Blockchain legal and regulatory issues and GDPR

Ludovica Durst

Blockchain Applications in Health WG @P16

Virtual Plenary, 9th November 2020, 5:00-6:30 pm UTC

Co-chairs: Edwin Morley-Fletcher, Ludovica Durst

research data sharing without barriers

rd-alliance.org

1. Legal framework and GDPR

- **GOLDEN RULE : ACCOUNTABILITY**

The main innovation introduced by the EU *General Data Protection Regulation* 2016/679 is the strengthening of the concept of accountability.

- The **data controller** must be able to **demonstrate compliance with the obligations imposed by the GDPR**,

with particular reference to the **principles** of

- (a) **lawfulness**, **fairness** and **transparency**;
- (b) **purpose limitation**;
- (c) **data minimization**;
- (d) **accuracy**;
- (e) **storage limitation**;
- (f) **integrity and confidentiality**.

Accountability… how?

- **DATA PROTECTION IMPACT ASSESSMENT (DPIA)**
- When a company **innovates** its processes, brings **new services to the market** or implements **new methods** of production or service, the **protection of personal data** must be taken into the utmost account and evaluated, **not as a bureaucratic formality**, but as a **structural phase of the innovation process** which contributes, together with more traditional ones, to the overall assessment of the **innovation's sustainability.**
- Designing, making use or marketing **new technologies** (e.g. homomorphic encryption, secure multiparty computation, **blockchain**, smart contracts, etc.) **obliges** their developer, manufacturer or seller**, as data controllers**, to **carry out a data protection impact assessment** aimed to weigh up and identify any **threats** which may arise from the underlying data processing activities, in order to **prevent** or at least **minimize** such **risks**.

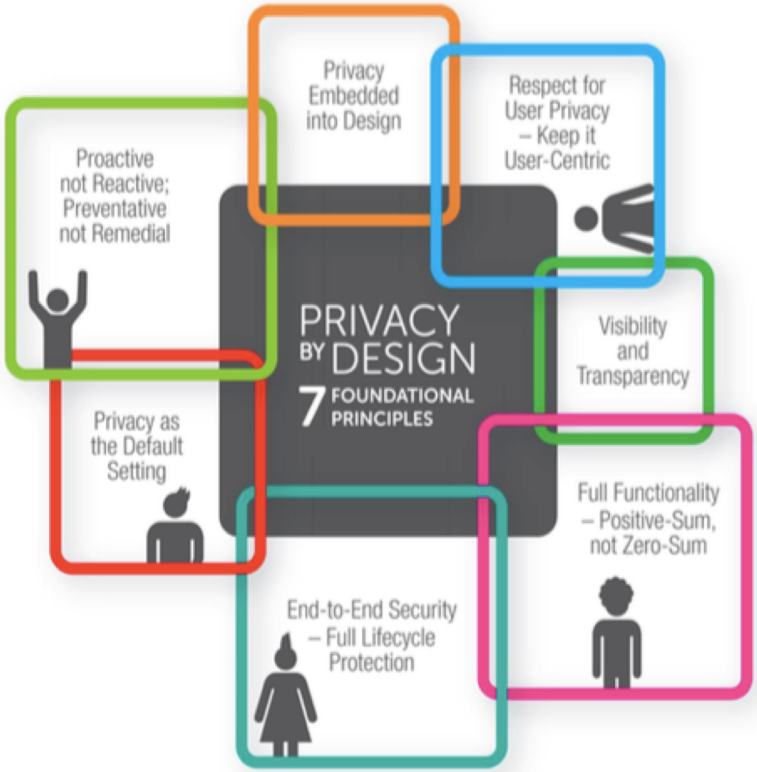
Basic elements of a DPIA

- A systematic evaluation of the processing operations and the purposes, including, where applicable, the legitimate interests pursued by the controller;
- An assessment of the necessity and proportionality of the processing operations in relation to the purposes;
- An assessment of the risks to the rights and freedoms of the persons concerned (data subjects);
- The measures provided to address the risks, with a specific indication of the safeguards, security measures and mechanisms adopted to ensure the protection of personal data and to demonstrate compliance with GDPR.

Privacy-by-design and privacy-by-default approach

PRIVACY-BY-DESIGN

Taking into account the state of the art, the cost of implementation and the nature, scope, context, purposes and risks of the processing, the controller must, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures aimed at ensuring compliance with the data-protection principles.



PRIVACY-BY-DEFAULT

The controller must implement appropriate technical and organisational measures for ensuring that, by default, only those personal data which are necessary to achieve the intended purposes are processed, in relation to, among other elements, the volume of data collected, the period of their storage, the scope of the processing activities and the permissions to access the datasets

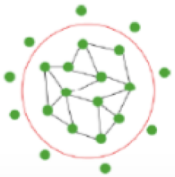
Blockchain & DPIA

- Is it a DPIA always required for Blockchain?
- It is presumed that a DPIA is necessary for that kind of blockchain system, where the **processing of personal data is the very purpose of the system**.

2. Challenges and opportunities for blockchain in Health – a legal perspective

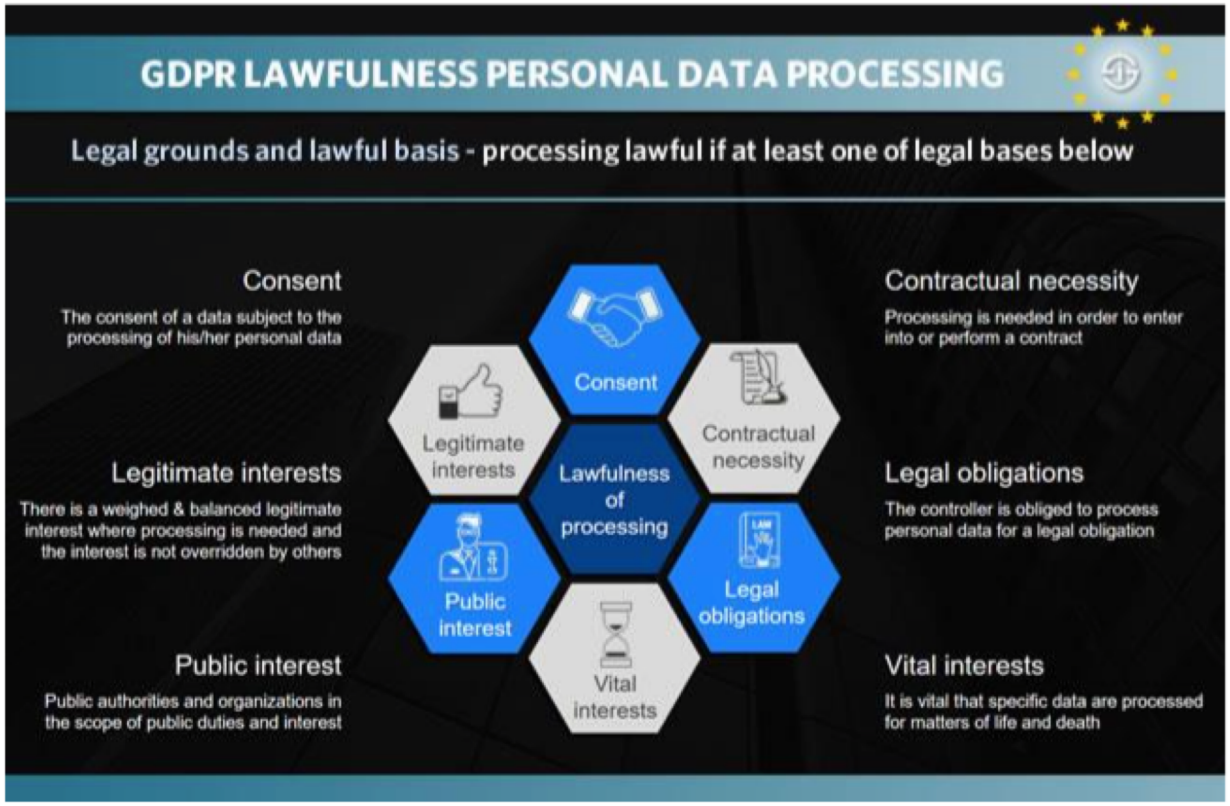
- **public, permissionless blockchains**: anyone is allowed to join the network and become a participating node or a validating node;
- **public and permissioned blockchains**: anyone can be a participating node and see all data, but only pre-approved actors can become validating nodes and add data to the ledger;
- **private and permissioned blockchains**: validating nodes and participating nodes must be preapproved by a governance of actors, generally in the form of a consortium of companies or government agencies. Furthermore, in some cases, there are rules in place that define who is able to see what data.
- As recommended by the *European Blockchain Observatory and Forum* ("**EU Observatory**") in its '*Blockchain and the GDPR*' report, in case of need to **store personal data**, it is **necessary to rely on private and permissioned blockchain**.

Blockchain types

Blockchain type	Explanation	Example	Visualization
Public permissionless blockchains	In these blockchain systems, everybody can participate in the consensus mechanism of the blockchain. Also, everyone in the world with a connection to the internet is able to transact and see the full transaction log.	Bitcoin, LiteCoin, Ethereum	
Public permissioned blockchains	These blockchain systems allow everyone with a connection to the internet to transact and see the transaction log of the blockchain, but only a restricted amount of nodes can participate in the consensus mechanism.	Ripple, private versions of Ethereum	
Private permissioned blockchains	These blockchain systems restrict both the ability to transact and view the transaction log to only the participating nodes in the system, and the architect or owner of the blockchain system is able to determine who can participate in the blockchain system and which node can participate in the consensus mechanism.	Rubix, Hyperledger	
Private permissionless blockchains	These blockchain systems are restricted in who can transact and see the transaction log, but the consensus mechanism is open to anyone.	(Partially) Exonum	

Lawfulness of processing

- **LEGAL BASIS WHICH ENSURE THE LAWFULNESS OF THE PROCESSING: ART. 6 OF THE GDPR**



Health Data as Sensitive Data

- **LEGAL BASES WHICH ENSURE THE LAWFULNESS OF THE PROCESSING: ART. 9 OF THE GDPR**
- **It is prohibited to process special categories of data**

  unless specific conditions are satisfied which ensure the lawfulness of the processing
- e.g. with reference to the **health sector**, when the processing is necessary:
  - ✓ **for reason of public interest in the area of public health**, such as protecting against **serious cross-border threats** to health or ensuring **high standards of quality and safety of health care** and of medicinal products or medical devices, on the basis of Union or Member State law which provides for suitable and specific measures to safeguard the rights and freedoms of data subject, in **particular professional secrecy**;
  - ✓ **for archiving purposes** in the **public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89.1** based on Union or Member State law which shall be **proportionate** to the aim pursued, **respect** the essence of the right to data protection and provide for suitable and specific measures to **safeguard** the **fundamental rights and the interests of the data subject.**

Reuse of medical data for research purposes

SCIENTIFIC AND MEDICAL RESEARCH UNDER THE GDPR

REUSE FOR RESEARCH PURPOSES

The GDPR establishes that **further processing for scientific research purposes must «*not be considered to be incompatible with the initial purposes*», so long as technical and organisational measures have been put in place** to ensure that: a) only those personal data will be processed which are strictly necessary to pursue the research purposes (data minimization) and, b) strict functional separation between participants in the research and outside stakeholders, meaning that data used for research purposes must not be made available for other kind of processing activities

Sensitive data

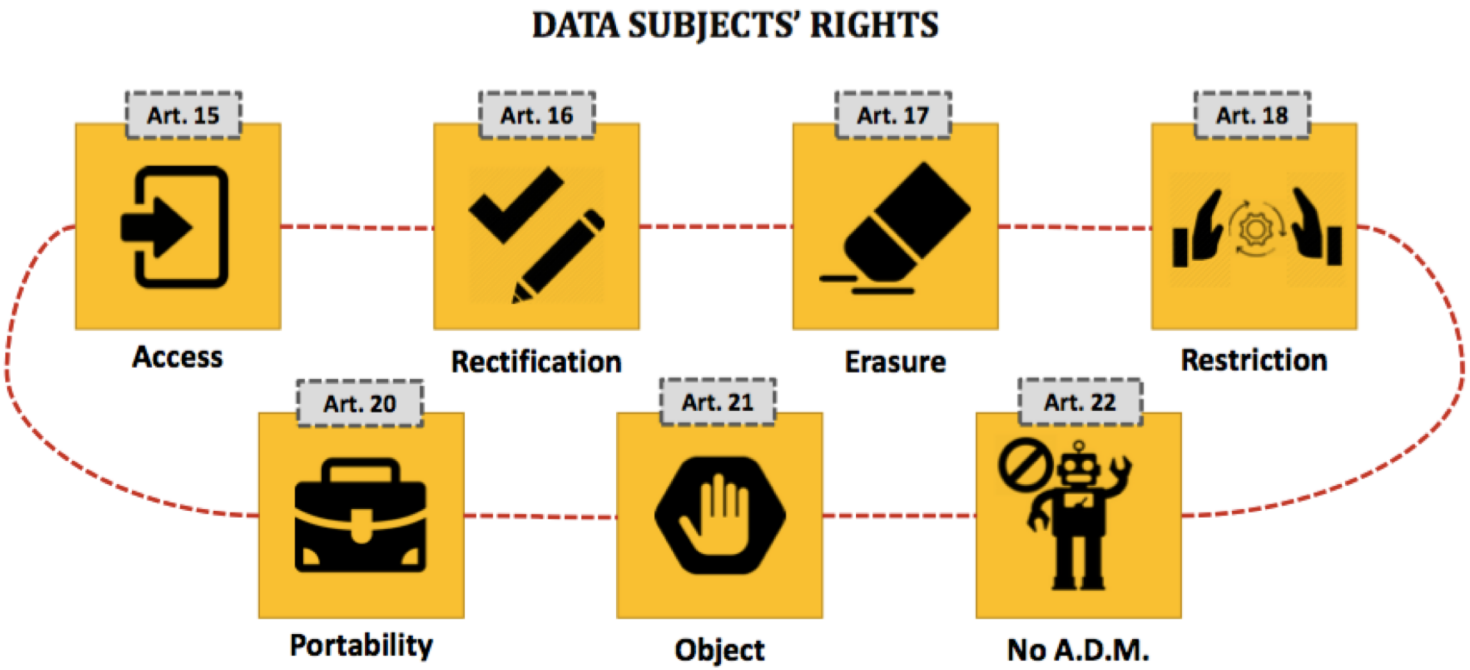
When the research involves any special category of data (e.g. health data), the application of flexible approach should be subject to a stricter interpretation, requiring a high degree of scrutiny

'Broad consent'

As it is often impossible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection, **data subjects should be allowed to give their consent to certain areas of scientific research** when in keeping with recognised ethical standards.

Although the choice to adopt a Regulation, in lieu of a Directive, was mainly aimed at preventing any fragmentation of the rules on the circulation and the protection of personal data within the European Union, many member States are however adopting their own laws to implement the GDPR in its entirety or, in some cases, to adapt the national frameworks to the new rules set forth in this Regulation.

Respecting data subjects' rights

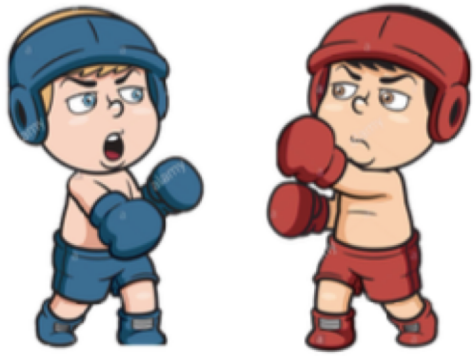


Where personal data are processed for scientific research purposes, Union or Member State law may provide for additional derogations from the rights referred to above, subject to the conditions and safeguards set forth by Art. 89 of the GDPR, when such rights are likely to render impossible or seriously impair the achievement of the research and such derogations are necessary for the fulfilment of the relevant purposes

3. Regulatory-compliant solutions: GDPR vs BLOCKCHAIN

*«GDPR AIMS TO **REGULATE** THE WORLD CENTRALISED DATA CONTROL,*

*WHEREAS THE AIM OF BLOCKCHAIN IS TO **CHALLENGE** IT»*



FONTE: Luis-Daniel Ibanez, Kieron O' Hara, and Elena Simperl, *On Blockchains and the General Data Protection Regulation*, University of Southhampton, 2018

Data subjects' rights and blockchain environment

- In a recent report, the French Data Protection Authority (*Commission Nationale de l'Informatique et des Libertés*, briefly "**CNIL**") explicitly revealed its **concerns with regards to the exercise of data subjects' rights in the blockchain environment**.
- Notwithstanding the choice of a **private and permissioned blockchain**, more in line with GDPR's principles and obligations, the CNIL focused its attention on the remaining **unsolved issues** at stake.
- As a matter of fact, given the **immutability** of the data retained on a blockchain, compliance with the GDPR has to be ensured by means of technical loopholes, with specific reference to the **rights of erasure, limitation and rectification**.
- Even if **no personal data are registered on the blockchain**, **actions** are to be envisaged in the event of exercise of individual rights, to guarantee full accountability

Risks of non-compliance

- Generally, even if **strong encryption** is applied on personal data, the result is very likely to prove **not fully anonymous** given that, as long as the decryption key exists somewhere, the data can still be singled-out, leading to a **reversal risk.**
- Another risk is that the **linkability** of encrypted data to an individual can be reached by further examining patterns of usage or context, or by comparison to other pieces of information.
- On top of this, cryptography technologies and science are subject to a **seamless evolution**.
- All such risks can be avoided by **preventing the registration of any kind of personal/sensitive data on chain** both in the Segregated Computation and the Secure Sharing Model, thus fulfilling the requirements of privacy and security in accordance with the GDPR and the EU Observatory's guidelines. The encrypted data are stored **solely off-chain**, in MHMD **distributed database**.

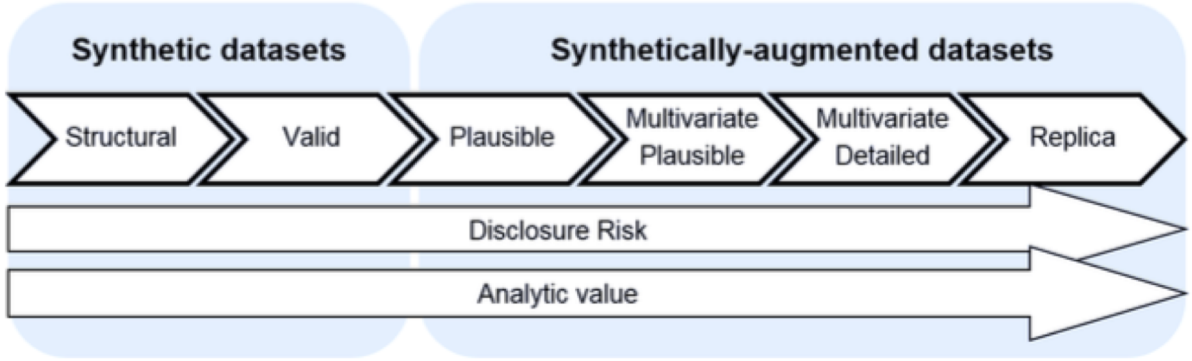
MHMD Blockchain solution

- Private and permissioned;
- Off-chain health data retention;
- Smart contract records the consents expressed by the interested parties;
- Inability to know the parties involved in a transaction;
- Each participant is solely responsible for the dataset they receive;
- Impossibility of hindering the reception of the dataset by the network participant.

Making use of synthetic data

SYNTHETIC DATA

These data are indeed totally made-up and do not contain any of the original identifiable information. They are generated after the density function of the attributes in the original dataset is identified and their parameters has been estimated. Then for each attribute, privacy protected series are generated by randomly picking up the values from said estimated functions. Multiple imputation and bootstrap methods are few classical techniques used to generate fully synthetic data. These data, automatically generated by making use of machine learning algorithms, are based on recursive conditional parameter aggregation, operating within global statistical models which, by definition, do not allow any personal re-identification of original individual datasets.



This image is taken from the *ONS methodology working paper series number 16 - Synthetic data pilot*

GDPR & BLOCKCHAIN: possible solutions

1. Start with the big picture: how is the value created for the user, how the data is used and ask yourself if **you really need a blockchain**?

1. **Avoid storing personal data on the blockchain**, making full use of data obfuscation, encryption and aggregation techniques to **anonymize data.**

2. **Collect personal data off-chain** or, if the blockchain cannot be avoided, on **private and permissioned blockchains.**

3. **Stay up to date** and propose innovative solutions, trying to be as transparent as possible towards users.