inDemand
Together for better health
MURCIA | OULU | PARIS REGIONS

inDemand: Demand driven co-creation for public entities

CHALLENGE 1: ANONYMOUS (GHU PARIS)

Pitch

Anonymization of patient data for research purpose

Global definition of the challenge

In the era of big data and advances in data analysis, the collection and storage of this data for exploitation and possible publication seems essential in order to produce and develop scientific knowledge in research.

However, there is always a risk of re-identification. The qualification of anonymous data represents a real challenge, insofar as it constitutes either a legal obligation (open data), that is to say a tool for compliance with the protection of personal data. This situation leads us to a process of protecting the use of personal data in a context of growing computerization of the processing of patient records. Our goal is to set up a reference methodology with an industrializable solution to guarantee anonymisation while respecting the laws of privacy of patients.

The technique of anonymizing the data destroys any possibility of being able to identify to which individual the personal data belongs. This process consists of modifying the content or structure of the data in question in order to make the "re-identification" of people almost impossible, even after treatment.

A new ISO/TS 25237: 2008 technical specification is applicable to institutions that claim to be trustworthy for their national or transnational pseudonymization services. The specification will serve as a general guide for implementation, as well as for quality assurance purposes, and will help the user to determine the level of confidence that can be placed in the services provided. The scope of application of this standard concerns, but is not limited to, the following areas:

- Research or other secondary use of clinical data;
- Clinical trials and post-marketing surveillance;
- Public Health Monitoring and Evaluation;
- Confidential patient records (for example, adverse drug reactions);
- Comparative reports based on quality indicators;
- Peer reviews;
- Consumer groups.

This data is a major source of information that can be reused for research purposes. Researchers are today faced with the issue of opening up research data (Open Data) in an Open Access logic.

Definition standard profiles

The objective of defining standard profiles is to better understand the current situation within the facility and more particularly the people related to the problem defined in the Challenge, and who will therefore eventually use the innovative developed. These standard profiles will enable interested companies to apply to fully understand all the contours of the current situation in the facility.

2 "patient" standard profiles related to the Challenge have been described (the standard profile of the "problem" patient and the standard profile of the "easy" patient). Also, 2 "health professionals" standard profiles related to the Challenge are described (the standard profile of the specialised physician and the standard profile of a nurse, for example).

STANDARD PROFILE OF THE « Problem patient »	
Name	n/a
Age	n/a
Status	n/a
Location in relation to the institution	Intern at the facility
Occupation	n/a
Short biography in three points	n/a
Pathologies	Mood disorders and bipolar disorders
Potential disabilities	n/a
Motivations (recognition, power, reward…)	n/a
Objectives in relation to illness and treatment	n/a
Constraints and frustrations	n/a
Personality traits (introverted, logical, reflective…)	n/a
Understanding of the illness	n/a
Observance	n/a
Internet skills	n/a
Mobile telephone skills	n/a
Social networks skills	n/a

STANDARD PROFILE OF THE « Easy patient »	
Name	n/a
Age	n/a
Status	n/a
Location in relation to the institution	Intern at the facility
Occupation	n/a
Short biography in three points	n/a
Pathologies	Memory and speech disorders
Potential disabilities	n/a
Motivations (recognition, power, reward…)	n/a
Objectives in relation to illness and treatment	n/a
Constraints and frustrations	n/a
Personality traits (introverted, logical, reflective…)	n/a
Understanding of the illness	n/a
Observance	n/a
Internet skills	n/a
Mobile telephone skills	n/a
Social networks skills	n/a

STANDARD PROFILE OF THE « Health professional 1 »	
Position	Researcher
Name	n/a
Age	n/a
Status	n/a
Location in relation to the institution	n/a
Short biography in three points	n/a
Motivations (recognition, power, reward…)	n/a
Objectives in relation to illness and treatment	n/a
Constraints and frustrations	n/a
Personality traits (introverted, logical, reflective…)	n/a
Understanding of the illness	n/a
Internet skills	n/a
Mobile telephone skills	n/a
Social networks skills	n/a

STANDARD PROFILE OF THE « Health professional 2 »	
Position	Psychiatrist
Name	n/a
Age	n/a
Status	n/a
Location in relation to the institution	n/a
Short biography in three points	n/a
Motivations (recognition, power, reward…)	n/a
Objectives in relation to illness and treatment	n/a
Constraints and frustrations	n/a
Personality traits (introverted, logical, reflective…)	n/a
Understanding of the illness	n/a
Internet skills	n/a
Mobile telephone skills	n/a
Social networks skills	n/a

Definition uses

Uses are defined as the description of the current situation in the healthcare organisation facility from a particular point of view (patient or health professional, for example).

In order to have a perfect understanding of the current situation, it is important that each use be defined step-by-step, starting with step 1, which is the starting point (which can be, for example, the first consultation at the hospital for the patient, admission to the emergency room, return to the room after surgery) and finishing with stage X, which is the end point (the patient is completely treated, an end of remote monitoring ...). The time between the starting point and the end point represents the moment when the new digital solution must be used in the patient's treatment. At each defined step, it is important to highlight the current problems encountered and which must be solved by the use of the new co-developed digital solution. For each step, please list the current issues and what should be done to solve them.

Find below the description of the case

- Anonymization of patients when their case is presented in a scientific publication.
- Permanent monitoring of the evolution of diseases and thus refining studies.
- Establishment of epidemiological cohorts.

- Translational research.

Technical and operational requirements

Pseudonymization consists of deleting the directly identifying fields from the records, and adding to each record a new field, called PSEUDONYME, whose characteristic is that it must make any link between this new value and the real person impossible.

When creating this pseudonym, the hash function is very often used; it will be applied to one of the fields of identifiers (for example the social security number), which is a particular type of function that makes it impossible (or at least extremely difficult) to derive the initial value.

Thus, the two entities having information on the same person, (identified by its social security number), could share this data anonymously by hashing this identifier. There is also the possibility of having a random function to generate a unique identifier for each person.

Feasibility

The organization is the key to an anonymization project and conditions its success, but currently issues remain concerning the protection of the data:

- Should we erase unused data?
- Are our data protected by a cryptographic process?
- Are our security techniques adapted to the use of our data?
- Do we regularly exchange sensitive data?

Financial Aspects

N/A

Expected Impact

The research teams of the University Hospital Group continue to be benchmarks in clinical research in areas such as: major psychiatric disorders (psychosis, schizophrenia, depression, mood disorders), addictions, behavioural disorders neuro-physiology and experimental neuro-pathology, neuro-imaging, stroke and cognitive disorders related to ageing and neuro-oncology/epileptology.

Such a tool will provide a simplification for the data-scientist, which will save time with a high level of data quality for research projects, bringing a major evolution for science and real progress in the general field of neuroscience

Scope of the challenge

Information on the specificity of the challenge, if it is defined specific to the hospital

Yes, this Challenge is initially the ambition of the Paris University Hospital Group - "Psychiatry & Neurosciences" with a goal first of all to improve the description of the different pathologies of the brain that may occur.

The goal is prevention and care for patients, through its strategic vision of integrating basic and clinical research in the field of psychiatry and neuroscience.

The solution will then be put in place and mastered, so that the opening of anonymous data can now be possible and open to patients