MY HEALTH
MY DATA

Encryption, Anonymisation, and Artificial Intelligence

Edwin Morley-Fletcher, Lynkeus

EMA – London 1st December 2017

Artificial Intelligence

a controversial philosophical issue (1)

René Descartes

The body is a machine, I am the thought: *cogito ergo sum* (1637)

Nam revera nunquam vidi au tercipi humana corpora cogitare, sed tantum eosdem esse homines, qui habent et cogitationem et corpus. (1641).

Where is the *res cogitans*?

Alan Turing

Thought is a process, which a machine can imitate (1939)

Ludwig Wittgenstein

The world is not simply a series of representations capable of being expressed by language, but a series of interpretations and communal understandings which take place through the playing of 'language-games'. Language-games are 'active' and are made comprehensible by the 'form of life' in which they are nested.

Eric Kandel

Our mind is a set of operations carried out by our brain (2013)

AI's controversial philosophical issue (2)

Karl Raimund Popper

Medaware Lecture (1986). "Evolution is not driven by the mechanical process of variation and elimination, but by the activity of organisms concerned with problem-solving and gaining new knowledge about the world … Cells also ... make use of unpredictable new knowledge about their environment. Their activity is also driven by a network of propensities. They too are partly unpredictable, but not irrational… because all of the inner processes are concerned with problem solving…. For each newly-discovered environment or … behaviour, the selection pressure supported the more creative cells or individuals and their suitable genes. In the end, humans learnt to speak and write. Life was able to build up huge quantities of exosomatic knowledge. Exosomatic knowledge is now continuously exchanged between all humans.

Denis Noble

The Music of Life: Biology Beyond Genes (2006): "Does life have a logic? Some evolutionary geneticists have argued that it cannot have. The evolutionary process is blind, imperfect, and subject to chance…. Evolution might have taken all sorts of different turns at any number of the many corners it encountered during the billions of years; and still what we have got makes its own kind of sense… It all has to emerge without there being a driver. The grand composer was even more blind than Beethoven was deaf!"

Daniel Dennett

From Bacteria to Bach and Back: The Evolution of Minds (2017) suggests that a natural part of the evolution of intelligence itself is the creation of systems capable of performing tasks their creators do not know how to do.

A successful biomimicry approach

- Some thought that intelligence would more easily emerge if machines took inspiration from biology, and learned by observing and experiencing.
- This meant turning computer programming on its head: instead of a programmer writing the commands to solve a problem, the program would generate its own algorithm, based on example data and a desired output.
- The machine-learning techniques that would later evolve into today's most powerful AI systems followed this path: the machine essentially programs itself.
- The essence of learning becomes the ability to detect, recognise, and eventually reproduce patterns.
- Algorithmic systems are increasingly able to learn from one another
- Any algorithm is able to learn and keep doing so indefinitely.

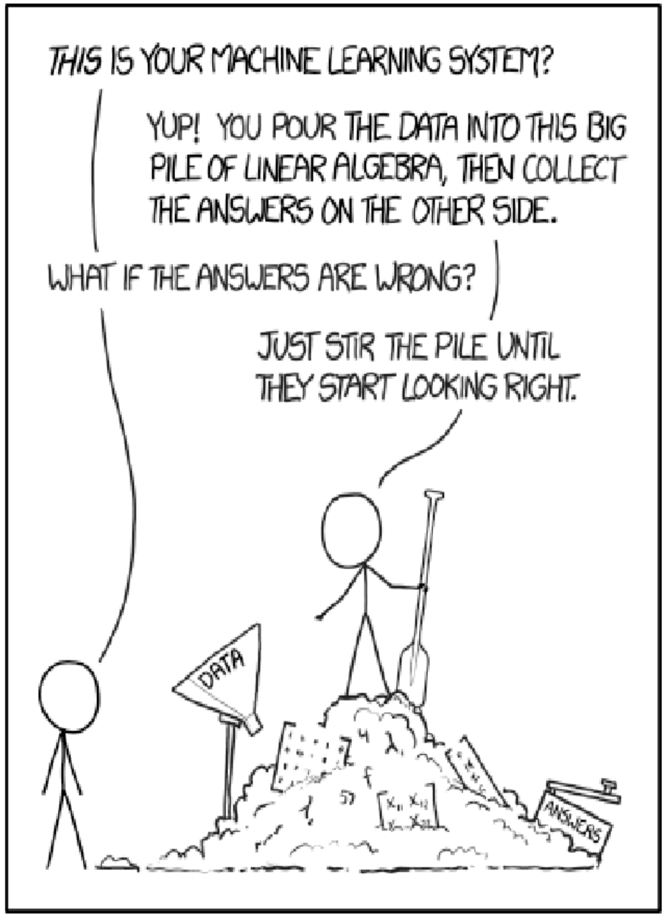
Artificial assistants

- Lower-end applications of AI allow already to augment human decision making performance.
- This process starts being commodified.
- It makes it possible to have skilled human operators focussing on ambitious and creative, instead of repetitive tasks.
- This rings a bell in response to a very old issue:
  - Aristoteles (384-322 BCE), *Politics*: “If every instrument could accomplish its own work, obeying or anticipating the will of others […] if, in like manner, the shuttle would weave and the plectrum touch the lyre without a hand to guide them, chief workmen would not want servants, nor masters slaves”.
  - Robert Aron and Arnaud Dandieu, *La révolution nécessaire (1933),* were distinguishing between repetitive work (*la besogne*) and creative activity.
- But raises societal concerns with regard to traditional employment.

Algorithmic production of knowledge (APK)

- Since 2015 computers are better than humans at image labelling:
  - The ImageNet Large Scale Recognition Challenge – whose accrued dataset is now more than 13 million images strong –, has reached an error rate which is incredibly low (around 2%), by having recourse to deep convolutional neural networks
  - Illumeo is the new imaging and informatics technology with adaptive intelligence developed by Philips, that redefines and enhances how radiologists work with medical images.
- AlphaGo (2016)
  - AlphaGo is a computer Go program which uses a Monte Carlo tree search algorithm to find its moves based on knowledge previously "learned" by machine learning, specifically by an artificial neural network, extensively trained both from human and computer play.
  - In 2016, it could beat a 9-dan professional, who found out that AlphaGo was incurring in apparent mistakes, later revealing how a machine can transcend the traditional definition of mastery: “it was not a human move, I had never seen a human play this move. So beautiful”, said the beaten master, Lee Sedol.
- DeepMind Health and Moorfields Eye Hospital in London (2016)
  - They have jointly started a medical research partnership through which approximately one million Optical Coherence Tomography scans will be used for deep learning on age-related macular degeneration (AMD).
  - DeepMind has committed to take rigorous measures to protect the security of the data, and avoid disclosing it to anyone other than the researchers and engineers working on the project.
- Deep Patient (2017)
  - A novel unsupervised deep feature learning method, based on a three-layer stack of denoising autoencoders, used to capture hierarchical regularities and dependencies in the aggregated EHRs of about 700,000 patients from the Mount Sinai data warehouse .
  - 76,214 test patients were analysed, comprising 78 diseases from diverse clinical domains and temporal windows.
  - The result is a representation evaluated as broadly predictive of health states by assessing the probability of patients to develop various diseases. Prediction models for severe diabetes, schizophrenia, and various cancers were among the top performing ones.

Further AI applications in clinical decision making


THIS IS YOUR MACHINE LEARNING SYSTEM?
YUP! YOU POUR THE DATA INTO THIS BIG PILE OF LINEAR ALGEBRA, THEN COLLECT THE ANSWERS ON THE OTHER SIDE.
WHAT IF THE ANSWERS ARE WRONG?
JUST STIR THE PILE UNTIL THEY START LOOKING RIGHT.
DATA
ANSWERS

SIEMENS Healthineers **DeepReasoner**, the generic and flexible clinical case reasoning tool , based on multi-task deep neural networks, which permit to learn abstract and compact representations.

- Its unsupervised training is driven by the aim of reconstructing input data after dimensionality reduction.
- It supports evidence-based decision making by retrieving similar patients, and enables different types of use cases, such as patient-like-mine as well as patient-like-me.

Watson for Oncology

- Watson uses the cloud-based supercomputer to digest massive amounts of data — from doctor’s notes to medical studies to clinical guidelines.
- Its treatment recommendations are not based on its own insights from these data. Instead, they are based on training by human overseers.
- Watson is improving rapidly: the system is offering guidance about treatment for 12 cancers that account for 80 percent of the world’s cases.
- This can save doctors’ time and ensure that patients get top-quality care.

Promises and paradoxes of APK

- APK learns from previous situations to provide input and automate complex future decision processes, making it easier to arrive at concrete conclusions based on data and past experiences.
- The first phase is normally ‘unpredictable by design’: it is based on Big Data Analytics, in which large number of algorithms are tested on data in view of discovering meaningful correlations.
- Once relevant correlations are found, new algorithms based on running machine learning techniques can be applied, aiming at learning their causality status.
- Deep learning implies feeding vast quantities of data through non-linear neural networks which classify the data based on hierarchical outputs from each successive layer.
- The complexity of this self-modelling is, as yet, inherently non-self-explicative.
- This can determine a black box effect, rendering automated decision-making altogether inscrutable: no one really knows how the deep learning algorithms get to do what they do.
- In as much as this remains so, the APK is built and operates in ways which appear as incomprehensible and seems to require paradoxically a ‘trust leap’, in order to let algorithms ultimately make decisions on your behalf.

MY HEALTH MY DATA

AI as a threat to the GDPR

- Of course, AI makes it easier and easier to re-identify data subjects.
- “It may be impossible to fulfill the legal aims of the Right to Be Forgotten in AI environments”: *Humans forget, machines remember: Artificial intelligence and the Right to Be Forgotten*, Computer Law and Security Review, Elsevier, 2017.

and also as a new tool for anonymisation

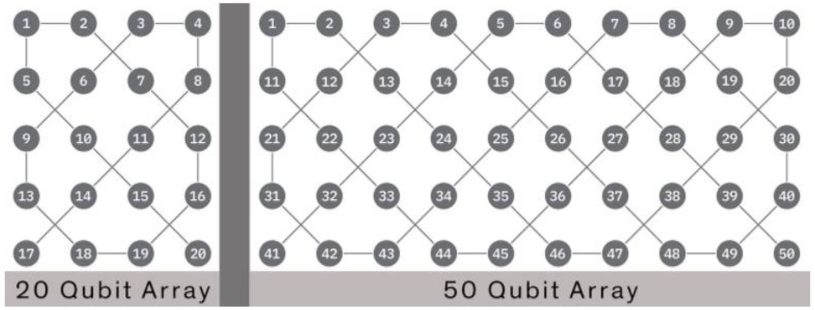
- Synthetic Data Vault uses machine learning to automatically generate artificial data on which scientists can test their algorithms and models.
- The algorithm itself is a form of recursive conditional parameter aggregation of real databases, which exploits their hierarchical nature.
- “Once we model an entire database, we can sample and recreate a synthetic version of the data that very much looks like the original database, statistically speaking…If the original database has some missing values and some noise in it, we also embed that noise in the synthetic version… In a way, we are using machine learning to enable machine learning”.

MHMD makes use of Synthetic Data

- MyHealthMyData (MHMD) initially found itself entangled within a sort of “Catch 22” condition with regard to its participating clinical institutions:
  - data could get mobilized only after the Ethics Committees would have given their green light,
  - the same Ethics Committees would not authorize the sharing of routine data until all MHMD solution details were fully clarified.
- As a a pragmatic alternative, Barts has offered to generate synthetic cardiac-oriented data sets (purporting to fictitious individuals) based on aggregate statistics of a population of 100,000 patients.
- These datasets have spurious correlations added to reflect the impact of cardiovascular risk factors on cardiovascular health.
- The datasets contain fake names, addresses, DOB, DOD, episode visits, anthropometry (e.g. weights, heights, BMI, BSA, etc.) and cardiac function parameters, etc.
- Examples of data types/sources targeted for early inclusion include Myocardial Infarction, cath lab data, demographics, CT images and text reports, MRI images and text reports, pacemaker data, echocardiography images and text reports, cardiac surgery data, data from the chest pain clinic and pathology data.
- Not only does this solution allow to get MHMD concretely unrolling, but it also makes it possible to test the major elements of MHMD development, including:
  - Evaluation of how standardized ontologies can be mapped onto such a (typical) data export format
  - Loading and processing of large datasets
  - Algorithm scaling (compute cost as a function of data size)
  - Multi-site compute (e.g. by chunking data and distributing over multiple sites, modelled for example as Virtual Machine instances)
  - Impact of pseudonymisation/obfuscation/aggregation techniques on a range of dimensional statistical measures
  - Allowing the hacking challenges and pen-tests, but, of course, for the reidentification tests, for which real data will be necessary (and are reasonably expected to be available and duly consented).

More fundamental challenges deriving from Quantum Computing

- Quantum computers can explore potential solutions to problems simultaneously versus sequentially, through their ‘quantum parallelism’.
- All traditional computer programming resolves into binary code, but quantum computers represents data using qubits.
- Qubits are two-level quantum systems that can be manipulated into complex intermediate states, including states where the spin of one qubit may be linked, via quantum ‘logic gates’, with the entangled states of one or more other qubits.
- In a database of N items, a classical computer needs N/2 tries for factoring a number, whereas a quantum computer would find it in a number of steps equal to the square root of N.
- The superposition of interconnected qubits translate to exponentially more computing power.
- An IBM 20-qubit machine, using super-conducting qubits, with increased ‘coherence time’, is being made available by the end 2017.



Quantum Key Distribution

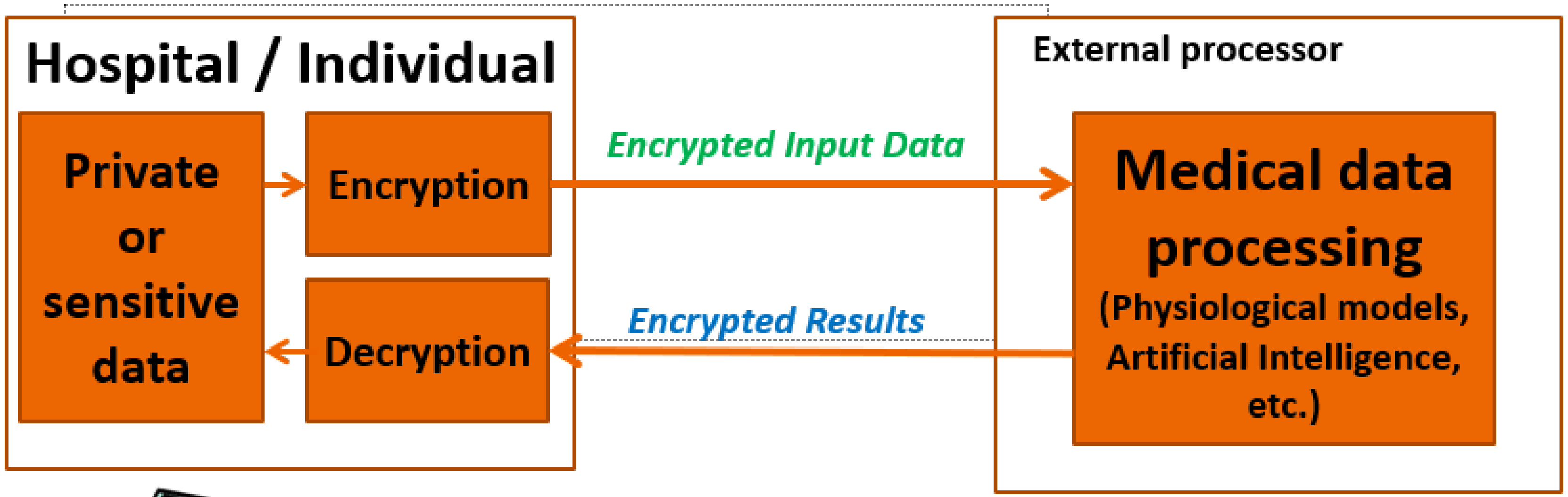
- The NSA has already developed a list of current encryption techniques that should no longer be used.
- Adversaries making use of quantum computers will soon be able to break current public-key encryption and signature schemes, unless encryption is also upgraded to quantum cryptography.
- Quantum key distribution offers mathematically proven security, guaranteed by the laws of physics.
- According to quantum mechanics, it is impossible to copy data encoded in a quantum state since the very act of reading quantum encoded data interferes with their superposition and modifies their state.
- Any attempt by eavesdroppers to intercept the quantum key would involve a measurement of some kind, unavoidably interrupting the qubits entanglement and alerting both sender and receiver.

Secure Multiparty Computation

- SMC allows a set of distrustful parties to perform the computation in a distributed manner, while each of them alone remains oblivious to the input data and the intermediate results.
- The computation is secure when at the end, no party knows anything except its own input and the results.
- Even though SMC has been known and researched for several decades, it was until recently considered too complicated and inefficient for practical use.
- Lately, several general SMC frameworks, distributed computational infrastructures and special purpose protocols, have been proposed that are able to support real world use cases proving the feasibility of SMC based data analysis.
- The SODA (Scalable Oblivious Data Analytics) H2020 project, led by Philips, parallel to MHMD (it is funded under the same EC call), is fully based on SMC and aims at data not needing to be shared, only made available for encrypted processing.
- The PEP (Polymorphic Encryption and Pseudonymisation) project, led by RadboudUMC, and largely sponsored by Verily (which is a subsidiary of Alphabet), is an opensource system which responds to the GDPR requirements of both data protection by design and by default, and is expected to be functionally available by Q1 2018.

Homomorphic Encryption

for secure distributed processing of medical data



Hospital / Individual

Private or sensitive data

Encryption

Decryption

Encrypted Input Data

Encrypted Results

External processor

Medical data processing

(Physiological models, Artificial Intelligence, etc.)

Homomorphic encryption encryption that allows for computations on encrypted data, generating an encrypted result which, when decrypted, matches the result of the operations performed on the non-encrypted data

→ **Main purpose: allow for computation on encrypted data**

Homomorphic encryption

Fully Homomorphic encryption:

- Standard approach: Gentry’s scheme
- 7 orders of magnitude slower than computing on plaintext data
- Performing operations on encrypted data adds noise → Limited number of operations that can be performed

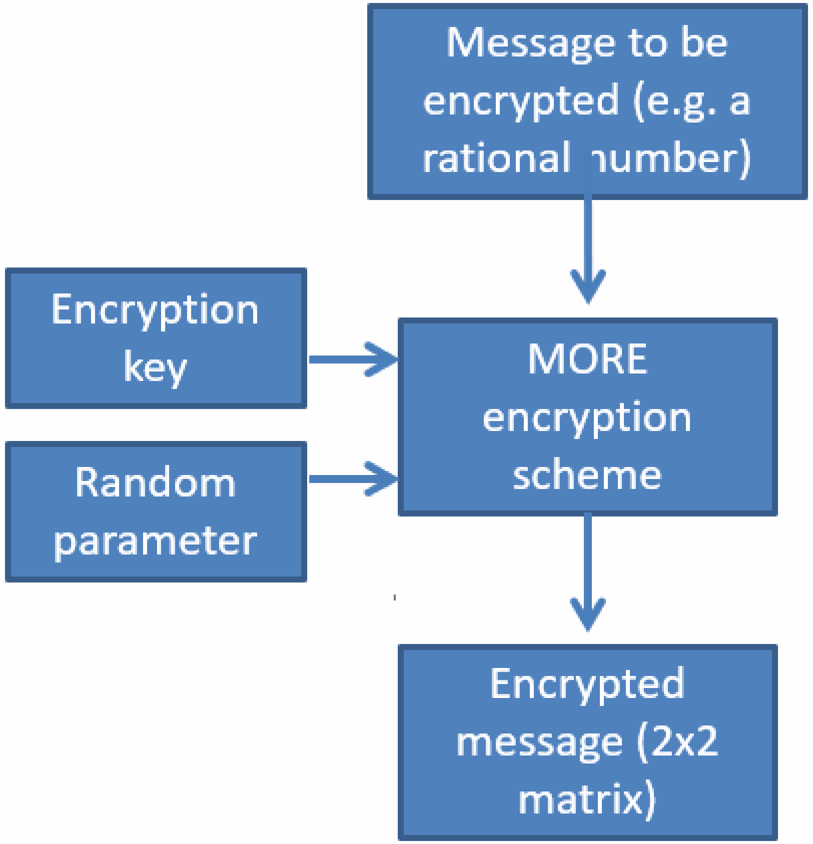
Partial homomorphic encryption:

- Allows for the evaluation of certain operations (e.g. addition, multiplication)
- Significantly faster than Fully Homomorphic encryption
- Can be used in practical applications with reasonable overhead
- There are different PHE schemes with different homomorphic properties.
- MHMD is using the MORE PH encryption system introduced by Kipnis and Hibshoosh in 2012.

Partial Homomorphic Encryption

MORE Encryption scheme used by MHMD

- Fully homomorphic w.r.t algebraic operations: addition, subtraction, multiplication, division.
- Noise free: unlimited number of operations can be performed on cyphertext data.
- Non-deterministic: multiple encryptions of the same message (and with the same key) result in different cyphertexts.
- Negligible computational overhead.


Message to be encrypted (e.g. a rational number)
Encryption key
Random parameter
MORE encryption scheme
Encrypted message (2x2 matrix)

$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

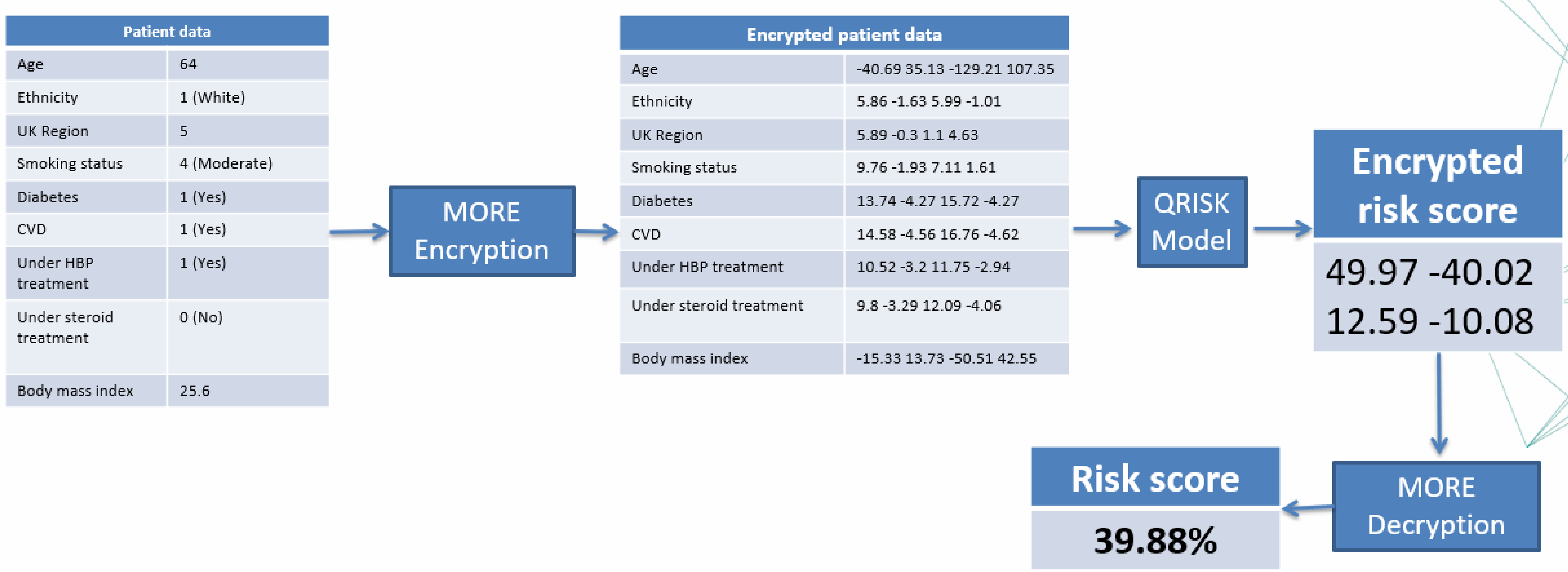
Secure and anonymised computation of health risk scores

Qrisk score (prediction algorithm for CVD)

- QResearch prediction algorithms based on:
  - Traditional risk factors (demographics, patient history, etc.)
  - Blood tests
  - Blood pressure measurements
- Derived from 24 million health records
- Predicts the risk to develop a certain pathology / condition during an interval of time of 1-10 years.
- Several risk scores: Cardiovascular disease, Diabetes, Kidney disease, Osteoporosis, Thrombosis, Cancer, Stroke, etc.

- **Goal: Develop a homomorphic encryption scheme that enables the computation of risk scores without decrypting the input information.**

Secure and anonymized computation of health risk scores
Qrisk score - Example



Patient data	
Age	64
Ethnicity	1 (White)
UK Region	5
Smoking status	4 (Moderate)
Diabetes	1 (Yes)
CVD	1 (Yes)
Under HBP treatment	1 (Yes)
Under steroid treatment	0 (No)
Body mass index	25.6

MORE Encryption

Encrypted patient data	
Age	-40.69 35.13 -129.21 107.35
Ethnicity	5.86 -1.63 5.99 -1.01
UK Region	5.89 -0.3 1.1 4.63
Smoking status	9.76 -1.93 7.11 1.61
Diabetes	13.74 -4.27 15.72 -4.27
CVD	14.58 -4.56 16.76 -4.62
Under HBP treatment	10.52 -3.2 11.75 -2.94
Under steroid treatment	9.8 -3.29 12.09 -4.06
Body mass index	-15.33 13.73 -50.51 42.55

QRISK Model

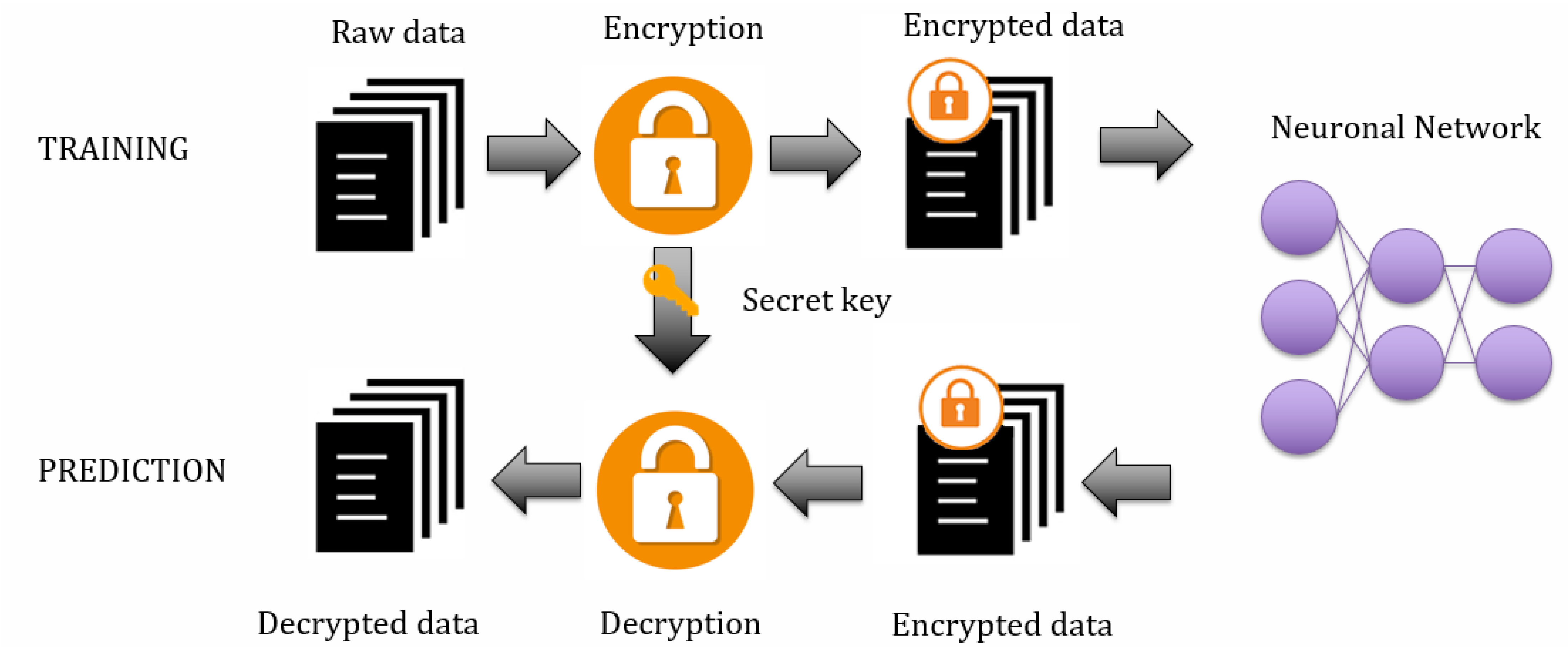
Encrypted risk score
49.97 -40.02 12.59 -10.08

MORE Decryption

Risk score
39.88%

Artificial Intelligence and Homomorphically Encrypted Data

Baseline Infrastructure

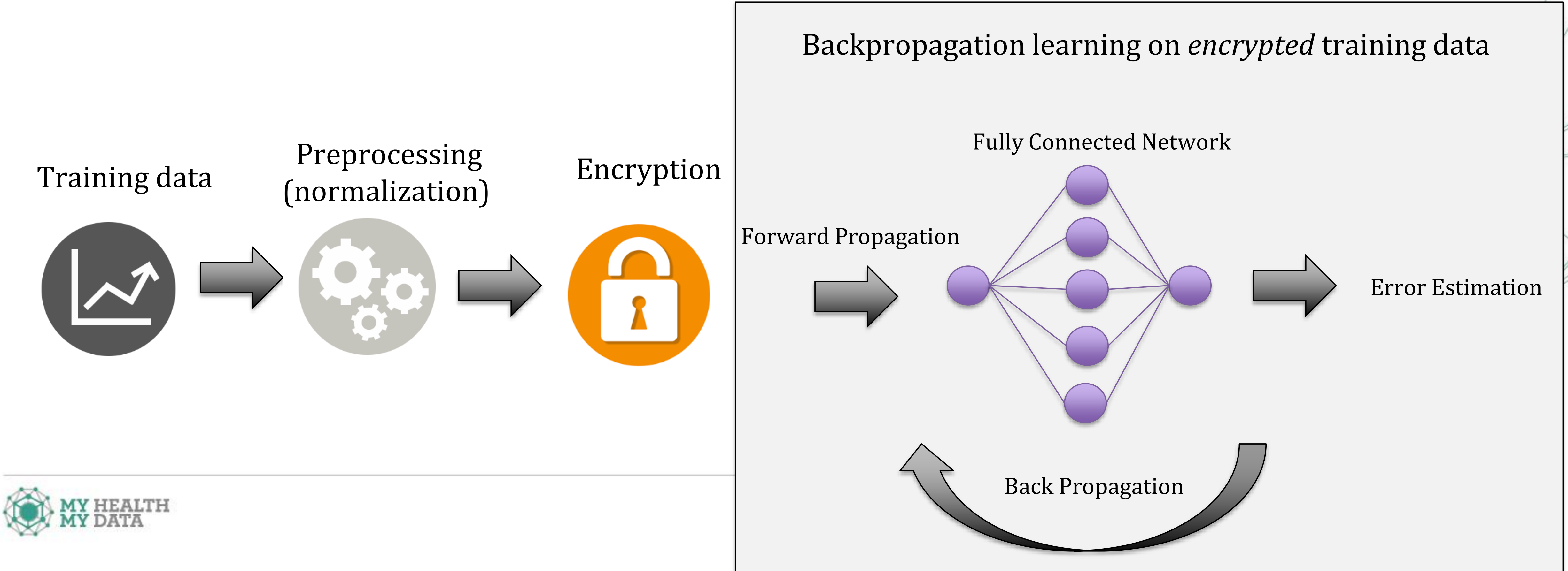


Privacy Preserving Regression

Fully Connected Neural Networks – Sine Function Modeling (Training phase)

A neural network is trained to approximate the sine function

$$x = \sin(y),\ \ y \in [0, 2\pi]\ ,$$ where y represents the network input



Training data

Preprocessing (normalization)

Encryption

Backpropagation learning on *encrypted* training data

Fully Connected Network

Forward Propagation

Error Estimation

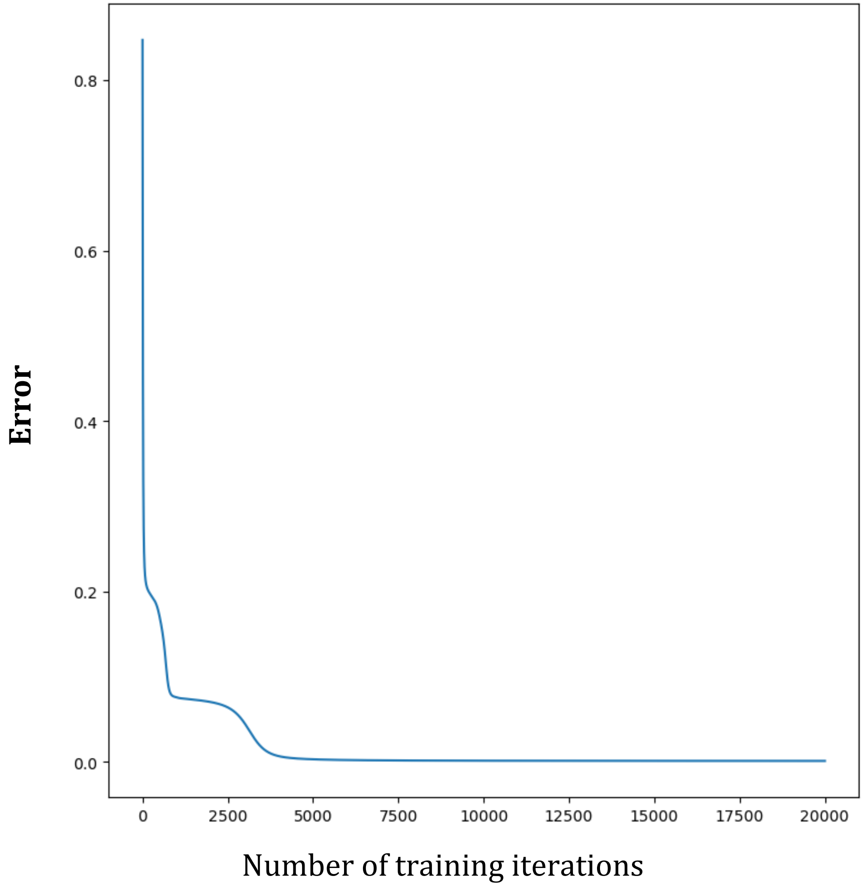
Back Propagation

Privacy Preserving Regression

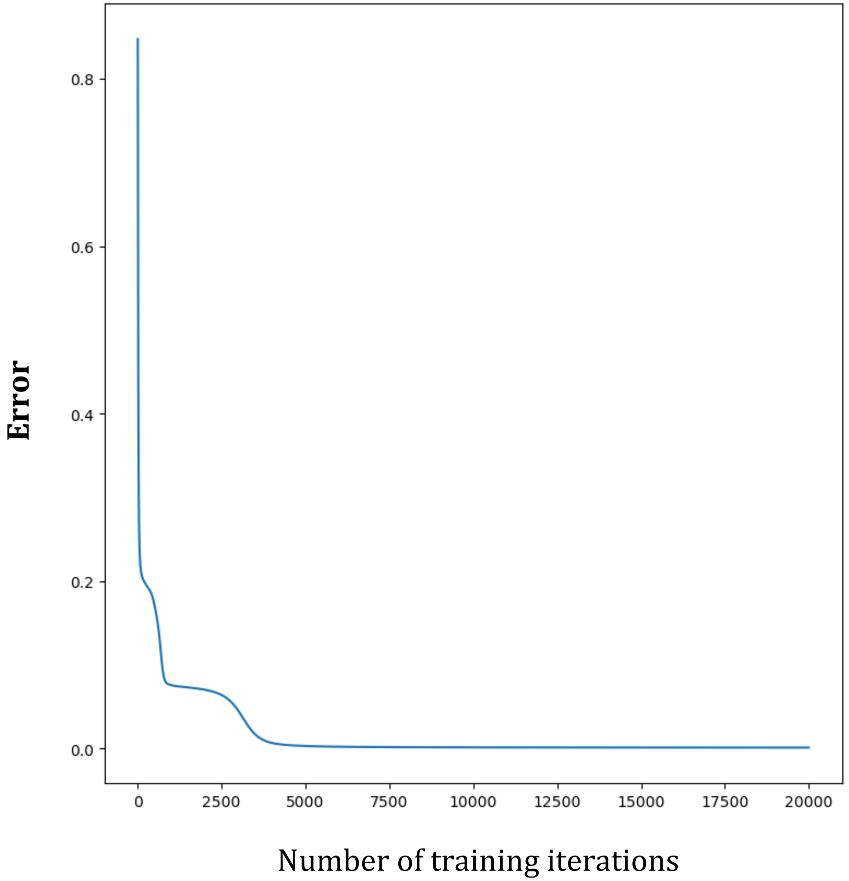
Sine Function Modeling – Encrypted vs. non-Encrypted (Training phase)

Training progresses similarly in encrypted and non-encrypted use-case

Training loss on non – encrypted data



Training loss on encrypted data

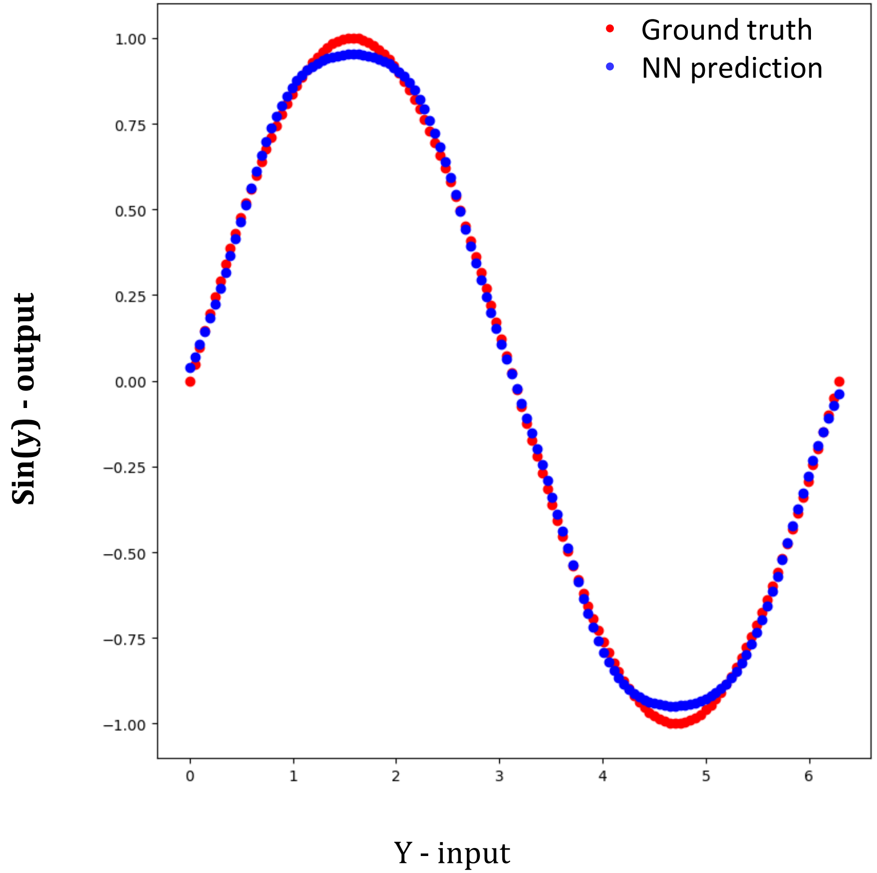


Privacy Preserving Regression

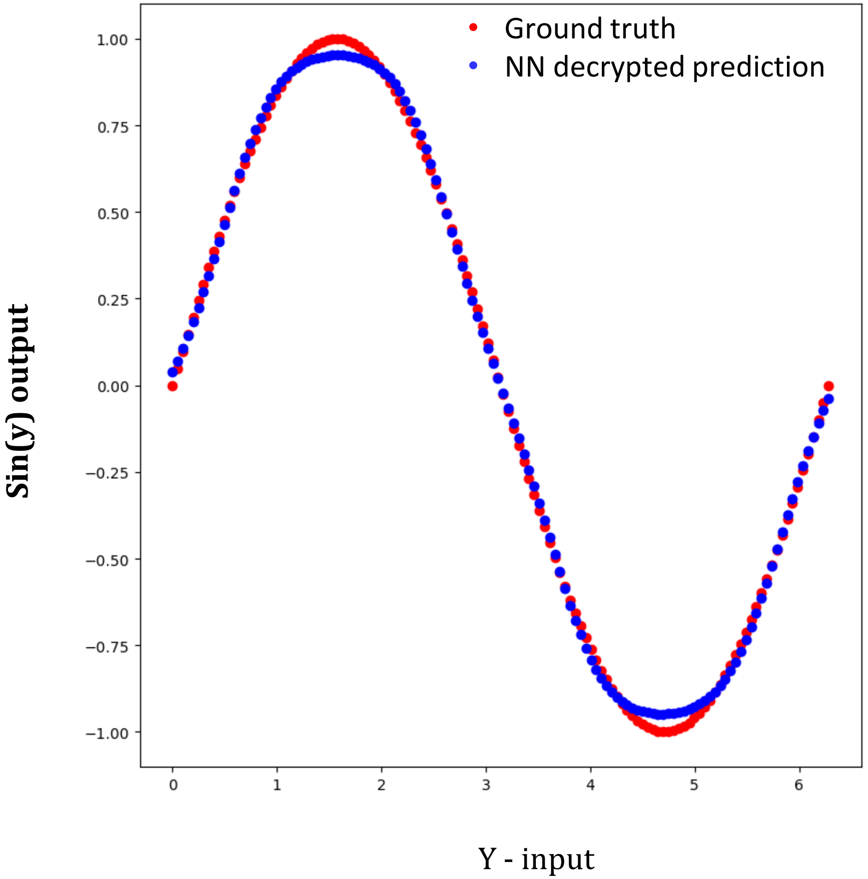
Sine Function Modeling – Encrypted vs. non-Encrypted Prediction

Training progresses similarly in encrypted and non-encrypted use-case

Prediction from **non–encrypted** trained model



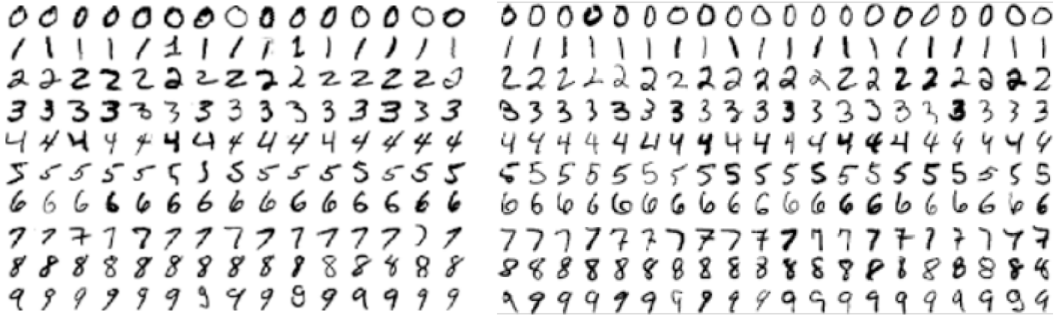
Prediction from **encrypted** trained model

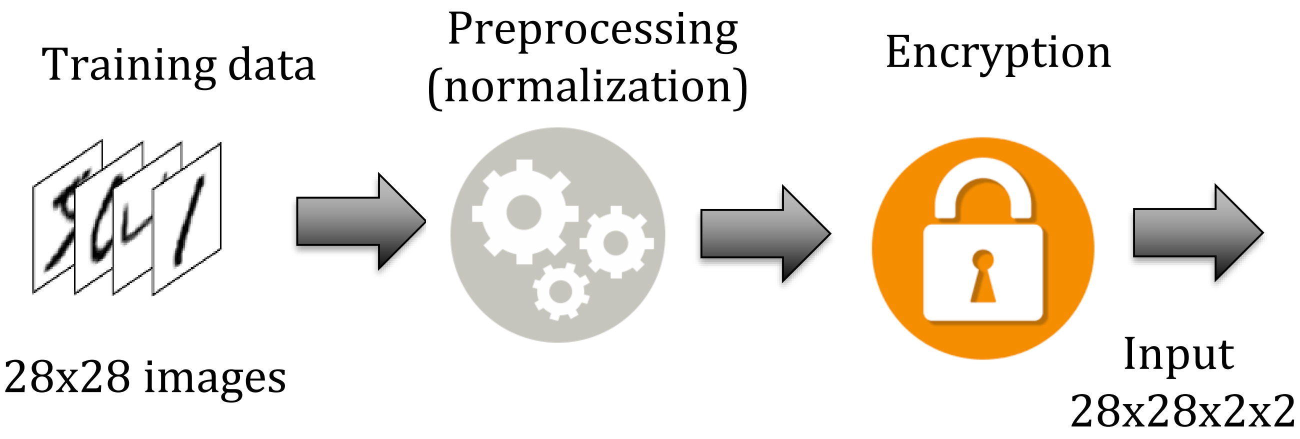


Privacy Preserving Classification

Convolutional Neural Networks – MNIST Digit Recognition

A neural network is trained to recognize hand written digits from 2-D images





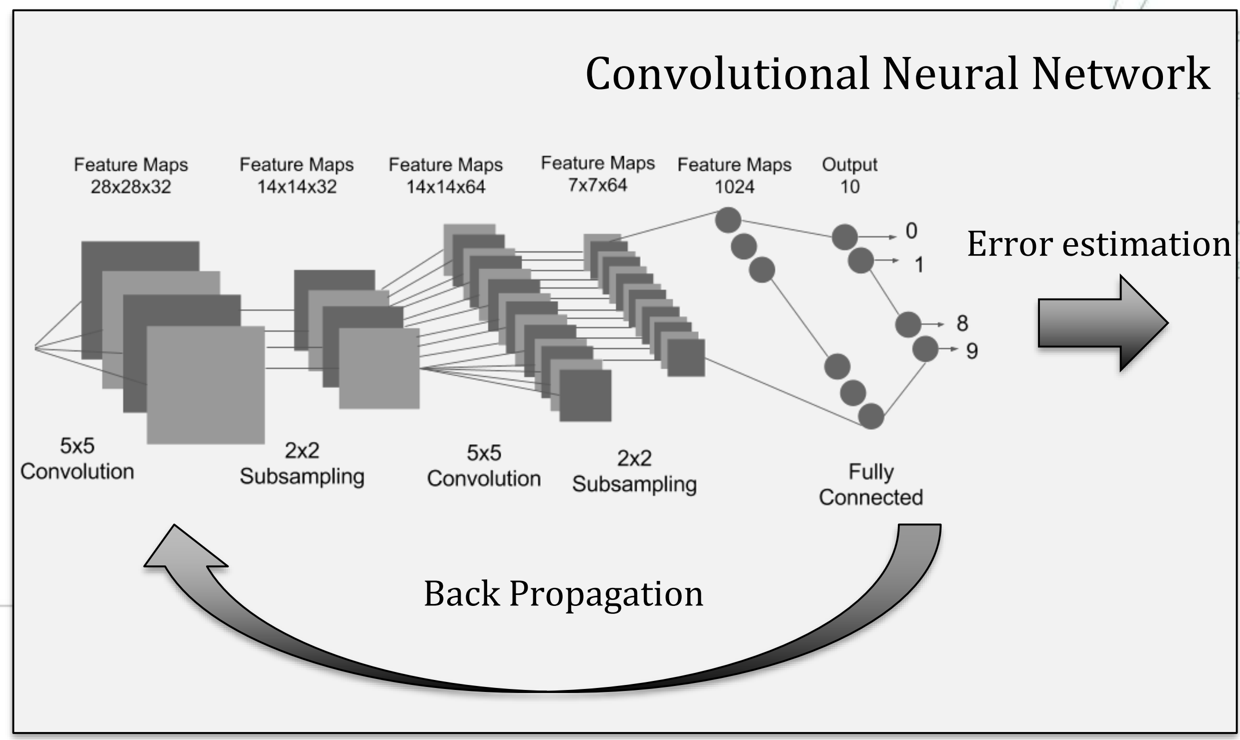
Training data

Preprocessing (normalization)

Encryption

28x28 images

Input 28x28x2x2



Convolutional Neural Network

Error estimation

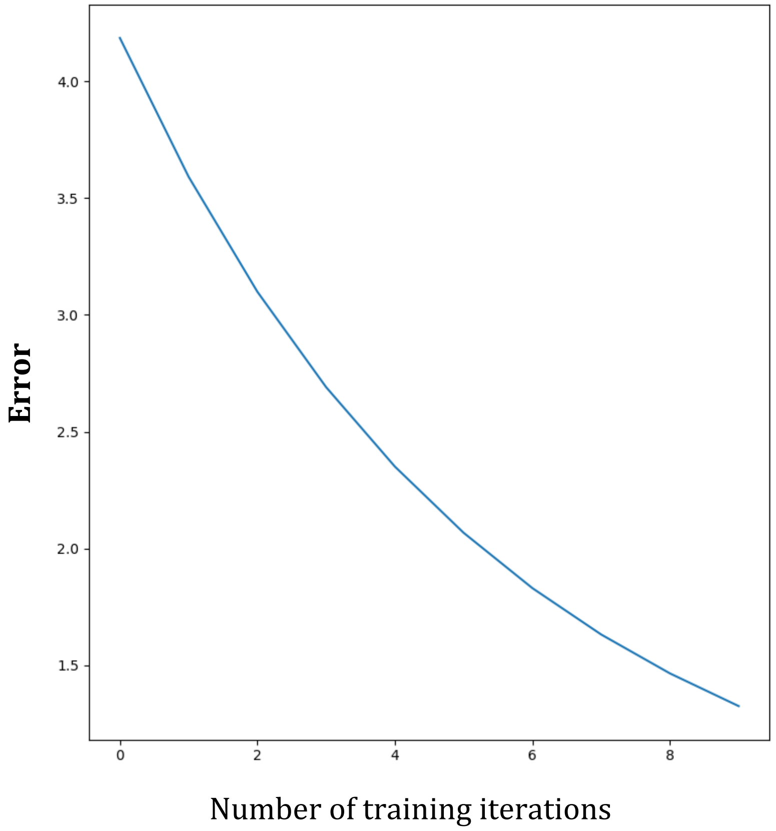
Back Propagation

Privacy Preserving Classification

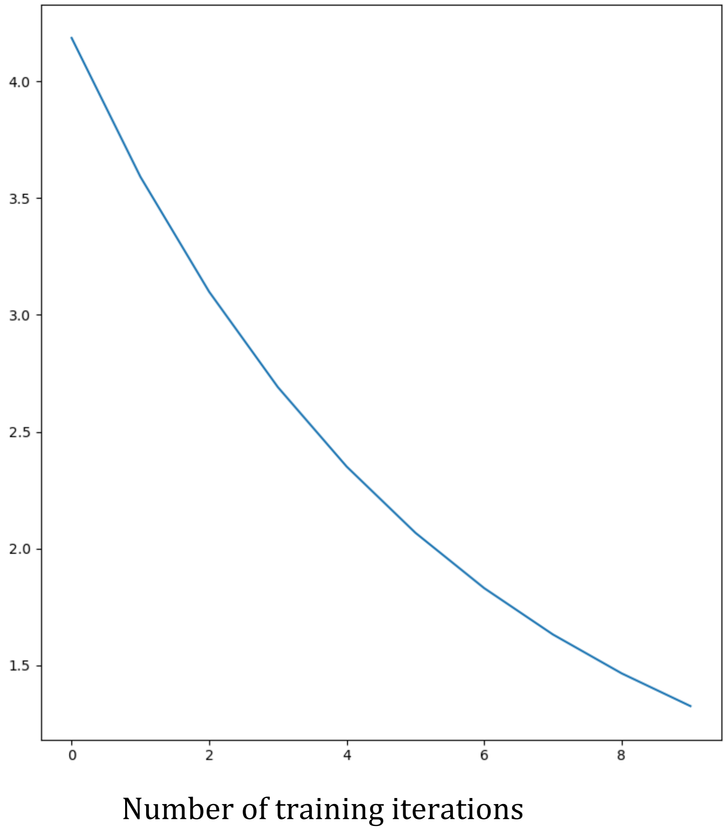
MNIST Digit Recognition – Encrypted vs. non-Encrypted Training

Training progresses similarly in encrypted and non-encrypted use-cases

Training loss on non – encrypted data



Training loss on encrypted data



Classification accuracy **non-encrypted** data: **97.45%**

Classification accuracy **encrypted** data: **97.22%**

Both SMC and HE are not directly taken into account by the GDPR

- The unique piece of regulation adopted so far is the **Opinion 05/2014 on Anonymization Techniques, issued by** the **Article 29 Working Party** (the technical body tasked with providing the European Commission with independent advice on data protection matters and supporting the development of harmonised policies for data protection).

- The WP29 points out that **anonymization must be irreversible in ordered to considered as such**.

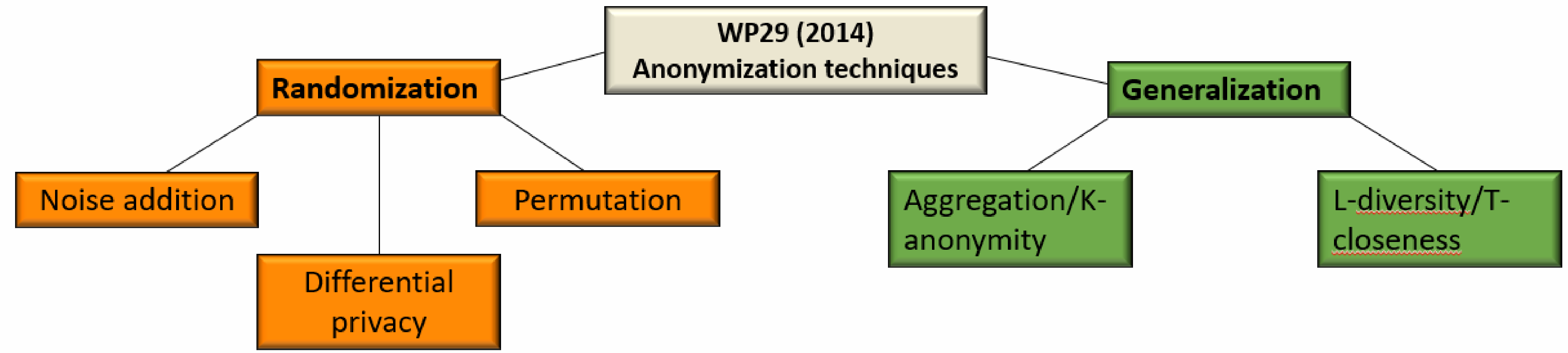
- Based on the applicable legislation, the Article 29 WP has highlighted four key features:
  - anonymisation is the result of processing personal data with the aim of irreversibly preventing identification of the data subject;
  - several anonymisation techniques may be envisaged, there is no prescriptive standard in EU legislation;
  - importance should be attached to contextual elements: account must be taken of 'all' the means 'likely reasonably' to be used for identification by the controller and third parties, paying special attention to what has lately become, in the current state of technology "likely reasonably" (given the increase in computational power and tools available);
  - a risk factor is inherent in anonymisation and must be evaluated in depth, also in terms of relevant severity and likelihood, when assessing the validity of any technique (including the possible uses of any data that is "anonymised" by way of such technique).

- **Being both reversible, SMP and HE must be considered as pseudonymisation techniques**

Ongoing evolution of anonymisation techniques

"**Research, tools and computational power evolve. Therefore, it is neither possible nor useful to provide an exhaustive enumeration of circumstances when identification is no longer possible**" (*Opinion 05/2014 on Anonymization Techniques*, par. 2.2.2).

According to the Article 29 WP, "*a specific* **pitfall is to consider pseudonymized data to be equivalent to anonymised data** *(…) pseudonymized data cannot be equated to anonymised information as they continue to allow an individual data subject to be singled out and linkable across different data sets*".



Each of the above techniques "**fails to meet with certainty the criteria of effective anonymisation** *(i.e. no singling out of an individual; no linkability between records relating to an individual; and no inference concerning an individual). However as some of these risks may be met in whole or in part by a given technique, careful engineering is necessary in devising the application of an individual technique to the specific situation and in applying a combination of those techniques as a way to enhance the robustness of the outcome*" (*Opinion 05/2014 on Anonymisation Techniques*, par. 5.2).

Cancellation of the encryption key, or of the initial identifiable data, may help to reach an adequate anonymisation standard.

Anonymisation and Pseudonymisation

The Data Protection legislation does not apply to anonymous/anonymised data.

Neither the currently applicable Directive 95/46/EC, nor the forthcoming *General Data Protection Regulation* 2016/679 (“GDPR”, which will repeal said Directive and will apply starting from the 25th May 2018), provide a definition of anonymisation.

Anonymisation

Recital 26 of the GDPR reads that: to determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used (including all objective factors, such as the costs of and the amount of time required for identification, the available technology at the time of the processing and technological developments), either by the controller or by another person, to re-identify (single-out) the natural person directly or indirectly.

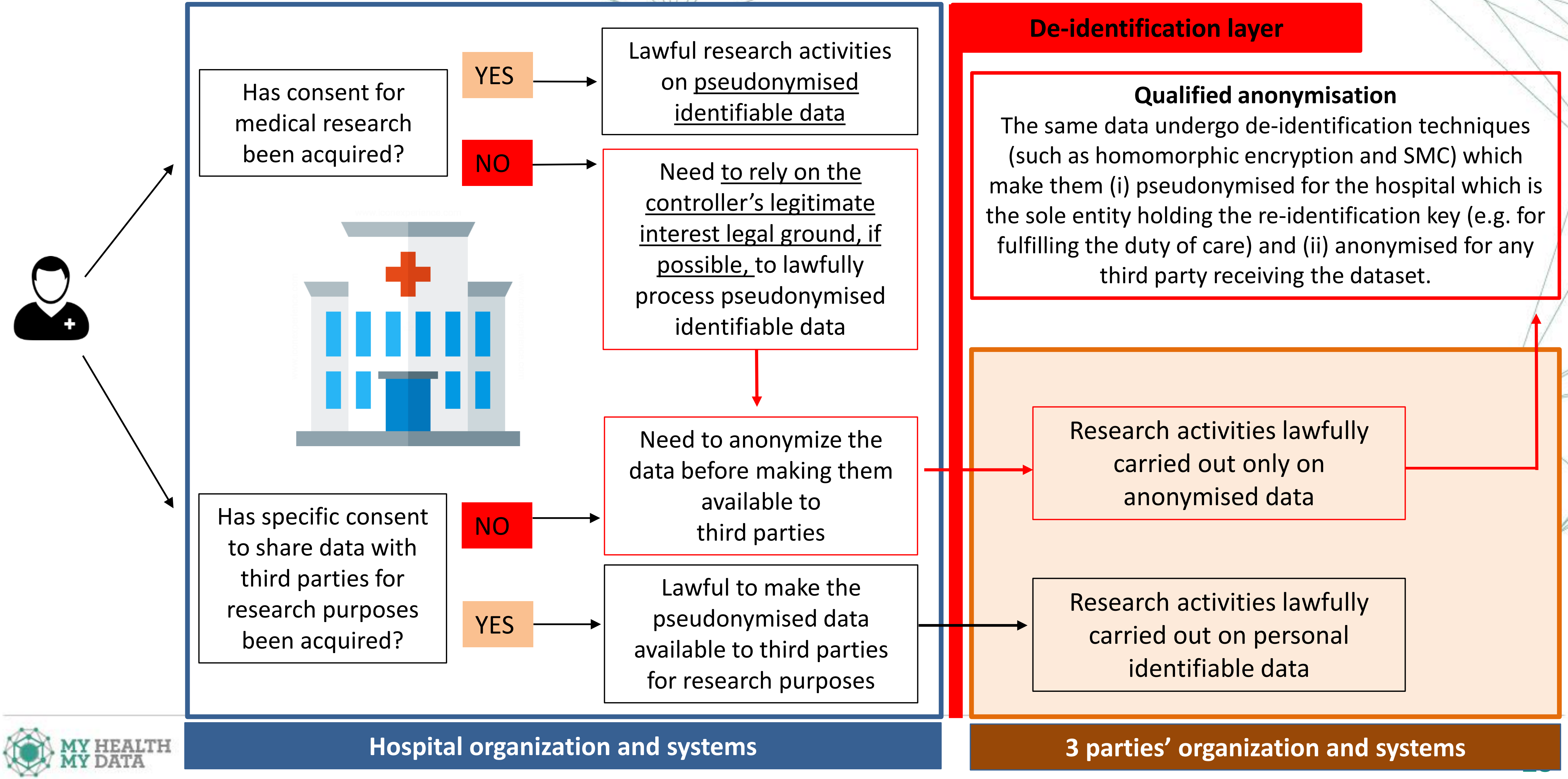
Pseudonymisation

Article 4(5) of GDPR defines pseudonymisation as processing personal data in such a manner that they can no longer be attributed to a specific data subject without the use of additional information to be kept separately and to be safeguarded by technical and organisational measures aimed to ensure that the personal data are not attributed to an identified or identifiable natural person. Pseudonymisation may therefore reduce general risks and help businesses to meet their data-protection obligations.

Anonymised information are not personal data – GDPR shall not apply

Pseudonymised information are personal data – GDPR shall apply

Can there be a “Qualified Anonimity” approach?



Has consent for medical research been acquired?

YES

Lawful research activities on pseudonymised identifiable data

NO

Need to rely on the controller’s legitimate interest legal ground, if possible, to lawfully process pseudonymised identifiable data

Need to anonymize the data before making them available to third parties

Has specific consent to share data with third parties for research purposes been acquired?

NO

YES

Lawful to make the pseudonymised data available to third parties for research purposes

De-identification layer

Qualified anonymisation

The same data undergo de-identification techniques (such as homomorphic encryption and SMC) which make them (i) pseudonymised for the hospital which is the sole entity holding the re-identification key (e.g. for fulfilling the duty of care) and (ii) anonymised for any third party receiving the dataset.

Research activities lawfully carried out only on anonymised data

Research activities lawfully carried out on personal identifiable data

Hospital organization and systems

3 parties’ organization and systems

The proposed paradigm of “Qualified Anonimity”

This concept has been introduced in the EU Horizon2020 **call DS-08-2017**, with deadline for submission 24 August 2017, relating to “**Cybersecurity PPP: Privacy, Data Protection, Digital Identities**”.

There are many cases in which researchers need to keep the capacity of re-tracing and singling-out specific participants into a study in order to assess the progression of diseases and the long-term outcomes of treatments, or simply to keep them informed, also about unexpected findings or life-saving discoveries (as well as in several other situations). Applying standard anonymisation rules do not constitute a viable solution in such cases, because if truly irreversible, it would prevent anyone to re-identify the data subjects, so hindering the objectives of the research and contradicting the basic principles of medicine.

The identification of individuals is not only something that may happen, rather it is something that must happen, **under specific circumstances defining a proper “qualification” granted by the law** (e.g. judges fulfilling their official duties, researchers finding a cure which may eradicate a disease, public authorities exercising their powers, etc.).

Should the response to this need be only left to national laws defining public interest issues?

Will it be possible to navigate between Scylla and Charrybdis?

- The GDPR is not only a fundamental European regulation.
- It also establishes some key 'civilisation principles' in the area of data protection.
- Privacy is a common good.
- Also anonymity should be considered as a common good.
- Currently, the are some significant risks:
  1. Problem regarding the **real role remaining for anonymisation**:
     - In absolute, data are always re-identifiable.
     - Research development implies also the need of a market capable of explicating the value of data.
     - Commercial transactions on data are lawful only if they are anonymised.
  2. The 'specific' **consent and re-consent** requirements implied by pseudonymisation **may be unpractical** and possibly highly counterproductive.
- One solution is to **reduce the transaction cost** of any such specification.
- These are some of the **reasons why MHMD**:
  - Is **blockchain-based**
  - Transforms the consent and permission choices into friction-free and permanently modifiable self-executable **smart contracts**
  - Makes it possible for any data provider to **fully track** the usages made of their data, while remaining **encrypted**.

MyHealthMyData (MHMD)

- MyHealthMyData aims to guarantee **privacy** and **security** of healthcare data by:
  - introducing a distributed architecture based on **Blockchain** and **Smart Contracts**,
  - serving both clinical institutions as well as individual data subjects, who will be making use of **Personal Data Accounts**.
- MHMD develops a comprehensive methodology to guide the implementation of data and identity protection systems, specifically defining approaches and tools to classify sensitive data based on their **medical** as well as **predictive**, and potentially **economic**, value, aiming to:
  - assess the most suitable and robust de-identification and **encryption** technologies needed to secure different types of information,
  - allow **advanced analytics** applied on such data,
  - evaluate the overall reliability of a generic multi modular architecture.
- MHMD also analyses users' behavioural patterns alongside ethical and cultural orientations, to identify dynamics related to events like **WannaCry**, the coming into force of the **GDPR**, and the **interactions of hospitals and individuals** within a system like MHMD.
- MHMD will check the ability of avoiding privacy & security breaches by having recourse to:
  - active **self-hacking**,
  - public challenges of **penetration testing** and **vulnerability assessment**,
  - testing **external re-identification possibilities** on patients consenting to being used as test-basis
- MHMD ultimately aims to:
  - improve the design of data-driven biomedical platforms,
  - foster the development of an information multisided-platform, in which a growing number of clinical institutions may find secure GDPR compliant ways of sharing data and leverage their value, as well as individuals, becoming able to easily access their personal data and control what use is made of them.

MHMD Partners

- **5 SMEs:**

  Lynkeus (Italy) [Coordinator], Digi.Me (UK), HWC (UK), Gnúbila (France), SBA Research (Austria)

- **4 Clinical partners:**

  Charité Berlin (Germany), Ospedale Pediatrico Bambino Gesù (Italy), St. Bart's-Queen Mary University London (UK), Great Ormond Street Hospital-University College London (UK)

- **4 Research centres and Academia:**

  Athena Research (Greece), Consiglio Nazionale delle Ricerche (Italy), HES-SO (Switzerland), Universitatea Transilvania din Brasov (Romania)

- **1 Legal consultancy:**

  P&A (Italy)

- **1 Industry:**

  SIEMENS Healthcare (Germany)

- Now additionally introducing also an **Icelandic** extension, composed of both Personal Data Accounts and multimodal datasets from the National Hospital in Reykjavik

Thank you

Contacts: emf@lynkeus.com

Lynkeus srl
Via Livenza 6
Roma 00198