RESEARCH DATA ALLIANCE

Blockchain Applications in Health WG @VP16

Comprehensive Guidelines on Blockchain Applications in Health

Costa Rica (Virtual) 9th November 2020, 5:00 - 6:30 PM UTC

Co-chairs: Edwin Morley-Fletcher, Ludovica Durst

research data sharing without barriers
rd-alliance.org

Blockchain Applications in Health WG

AGENDA

1. Introduction on blockchain basics and blockchain in health for this virtual session of the WG at the time of COVID, Edwin Morley-Fletcher, co-chair
2. Q&A
3. Blockchain legal and regulatory issues within the General Data Protection Regulation”, Ludovica Durst (Lynkeus)
4. Q&A
5. Blockchain and COVID initiatives, Mirko de Maldé (Health WG co-chair of the International Association for Trusted Blockchain Applications)
6. Q&A
7. Next steps: looking forward to the conclusion of the WG activities and producing a set of comprehensive guidelines

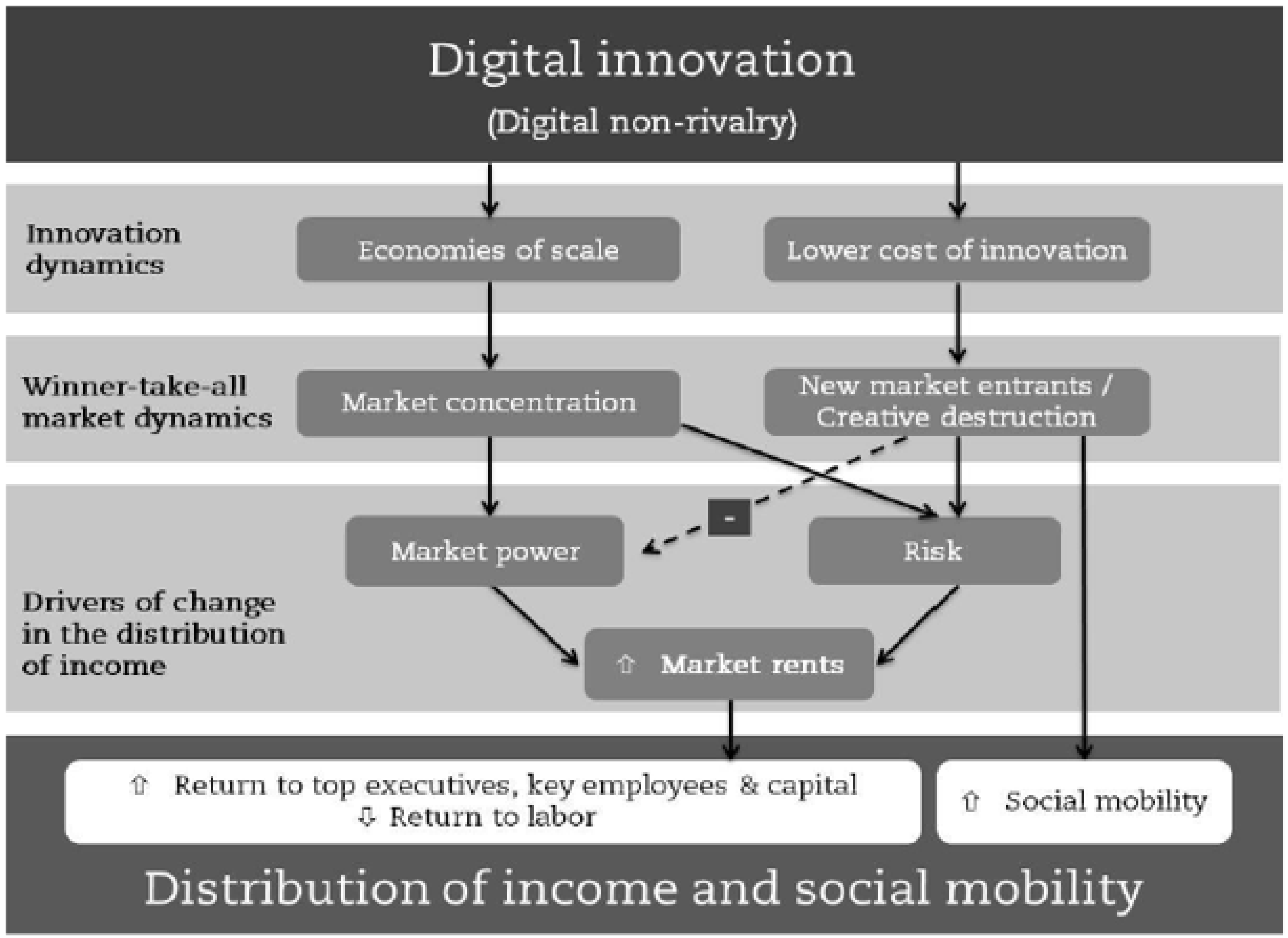
A delay in concluding the activities of the WG

- Couple of on-line meetings before COVID broke through
- Cancellation of the meeting scheduled in Melbourne in Spring
- Ongoing increase in the number of members of the WG
- Difficult situation again because of COVID in these last months
- A useful initiative by RDA France (Isabelle Perseil): Atelier Blockchain pour la Santé – 21 October 2020A

Blockchain's basics

- Blockchain is a disruptive innovation regarding uncertainties that traditionally have implied the need of relying on some amount of trust for coping with them.
- It is a technology providing transparent and secure storage and transfer of data without having recourse to a central authority. With blockchain all data transfers become traceable and auditable by participants to the ledger.
- Blockchain allows to envisage a distributed rather than a hierarchical foundation of trust. It allows "trust-less certainty".
- If internet has dramatically reduced transaction costs on information, blockchain can do the same regarding the exchange of data incorporating value.
- Blockchain allows to digitize value transfers and use self-enforcing "smart contracts" for automating the enactment of contractual rules (*code is law*) .
- Blockchain allows to have recourse to issuing digital tokens for crowdsourcing (*tokenomics*).

A substantial shift in the intangible economy


Digital innovation
(Digital non-rivalry)
Innovation dynamics
Economies of scale
Lower cost of innovation
Winner-take-all market dynamics
Market concentration
New market entrants / Creative destruction
Market power
-
Risk
Drivers of change in the distribution of income
⇧ Market rents
⇧ Return to top executives, key employees & capital
⇩ Return to labor
⇧ Social mobility
Distribution of income and social mobility

Source: D. Guellec and C. Paunov, Digital Innovation and the Distribution of Income, NBER Working Paper 23987, 2017

Blockchains change the balance between hierarchies and markets

- What Internet did to transaction costs regarding information, blockchain can do regarding trust.
- Reliable data-rich information systems become possible without being paralysed by excessive transaction costs.
- The limitation of decentralised solutions was in the past the all-pervasive (and excessively reductive) role played in market-based solutions by just one synthetic information: the price.
- The advantage of hierarchical solutions was based on the capacity to tackle uncertainty by centrally ordering a multiplicity of information about which actions to follow.
- The blockchain methodology, by allowing solutions based on decentralised protocols,
  - removes the friction and costs of current intermediaries
  - makes it possible to develop distributed and transparent systems
    - where empowerment can be shared
    - asymmetries can be balanced
    - qualitative aspects can be taken into account.

What is the potential of blockchains as social technologies?

- Blockchains are technical solutions that facilitate the smooth functioning of an ecosystem by:
  - managing and implementing decision-making through automated consensus
  - creating incentives that nudge participants into behaving constructively
  - generating trust and transparency
- By enacting decentralised information systems with inherent data integrity, blockchains constitute strong anti-corruption tools and feature as relational software
  - enabling new trust mechanisms capable of transforming social relations
  - reducing transaction costs
- There are close similarities between blockchains and bureaucracies, though bureaucracies are centralised and blockchains are distributed:
  - both are defined by the rules and execute predetermined rules
  - both work as information processing machines
  - both work as trust machines

Bureaucracies are thus natural candidates to have centralisation being replaced by federated blockchain systems.

- Sustainable growth implies doing more with existing resources and attracting more resources to expand the scale of operations: blockchains reduce costs and increase the flow of funds, helping social innovation organisations to scale up, by enabling marketplaces and the issuance of alternative currencies and tokens.

How can blockchains translate into innovative social policies?

- Social protection systems are a fundamental element of pride for EU countries, though differing between bismarckian and beveridgean models (including also Mediterranean varieties) and experiments combining both approaches.
- Traditionally the welfare states are characterized as mainly exerting a “piggy bank” or a “Robin Hood” role:
  - the first helping people to insure against social risks and redistribute resources over the life cycle
  - the second entailing measures to reduce social exclusion by redistributing income and wealth
- This varying mix is translated into a combination of four key functions: regulatory, redistributive, insurance, production
- Historically, these functions have been largely centralised, but could be reorganised in different manners, for instance as Personal Welfare Accounts, operating in blockchain-regulated social markets, where beneficiaries could have incentives to:
  - check that the universal (but not unlimited) coverage is responsibly guaranteed by public agencies and private organisations
  - play a role in changing the market through informed decision and collective actions

Can blockchains translate into opportunities for healthcare?

Blockchains reside at the nexus of several disciplines which are key for providing healthcare solutions: cryptography, game theory, tokenomics, network theory.

There is a huge potential for:

- employing cryptographic and algorithmic methods to record and synchronise health data transactions across distributed networks in an immutable manner.
- using smart contracts as coded instructions which execute on the occurrence of an event and extend the functionality of blockchains from storing transactions to performing computations.
- developing multi-sided platforms where data providers (being both clinical institutions and individuals), researchers and industries can all rely on data integrity and security and mutually reinforce network effects.
- allowing to manage data flows and usage, based on individual free choice and self-determination, making dynamic data portability in real time possible for individuals and companies, along with various compensation models.
- applying Health Big Data to Artificial Intelligence and Machine Learning for medical knowledge discovery.

What are the basic architectural features?

- A private permissioned blockchain recording all transactions related to Off-chain data stored by multiple hospital repositories and by individuals
- A Metadata Catalogue allowing to safely inspect what health-data are available
- The possibility of making use of Smart Contracts for automatically checking the needed consent and enacting the permissioned transactions.
- Privacy-enhancing technologies for assuring GDPR compliance and advanced ways of handling data.
- An overall Privacy-by-Design and Compliance Assessment

Lessons learnt on data sharing

Despite much talking on the need to foster data pooling and data sharing, there are significant hurdles and constraints in doing so in healthcare:

- There is a discrepancy in the way anonymisation is referred to in the GDPR and by the European Data Protection Board.
- The subsequent regulatory uncertainty makes it difficult to obtain anonymised data from clinical institutions.
- Pseudonymised data require on principle a specific legal ground, such as an explicit personal consent, or a personal notification, for being shared with third parties.
- Big Data and AI remain difficult to apply at scale in medicine, given that effective data sharing is still the exception in healthcare.

Two key modalities

1. “Visiting mode”: bringing the algorithms to the data

Secure computations, which permit running AI without disclosing neither data nor algorithms, are performed through three tools:

1. Homomorphic Encryption
2. Secure Multiparty Computation
3. Federated Deep Learning with an untrusted Black Box

A further lesson learnt

The “visiting mode”, and more generally Privacy Preserving Machine Learning, is an emerging field in data science.

As yet, the foundation on either HE or SMPC still implies a large communication and computation overhead

Which makes it hard to use where very large amounts of data are required

- since communication and computation costs are greatly affected by the increase of the number of involved parties or of the model’s complexity.

2. Synthetic Data

- Synthetic data are fully artificial data, which achieve anonymity by breaking the link between private information and data’s information content.
- They are automatically generated by making use of Generative Adversarial Networks (GANs), based on two models playing recursively against each other.
- High-quality synthetic data closely resemble the real data and are a suitable substitute for processing and analysis.

Generating differentially-private synthetic data

- Differential privacy provides an until-now lacking mathematical foundation to privacy definition:
- “Differentially Private Synthetic Data Generation is a mathematical theory, and set of computational techniques, that provide a method of de-identifying data sets—under the restriction of a quantifiable level of privacy loss. It is a rapidly growing field in computer science”

[National Institute of Standards and Technology Differential Privacy Synthetic Data Challenge 2019: Propose an algorithm to develop differentially private synthetic datasets to enable the protection of personally identifiable information while maintaining a dataset's utility for analysis]

AI in Healthcare Whitepaper

Big Data Value association – BDVA Task Force 7 - Sub-group Healthcare - November 2020:

- “Many researchers find unclear the approaches that are required to collect anonymized data to ensure final users’ privacy”.
- “Generative Adversarial Networks (GANs) can generate meaningful synthetic data which do not suffer from the confidentiality constraints of the source data”.
- EU-funded projects are encouraged to make available synthetic data sets that sufficiently resemble source data while avoiding privacy issues. Generative Adversarial Networks (GANs) have demonstrated potential, but their general applicability has not been established yet.

Conclusion

- The time is ripe for initiatives that fill the void of legal and technical definitions and ease the difficulties of coping with health data sharing while fully implementing the GDPR.
- Initiatives prompting private and institutional centres to work on the potential of synthetic data and secure computation systems with the aim of suggesting a roadmap for their further implementation and market adoption, could be a stepping-stone to activate a thriving digital ecosystem for the biomedical sciences.