Data Interoperability

05/10/2021

Giada Lalli
giada.lalli@kuleuven.be

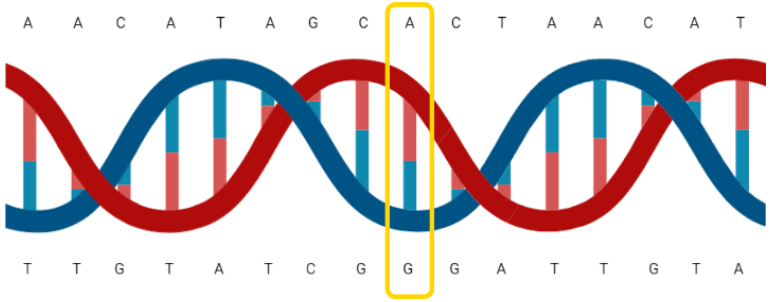
Background

Genotype is an organism's full hereditary information.

Phenotype is an organism's actual observed properties, such as morphology, development, or behavior.

Allele any one of two or more genes that may occur alternatively at a given site (locus) on a chromosome. Alleles may occur in pairs, or there may be *multiple alleles affecting the expression (phenotype) of a particular trait. The combination of alleles that an organism carries constitutes its genotype.*

Haplotype is a group of genes within an organism that was inherited together from a single parent.



More information can be found here.

Data Interoperability

Data interoperability addresses the ability of systems and services that create, exchange and consume data to have clear, shared expectations for the contents, context and meaning of that data.

Data

Interoperability

What about DATA?

No fixed definition → **semantic problem**:

- improper / too common use
- goal-driven assignment of the type
- data in healthcare: big data

METADATA

volume
velocity
variety

The three major challenges to bring the concept of Data Interoperability from theory to practice are related to semantic (*there is neither a definition of interoperability nor one of data and data interoperability that is shared across different communities and domains*), variety (*Variety is a characteristic spanning the entire spectrum of data features when they come from different and independent data sources. In a (global) research data infrastructure, data to be managed might be heterogeneous with respect to their type, accuracy, size, semantic, etc.*",) and data re-use (*In some contexts data have value if and only if they can be re-used. In order to make it possible for an entity to actually re-use data that have been collected or produced by a different entity, it is fundamental that a rich set of contextual information about such data be made available*).

The resolve on which more or less all DI scientists agree is a multistage and multilevel approach, integrating various existing methods and assuming the synthesis of different technologies and the adaptation of existing systems with an innovative solution between data providers and data users.

More information can be found here: https://doi.org/10.2481/dsj.GRDI-004

What about INTEROPERABILITY?

"The ability of two or more systems or components to exchange information and to use the information that has been exchanged", IEEE Glossary, 1991

Various level of interactions among entities:

- Organizational
- Semantic
- Technical

The definition originally assigned to the term 'interoperability', which was more likely to be associated with the integration of digital protocols or economic trends, holds certain transversality and there has recently been excitement in applying it to healthcare as well since it is a problem affecting the interaction of entities at various levels including: organizational, semantic and technical.

To date, data interoperability in healthcare still remains a very complex theoretical concept to put into practice, due to the diverse and fragmented nature of this sector.

More information can be found here.

Why

Precision Medicine

the focus is on identifying which approaches will be effective for which patients based on genetic, environmental, and lifestyle factors.

Stratified Medicine

based on identifying subgroups of patients with distinct mechanisms of disease, or particular responses to treatments. Stratified medicine allows us to identify and develop treatments that are effective for particular groups of patients.

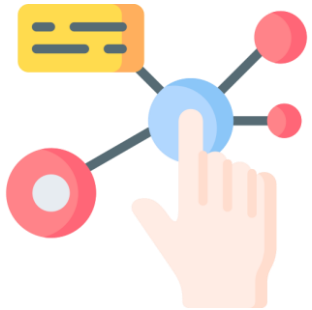
Point of Care

with-patient testing allows physicians and medical staff to accurately achieve real-time, lab-quality diagnostic results within minutes rather than hours.

“Point of care“ testing (POCT), also known as bedside testing, involves any type of diagnostic test that isn't done in the laboratory. More specifically, this type of testing is performed as close as possible to the patient, be it at their bedside or near them — hence the name “point of care”.

What for

GWAS



GWAS is the **association of individual markers** or groups of markers across the genome with phenotypic data. Markers are allele calls of a representative set of loci across the genome.

WGS

WGS provides a **very precise DNA fingerprint** that can help link cases to one another allowing an outbreak to be detected and solved sooner. Scientists take bacterial cells from an agar plate and treat them with chemicals that break them open, releasing the DNA. The DNA is then purified.

PHARMACOGENOMICS

Pharmacogenomics is a part of precision medicine. Pharmacogenomics is the **study of how genes affect a person's response to particular drugs**. This relatively new field combines pharmacology (the science of drugs) and genomics (the study of genes and their functions) to develop effective, safe medications and doses that are tailored to variations in a person's genes.

More information can be found here: https://doi.org/10.1038/s41598-019-53111-7

How

Deep Learning

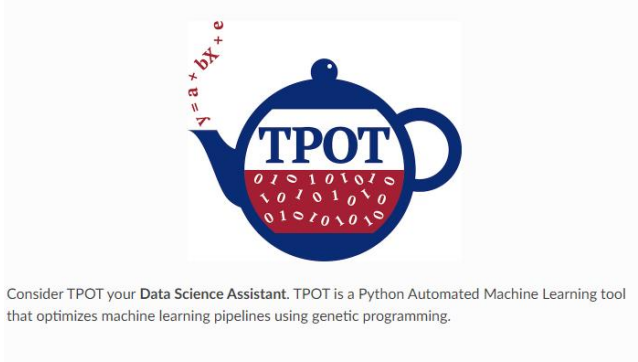
or **deep neural learning**, is a subset of machine learning, which uses the neural networks to analyze different factors with a structure that is similar to the human neural system.

Machine Learning

is a subset of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. In ML, there are different algorithms (e.g. neural networks) that help to solve problems.

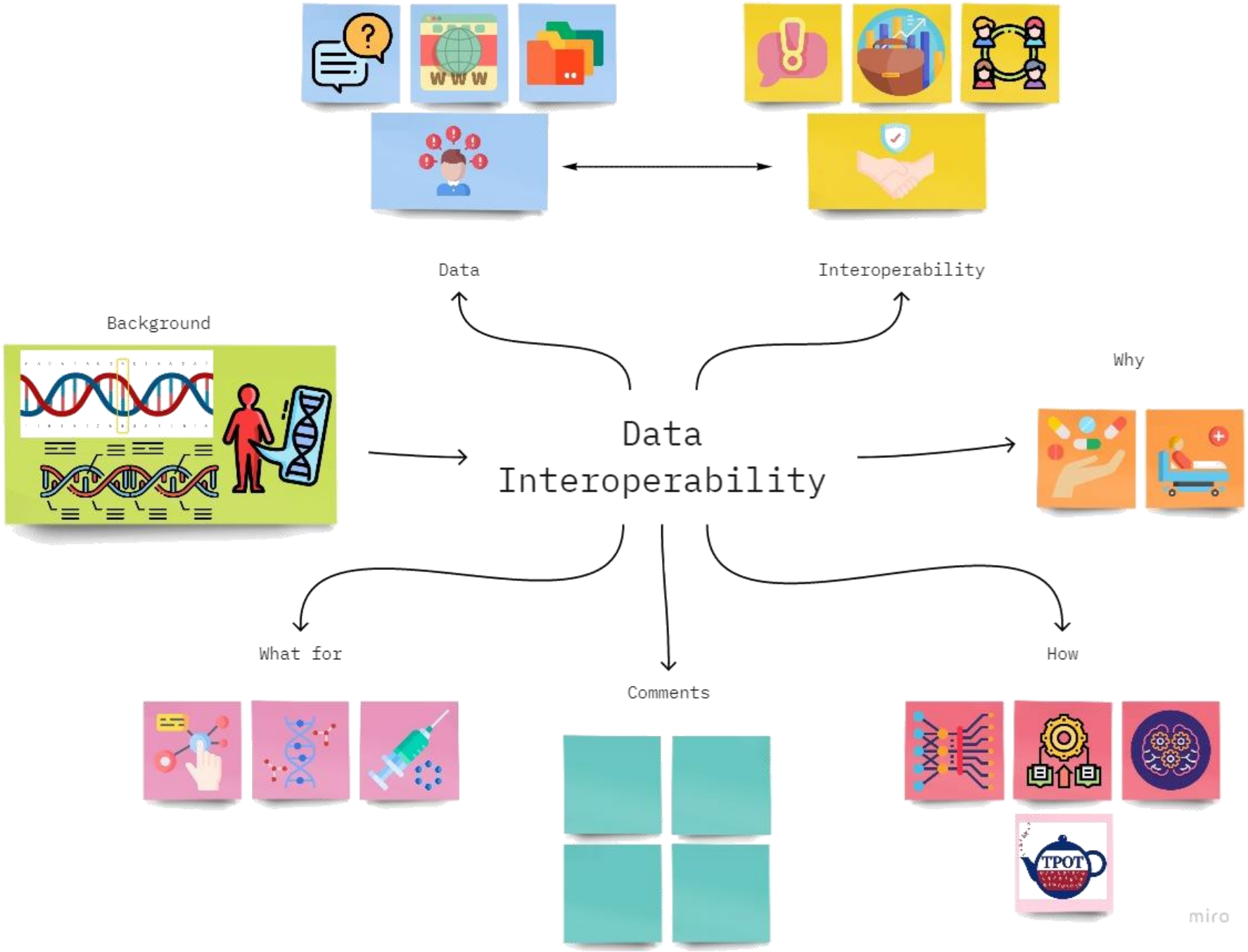
Artificial Intelligence

is a science like mathematics or biology. It studies ways to build intelligent programs and machines that can creatively solve problems, which has always been considered a human prerogative.



More information can be found here, here and here.

Overview


Data
Interoperability
Background
Why
Data
Interoperability
What for
Comments
How
TPOT
miro