security, and regulatory compliance, blockchain technology and encryption will drive innovation in healthcare services and administration”

IBM Global Business Services Public Sector Team (2016), Use of Blockchain in Health IT and Health-related Research, proposal submitted on August 8, 2016, to the Ideation Challenge launched by the Office of the National Coordinator for Health Information Technology in the USA.
Blockchain Hype ? Deloitte

• Healthcare pain points and potential blockchain solutions were similarly indicated by IBM as well as by Deloitte, in whose White Paper, however, they appeared to be more conveniently summarised as shown in the next table, taken from:

Deloitte (2016), Blockchain: Opportunities for Health Care, White Paper developed in response to the Department of Health and Human Services’ Office of the National Coordinator for Health Information Technology (ONC) ideation challenge on “The Use of Blockchain in Health IT and Health-Related Research”.

# Blockchain Value Propositions for Healthcare

<table>
<thead>
<tr>
<th>Health Information Exchange (HIE) Pain Points</th>
<th>Blockchain Opportunities</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Establishing a Trust Network</strong> depends on the HIE as an intermediary to establish point-to-point sharing and “book-keeping” of what data was exchanged.</td>
<td><strong>Disintermediation of Trust</strong> likely would not require an HIE operator because all participants would have access to the distributed ledger to maintain a secure exchange without complex brokered trust.</td>
</tr>
<tr>
<td><strong>Cost Per Transaction</strong>, given low transaction volumes, reduces the business case for central systems or new edge networks for participating groups.</td>
<td><strong>Reduced Transaction Costs</strong> due to disintermediation, as well as near-real time processing, would make the system more efficient.</td>
</tr>
<tr>
<td><strong>Master Patient Index (MPI)</strong> challenges arise from the need to synchronize multiple patient identifiers between systems while securing patient privacy.</td>
<td><strong>Distributed framework for patient digital identities</strong>, which uses private and public identifiers secured through cryptography, creates a singular, more secure method of protecting patient identity.</td>
</tr>
<tr>
<td><strong>Varying Data Standards</strong> reduce interoperability because records are not compatible between systems.</td>
<td><strong>Shared data</strong> enables near real-time updates across the network to all parties.</td>
</tr>
<tr>
<td><strong>Limited Access to Population Health Data</strong>, as HIE is one of the few sources of Integrated records.</td>
<td><strong>Distributed, secure access</strong> to patient longitudinal health data across the distributed ledger.</td>
</tr>
<tr>
<td><strong>Inconsistent Rules and Permissions</strong> inhibit the right health organization from accessing the right patient data at the right time.</td>
<td><strong>Smart Contracts</strong> create a consistent, rule-based method for accessing patient data that can be permissioned to selected health organizations.</td>
</tr>
</tbody>
</table>
Who is Ronald Coase?

• Ronald Coase (1910-2013): Nobel Laureate for Economics on 1991
  More than 50 years after his disruptive *The Nature of the Firm* (1937), and

• It is now possible to reverse Ronald Coase’s Transaction Costs.

• What Internet did to transaction costs regarding information, blockchain can
do regarding trust.
Assumptions and Expectations

• Public and private initiatives, both in Europe and in the US, are currently addressing the potential of applying the blockchain approach to health data.

• This is related to great general expectations ("what Internet did to transaction costs regarding information, blockchain can do regarding trust") and to the assumption that what is needed for health data is a Distributed Empowerment system, providing secure access from anywhere on any device.

• There is the need to develop new mechanisms of trust and of direct, value-based relationships between people, hospitals, research centres, and businesses, leading to an open biomedical information network centred on the connection between organisations and the individual.
Blockchain Ledger

• A Distributed Empowerment system having the Blockchain ledger as secure, non-editable record, where all transactions are confirmed by the network as entries forming blocks of transactions, and the whole network monitors the legitimacy of each transaction, guaranteeing distributed control.
Smart Contracts

• The blockchain is expected to be based on portfolios of Smart Contracts.
• Smart Contracts are the executable pieces of code, stored on the blockchain for future execution.
• These bind people and transactions to specific actions and outcomes and require no further direct human involvement after the smart contract has been made a part of the distributed ledger (which is what makes these contracts "smart" or self-enacting).
HD-IG GOALS
with regard to Privacy, Security, and Blockchain

• **Profile and classify** sensitive data based on their informational and economic value
• Assess the most suitable and robust **de-identification** and **encryption** technologies needed to secure different types of information
• Allow having **advanced analytics** running on anonymised or pseudonymised data
• Evaluate the overall security of MHMD multi-modular architecture by testing it through dedicated **self-hacking** simulations and **public hacking challenges**
• Analyse **users’ behavioural patterns** alongside **ethical and cultural orientations**, to identify hidden dynamics in the interactions between humans and complex information services
• Improve the design of **data-driven platforms**
• Foster the development of an **information marketplace**, in which both individuals and clinical institutions will be able to exert control on their health data and leverage their value
Strategically relying on the ongoing MyHealthMyData EU-funded project and on its Four Leading Hospitals

• Following the example of routine data inflow by the OPBG PCDR, and the interoperability system established in Cardioproof and MD-Paedigree

• Taking into account the less restricted data processing allowed by the GDPR when it is aimed at scientific research, and the proviso that the data protection legislation does not apply to duly anonymised data

• Guaranteeing that all health and personal data will:
  • Be duly anonymised before been uploaded on MHMD Infostructure
  • Be processed, should the use of partial anonymisation techniques be indicated for the intended use of data, on the ground of a Dynamic Consent provided by the data subjects.

• Exploring different open data implementation approaches

• Evaluating, to the extent permitted by national and European regulations, solutions providing some concrete acknowledgment of data value
Two layers of data flow

• A semi-automated data profiling and cleaning engine that:
  o Ensures and assesses data quality
  o Guarantees the most appropriate de-identification or encryption mechanism, according to each type of data or modality

• A privacy preserving and security layer that combines:
  o A privacy preserving data publishing engine (providing anonymisation tools)
  o A privacy preserving complex data flow execution engine (i.e., differential privacy, SMPC, homomorphic encryption)

• The joint goal is to allow:
  o Classifying medical data and correspondent security and privacy provisions in each category
  o Assessing relevance, sensitivity, risk for the individual and practical value
  o Selecting the most appropriate security and privacy preserving technique in each case
Institutional Data providers requirements

All MHMD hospitals commit to:

• Acquire a server to host the local data catalogue and contribute to blockchain transactions
• Have their research data indexed using DOIs
• Integrate the server in their Healthcare Information System (HIS) (thus allowing later production)
• Consent to be consulted on future cohort requests
• Ensure Veracity, Validity & Integrity of their data
• Only make pseudonymous data available
• Handle the responsibility of patient identity mapping with local anonymous data
Individual MHMD User Entitlements (1)

• **Aggregate personal data from disparate sources:**
  • Social media accounts, clinical data repositories, personal drives, wearable devices, etc., in a single, user-owned account (PDA).

• **Assign data access rights**
  • Within an efficient workflow, based on stakeholders’ permissions and addressing simple questions:
    o Type of data requested
    o Intended use
    o Data that will be retained
    o Data that will be shared with 3rd parties and intended use
    o Implementation of the Right to be forgotten.
Individual MHMD User Entitlements (2)

• Stay informed of, and enquiry on, relevant data transactions after access has been granted
• Be able to revoke data access rights, or extend them
• Be able to receive requests from stakeholders for data access permissions.
  • Requests may also include incentives offered by stakeholders in exchange for data
• Define post-mortem usage or donation of personal data
Blockchain: no recourse to Trusted Third Party

- Applying the blockchain approach to health data guarantees secure access from anywhere on any device
- The Blockchain ledger is the secure, non-editable record where:
  - All transactions are confirmed by the network as entries forming blocks of transactions
  - The whole network monitors the legitimacy of each transaction, guaranteeing a distributed control system
- Each stakeholder can enact anonymous transactions through the ledger:
  - Employing public key encryption for identifying owners in the ledger, recording one half of the public key pair
  - Only the person or institution holding the corresponding private key can decide what happens next to their data
- Each stakeholder is equipped with a ‘wallet’ containing:
  - An encrypted identifier
  - His/her Dynamic Consent
  - His/her data access policy file
DOI System: Handle.net

• The Handle System is a comprehensive system chosen for assigning, managing, and resolving persistent identifiers for digital objects and other resources on the Internet.

• The protocols enable a distributed computer system to store identifiers of digital resources and resolve those identifiers into the information necessary to locate and access the resources.

• The handle.net 8.1 software includes:
  o a RESTful JSON-based HTTP API
  o a browser-based admin client
  o an extension framework allowing Java Servlet apps
  o authentication using handle identities without specific indexes
  o multi-primary replication
  o Security improvements
Applied Analytics: Probabilistic Modeling for statistical simulation
Applied Analytics: DeepReasoner by Siemens
Multi-modal patient modelling in MD-Paedigree

Create patient representation combining information from multiple sources

Microbiome
Clinical information
Fat information
Cardiac function
Circulation
Strains

- Microbiome:
  - Small intestine
  - Stomach
  - Cecum
  - Appendix
  - Colon
  - Rectum
  - Anus

- Clinical information

- Fat information

- Cardiac function

- Circulation
  - Strains