HBP Innovation Market Analysis Series - UPM-2021

The Medical Informatics Platform (MIP): Analysis and options for its exploitation in the healthcare market

POLITÉCNICA

UNIVERSIDAD POLITÉCNICA DE MADRID

Human Brain Project

EBRAINS

September 2021, V1

Human Brain Project
UPM
Innovation Team

Duran T., León, G., Velasco, G., Strange, B.

ACKNOWLEDGMENTS

This report has been directed by the Innovation Team of Universidad Politécnica de Madrid (UPM) within the task 8.5 of SGA3.

With special thanks to the Human Brain Project, for providing a space for collaboration and knowledge generation that inspires neuroscientific continuity and multidisciplinary teams in developing research and solutions for the future. Also to the Medical Informatics Platform Team, for their kind support in providing us with the latest and most accurate information possible, enriching the present analysis.

This work has been funded by the EC under grant 945539 (HBP SGA3).

Project Number	945539	Project Title	HBP SGA3

Document Title:	The Medical Informatics Platform (MIP): Analysis and options for its exploitation in the healthcare market.
Document File Name:	MIP Exploitation Analysis
WP(s) / Task(s):	WP8, T8.5
WP Objective(s):	WPO8.1
Output(s):	OP8.19
Author(s):	Teris DURAN, Gonzalo LEON, Guillermo VELASCO, Bryan STRANGE (P68).
Compiled by:	Gonzalo LEON, Guillermo VELASCO, Bryan STRANGE (P68).
Abstract:	The MIP is an outstanding technology for training and analysis for the research clinician of tomorrow. It can be stated that is the convergence between the neuroscientific capacity, medical practice, and accelerated computer development, upshot of the effort of a multidisciplinary human resource. Integrating the exorbitant amount of data in health and organising it in a federated scheme with AI, has been an arduous and extraordinary task, currently presented as one of the EBRAINS research tools with technical potential to grow in the market, the main message of this report. Therefore, the main outputs have been identified in which the MIP stand out above other platforms, as well as those hospital softwares that can benefit directly or indirectly from its explorations. Another aspect to highlight is its current AI analytical tools, with useful algorithms and frequently used by the clinician. The current and forthcoming initiatives were described in a comparative scheme that allowed to define the limitations of other projects and combinations with competitive technologies such as blockchain. This led to identify in a practical way the participants in the industry, considering strategies towards their first stage as a startup, final decisions that will obviously depend on the leaders of the technology. We conclude with a SWOT analysis sharing final observations and some suggestions for business model and product innovation.
Keywords:	Brain health technologies, federated data platform, market analysis, strategic planning, artificial intelligence.
Target Users/Readers:	Non-experts interested in large-scale brain networks; investors interested in neurotechnologies and mental health; Experts looking for research and/or industrial actors in the field.

SUMMARY

Background: Lausanne University Hospital.
Copyright: SAM-CHUV

PART I.

CURRENT CONTEXT OF THE HEALTHCARE INDUSTRY AND THE MIP

MIP Medical Conditions:

Dementia*
TBI*
Epilepsy*
Mental health*
Stroke
Parkinson
Multiple Sclerosis
Depression
Coma

**Based federations.*
MIP Data Factory (2020)

The Medical Informatics Platform (MIP) was conceived as a federated data platform and advanced analytical techniques powered by AI, with the aim of providing the neuroscientific community and related professionals with a virtual workspace for experimentations, in order to improve some of the problems that arise in the prevention, diagnosis and treatment of neurodegenerative diseases, mainly. As indicated (Venetis et al, 2015): "**The overall results will be used for diagnosis, more accurate prognosis and new types of drug discovery for the development of new medicines**".

The healthcare industry needs neurotechnologies to overcome the current barriers in medicine, and furthermore we need physicians and researchers who believe in these technologies, those who at the end have a greater decision-making power as lead users. Starting from the high prevalence of brain diseases globally, neuro-degenerative conditions associated with ageing, misdiagnosis, insufficient specialists in mental health between regions, even the expenditure in healthcare by governments and its prevention programs; are only some of the main reasons that justify further progress in better solutions for brain disorders. Clearly, the current advances including AI do not have the capacity to intervene with total precision to overcome these conditions.

Importantly, it is necessary to have digital structures that provide reliable support to the medical service and reduce the incidence of misdiagnosis and mistreatment. In this respect, the Human Brain Project and EBRAINS, have developed a digital platform that combines the data of thousands of patients, a platform currently deployed in more than 31 hospitals in various parts of Europe, called the Medical Informatics Platform, described in an upcoming section. Some reflections have also been raised which will be described throughout this report:

- How can the MIP with cutting-edge technologies built into its platform and generated within the Human Brain Project **transform diagnostics** and healthcare as we know it?
- Can the MIP become an extraordinary and competitive global healthcare platform that is able to **exploit the information** considering the amount of data generated?
- What are the **advantages** and **disadvantages** of this model over the technologies generated? What **strategies** should be considered to be a competitive model?

The key concepts to be considered are highlighted, summarizing most of the information presented.

Also, complementary ideas are included to enrich the context.

- What **business models** can be generated from the MIP that can create new jobs and improve the quality of life of hundreds of thousands of people? What jobs could be threatened?

BRIEF SUMMARY OF THE MIP OPERATION

MIP federation
available records:
- TBI: 10,500.
- Mental Health: 7,300.
- Dementia: 6,348.
(EBRAINS, 2021)

The following diagram simplifies the current structure of the MIP (taken from the Human Brain Project portal). Evidently, the current report is not intended to repeat the information contained in manuals and current publications, but rather offer an integral and practical perspective about its function to reinforce the virtues and weaknesses of the platform, reflecting as far as possible some of the informatics systems and current systems that compete or can complement it, answer some of the concerns that arise to transfer them to the market, considering the opinions of possible lead and final customers, as well as identify and justify possible business models derived from it.

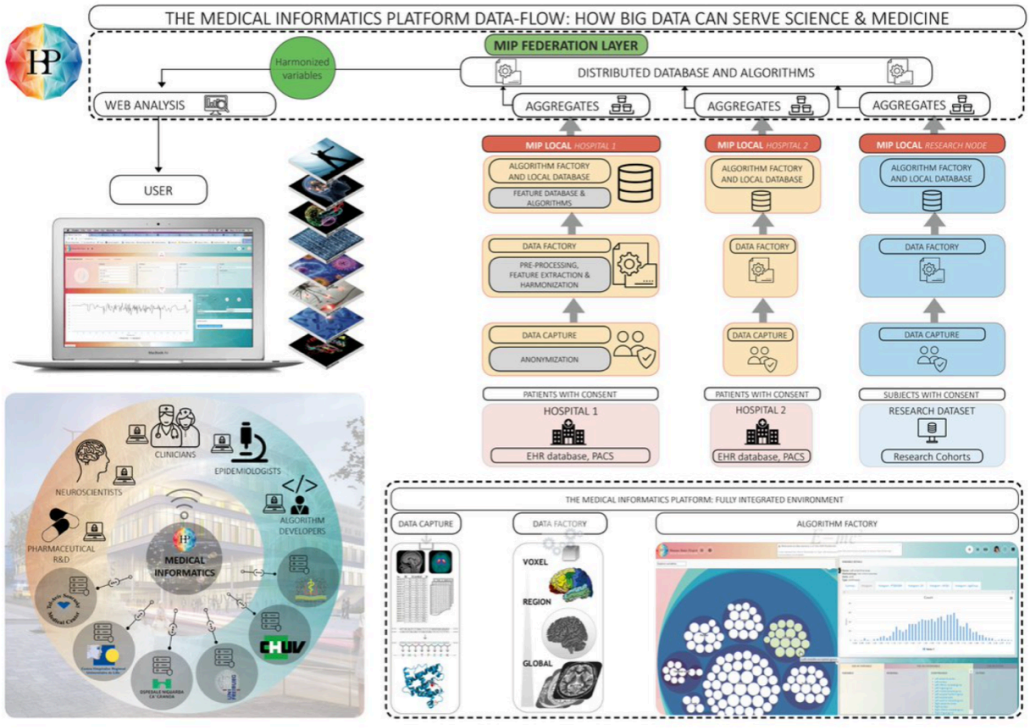


Figure 1. "How does MIP work?". Human Brain Project (2021)

From a general perspective, the MIP is structured in four main levels (Demiraj et al, 2019). At the closest level to the user is found the **front-end or the portal**, where the digital structure is displayed with the

Data types available in the **MIP:** Clinical information, regional brain volumes, intracerebral EEG (EBRAINS, 2021)

diverse functionalities and elements that allow exploration of data for the experiments. This segment includes a second level, the **MIP Federated**, which integrates the anonymised data from hospitals and centres and is, therefore, widely open and accessible for other clinicians who wish to perform experiments or research and also to have a structure that facilitates for example, their clinical trials. In the third level or **MIP local,** the data is managed by hospitals and centres in internal level, meaning only they have access, and the data is pseudonymised.

In this platform, specific algorithms are implemented as instruments of AI to perform predictions and experiments based on machine learning. Fifteen algorithms are so far included at the federated level[1] that are automatically enabled or not automatically in the same portal, depending on the variables selected during the explorations within the platform in an agile manner. For the local version, according to the development team, clinicians suggest to computer scientists personalise certain algorithms so that they can be adapted to specific examinations or diseases, which constitutes a powerful strategy for bridging the global gap between medical expertise, their needs and the latest advances in computer science.

The **MIP** allows to analyse hospital raw data such as images from MRIs, CSV files for genetics and clinical data. (Venetis and Vassalos, 2015)

At the fourth level is the **MIP Data Factory**, which is not viewable in the front-end but rather it is a facilitating tool for transferring certain data (Brain Scans, EHR, and Metadata mainly) and is managed by hospitals managers and centres, in order to make them compatible for the MIP, and as a base for this structure. These levels are also complemented by the **MIP Data Catalog**, which integrate the medical conditions that are handled in this platform, through Data Models for a local level, and with the **Common Data Elements (CDE)** in a federated form, which serve as reference and have been predefined by the federated community (Data Processing User Guide, 2020).

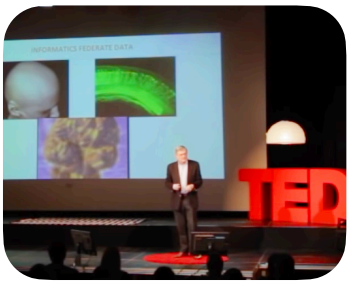
[1] https://github.com/madgik/exareme/tree/master/Exareme-Docker/src/mip-algorithms

THE MIP AND THE MAJOR HOSPITAL IT SYSTEMS

The **European FET** research was released in 1989 to support scientists, engineers and researchers from different professional areas, in a multidisciplinary initiative for the development of innovative ideas and projects in the technological, industrial and social fields in Europe.
The European Commission (2009)

MIP advances as they are currently presented would not have been possible without **the cohesive efforts of the scientific and multidisciplinary community at European level**, promoted by the European Commission in the Human Brain Project (HBP), as Future and Emerging Technologies (FET), the latter approved by the Competitiveness Council since 2009 and formally initiated four years later (UPM, 2011). In the interest of taking a step toward greater integration of the fragmented efforts of neuroscientists and, therefore, empowering them with computational tools that would allow them to perform brain simulations (Markram et al, 2011; Amunts et al, 2016), the initiative aimed at integrating data from the International Neuroinformatics Coordinating Facility (INCF) initially, with programmes that already stored data and scientific resources in order to provide them under a federated form (Ibid, 2011). Currently the platform stand out as part of the services offered by EBRAINS Research Infrastructure.

Among other reasons, it was desired to provide researchers with tools to simulate brain functions, redefine models where the application of specific data is needed, together with the possibility of testing those hypotheses contingent on a certain technological capacity. Thus, the platform is born from the need to transform the vast amount of patient data and their different diseases, in order to make them available for medical research, to achieve better diagnoses and treatments, combining the latest advances in the field of computer science and medicine. Since then, **the platform has evolved as a result of continuous changes and structured plans at different levels**, coupling with the hospitals' IT systems -regarding the nature of the different medical equipment-, to the requests of healthcare staff, in addition to compliance with data protection and patient privacy regulations.



Frackowiak, R. (2017), Future Medicine: Modern Informatics. TEDxYouth@Zurich.

In the words of Professor Frackowiack, neurologist and one of the MIP founders, in a talk about brain diseases and the moderns informatics: "The real thing that's missing is the impact of computers and informatics on this", and evidently improve them. In this sense, the MIP could find new possibilities in meeting this challenge, however, it has involved a laudable effort in collaboration with hospitals and

scientific teams, researchers and epidemiologists; naturally with the aim of providing accurate diagnoses, improving the quality of life of patients and their families, as well as reducing the costs associated with health services.

In recent years **Deep Learning (DL)** algorithms have demonstrated their enormous potential especially in genomics, for the ease of processing specific and relevant information with endless possibilities in prediction models.

Therefore, the main challenge of the platform is to provide a comprehensive interface with the ability to combine patient data such as information in EHR, laboratory tests, mainly MRI brain scans, as well as their features, demographic data (Demiraj et al, 2019), equipped with the latest tools in artificial intelligence and big data at the service of the neuroscientific community. Some of the virtues of this platform is to facilitate the integration of this data in a federated form -that is basically, data shared between different locations- anonymously, harmonised, complying with its regulations (some will be addressed in this report), as well as employing **the most advanced ML and DL algorithms** for the deployment of experiments with the highest possible reliability in their relationships and predictions.

Image: Freepik

According to recent estimates (Huesch and Mosher, 2017), a single patient can generate an average of **80 megabytes** each year in imaging and EMR.

To make a future of accurate, personalised diagnostics a reality and further transform the medical service as we know it today, numerous global initiatives and software have been developed, some with greater specialisation than others. In addition, the volume of data from the health sector is growing exponentially: By 2013 it was estimated that by 2020 the volume would reach about 2,314 exabytes[2], with an overall increase rate of 48% per year (Zwolenski and Weatherill, 2014, mentioned by Stanford Medicine, 2017). This urges not only capturing more specific data, but making it a useful resource.

On the other hand, various published articles explain the workings of the platform and what are the dynamics that make it particularly different. This report addresses them from a multidisciplinary perspective, that is, in addition to the analysis of the technology created, the intention is to contribute to its scope from a business point of view which, together with the vision of medical specialists inside and outside the project, could enrich its journey towards the continuous and exciting digital transformation of the healthcare sector.

[2] To understand the current units: Megabyte (10^6), Gigabyte (10^9), Terabyte (10^12), Petabyte (10^15), Exabytes (10^18), Zettabyte (10^21) and Yottabyte (10^24). Tavazzi, 2019.

SYSTEMS USUALLY FOUND IN HOSPITALS CLOSER TO THE PLATFORM

The figure 2 shows the basic structure of software and systems usually applied in hospitals (it is understood that the image may contain higher specifications, however those with the most technical closeness to the MIP are shown). This schematic is useful because it summarises the current context before explaining the dynamics of the platform and explains how some of its internal systems relate, how it fits into the current structure, as well as comparing the basic dynamics with some similar technologies.

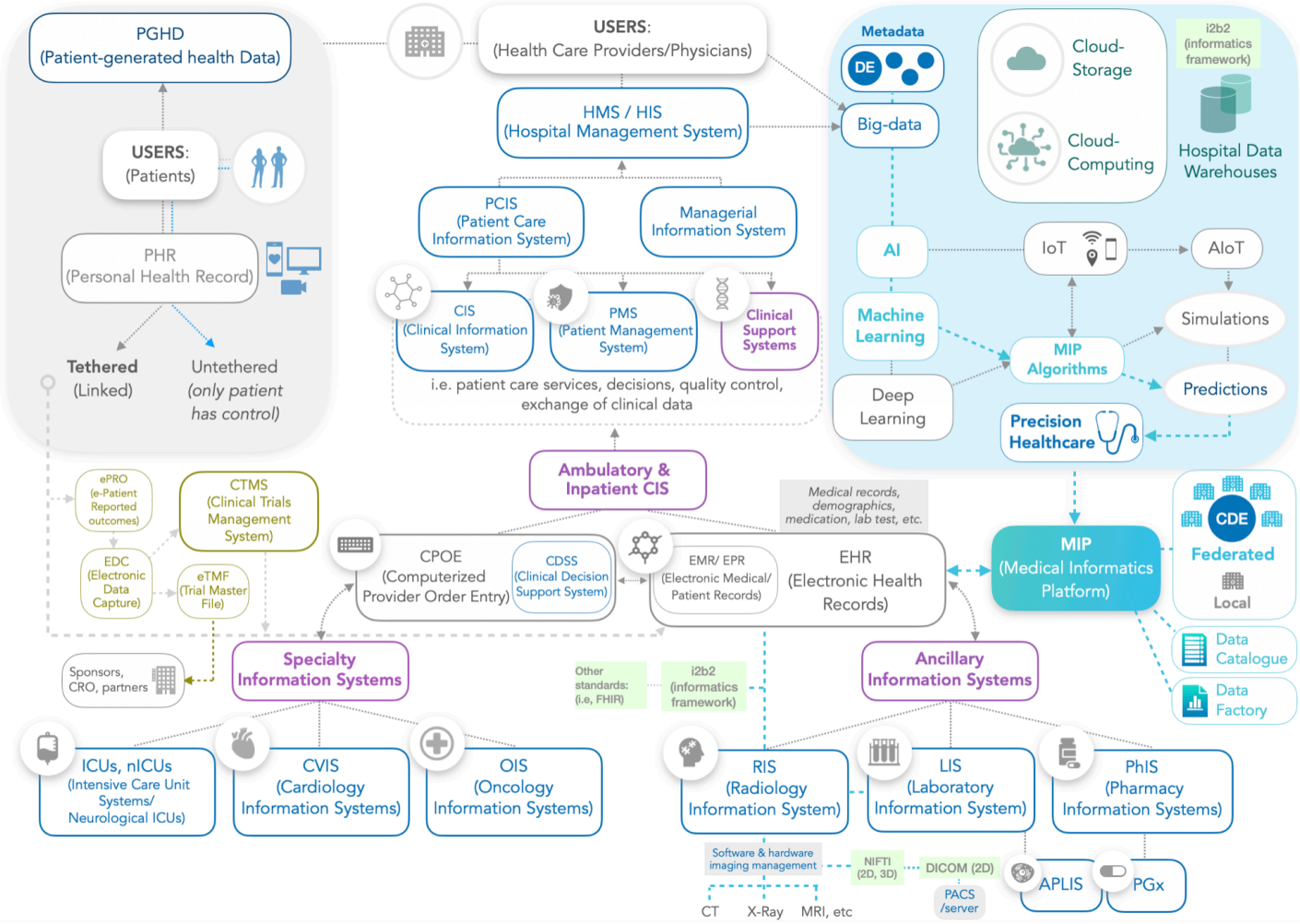


Figure 2. Some of the systems found in hospitals according to literature (described below) and the MIP

The Healthcare Providers and Physicians are the initial focus and represent one of the lead users of the MIP, they clearly have a greater prominence and control of the health system. Given the complexity of the administrative and operational system of health care centres and

Image: Freepik -DCStudio

Hospital **I**nformation **S**ystem is defined as the **sociotechnical subsystem** of a hospital -e.g. enterprise functions, business processes, application components, and physical data processing components-. (Haux et al, 2004)

their communication networks between different departments, **increasingly sophisticated tools have been required to better control and organise the service offered to patients.** Nevertheless, establishing interconnectivity and harmonisation throughout its structure is a challenge for health workers, managers, as well as creators of new technologies.

Therefore, applications such as hospital management systems (HMS) are needed that facilitate the registration and control of many of the activities that are carried out in the patient evaluation (Tiwari et al, 2019). Variants of this type of software are often diverse, ranging from data integration to facilitate predictive analysis in research, data storage, monitoring of the available workforce, linkage with specific departments as Laboratory Information Systems (LIS), and account and invoice management (Gupta and Niranajan, 2020); also they can be found in literature as HMS or Hospital Information System (HIS). However, in practice they have multiple limitations such as not being able to transfer medical records in specific formats, lack of integration of new modules within hospitals, availability of specialised technical staff, for example, (ibid); which requires these systems to continuously update and integrate new modules meeting the ever-deeper needs of medical specialties.

Subsequently, as part of HIS, are found the Patient Care Information Systems (PCIS), which are composed of different subsystems, including the Clinical Information System (CIS), Patient Management System (PMS), as well as the Clinical Support System (CDSS) (Winter, 2010, mentioned by Kazancigil, 2019). The CIS, integrate the set of techniques that support the needs of the medical service, including hospitalisation, patient management and disease prevention activities (Combi and Pozzi, 2019). It also supports the integration of **Electronic Health Records (EHR)** with the Computerised Order Entry (CPOE), defined later. In addition, it manages a large amount of data that is also used in the field of big-data and AI, as well as interconnectivity with other internal systems no less important as laboratories, radiology, pharmacy, among others (Islam et al, 2018); therefore, it can be deduced that there is a continuous flow of data between the main systems including those forming part of the Clinical Support Systems, in purple- that compose the different subsets described in the literature.

***P**atient **M**anagement **S**ystem* allows integration of clinical applications and administrative application components. (Haux et al, 2004)

Regarding Patient Management Software (PMS), these are programms that link databases to patient registry management which, as a complement to CIS, also record the consultation with specialists' doctors associated to the hospitals, and therefore it should also be linked with HMS/HIS (Degoulet and Fieschi, 2012), the upper structure in the diagram. PMS can be found in different modalities such as CPOE, PACS and even the EMR (Weber-Jahnke, 2011) and regulated as medical devices[3], however, these have shown relevant weaknesses that could hardly been published given certain confidentiality clauses in software agreements, as well as not being mandatory to request reports on failures in their systems.

This study describes -among other aspects- the main causes of failure of these systems, the seven main ones being: **"Misidentification, end-user customisation, Human-Computer Interaction (HCI) problems**, missing warnings, failures caused due to software upgrades, incorrect reporting, and interoperability failures (…)" (Ibid, pp. 7). Some of them could be improved by the current advances of the MIP given the rigorous processes that are part of the platform, both for the transfer of data in its different stages and in its subsequent management explained later.

Recommended **C**linical **D**ecision **S**upport **S**ystems are those that integrate clinical information systems and computerized provider order entry. (Wright et al, 2009)

Continuing with some of the hospitals systems shown in diagram 2, the Computerised Provider Order Entry (CPOE) are systems that have enabled hospitals and healthcare facilities to reduce errors in patient prescribing, reduce repetition of activities among staff, and cut down on the need to hire staff to meet objectives (Ohsfeldt et al, 2005). It is therefore the management system that replaces traditional methods of prescription on paper, telephone, etc. (AHRQ, w.d); often integrated with PMS (HealthIT, 2018), or even improving the quality of medical service by complementing it with the Clinical Decision Support Systems (CDSS), the latter as web applications that amplify medical expertise by combining digital patient records with the power and agility of computing, facilitating staff decisions and may or may not be integrated with EHRs (Sutton et al, 2020).

In addition, all patient information acquired at these centres should be integrated into Electronic Medical Records (EMR) internally, that is, not interchangeable between departments defined as intra-organisational (Heart et al, 2016), and constitutes the main source of

[3] Require the ISO 13485 certification and maintain a quality management system, Weber-Jahnke (2011).

Electronic Health Records (EHRs) (Garets and Davis, 2006). Hospitals that have these systems have made a gradual effort to integrate the EMR as part of EHR, (Aldosari et al, 2017; Gibbins and Wickramasinghe, 2018), however, in both cases, the information extracted is diverse and includes demographics, medications, diagnostics, laboratory data, x-rays, among others (Heart et al, 2016; Demiraj, 2019), **fundamental components of the MIP**.

One of the most MIP-related systems is the **Radiology Information System (RIS)** also managed through software and hardware, usually through servers as **Picture Archiving and Communication Systems (PACS).** The latter allows the storage and sharing of assessments by images evaluation between specialists (Garets and Davis, 2006). It is necessary to specify the usual formats of diagnostics per image considering the number of scanners and variants of use between centres, as well as to understand whether this represents a limitation or an advantage for the MIP.

MAIN VULNERABILITIES OF THE CURRENT SYSTEMS

However, one aspect to highlight is the vulnerability of PACS and Medical Imaging to cyberattacks, as their systems have maintained similar formats for a long time, do not have protections to deal with new and more complex and ineludible digital threats. Sometimes they carry malware in images that are downloaded, as well as in communications or reports, others involve intentional image manipulation compromising the correct patient's diagnosis, or even requesting rewards with cryptocurrencies as a threat to hospital systems (Eichelberg et al, 2020). Even with current protocols to prevent these scenarios globally, cyberattacks continue to prevail: Only in the United States 4,000 ransomware attacks occurred daily in different sectors in 2016, being the healthcare industry the third most affected worldwide (Argaw et al, 2019). Additionally, in 2019, 66% of healthcare organisations suffered from ransomware attacks and 45% of organisations paid the ransom (ENISA, 2020).

Although it may seem a truism, they expose the informatics systems of the healthcare services, increasingly necessary, as well as in health emergency scenarios, affecting the privacy and treatments of patients in medium and long term. As an example, by 2020, given the health crisis caused by the virus COVID-19, cybercriminals took advantage of

Ransomware is malicious code or a form of malware and is most commonly used to attack healthcare systems. The different types of malware according to self-replication could be: Worms (informal), viruses (require a host for replication) and malware-carrying (Kramer and Bradfield, 2010).

Image: Adobe Stock

the vulnerability of healthcare systems distributing ransomware into the internal infrastructure of hospitals, and other no less important cybersecurity cases were also reported worldwide by INTERPOL (2020): "two-thirds of member countries from Europe reported a significant increase in malicious domains registered with the key words 'COVID' or 'Corona'" (pp. 7). Other methods such as phishing [4](León, 2020) also were detected in European countries, which is evidence of a significant vulnerability of healthcare systems.

Ransomware Best Practices: "The processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information (…)". Article 4, (5). General Data Protection Regulation (GDPR, 2016).

Once the ransomware has invaded the system, users have the following options (Sittig and Singh, 2016): "1) try to restore their data from backup; 2) pay the ransom; or 3) lose their data". In these critical scenarios, the CISA[5], FBI and HHS, also provide a large number of Ransomware Best Practices, some of them recommend not paying the requested ransoms, among other reasons because it does not guarantee that data will be recovered, also to have data backup, training healthcare workers in this type of situation, etc. In this sense, **the MIP could be useful as an additional complement to hospitals to safeguard the patient data** (because independently, protection against cyber-attacks must be stablished in hospitals) -*although pseudonymised in the MIP local server*- and discourage as far as possible to cybercriminals that ask for rewards in possible attacks by freezing the systems and patient data in return of unlocking them.

Four transfer stages within the platform: Data Factory, Local Layer, Federation Layer and the MIP portal.

Still, the MIP evidently has no possibility of avoiding the degree of vulnerability in DICOM or PACS from its origin -because it is a fragility of these systems, facing increasingly sophisticated attacks (Mirsky, 2019, mentioned by Eichelberg et al, 2020)-, instead has the possibility to evaluate in advance the nature of the data entered. This means that the platform struggles in offering an evaluation and protection of the information already managed by hospitals and centres, in its four transfer stages and identified as follows (in a bottom-up perspective): **MIP Data Factory, Local Layer, Federation Layer and the MIP portal** (Demiraj et al, 2019).

Therefore, from the initial moment of data transfer or in MIP Data Factory, there is a pseudonymisation process, together with a subsequent distribution and classification of the content according to

[4] Through *phishing* private information can be obtained from individuals with fake messages and usurping the identity, registering a strong increase during the pandemic months (León, 2020).

[5] Cybersecurity and Infrastructure Security Agency (CISA, 2020).

MipMap: A tool used for designing the mapping tasks that are going to be used in the Schema mapping. MIP Deployment and Support Team (2020).

diseases catalogued and available in the platform through a tool called MIPMap. The data entered comes from Brain Scan (MRI in DICOM generally), EHR and Metadata mainly (Ibid, 2019); while those that do not comply with the quality standards requirements - e.g. by inconsistency of data from origin- are rejected, allowing a better quality data and metadata available in the whole structure.

In general, image diagnostics are usually presented in DICOM, NIFTI, MINC, Analyse, NRRD, PAR/REC format; the most common are the first four (Larobina and Murino, 2014). They have multiple differences between them ranging from the image dimension –2D, 3D- to supported pixel levels. Thus, NIFTI and DICOM formats tend to be more widely accepted, however, NIFTI is a format that supports 3D and 2D, unlike DICOM, for example. **For the MIP they need to be switched to LORIS[6] first, then NIFTI format**, which implies a higher demand on the quality of data entered as mentioned above, through the Quality Control Tool (QCT) available for the hospitals and centres, which also enable their harmonisation.

Another detail to consider is that, in the same form in which Health Care Providers generate data, patients or the final users also generate data, commonly known as Patient Generated Health Data (PGHD) -simplified in the upper left box in green in the diagram 2-. The latter generate data in a tethered or untethered form, depending of the use they make with their applications or the relationship that they have with the healthcare provider. For example, for a patient with epilepsy for whom the use of a wireless device for seizure detection and prediction is recommended– a hypothetical example- these devices have sensors whose interface is connected to the hospital's software, therefore, would be the path marked in grey that feeds the EHR or even to PMS.

PGHD: All information related to history, symptoms, biometric data, treatments, etc. obtained from several sources for example patient portal, wireless blood pressure, smartphone applications and other digital devices and sensors. (Demiris et al, 2019)

In this sense, considering that PGHD is collected by sensors, data inputs, and other external platforms as social networks, for example, **the MIP has the potential to clarify the differential diagnosis on receiving data** (not real-time data but synthetic data for now). This process could be an alternative utility for the platform; however, the current dynamic of the platform is marked by the RIS, EHR and Metadata, mainly.

[6] https://github.com/aces/Loris/tree/main

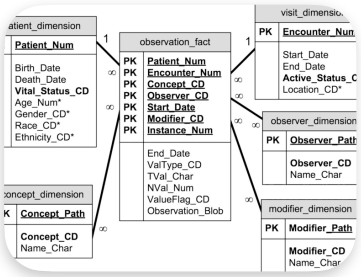


Image: i2b2 Research Data Warehouse

On the other hand, in the upper right blue box of the figure 2 is the informatics system, whose transfer of data is initially given by a process of harmonisation or standarisation created years ago as medical industry initiatives, updated periodically and shared between different countries, in order to homogenise the data or at least establish a bridge that allows the exchange of data between the centres, making them more universal. These include the i2b2, one of the clinical data warehouses widely used or the recent FHIR framework, which are the small fluorescent green squares in the diagram 2. The MIP uses the i2b2 specifically and also allows to apply AI tools for the algorithms. In this sense, **we highlight the singularity of the MIP with respect to other platforms in addition to its inherent ability to offer data in a federated way**, the development of ML and DL algorithms that are automatically enabled according to the variables and metadata chosen for the models.

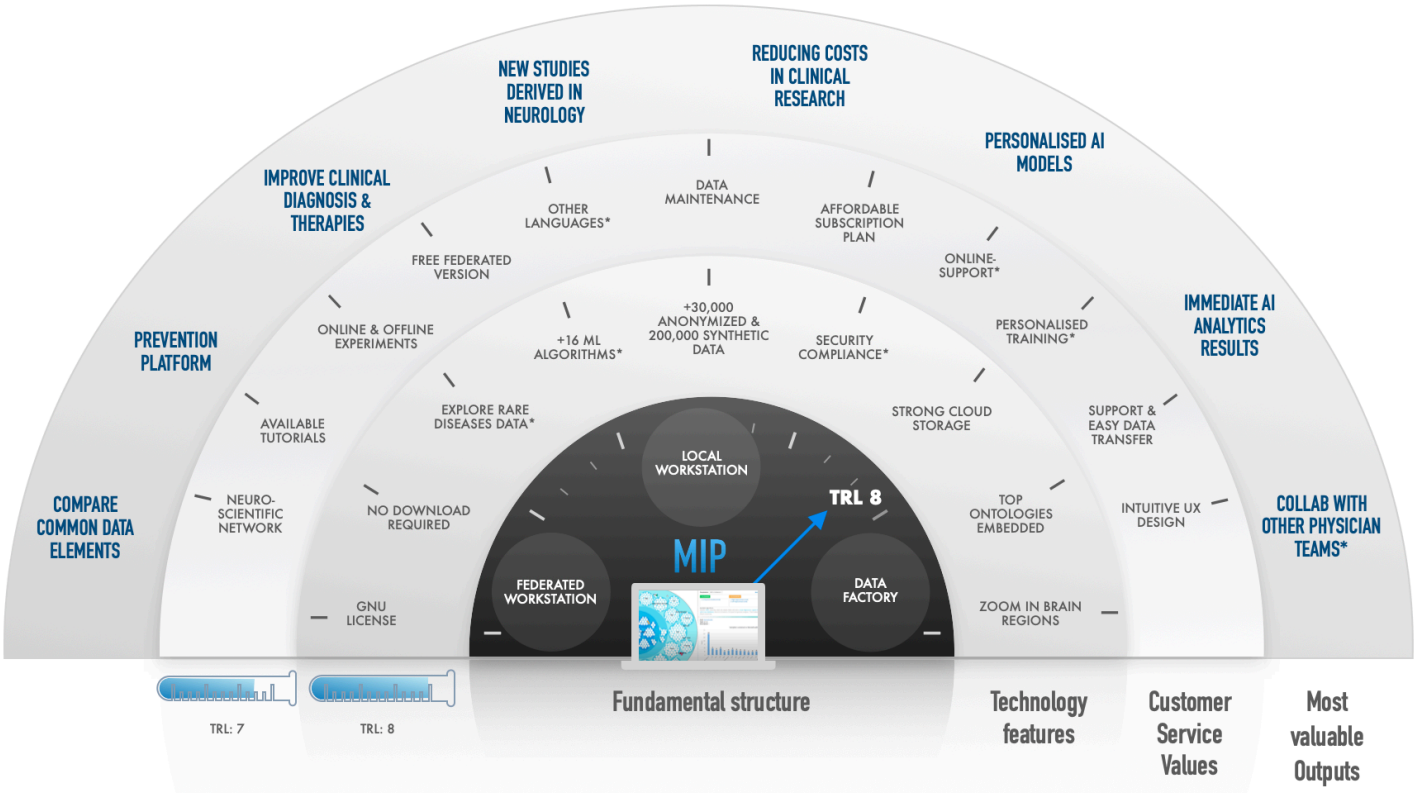


Figure 3. MIP identified outputs.

POSSIBLE INFLUENCES OF THE MIP ON HIS

In recent years, a large number of studies describing the inter-operability of hospitals systems and their functions have been

published. At a technical level they become increasingly complex due to the specialisation achieved, requiring a specialised personnel in medical and informatics areas, which has contributed to a gap in knowledge and its practical application, in addition to the technical limitations and the differences presented by the systems of each centre. At present, the MIP continues to be deployed in several hospitals in Europe mainly, adapting to the requirements of the medical service, so **it is therefore premature to state that the performance of these systems is currently dependent on this technology**. Which also does not imply a withdrawal from being applied in conjunction with other software, in fact, it is the platform that has been adapted to the request of the staff and the varied data formats in hospitals, who will be the protagonists in the design of new studies, for which it would be in the best interests of the medical community and their current creators to find needs that can be met reciprocally in terms of data and AI models.

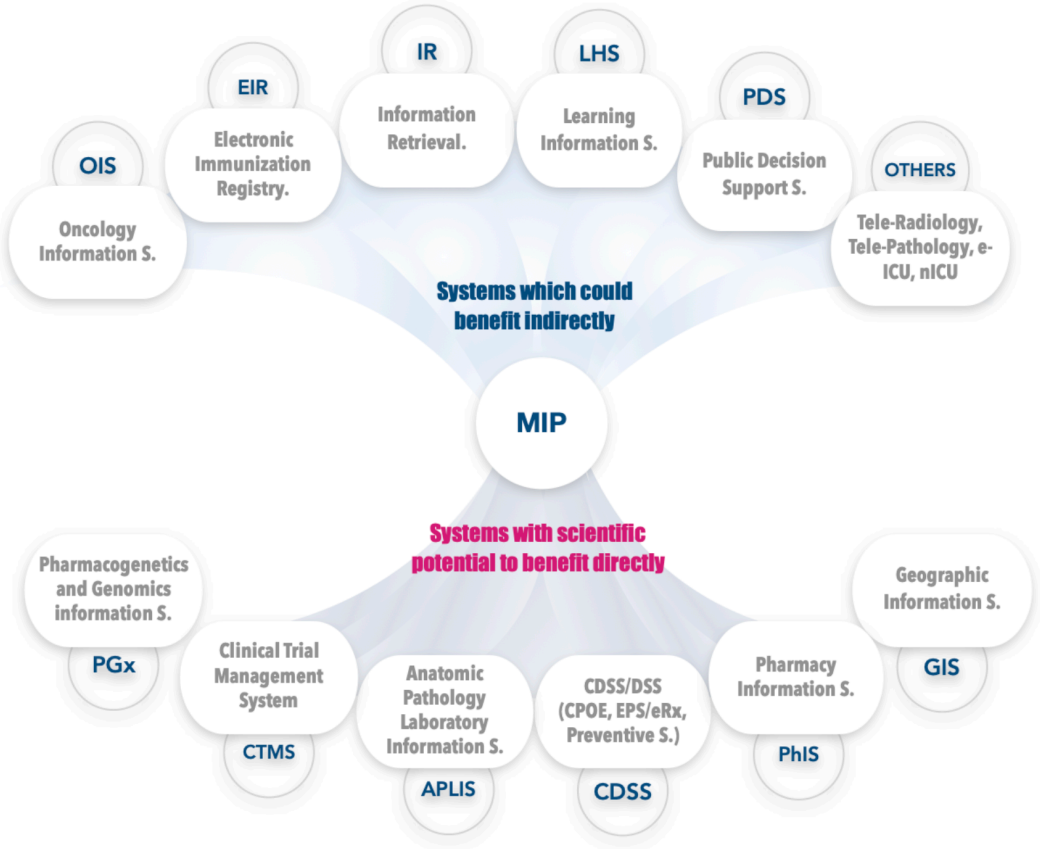


Figure 4. Possible influence of the platform on other relevant HIS *(in addition to the current ones: EHR and imaging systems)*

In short, what difference could the use of this technology make considering the current computer systems of hospitals?

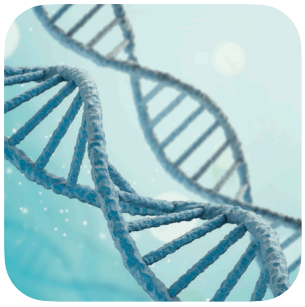
In a scenario in which the participation of hospitals and their research centres can be increased, promoting continuous use, **the systems that could benefit directly**, i.e, those in which the results could potentially be influenced by the MIP are as follows: Clinical Decision Support Systems (CDSS) *-including DSS, CPOE, Electronic Prescription Systems (EPS or eRx), Preventive Systems-*, Pharmaco-genetics and Genomics Information Systems (PGx), Pathology Information Systems (APLIS), Pharmacy Information Systems (PhIS), Geographic Information Systems (GIS), in addition to the systems already used by the platform to obtain the data (EHR and imaging systems mainly). Whereas those **systems which can benefit indirectly** or using as a reference the AI techniques to complement their studies are as follows: Oncology Information Systems (OIS), Electronic Immunisation Registry (EIR), Information Retrieval (IR), Learning Information System (LIS), Public Decision Support System (PDSS), Others (Tele-Radiology/Tele-Pathology, e-ICU, neurological intensive care unit -nICU-), as reflected in the diagram 4. New studies with more specified perspectives for each specialty could be expanded with the use of the MIP, however, potential benefits and their relationships with available software are simplified:

- *Increased safety and prediction in future diagnostics and therapies, this relates to systems:* **EHR, CDSS *-including DSS, CPOE, EPS, Preventive Systems-*,** mainly. Although they are specially designed to facilitate medical decisions with a quick access to digital records and try to reach an intervention to the patient as timely as possible, it is not excluded possibility of errors in the diagnosis or omission of information either by system limitations or greater communication among staff. In the neurological field these diagnoses can be more accurate with the use of the platform, allowing the integration of clinical data from the attached hospitals, along with analytical tools to identify biological components and the prediction of functions at the brain level (Karozos et al, 2018). With an easy-to-use variable deployment system, it allows the clinician a flexible workspace that complements the functions of the usual programs.

  Another aspect to highlight is related to the alarms or continuous warnings in the CDSS, widely analysed for the impact they have on the medical service (Kesselheim et al, 2011). On the one hand, the companies that provide them cautiously evaluate whether or not they can reduce alarms, which are sometimes repetitive, in order to avoid possible failures and serious damage to the patient that may be attributable to

their systems, on the other hand, the overload of warnings that the medical staff may be exposed during the entire therapy with the same patient. In a nutshell, it is called *cognitive overload (Ancker et al, 2017)*, the fatigue caused by receiving a large amount of information in a very short time, affecting medical personnel's ability to discern between information that may or may not be relevant. In this sense, considering alternatives that facilitate medical decisions given the overload of information, virtual experiments within the MIP can provide **alternative support in data analysis quickly, reinforcing current therapies of promoting new opportunities in healthcare**, increasing the effectiveness of results and reducing the incidence of errors.

Image: The U-PGx Consortium



- *Evaluate drug effects and promote advances in gene therapies virtually*: Up to now, the platform has the ability to relate a limited variety of medications to a set of neurological conditions, therefore in addition to expanding the range of its components, it also contemplates the possibility of linking them with the genetic characteristics of patients, in systems such as **Pharmacogenetics and Genomics Information Systems (PGx).** Although doctors have pharmacogenomics results available in some patients, it is difficult to apply given the in-depth knowledge in genomics and the priority in the choice of therapies based on evidence rather than on these results. This makes it harder to introduce PGx into the clinical routine as a differentiated system, thus in many cases the PGx have been included into EHR to facilitate CDS, which integrates these details automatically (Hicks et al, 2016). In this sense, the role of the MIP **would contribute to new perspectives for the integration of these components and design therapies tailored to groups of patients with similar phenotypes**. Another aspect that could be considered is **the inclusion of the doses supplied and the availability of their drugs** in the simulation that, in conjunction with the deployment of their algorithms, would offer better predictions in current supplies and the design of alternative or novel treatments, so we also consider it would be of support for systems such as **Pharmacy Information Systems (PhIS).**

In Europe relevant initiatives such as **U-PGx** are still in progress in order to contribute to their application at the clinical level. See https://upgx.eu

In relation to these treatments, is widely known to neurologists the limitation of Blood-Brain-Barrier (BBB) and the delivery of current drugs. This has strongly led physicians and data scientists to develop methods that move toward more effective therapies, one of them through AI. ML algorithms such as Support Vector Machine (SVM), Decision Tree (DT) and K-Nearest Neighbor (KNN) are commonly used to search for the best drugs considering these barriers. However, despite advances in analytical tools, an optimal range of accuracy has not yet been achieved. This has

led to the application of algorithms in order to save costs and time in finding the ideal drugs and more recently these results have been surpassed by new DL techniques (Miao et al, 2019). Therefore, the MIP with its algorithms available in both ML and DL, in addition to predicting effects of drugs in response to their indications could **include the components of drugs available for experimentations in the face of these barriers,** moving towards new scenarios to analyse the relationships between chemical components at the molecular level along with the clinical phenotypes of patients already available[7].

- The branch of pathology -including neuropathology- has shown a major transformation in recent years, already known as his third revolution (Salto-Tellez et al, 2018) and powered by AI analytics tools. Especially DL algorithms have increased the level of accuracy and the identification of elements that are normally difficult to perceive by specialists, contained in data on brain structures or on digital pathology slides images, thus systems such as the **Anatomic Pathology Laboratory Information Systems (APLIS)** and the MIP, could be complement each other.

- Other systems such as ***Geographical Information Systems (GIS)*** can rely on the MIP, considering the incidence and impact of neurological conditions by location and the ability to analyse big data, we believe that it could also indirectly influence better planning in systems such as the **Electronic Immunisation Registry (EIR)** and at the governmental level by the **Public Decision Support System (PDSS)**, of utmost relevance considering the differences in medical services, its technologies as well as the availability of neurologists.

[7] Some of these phenotypic features included in the Referential Ontology Hub for Applications within Neuroscience, specifically in Human Disease Ontology (ROHAN); https://rohan.scai.fraunhofer.de/ols/ontologies/doid

Supervised: The algorithm is trained with labeled data.

Unsupervised: The algorithm is trained with unlabeled data.

Semi-supervised: Unknown labeled data and some supervision information. (Chapelle et al, 2009)

STRATEGIC DESIGN IN THE MIP ALGORITHMS

The following table (N°1) simplifies the analytical tools available in the MIP platform, together with a brief description, use, origin, purpose, type (S) Supervised, (UN) Unsupervised, (SS) Semi-supervised, mainly. All of this towards a practical vision that relates its algorithms both with the main computer systems of hospitals (discussed in the previous section), describing its main tools to new researchers and reaffirming its innovation potential from the infrastructure to the deployment of new and improved AI techniques in neurological research.

Table 1. An up-close look at the federated MIP algorithms

	MIP Algorithm:	(*) Based on/ Proceeding from:	Method & Main purpose:	Algo-rithm Type:	Examples of HIS and other systems closer to neurological care that can use it (*within the platform*):	Some research techniques in neuroscience that may use it: *(Enlarged in the figure 5)*
1	**Calibration Belt**	Shows expected to observed probabilities across classes of risk. Evaluate quality care in ICU, mortality prediction / Proceeding from: Finazzi et al, (2011) Bergamo hospital	ML-*Regression*	S	HIS, PCIS: CDSS (CPOE, eRx/EPS, including preventive systems), ICUs/nICU	Structural techniques; functional techniques, manipulation of neural activity
2	**Classification and Regression Trees (CART)**	Creates binary decision tree. Predicts values of continuous variables / Proceeding from: Breiman, Friedman, Olshen & Stone (1984)	ML-*Classification and regression*	S	HIS, PCIS: CDSS (CPOE, eRx/EPS, including preventive systems), OIS, ICUs/ nICU	Structural techniques; functional techniques, manipulation of neural activity
3	**Iterative Dichotomizer 3- (ID3)**	Creates decision tree. Allows to predict outcomes, also for clinical decision / Proceeding from: Quinlann (1986)	ML-*Classification*	S	HIS, EHR, PCIS: CDSS (CPOE, eRx/EPS, including preventive systems), OIS, RIS, Tele-Radiology, Tele-pathology, APLIS, EIR, ICUs/nICU	Structural techniques; functional techniques, manipulation of neural activity
4	**Kaplan-Meier Estimator- (KM)**	Estimates the survival distribution function from lifetime data (in a particular time)/ Proceeding from: Kaplan, E.L. & Meier, P. (1958)	Statistical method used for ML-*Estimation*	S	HIS, EHR, PCIS: CDSS (CPOE, eRx/EPS, including preventive systems), RIS, Tele-Radiology, EIR, ICUs/ nICU, CTMS: EDC	Structural techniques; functional techniques, manipulation of neural activity

	MIP Algorithm:	(*) Based on/ Proceeding from:	Method & Main purpose:	Algo-rithm Type:	Examples of HIS and other systems closer to neurological care that can use it (*within the platform*):	Some research techniques in neuroscience that may use it: (*Enlarged in the figure 5*)
5	**k-Means Clustering**	Partition data into k clusters, categorize unlabelled data and find groups represented by k/ Proceeding from: Stuart Lloyd (1957)	ML-*Clustering*	UN	HIS, EHR, PCIS: CDSS (CPOE, eRx/EPS, including preventive systems), RIS, Tele-Radiology, PGHD, PDS, LHS, GIS, PGx, PhIS, CTMS: EDC	Structural techniques; functional techniques, specific brain region/ cognitive, manipulation of neural activity; genes and proteins identification; isolation of DNA fragments; gene delivery strategies
6	**Linear Regression**	Statistical model for analyze the relationship between variables: predictors (x) and one or more outcome variables (y)/ Proceeding from: F. Galton (1894)	Statistical method used for ML (Glm)-*Regression*	S	HIS, EHR, PCIS: CDSS (CPOE, eRx/EPS, including preventive systems), OIS, RIS, Tele-Radiology, Tele-Pathology, APLIS, PGx, PhIS, eICU/nICU, PGHD, PDS, LHS, GIS, CTMS: EDC	Structural techniques; functional techniques, specific brain region/ cognitive, electrophysiological techniques, manipulation of neural activity; genes and proteins identification; isolation of DNA fragments; gene delivery strategies
7	**Logistic Regression**	Combination of x-variables to predict the risk of a medical event/ Proceeding from: Berkson (1944)	Statistical method used for ML-*Classifica-tion*	S	HIS, EHR, PCIS: CDSS (CPOE, eRx/EPS, including preventive systems), OIS, RIS, Tele-Radiology, Tele-Pathology, APLIS, PGx, PhIS, eICU/nICU, PGHD, PDS, LHS, GIS, CTMS: EDC	Structural techniques; functional techniques, specific brain region/ cognitive, electrophysiological techniques, manipulation of neural activity; genes and proteins identification; isolation of DNA fragments; gene delivery strategies
8	**Naive Bayes Training**	Analytical model using Bayes' theorem to predict, applying probability/Proceeding from: T. Bayes (1763)	ML-*Classifica-tion*	S	HIS, EHR, PCIS: CDSS (CPOE, eRx/EPS, preventive systems), OIS, RIS, Tele-Radiology, Tele-Pathology, APLIS, PGx, PhIS, ICU/nICU	Structural techniques; functional techniques, electrophysiological techniques, manipulation of neuronal activity
9	**Naive Bayes with Cross Validation**	Analytical model using Bayes' theorem to predict, applying probability / Proceeding from: T. Bayes (1763)	ML-*Classifica-tion*	S	HIS, EHR, PCIS: CDSS (CPOE, eRx/EPS, preventive systems), OIS, RIS, Tele-Radiology, Tele-Pathology, APLIS, PGx, PhIS, ICU/nICU	Same as previous

	MIP Algorithm:	(*) Based on/ Proceeding from:	Method & Main purpose:	Algo-rithm Type:	Examples of HIS and other systems closer to neurological care that can use it (*within the platform*):	Some research techniques in neuroscience that may use it: (*Enlarged in the figure 5*)
10	**Principal Components Analysis (PCA)**	Reduce dimensionality in large datasets/ Proceeding from: Hotelling (1933)	Statistical method used for ML- *Dimensionality Reduction*	UN	HIS, EHR, PCIS: CDSS (CPOE, eRx/EPS, including preventive systems), RIS, Tele-radiology, PGx, PhIS, eICU/nICU, CTMS: EDC	Structural techniques; functional techniques; specific brain region-cognitive function techniques; electrophysiological techniques; genes and proteins identification; *molecular cloning and recombinant DNA technology;* gene delivery strategies
11	Pearson Correlation	Reflects Pearson correlation coefficient between two vectors/ Proceeding from: Pearson (1896)	Statistical method used for ML- *Classification*	S	HIS, EHR, PCIS: CDSS (CPOE, eRx/EPS, including preventive systems), OIS, RIS, Tele-Radiology, PGx, PhIS, eICU/nICU, PGHD, PDS, LHS, GIS, CTMS: EDC	Structural techniques; functional techniques; specific brain region-cognitive function techniques; electrophysiological techniques; manipulation of neuronal activity; genes and proteins identification; molecular cloning and recombinant DNA technology, gene delivery strategies
12	Analysis of Variance-(ANOVA)	Statistical model to analyse differences among groups means in a sample. Compares the arithmetic average between two or more groups / Proceeding from: R.A. Fisher, (1925)	Statistical method used for ML- *Classification*. (For verification, comparison)	S	HIS, EHR, PCIS: CDSS (CPOE, eRx/EPS, including preventive systems), RIS, PGHD, OIS, EIR, LHS, PDS, Tele-Radiology, Tele-Pathology, eICU/nICU, PGx, GIS, CTMS: EDC	Structural techniques; functional techniques, specific brain region-cognitive, manipulation of neural activity; genes and proteins identification; isolation of DNA fragments; gene delivery strategies
13	T-Test Independent	Used to *test* the null hypothesis that two groups have the same mean.	Statistical method used for ML- *Comparison*	-	PCIS: CDSS (CPOE, eRx/EPS, including preventive systems), RIS, Tele-radiology, OIS, EIR, eICU/nICU, PGx, PhIS, CTMS: EDC	Structural techniques; functional techniques; specific brain region-cognitive function techniques, electrophysiological techniques; manipulation of neuronal activity, genes and proteins identification

	MIP Algorithm:	(*) Based on/ Proceeding from:	Method & Main purpose:	Algo-rithm Type:	Examples of HIS and other systems closer to neurological care that can use it (*within the platform*):	Some research techniques in neuroscience that may use it: *(Enlarged in the figure 5)*
14	T- Test One-Sample	Used to *test* the null hypothesis that the true mean is equal to a particular value	Statistical method used for ML-*Comparison*	-	PCIS: CDSS (CPOE, eRx/EPS, including preventive systems), RIS, Tele-radiology, OIS, EIR, eICU/nICU, PGx, PhIS, CTMS: EDC, LHS, PDS, GIS	Same as previous
15	T- Test Paired	Used to *test* the null hypothesis that the difference between pairs of measurements is equal to zero.	Statistical method used for ML-*Comparison*	-	PCIS: CDSS (CPOE, eRx/EPS, including preventive systems), RIS, Tele-radiology, OIS, EIR, eICU/nICU, PGx, PhIS, CTMS: EDC	Same as previous
	Others					
	Iteratively Reweighted Least Squares (IRLS)	Derived from Logistic Regression applying Maximum Likelihood Estimation (MLE) and Newton's method/ Proceeding from: Weiszfeld (1936)	Statistical method used for ML-*Estimation*	S	PCIS: CDSS (CPOE, eRx/EPS, including preventive systems), RIS, Tele-radiology, CTMS: EDC	Structural techniques; functional techniques; specific brain region-cognitive function techniques, electrophysiological techniques
	3-C Strategy	Method that combines clinical data, patients' diseases and potential biomarkers/ Proceeding from: Galili et al (2014). Tel Aviv University	ML-*Categorization, Clustering & Classification*	SS	HIS, EHR, PCIS: CDSS (CPOE, eRx/EPS, including preventive systems), OIS, RIS, Tele-Radiology, PGx, PhIS, CTMS: EDC	Structural techniques; functional techniques; specific brain region-cognitive function techniques, electrophysiological techniques; manipulation of neuronal activity, genes and proteins identification

Abbreviations: HIS/HMS = Hospital Information Systems, EHR = Electronic Health Record, PCIS = Patient Care Information Systems, PMS = Patient Management Systems, CDSS = Clinical Decision Support Systems, CPOE = Computerised Order Entry, eRx/EPS = electronic Prescription System, RIS = Radiology Information Systems, PGHD = Patient Generated Health Data, OIS = Oncology Information Systems, EIR = Electronic Immunisation Registry, LHS = Learning Health Systems, PDS = Patient Decision Systems, eICU/nICU= electronic Intensive Care Units/ Neurology Intensive Care Units, PGx = Pharmacogenomics and Genomics Information Systems, APLIS = Anatomic Pathology Laboratory Information Systems, PhIS = Pharmacy Information Systems, GIS = Geographic information Systems, EDC = Electronic Data Capture, CTMS = Clinical Trial Management Systems.

(*) Some references:

Shwartz and Shai, 2014; SPSS Statistical Algorithms, 1991; EBRAINS Technical Documents 2020; Cleopas and Zwinderman, 2015; Medical Informatics Platform Releases- SOFTWARE & REPORT, 2020; Finazzi et al, 2011; Rosenfeld 2015; Stanton, 2001; Bayes (1763); Nattino, 2017; Saxena, 2020; Kümmerle, 2020; Galili et al, 2014; Galbraith, 2002; Bartholomew 2010; Analytics Vidhya (2018); Carter and Shieh, 2015; Ganesana et al, 2016; Mangesius et al, 2020; Schaar (2017); Chen et al, 2011; Li et al, 2011, Fleming and Lin (2000); Bourke et al, 2020; Alashwal et al, 2019; Ding et al, 2014; Florian et al, 2020; Jung et al, 2012; Adamos et al, 2008; Rashidi et al, 2019; Ahmed et al, 2015; Barker et al, 2016; Mitelpunkt et al, 2020.

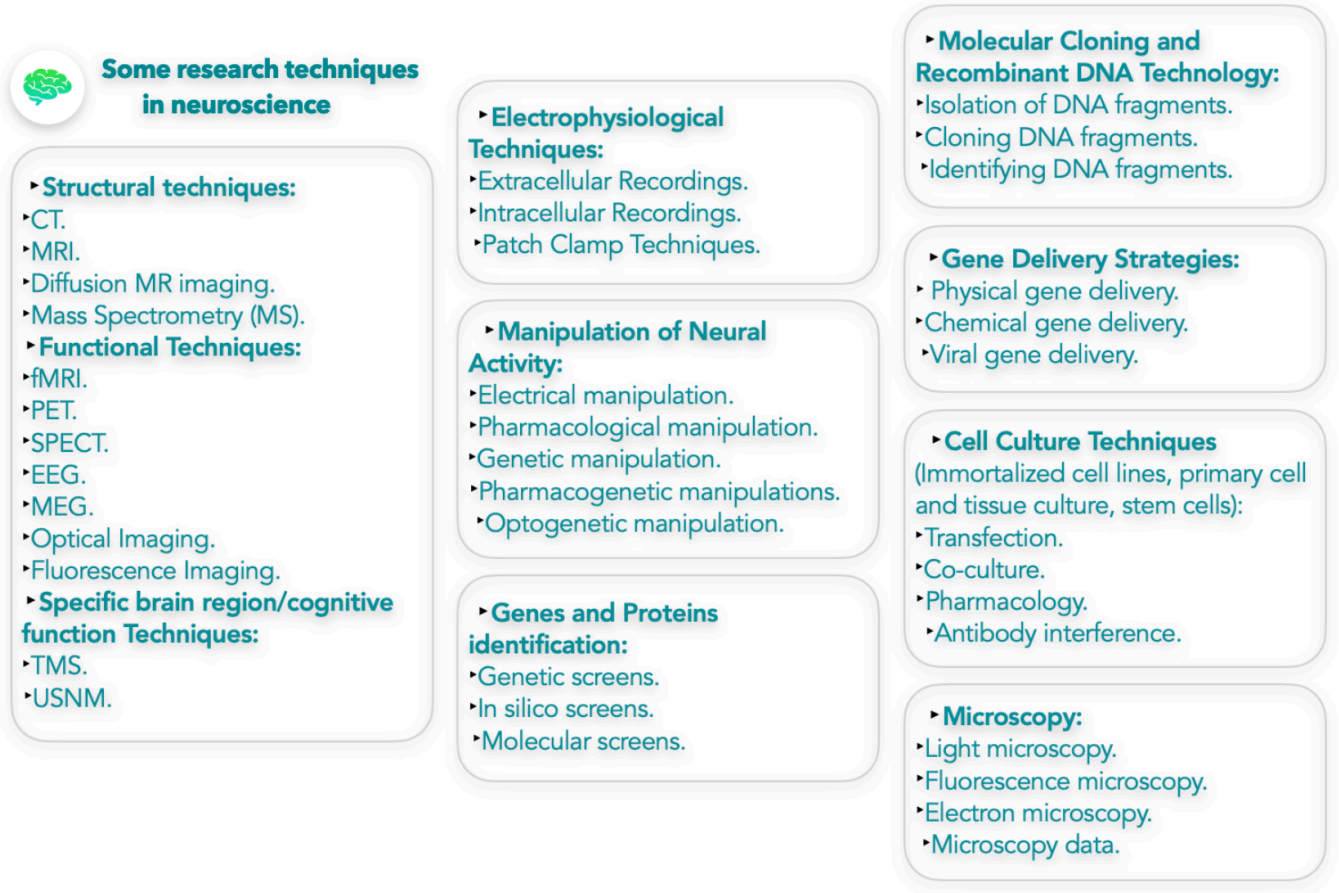


Figure 5. Some research techniques in neuroscience. Adapted from (Carter and Shieh, 2015); (Ganesana et al, 2016).

It is understood that there is no single method or algorithm that fits perfectly with the data of its researchers, its selection and combination will depend on multiple factors such as the condition to be studied, the availability of the data, the research questions, the relationships between variables to be analysed, etc. However, **this perspective has allowed a complementary evaluation of those algorithms that can increase the technological value of the platform as part of its product innovation** (Table N°2), based on some potential competitors (table N°3) and other studies that have shown advances in favour of precision medicine, related to those systems that could benefit directly from experiments in the MIP, which we reflect in Figure 4. After conversations with an AI expert (Pedro Larrañaga; Bielza & Larrañaga, 2020), one suggestion is to complement the MIP with algorithms that allow the platform to increase its differentiation among the various existing softwares that deploy other analysis tools, e.g., include the multi-level classification or the multi-dimensional such as Bayesian network classifiers (MBC), as well as those related to Neural Language Processing (NLP) taking advantage of the information potential of data from medical records.

Natural **L**anguage **P**rocessing gives machines the ability to infer meanings from textual and speech information, -included EHR- with the help of Recurrent Neural Networks (RNNs). (IBM, 2021)

Table 2. Algorithms that could be offered in the Federated MIP as part of its product innovation strategy:

Systems with scientific potential to benefit directly from MIP	Algorithm	Method & Purpose:	Type:	Recent achievements in neurology conditions, associated to MIP	EBRAINS's technology support platforms
PGx, PhIS, CDSS, APLIS, EHR, CTMS	Multi-dimensional Bayesian network classifiers (MBC)	ML- Classification	S	Biomarkers for neurodegenerative disorders (such as Alzheimer's Disease), neurodevelopmental disorders (autism, learning, intellectual disabilities, etc) and mood disorders (such as depression, anxiety, bipolar disorder). Also diagnosis and treatment predictions	Data & Knowledge/ Simulation/ Community
	Support vector machine (SVM)	ML- Classification	S		
	Random forest classifier (RFC)	ML- Classification/ Regression	S		
	Long short-term memory (LSTM)	ML- Classification/ Regression	S		
	k- nearest neighbor (k-NN)	ML- Classification/ Regression	S		
	Convolutional neural networks (CNNs)	ML/DL	Both		
	Multilayer perceptron (MLP)	ML/DL/*Recurrent Neural Network* -RNN. *Regression*	S		

References: *Bielza & Larrañaga, 2020; Maciukiewicz et al, 2017; Nunes et al, 2020; Jamal et al, 2016; Kim et al, 2019; Lin et al, 2020; Chang et al, 2021; Sewani and Kashef, 2020; Eraslan et al, 2019*

Table 3. Main AI algorithms used by potential competitors.
(For further information see the next section "A comparator of The MIP")

Company Name	Main Algorithms	Method & Purpose:
MUSKETEER Musketeer	Linear models (Logistic classifier, linear regression), Kernel methods (classification: SVMs, regression: Gaussian Processes), **Trees***: (Decision tree classification), **Deep Neural Networks***, Clustering (k-means), Principal Component Analysis (PCA). (1)	ML- DL / Classification- Regression, Dimensionality Reduction
TriNetX TriNetX	Logistic Regression - and multivariable-, **SVMs***, Kaplan-Meier analysis, **Nearest Neighbor Greedy Matching***, t-test, z-test, **Propensity Score Matching***, **Log-rank test***, Pearson's correlation, analysis of variance (ANOVA), **Analysis of Covariance (ANCOVA)***, Cox Proportional Hazards Model. (2)	ML- DL / Classification- Regression
BC Platforms BC Platforms	Based mostly on **Microsoft Azure algorithms***. Pre-built models: classification, regression, clustering), pre-processing for feature extraction, Normalization, Multiclass Classification, Two-Class Classification, Image Classification, Anomaly Detection, Clustering, Text Analytics, also for dimensionality reduction, among others. (3)	ML- DL / Classification- Regression, Segmentation, Clustering, Dimensionality Reduction

OWKIN **Owkin**	Based on Substra, their own open-source framework to train combined, on demand models (**Neural Networks***, Classification analysis for images, algorithms to create predictive models). (4)	ML- DL / Classification-Regression
Rhino Health **Rhino Health**	Training framework for AI models supported by **NVIDIA Clara Train SDK*** (enables MONAI, an open-source PyTorch-based for deep learning in healthcare imaging). The learning of MONAI is based mostly in image analysis (registration, lesion section in digital pathology). The analytics algorithms are focused on supervised training (classification models, segmentation models). (5)	ML- DL / Classification, Regression, Segmentation
LYNXCARE **LynxCare**	Based mostly on **Microsoft Azure algorithms*.** Text Analytics are used such as **Natural Language Processing (NLP)***, correlation analytics. (6)	ML- DL / Classification-Regression, Segmentation, Clustering, Dimensionality Reduction
MedCalc **MedCalc**	Correlation, Regression Analysis, T-test, Analysis of Variance, **Crosstabs***, Survival analysis, **Rank sum tests***, Reference Intervals, **Comparison***, **ROC curve analysis***, etc. (7)	ML- DL / Classification, Clustering, Regression, Segmentation
IBM **IBM Federated Learning**	Logistic regression and classifier, **SPAHM with K-Means***, Differential private Naive Bayes models, ID3 decision trees, **Cartpole*, Pendulum*, Keras Classifier*, Reinforcement learning Cartpole*, linear SVM*, Neural Networks*, SPAHM*, Krum*, Zeno***, etc. (8)	ML- DL / Classification, Clustering, Regression, Segmentation
IBM **SPSS - IBM**	Analysis of variance (ANOVA), Cluster algorithms, Correlation algorithms, Correspondence algorithms, **Crosstabs***, Descriptives, **Ensembles (Bagging and Boosting)***, Survival analysis, **k-nearest neighbor (k-NN)***, Linear models, Logistic Regression, Loglinear Algorithms, **MANOVA***, MEANS, Mixed algorithms, etc. (9)	ML- DL / Classification-Regression, Segmentation, Clustering, Dimensionality Reduction
Microsoft Azure **Microsoft Azure**	Clustering (k-Means), **Anomaly Detection (One Class SVM)***, Multiclass Classification, Two-Class Classification, Image Classification, Regression (Fast Forest Quantile, Poisson, Linear, Bayesian Linear, **Decision Forest, Neural Network*, Boosted Decision Tree Regression algorithms***, etc. (10)	ML- DL / Classification-Regression, Segmentation, Clustering, Dimensionality Reduction
EHDEN **EHDEN** (Innovative Medicine Initiative, IMI)	Based on OHDSI collaborative framework. Algorithms: **Lasso regression***, **random forest***, logistic regression, large-scale regression models, predictor models, regularized logistic, **Poisson***, Cox regression, **Natural Language Processing (NLP)*.** (11)	ML- DL / Classification, Regression, Segmentation

***Some analytics currently not available in MIP**

Main references:
(1): Diaz et al (2019); (2): ICPE (2020), Patel et al (2019), Stacey (2020), Kashambwa (2020), Zhang et al (2020), Taquet et al (2021), Holwerda et al (2021), Mazurka, et al (2021). TriNetX (2021); (3): BC Platforms (2016), Microsoft Azure (2021), (4): Galtier and Marini (2019), (5): Github (2021), MONAI (2021). (6): LynxCare (2021), Microsoft Azure (2021), (7): MedCalc (2017); (8): IBM Federated Learning (2021), Ludwig et al (2020), GitHub (2021); (9): IBM SPSS Statistics 20 algorithms (2011); (10) Microsoft Azure (2021). (11): Ryan et al (s.f), OHDSI (2021).

FEDERATED MIP MACHINE LEARNING ALGORITHMS

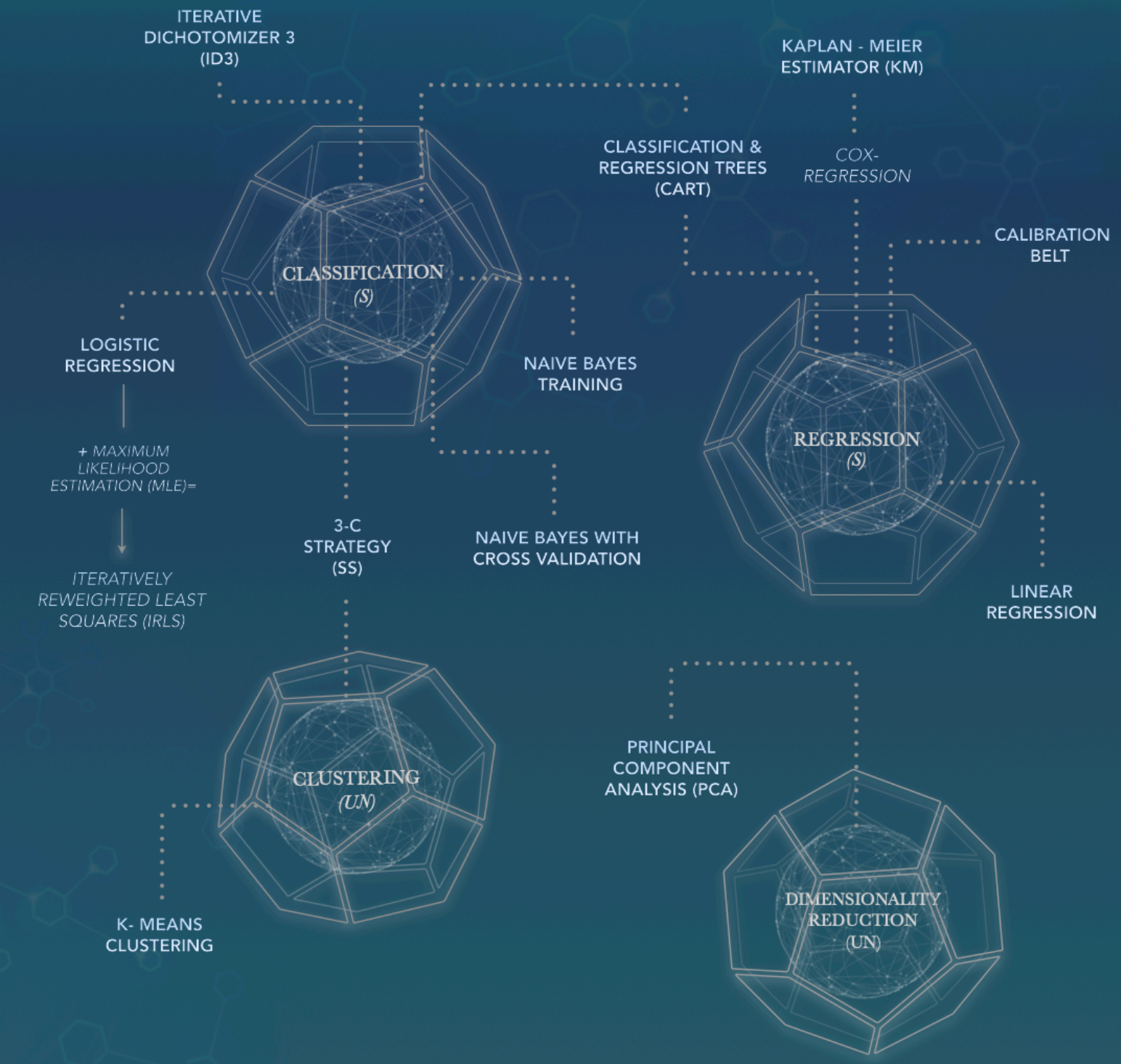


Figure 6. A summary of the federated MIP machine learning algorithms.

Below are shown an updated compilation of the companies specialised in brain conditions with algorithms approved by the FDA, relevant for the MIP startup for future partnerships:

Table 4. Companies with algorithms approved by the FDA

Company	Headquarters	Specific value
Qynapse	Canada	Alzheimer's, Parkinson's and Multiple Sclerosis
Aidoc Medical	USA	Stroke, intracranial hemorrhage, cervical-spine fracture algorithms
Apollo Medical Imaging	USA	Analysis of CT/MRI images
AlgoMedica	USA	CT image quality
BrainScope	USA	Stroke, intracranial hemorrhage
Cognoa	USA	Autism diagnosis
CorticoMetrics	USA	MRI analysis
EnsoSleep	USA	Seizure Monitoring System, sleep disorders
GE Medical System	USA	MRI analysis
Hyperfine Research	USA	MRI analysis
Imaging Biometrics	USA	MRI, CT scans analysis
Rapid AI	USA	Stroke, intracranial hemorrhage
Keya Medical	USA	CT scans analysis
Neural Anlytics	USA	Robot assistance
Nines	USA	CT scans, intracranial hemorrhage
Subtle Medical	USA	PET, MRI analysis
TeraRecon	USA	Imaging processing software
Vital Images-Canon	USA	CT scans, intracranial hemorrhage
Viz. AI	USA	Stroke, intracranial hemorrhage
Vysioneer	USA	Brain tumor
ZepMed	USA	CT imaging
Icometrix	Belgium	MRI, CT scans analysis
Combinostics	Finland	Memory disorders
Avicenna AI	France	Stroke, intracranial hemorrhage
Siemens Healththineers	Germany	MRI analysis
Empatica	Italy	Seizure Monitoring System
Cortechs AI	Nederlands	MRI analysis
Nico-Lab	Nederlands	Stroke, intracranial hemorrhage
Quantib BV	Nederlands	MRI, CT scans analysis
QbCheck	Sweden	ADHD tests
MindMaze	Switzerland	Neurorehabilitation therapy system
Cambridge Cognition	UK	Amnestic mild cognitive impairment
Qure AI	India	C-rays, MRI, CT scans
Zebra Medical	Israel	Stroke, intracranial hemorrhage
Deep01 Limited	Taiwan	Stroke, intracranial hemorrhage

Main references: U.S. Food and Drug Administration (FDA); (2021); Benjamens et al, 2020; ACRDSI (2021)

FEDERATED LEARNING EVOLUTION: A GROWING TREND IN AI APPLIED TO THE HEALTHCARE SECTOR

To understand the current development of healthcare federated platforms globally and associated entrepreneurship projects, it is necessary to describe at least briefly which have been some of the antecedents that have shaped a growing technological trend of great interest to the sector, known as federated learning. It should be noted that data repositories in medical research were designed for the advancements of studies in chronic diseases such as Alzheimer's, marking one of the first steps in the path to a global shared data revolution, clearly preserving privacy, that is, keeping them in the local repositories of their collaborators. One of the oldest initiatives born in 2004 in the United States known as the **Alzheimer's Disease Neuroimaging Initiative (ADNI)** promoted by the National Institute of Health (NIH), widely recognised in the medical research sector. This repository laid the foundations at the technological and organisational level for a subsequent platform also in the US under federated framework known as the **Global Alzheimer's Association Interactive Network (GAAIN)**, ten years later.

For the data shared by ADNI are the imaging diagnostics -MRI, PET, etc- also includes genetic data, biomarkers, among others. Although both Alzheimer's Disease (AD) repositories have evolved over time, GAAIN specialises in the design and selection of patient and cohort data studies, as well as to support the analysis of its researchers, integrating into their platform different repositories from other centres which make these digital resources a powerful means for the advancement of experimental medicine. However, a double-edged sword arises given the possibility offered by GAAIN to the collaborating repositories to 'switch on' or 'switch off' their online data, that is, on the one hand it offers autonomy to control the availability of the data that is shared, but on the other hand, given the value of health data and the effort that is invested to obtain them, can influence more collaborators to decide to temporarily disable them given whatever reasons, so it might be thought that the collaborative concept would be contingent on an uncertain will.

In this sense, some of the reasons why specialists and centres decide -or not- to share their data in general, according to Toga (2018), Professor of Neurology at the University of California, has to do with the privacy of patients, in other instances given the FDA's controls and restrictions to avoid bias and possible manipulation of the data, or because the intellectual property that pharmaceutical companies want to maintain. Although this information is nothing new, what is changing is precisely the way in which data privacy is protected, reduce research time and therefore financial resources, without sacrificing the quality of results.

A number of complex challenges also remain to be addressed and related to the sharing and publication of data, even if patient privacy is ensured. Being able to formulate strategies that promote and encourage sharing by recognising the efforts of doctors and researchers, possibly after years of study, as well as retribution for new trial participants, to name a few. Actions that will be decisive in promoting business models and platforms that offer added value with reliable information, in which it seeks to benefit as far as possible the parties involved, in this sense, new companies at a global level are rethinking these mechanisms. If the prevention of neurodegenerative diseases worldwide -which will continue to increase in the coming years[8]- will depend mostly on progress in precision medicine, and at the same time their data and diagnostics will continue to be supported on these platforms, therefore efforts in this area need to be duplicated[9].

Subsequently, other American repositories promoted by publicly accessible NIH such as **AllofUs**, inclined to recruit participants online, focusing only on national participants and their diversity. Some obstacles they face have to do with incentives, offer $25 to access the program, and participants can at any time quit without having to return them. Although an incentive promotes participation and there is a margin of loss assumed by the state, are complex scenarios to assume in private projects at least for large studies. As a result, data

All of Us RESEARCH PROGRAM

[8] As an example, the number of people aged 60 years and over expected to growth worldwide to 2 billion by 2050, neurodegenerative diseases will also increase. WHO (2018)

[9] Of the 95.5 billion euros approved for the European Commission between 4-6 billion will be used in data strategy, including European data space and federated cloud structures, assumed between the Member States and the Commission, a lower amount considering what these platforms are currently demanding and the business ecosystem that is being emerging from it.

and recruitment strategies are sometimes approached with skepticism.

In Europe, although it took longer to land and define a cohesive integration and data processing strategy at the health and competitive level as it reflects today, similar repositories emerged such as Biobank UK, which, although it began its activity in the 2006, the data has been published for research since 2012[10]. Other no less important initiatives such as the **European Medical Information Framework (EMIF)** -focused also on AD and obesity- in 2013, as well as the **European Prevention of Alzheimer's Dementia Consortium (EPAD)**, to name a few.

The scope of the data provided by each initiative is detailed in the comparative table of organisations and projects related to the MIP. It is clear that the scope and depth of available data for research between repositories depend mainly on the condition in which the project specialises, the willingness of partners and collaborating centres to contribute part of their data, the time defined for sample collection -at least for those repositories such as EPAD, initially developed as a longitudinal cohort study (EPAD, 2021)-, also the management of financial resources, as well as in the harmonisation of the standards considered in the processing of data by specialists.

For example, GAAIN, one of the largest in the US, exceeds 500,000 subjects for analysis[11], in Europe repositories as EMIF have data from 62 MM people and others like EPAD about 2,096; it is understood that its organisational dynamics standardisation and tools available for the study of conditions are deeply variable. Therefore, those mentioned in this section have developed a federated data distribution, whose contributions have driven valuable scientific publications over the years (taken as a reference since vast published information is available). Other well-known have also emerged, such as **genome-wide association study (GWAS)**, **Dementias Platform UK (DEPUK)**, or **RD-Connect** platform deployed from Barcelona Supercomputing Centre (BSC) to integrate clinical and genetic information from data on rare diseases[12].

[10] https://www.ukbiobank.ac.uk/enable-your-research

[11] See http://www.gaain.org

[12] See https://www.bsc.es

In Europe, another Project was launched in 2017 to offer consolidated biological data services for scientific research, called **ELIXIR** (Durinx et al, 2017; Navale et al, 2021). This biomedical infrastructure is composed of geographically distributed virtual nodes offering a series of free tools that allow to speed up the search, storage, analysis and transfer of data, also mobilising the required services through the communities integrated into its infrastructure. In addition, facilitate researchers to work with specific data, under a greater organisation of standards, saving time to specialists in formulating new studies, combining their private and freely available data and thus generate new results (Elixir-Excelerate, 2017).

Another more recent European initiative of a federated non-profit type is **GAIA-X**, created by Germany and France in 2020, currently composed of 300 organisations (also operative in both countries and with international scope), with the ambition -among other aspects- to form a data infrastructure in Europe that promotes an ecosystem of collaborative, trusted and available to its users. Likewise emerges as sovereignty strategy for data control and protection, in front of the global digital competitiveness of some leading providers, also driven by tensions arising during the COVID-19 crisis given technological dependence and the political component derived from digital democracies and techno-authoritarian regimes (Braud et al, 2021). However, it has been criticised by analysts given the ambition to have a strictly European scope when some of its companies have international operation, limiting their competitiveness (Sánchez, 2021). This project rests on cloud-bases services, promoting a scenario of greater trust and control of the amount of information available to its members.

Although it is more than ten years since AI with ML algorithms as currently applied in medical records was conceived as a distant possibility, terms such as learning healthcare systems (LHSs) began to spread in the medical community in the US, subsequently promoted in Europe given the digitisation of data in the EHR (Friedman and Rigby, 2013). I.e., it was initiated as a strategic objective for knowledge generation and management of knowledge derived from informatics systems of healthcare centres, providing monetary incentives to selected medical workers who make significant use of it, to this end the Institute of Medicine (US) -now the National Academy of Medicine (NAM)- set the goal that 90% of clinical decisions would be supported by up-to-date and accurate information with the best

evidence available by 2020 (Institute of Medicine, US, 2009). However, more than ten years later, its implementation remains limited (Menear et al, 2019), in addition to heterogeneous, each country has been adapting this strategy according to its own infrastructure capabilities and needs in other related areas, including federated storages.

Subsequently, a ***federated learning* concept** is introduced formally by McMahan et al (2016), deployed for mobile devices and tablets: "We term our approach Federated Learning, since the learning task is solved by a loose federation of participating devices (which we refer to as clients) which are coordinated by a central server. Each client has a local training dataset which is never uploaded to the server" (Ibíd, 2017), scheme implemented by Google since 2015, which made a huge echo worldwide in addition to the breadth that technology reaches across all its devices.

However, it is understood that the dynamics characterised the federated systems were already applied in the healthcare sector through the aforementioned repositories and with more advanced analytics on digital platforms such as the MIP, this means that in 2015 the latter was already carrying out its activities and protocols with the Centre hospitalier universitaire Vaudois (CHUV) their first participating official hospital (Venetus and Vassalos, 2015). Therefore, the essence of federated learning is precisely to consolidate the mechanisms already described in data analysis and AI models at a collaborative level without compromising the transfer of these, as one of his first publications explains (Ibíd, 2015, pp. 28): "The goal of the Medical Informatics Platform (MIP) is to federate clinical data, such as genetics, imaging, and other, currently locked in hospital and research archives and make them available to relevant research communities (…) Subsequently, these data will be used with the aid of sophisticated Data Mining algorithms and techniques in order to identify biological signatures of diseases".

In this sense, to distinguish the different modalities that in practice arise from federated learning, Kairouz et al (2020) describe and compare two main settings such as cross-device and cross-silo. The first is usually deployed on mobile devices, tablets and also IoT, so its scale is massively so the challenges are more complex, while the second through distributed data-centres, starting from several organisations involved generally less than 100 clients -such the MIP case, for example- where the data is kept local, and a central server is

identified that allows the training of the algorithms without accessing to the raw data from the participating centres. Is evident that in the recent years these schemes have progressively gained an interest in the different industrial sectors, given the facility it provides its users to access analytical tools and predictive models with AI, without exposing sensitive information.

Although multiple initiative and business models are growing with the philosophy of federated learning as shown the figure 7, this report focuses on those related to medical research, comparable or close to the MIP. Most of these have a cross-silo type setting, because participants are hospitals or R&D medical centres for drug development mainly, whose records by its specialised staff and medical devices are already stored in their own data-centres. However, this does not imply that there is no possibility of complementing or combining with cross-device characteristics, given the breadth of patient-oriented health devices and applications, as well as the advent of IoT in healthcare sector leaving a relevant margin for AI training in these.

Figure 7. Map of main federated learning initiatives and startups in healthcare.

For a timely and practical distinction from a business point of view, in order to formulate strategies according to the creation of a startup and its subsequent market prevalence, we previously identified two main platform models: On the one hand as a repository or database for non-profit medical research for partner clients -as currently offered by the MIP and some of the above-mentioned platforms also reflected in the white box of the figure 7-; and on the other hand, lucrative companies, which promote greater advantages and facilities aimed at meeting the specific needs of customers to whom they target in a personalised way. In this subgroup can be found those with a limited participation or specific partners for concrete projects (**Melloddy**, for example, targeting drug development), and others with broader participation (**Musketteer, Owkin, BC Platforms, Trinetx, Apheris AI, Rhino Health**, among others).

Projects with varied flexibility in the integration of new participants can be found, as well as aimed at promoting collaborative actions such as through COVID-19, deploying tracking apps -Apheris AI with OpenMined- or in case of algorithms to predict its outcomes -BC Platforms-. Still, beyond the legal aspect of data -which will be addressed soon along with intellectual property aspects-, other concerns emerge from participants regarding the difference in the shared data and its quality compared to other collaborators within the same project and may even be in competition with each other (Kairouz et al, 2021), examples such as the Melloddy consortium. As business strategy, new platforms offered rewards to their participants by enabling the commercialisation of certain AI models, in addition to a periodic (usually annual) **subscription for users who wish to access the platform**.

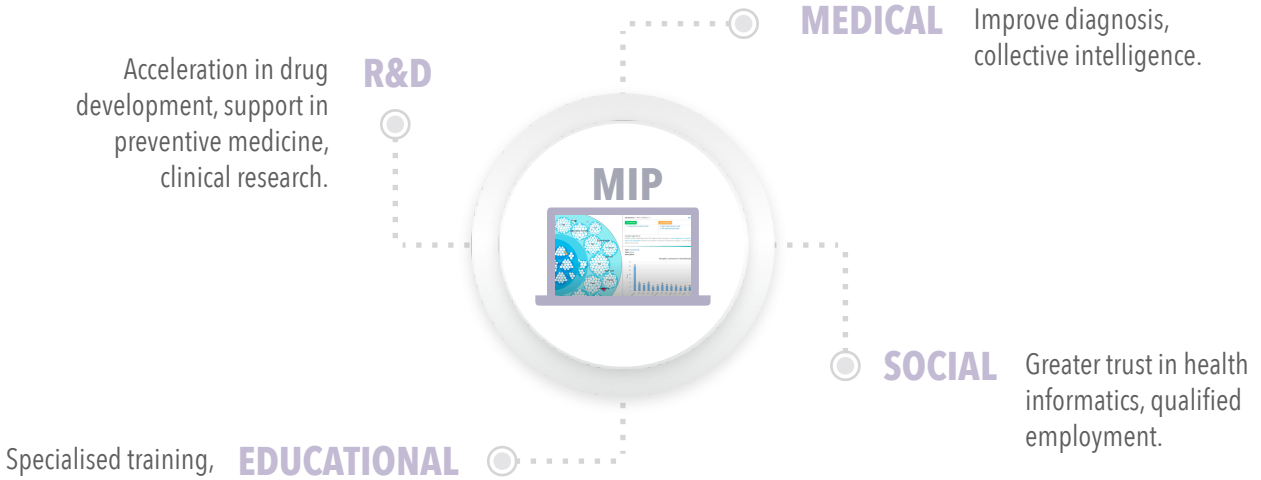


Image: "Summary of positive impact from MIP". Duran, T., León, G., Velasco, G., Strange, B. (2021)

A COMPARATOR OF THE MIP

A comparative exercise was carried out in detail on the characteristics of the MIP together with other related projects, however, the most relevant elements are incorporated and simplified as follows:

TABLE 5. MIP COMPARATOR WITH SIMILAR PROJECTS. BUSINESS, MARKETING AND FINANCIAL FACTS (1/2)

Company	Description	Country (Headquarters)	Year Founded	Geographic Scope	Mission & Objectives
MIP	FEDERATED HEALTH NETWORK	Switzerland	2013	12 countries	Help clinicians, clinical scientists, and clinical data scientists aiming to adopt advanced analytics for diagnosis and research in clinics
BC Platforms	FEDERATED HEALTH NETWORK	Switzerland	1997 (federated since 2018)	25 countries	Global, federated data network for analytics and personalized medicine
TriNetX	FEDERATED HEALTH NETWORK	Belgium	2019	30 countries	Streamlines the design of clinical studies and trials, operations and post approval research, providing health data to biopharmaceutical companies, study sites and researchers
OWKIN	FEDERATED HEALTH PLATFORM	France	2018	2 countries	Implement machine learning tools for drug development
MELLODDY Consortium (Deployed in 2022)	FEDERATED HEALTH PLATFORM	France	2019	8 countries	Help pharmaceutical R&D to develop drugs with ML federated platform
Apheris AI GmbH	FEDERATED HEALTH PLATFORM	Germany	2019	4 countries	Collaborate and analyse distributed data of multiple parties preserving privacy. Federated machine learning platform
Rhino Health	FEDERATED HEALTH PLATFORM	United States	2020	United States and R&D Israel	Federated Learning platform between hospital systems to enhance AI developments
AllofUs (by NIH)	RESEARCH PROGRAM	USA	2015	USA (only)	Advance in precision medicine trough sharing online health data from one million people to support researchers to develop new breakthrough medical findings
SENSETIME	AI technology company (Smart Health Platform not federated)	China	2014	Asia mainly (Singapore, UAE, Japan, China, South Korea, Malaysia, Thailand, Indonesia, the Philippines, Saudi Arabia)	Provide advanced solutions in AI, computer vision and deep learning algorithms through an integrated platform aimed to support researchers and organisations

TABLE 5. MIP COMPARATOR WITH SIMILAR PROJECTS. BUSINESS, MARKETING AND FINANCIAL FACTS **(2/2)**

Alliances & Collaborations	Main covered diseases	Lead Users for business model	Market Range and Target	Company
HBP, EBRAINS, IMI Neuronet, IMI Sophia	Federated: Neurology (Epilepsy, dementia, TBI, mental, Parkinson, depression, multiple sclerosis). Local:	Look MIP Integral Scheme*	31 Healthcare Organizations (12 EU, Switzerland, Russia, Israel)	**MIP**
IQVIA, Google Cloud, Microsoft Azure, Illumina, EHDEN	Autoimmune-GI, autoimmune-non GI, cardiovascular, men' s health, mental health, metabolism, endocrinology, neurology, pulmonary, solid tumor, women's health, oncology, population biobank, hematological malignancy	Researchers, pharmaceuticals and biotech, healthcare organizations, large scale research	100 healthcare systems and 25 biobanks globally. By 2020, 60% of business operation in North America and 40% in European and Asian countries, the targets are: Life sciences and research, healthcare	**BC Platforms**
Initially EHR4CR, after Insite, acquired by TriNetX and recently a major investment by the Carlyle Group	Oncology, genomic data, pulmonology, ophthalmology, gastrointestinal diseases, ENT, Infectious diseases	Healthcare organizations, BioPharma and CRO	113 Healthcare Organizations (67 in USA, 41 Europe and Middle East, 3 Brazil, 22 South Korea and others)	**TriNetX**
Melloddy Consortium	NSCLC, Breast Cancer, EGFR, and NTRK. Hepatocellular Carcinoma (HCC) drug development	Researchers in hospitals, universities and biotech companies	10 pharmaceuticals + 17 research organizations (France, United States)	**OWKIN**
IMI, H2020, EFPIA	Fewer drugs. Without specifications yet	Pharmaceuticals	10 pharmaceuticals + 6 academic partners + 4 SMEs + AI computing company (EU, Switzerland, UK)	**MELLODDY Consortium** (Deployed in 2022)
JLABS, OpenMined, Pistoia Alliance	Genomics, detection of anomalies in medical images, Covid-19 tracing app in collab with OpenMined	Healthcare organizations, pharmaceuticals, researchers, manufacturers, battery material suppliers, chemistry industry, automotive	Not specified	**Apheris AI GmbH**
NVIDIA Inception program (to accelerate AI startups)	Related to genomics, proteomics, microscopy with diagnostic imaging apps and digital pathology	Healthcare centers, AI developers, biopharma	Collaboration started with 20 healthcare organizations during SARS-COV-2 (South America, Canada, Europe and Asia)	**Rhino Health**
NIH	In the data browser (limited): Pain, disorder due to infection, metabolic disease, disorder of joint region, pain in limb, arthropathy, musculoskeletal pain, traumatic injury, hypertensive disorder, mental health.	Academic institutions, health care organizations, community partners, researchers	340 Recruitment Sites	**AllofUs (by NIH)**
MIT, Qualcomm, Honda, Alibaba, among others	Radiology, pathology, cardiology, orthopedic, surgery, radiotherapy AI applications. Also, clinical applications for 13 organs and body parts	Drugs and R&D centers, hospitals and clinics, rehabilitation centers	More than 700 global customers	**SENSETIME**

TABLE 6. MIP COMPARATOR WITH SIMILAR PROJECTS.TECHNICAL FEATURES OF THE PRODUCT & SERVICE

UX Design & Tools	Developed algorithms	Brief description of the structure	TRL	Company
2 main tools: Local and Federated Workstation	YES (15 since May 2020) + PERSONALIZED (?)	Web Platform, Local layer, Federation layer, Data Factory	8	**MIP**
Web application to upload or drag and drop the raw data, protocol selection to run parameters. In GeneVision, for example, as their software for precision medicine, biogenomics tools and algorithms are available such as the Burrows-Wheeler Aligner (BWA) and the Genome Analysis Toolkit (GATK), genetic risk measurements	Unspecified number. Joined forces with RIKEN to use their algorithms during COVID (19 prediction procedures) and other organisations globally	Web Platform (data research management and raw sequence data -BC Insight-), a private web interface for query building and analysis -BC Request-, data management and analysis platform for healthcare organisations and genomics labs -BCI Genome-	10	**BC Platforms**
2 main tools: TriNetX Download (only available for members of their research network) and TriNetX Analyze (basic analyses in their workstation)	YES (LR/Kaplan-Meier) and other measures (Odds Ratio/Risk Ratio)	Web Platform, TriNetX Virtual Private Cloud (as a local layer), TriNetX Appliance (data harmonization process to secure data center with TriNetX data model)	10	**TriNetX**
ML platform: Owkin Studio. Based in Substra* (a privacy-preserving opensource framework developed by Owkin to train models in Federated Learning integrated by: Command Line Interface -CLI-, Python Software Development Kit -SDK-). Also, semantic segmentation for complex images U-Net, SegNet, DeepLap or other	70 Predictive *Models (we infer can be equal to or less than 70 algorithms, models are not equal to algorithms, but they are needed in each model)*	Web Platform connect medical researchers with datasets (Owkin Loop), Local modeling to apply AI (Owkin Studio), Federated modeling and training of AI (Owking Connect). Substra: Command Line Interface (CLI) + Python Software Development Kit (SDK)	10	**OWKIN**
Blockchain platform	YES (Unknown quantity)	Web platform (ML algorithmic layer for drug discovery), Federation layer (Owkin Connect), Kubermatic -Kubernetes for infrastructure management-, Amazon Web Service (AWS)	7	**MELLODDY Consortium** (Deployed in 2022)
Encrypted data, Multi-task Machine Learning	YES (Unknown quantity)	Web platform and layers for different tasks: visualisation of the owner's data models (without exposing sensitive information), privacy approval step, federated training on multiple datasets (different owners and model released according to their contractual agreement)	10	**Apheris AI GmbH**
NVIDIA Clara Train SDK (healthcare app framework): Genomics, Imaging, NLP, protein structure, cheminformatics,	YES (Unknown quantity). Some of them: Clinical Decision Support (CDS), MissForest,	Web platform (NVIDIA Clara), Federation Layer	10	**Rhino Health**
Workspace, Cohort creation, Jupyter Notebooks, Dataset Creation	YES (For now phenotyping algorithms, deterministic or probabilistic)	Web Platform (website and mobile apps for participants) and trained program staff platform to review participant data	9	**AllofUs (by NIH)**
Software: Deep Learning Training Visualisation System, task scheduling system, customized communication library for deep learning. Platform: High speed storage system, operation/maintenance/monitoring system, heterogeneous deep learning supercomputer	100 algorithms and 600 pre-trained models (only in OpenMML an open-source framework), the total is unknown	Sensecare software: Web Platform, research task management to semi-automatic AI algorithms, 3D image processing	10	**SENSETIME**

TABLE 7. MIP COMPARATOR WITH SIMILAR PROJECTS.LEGAL & IP

Tutorials	Language	Published scientific articles & other sources	Certification and Protection Tools	Licenses	Patents	Company
YES	English		GDPR	YES	NO	**MIP**
NO (Only on-demand)	English	In progress	ISO/IEC 27001, ISO 13485, GDPR, HIPAA	YES	NO	**BC Platforms**
YES	English	200 publications	HIPAA+ GDPR+ ISO/IEC 27001. Also, Nessus vulnerability scan.	YES	NO	**TriNetX**
YES	English	19 publications	HIPAA, GDPR, Federating Learning as a protection method, audits executed by external companies specialized in cyber security. Black box and white box	YES	Pending	**OWKIN**
NO	English	In progress	GDPR, HIPAA	YES	NO	**MELLODDY Consortium** (Deployed in 2022)
NO	English	65 publications	GDPR	YES	NO	**Apheris AI GmbH**
NO (but NVIDIA Clara does)	English	Not specified	Local firewall, regulatory grade data packages from model validation, SSL encryption and communication protocols according to articles, partial weight-sharing scheme (for COVID-19 study), Ethical Principles and Guidelines for the Protection of Human Subjects of Research. HIPAA, GDPR	YES	NO	**Rhino Health**
YES	English	838 active projects	HIPAA, Certificate of Confidentiality, the 21st Century Cures Act for enhance protection of participants, Genetic Information Nondiscrimination Act of 2008*	NO -a Certificate of Confidentiality must be followed by researchers-	NO	**AllofUs (by NIH)**
NO	English/Chinese	300 publications	ISO/IEC 27701:2019, ISO/IEC 29151:2017 certification for Privacy Information Management System (PIMS) from the British Standards Institution (BSI), China' s Multi-Level Protection Scheme (MLPS) 2.0 for cybersecurity, CE mark in the EU market	YES	YES (filed 500 patents*)	**SENSETIME**

TABLE 8. MIP COMPARATOR WITH SIMILAR PROJECTS. DATA PROCESSING

Standardization	Initial data transfer process	Data Sources	Average time to move data	Company
Ontology: ROHAN/OLS (263 ontologies). DOID, PTS, BPDO, ECTO, HUPSON, EPO, BRCT, ADO, NIFT, MSO, PDON, SCHIZO, AMDP, CTO.	Only metadata are shared. Through the MIP Data Factory (MIPMap/MRI feature extraction and pipeline/LREN pipeline/NMM pipeline for CSV/ NIFTI organiser/data storage i2b2/Anonymization module), EHR Mapping Task	EHR (EMR)+ DICOM, NIFTI+ JSON (metadata)+CSV files	Unknown	**MIP**
Ontology Browser and BC MATCH curation solution. ICD codes, specific bank genome datasets, genetic markers. OMOP CDM standardized vocabularies, EHDEN certificate	Tools for genomic filetypes (VCF, Plink etc.), API support for distributed data access, HL7, FIHR, LIMS, EHR through their pipeline automation	Genotyping, NGS devices, hospital databases, clinical trials, eCRFs, apps, wearables, other devices, public and private annotation data. File storage for NGS data (BAM, FASTQ), patient data (images, PDFs, reports), the GeneVision platform for example uses the Microsoft Genomics service, by Microsoft Azure	Under a week	**BC Platforms**
Ontology: TriNetX Master Terminology + SNOMED+ ICD 9,10 (diseases)+CPT (procedures)+ RxNorm (medications)+ LOINC (lab)+ ICD-O (oncology)+ HGNC (genes)+ Human Genome Variation Society syntax	Advanced harmonization process that replaces the use of i2b2 + a hardware tool. A match is made between the codes of the previous diagnostics and the nomenclature integrated in the software (mapping). Query Cell, HOME Cell, Extracts.	EHR (EMR)+ PGHD+ PHARMA CLAIMS	2-3 months	**TriNetX**
Mesopath- Mesobank databases for histological samples that previously carry out the standardization process	Only metadata are shared. Through an *Opener* script. Some guidelines: Pseudonymisation data before registering in Substra (made by users themselves) and design non-identifying models. Then, there is no need to data transfer based in their software -stays in owner's node-	Clinical data: Slides of tumor cells, demographics, biopsy prep., whole slide imaging -virtual microscopy-. Mesopath- Mesobank datasets	Not specified	**OWKIN**
If applicable in some cases, python open-source RDKit and RDKit functions (MolStandardize, for example)	Developed by Owkin, a central dispatcher allows share data models between institutions within a private blockchain. Two phases are deployed: an individual code for each pharma to export samples and then structure processing with a common phyton script "Common Code". Public data will be shared as an open-sourced script via GitHub.	The type can vary according to their partners as preclinical data from their academics partners, PubChem public data (freely accessible chemical database), ChemXriv (open access for chemistry), etc. Then public data for researchers will be published as numerical data in an ASCII text file and CVS formats)	Not specified	**MELLODDY Consortium** (Deployed in 2022)
Not specified	Cryptographic methods as private set intersection (PSI) that allows share data without exposing according to OpenMined (2020) and mathematical techniques	Genomic, EHR data mainly	Not specified	**Apheris AI GmbH**
Not specified	Transfer Learning and Federated Learning	EMR, Imaging data (CT, MRI, X-ray, for example), vital signs, demographic, lab, pathology, clinical structured, clinical notes and genomics (we infer that also those using Nvidia Clara)	Not specified	**Rhino Health**
OMOP, SNOMED, LOINC, ICD, CPT, TxNorm	Observational Medical Outcomes Partnership (OMOP) Common Data Model Version 5	EHR, Surveys, physical measurements, bio samples, PGHD	Data is taken directly from participants through the Participant Portal: 20 minutes to answer surveys	**AllofUs (by NIH)**
Classify lesions according to international diagnostic guidelines without more specs	Supports small dataset training on medical images and automatically perform quantitative analysis of key parameter measurement. Then diagnose malignancy and severity	Data from CT, MRI, PET, digital pathology. Training images in a two-stream manner: seg-stream and int-stream.	Sources with inaccurate information. Days, but not specified	**SENSETIME**

TABLE 9. MIP COMPARATOR WITH SIMILAR PROJECTS. DATA STORAGE

ALLOW COPY, PASTE OR DOWNLOAD DATA	Data Maintenance and Quality	Federated Data Storage (CLOUD)	Local Data Storage (CLOUD)	Data Range	Company
NO	Data Quality Control tool/ LORIS for DICOM quality control	YES (FENIX consortium within HBP/EBRAINS)	YES	200,000 SYNTHETIC DATA, (PATIENTS ?)	**MIP**
NO	Quality control analysis for array data and integration of NGS. Reports following guidelines from American Society of Clinical Oncology (ASCO), American College of Medical Genetics (ACMG)	YES (Google Cloud, AMD processors and for genomic data multiple douds GCP, AWS, Azure)	YES	4 MM patients/ 400,000 Genomic data	**BC Platforms**
NO	Their own methodology: 4C (cleanliness, consistency, correctness and completeness)	YES (AMAZON WEB SERVICES -AWS)	YES	84 MM patients in scientific articles by 2018*/400 MM currently according to webpage	**TriNetX**
NO	Distributed Ledger Technology /Local calibration to optimize weights models (parameters), e.g., to balance the discrepancies in data between hospitals and academia, given the differences in the number of sick people among them	YES (GCP, Azure, AMAZON WEB SERVICES -AWS)	YES	40 datasets, 8700 patients in Hepatocellular Carcinoma (HCC) data	**OWKIN**
NO	All partners should introduce their data with the same protocols and principles, verifying their internal practice guidelines to avoid incorrect samples, also a quality control tool is applied	YES (AMAZON WEB SERVICES -AWS)	YES	1 billion data points for drug development (surpassing the average pharmaceutical company) + 100 terabytes of image data	**MELLODDY Consortium** (Deployed in 2022)
NO	Differential privacy, random falsification to reach quality, in chemical cases explain the data can be pre-processed with preshipment inspection (PSI) to discard duplicate data	YES (Do not specify cloud provider)	YES	Not specified	**Apheris AI GmbH**
NO	Not specified	YES (Not specified, we infer AWS, Google	YES	During COVID-19 study about 16,000 cases. Global, unknown	**Rhino Health**
YES. LIMITED, USER CODE OF CONDUCT (export for analysis to Jupyter notebook)	Sync for Science protocol by NCHIT	In a Raw Data Repository. Common cloud (online) environment	NO	More or less 230,000 participants	**AllofUs (by NIH)**
Unknown	Smart AI-assisted method, SenseCare®-Path Intelligent Diagnosis System	Does not appear to specify federated learning	Sense Time's AI Cloud platform and Supercomputing Centers	1.4 Billion people	**SENSETIME**

TABLE 10. MIP COMPARATOR WITH SIMILAR PROJECTS. DATA STORAGE

COMING SOON	Limitations Found	Strong Points	Company
Deployment in more hospitals, personalised algorithms, new version releases (March 2021)	One language/ UX design/ tutorials to improve details	Health World Data, complex and personalised algorithms, Rare and specific diseases could be explored/online mode and offline use, It does not require any downloads and does not use any storage space, It can compare diagnostic procedures and CDEs between countries. SEE STRENGHTS IN SWOT ANALYSIS	**MIP**
Partnership with leaders in clinical genomics, allowing AI/ML applications	Relies on a wide range of vendors for data storage and distribution. No algorithms used are specified, however they have a considerable amount of data over the years. More reserved in terms of publishing information on specifications and procedures.	Strong experience in genomic data, offer a global network combining the data of biobanks, rapid response on queries according to articles, also reduce harmonization processes. The platform shows researchers real time data on biobanks with parameters selected. For drug development allows to select previously the objective of the study (e.g., discovery, preclinical, clinical, for first approval) allowing an appropriate design scheme from the beginning, also select animal models, select risk biomarkers, customize workflow, re-use data for research, etc.	**BC Platforms**
Algorithms from provider notes, medical and pharmacy requests, children analyses, genomics data, improvements on analytics methods, prospective monitoring	One language, depends on third party for storage AWS, 2 algorithms by now (limited ML), Few published tutorials, limited to demo request	Health world data, design protocols, patient recruitment for pharma, clinical trials management, compare cohorts, map distribution patient, summary statistics, demographics, time or period facts by cohorts, charts predictions, compare common treatments and outcomes, synonyms in the search process and codes	**TriNetX**
Patents pending	Models have (as in all technologies) limits in predictions, which increases the need to look for representative datasets, their software also needs the support of complementary privacy protocols (to avoic origin identification, for example), common reluctance of centres to share data, so they share data models and rely in the good faith of other organisations associated	Data collaboration between partners without exposing privacy, perform machine learning algorithms remotely, private network with nodes connected allowing controlled by own organisations. Metadata stays in their nodes, only algorithms and models are exchanged.	**OWKIN**
Developing sophisticated AI tools based on pharmaceutical samples, strong communication strategy	According to IMI: Too many complex and technical concepts that require to be communicate simultaneously, might provide information to a small audience, given the diversity of partners at times there may be conflicting priorities	Multi-task ML data, chemical applicability, a federated training platform, strong security and privacy preserving, offer predictive models, associated with major pharmaceutical companies in the industry, personalised algorithms for evaluations	**MELLODDY Consortium** (Deployed in 2022)
Recruiting human resources and boosting business development	Low experience in the medical and business sector compared to the market, reserved with details of its processes and manuals	Support of public sector, open to business partners, wide application to other competitive sectors -automotive, manufacturing, etc.-, privacy-preserving, collaboration workflows	**Apheris AI GmbH**
Recruiting to accelerate the company growth, support clinical research to its different stages	Limited information on their portal, they rely on a third party to deploy the most important technology (NVIDIA) unlike the MIP platform	Training models with machine learning, protect privacy, data from multiple countries, supported by NVIDIA, recent financial support from investors, avoid moving data as other federations do, accelerate AI developments	**Rhino Health**
1 MM Participants by 2024, transfer data by IoT. A dashboard that compares the participant's data with the data acquired in the programmes	Many videos without depth in use, to a less specialised audience, some vulnerability of participant's data although with warnings for misuse, no algorithms showed, wide range of diseases but limited depth. According to operation protocol, 25$ dollars are offered for participation, but if a participant decides to withdraw (at any time), no refund is expected	Supported by NIH, focused on diversity to study different samples, precision medicine, diversity-focused fir effectiveness testing in mixed groups, matching patients for trials, objective in precision medicine, bring researchers to the data rather asking them to download it	**AllofUs (by NIH)**
Generate an impact in sensitive sectors such as health -decreasing disease diagnosis rates-, help to reduce time spent interacting with healthcare providers and government-, promote more creative jobs, decrease emergency responses, safety environments, among others.	Limited medical conditions for the moment, because they diversify into very different sectors. According to their publications, they face several challenges such as interconnects data at different levels, failures of individual nodes, find human resources who understand both deep learning and HPC	They have prominent breakthroughs in the AI field, their technology covers a wide range of activities inside and outside the healthcare sector, i.e.: urban management, food security, virtual avatars, hospital security; most of them with the power to recognize objects, events and people. They have also launched their education sub-brand, SenseTime Edu and a learning platform, SenseStudy	**SENSETIME**

COMPETITOR ANALYSIS & STRATEGY

The above comparative analysis shows the position of the MIP vis-á-vis competitors trying to enter the market with federated solutions or those companies with a strong experience curve that have migrated to this modality given the associated advantages. The reasons to adopt federated learning start from the protection of patient data, its preservation and control in specialised centres, the progressive impact on lower costs for new studies, support in the confirmation or discarding of hypotheses with AI, possible impact on the planning of prevention programs, accuracy in diagnoses, generation of business opportunities, to training in new technical skills, and more. Thus, additional factors arise that should be evaluated before structuring the business model, as well as in its innovation strategy considering the technology level and the service, as we summarised below:

- Business, Marketing & Financial Facts:

Starting with available industry data, recent reports project federated learning platforms to grow by USD 117 million in 2023 to 201 million by 2028, at a Compound Annual Growth Rate (CAGR) of 11.4% during the forecast period, expecting greater involvement of the healthcare and life science field (MaketsandMarkets, 2021; Bloomberg, 2021). In addition, in terms of AI tools for the health sector, funding reached a record of $2.7B for the sixth quarter (CB Insights, 2021). In Europe, increasing the possibilities with AI techniques, their management and learning has been established as a priority in government programs such as Horizon Europe or Digital Europe programme, and also highly attractive to VC investors who already reached $8.5 billion among the 50 largest cumulative investments at the end of 2019 (EIT Health, 2020).

It should be clarified, there is no definite data published on the quantity of startups specialised in the diagnosis of neurological conditions under the federated learning scheme[13], so it is not specified in this report at the moment. However, those with greater visibility in the leading countries in the diagnostics sector and technically closer to the platform have been compiled from many

[13] Until less three years ago, industry publications were unable to identify commercial vendors of federated learning solutions. At present, other recognised platforms on startup analysis reflect a limited number of companies with this technology in the neurology field or include other types of technologies.

sources on the map (Figure 7), thus may or may not include activity in brain diseases. Despite this technology has passed the initial phases that characterise the product life-cycle[14], its growth in the diagnostic sector is relatively recent comparing with other competitive privacy-preserving technologies such as cryptography with machine learning or blockchain, also complemented with federated learning, introduced approximately in 2008[15]. An illustrative diagram for the formulation of market strategy is shown below with the current position of the platform:

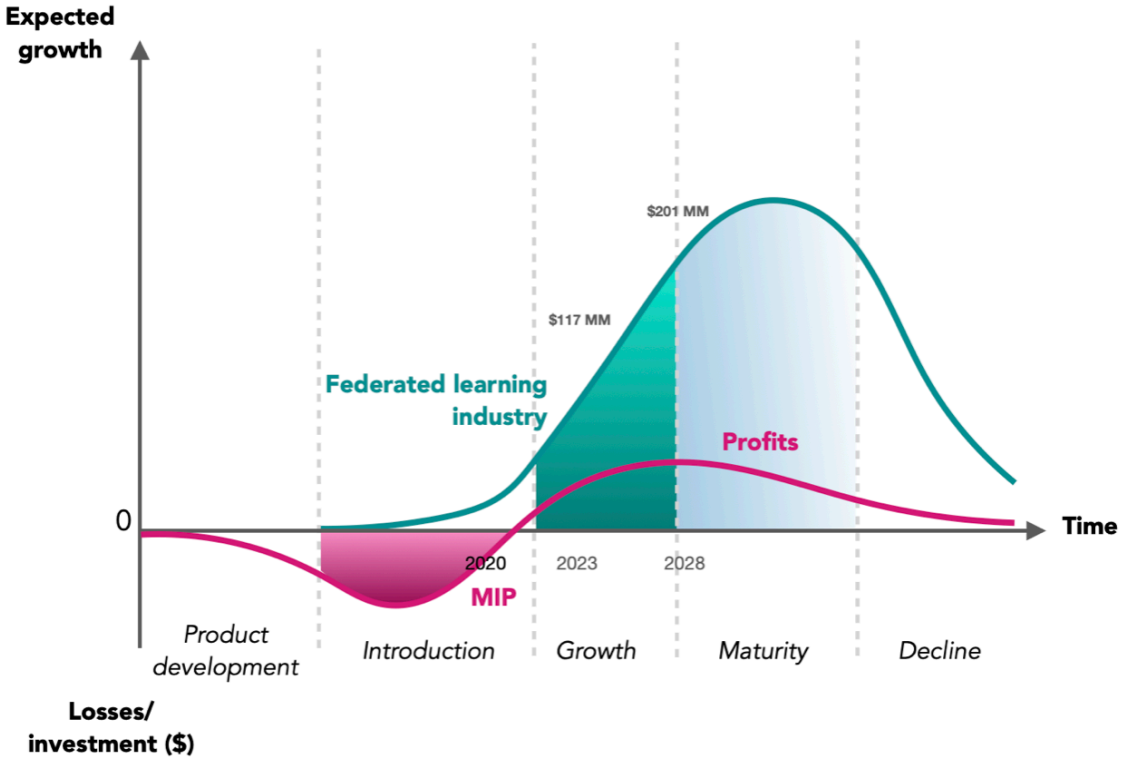


Figure 9. Diagram of the product-development cycle and the MIP.
Adapted from *Kotler and Armstrong, 2018) market projections (Marketsandmarkets, 2021)*

It is understood that the software has been developed early, however, its introduction in the market has shown a slowdown with respect to the rest of the industry that was already gaining advantage using the technology. Even so, key insights from the healthcare sector can be considered that will allow to consolidate its first steps in the market, for example, in a survey of 175 specialists in the field (Ibid, 2020), were consulted the areas where AI is most used, as shown in the

[14] Originally created by Gort and Kepler, 1982.

[15] Nakamoto (2008)

figure 10. Is noticeable that platforms such as the MIP can enter to the European market, in the areas where these tools are most frequently applied: Diagnostics (imaging, pathology, sequencing) (21.5%) Clinical decision making (18.1%), Data management (15.3%), Education (8.3%) through online tutorials and on-demand training programs. Another segment in which the MIP would fit is Pharma (drug development and clinical trials) although the survey suggests a small percentage (3.5%), we consider the platform quite useful for users who want to start clinical trials in neurology, a market of interest to the team described in more detail in Part II. Clearly, some available publications share projections that confirm the growth of federated learning and AI investment in health industry and also contributes to complement the previous analysis and the growing generation of startups in recent years.

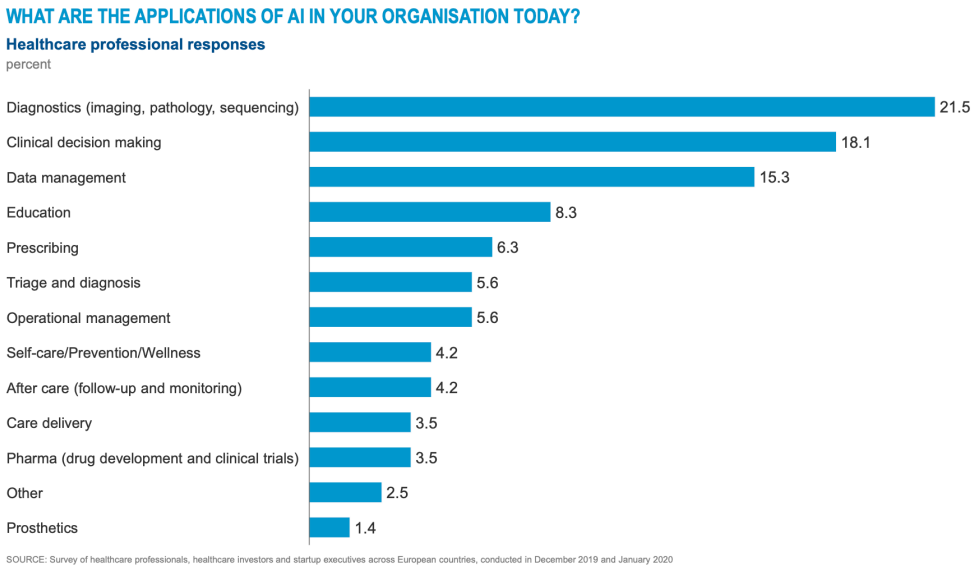


Figure 10. EIT Health McKinsey & Company (2021)

- Alliances & Collaborations:

Another aspect that arises from the comparative analysis concerns the types of competitors of the MIP and the relationship they have with other giants of the sector in terms of collaboration or strategic alliances, providing them with a cost advantage, support in innovation management as well as know-how, to name a few. However, it is important to identify the types of competitors and these strategies that, for practical purposes, we have differentiated them according to the level of maturity and size, within an industry of constant growth

accompanied by shorter service life-cycles[16]:

- Companies with high technological capacity, bargaining power and positioning in the sector[17]*:* Multinationals and firms with a relevant experience curve, enjoy a brand prestige, provide specialised services in AI and offer solutions in different industrial areas including the health sector through their marketplace, offering both the technologies of their own subsidiaries, as well as those of its associated partners (*e.g. Microsoft Azure, NVIDIA, IBM Federated Learning, Google Cloud, Alibaba Cloud, AWS, Intel Federated Learning*). Another group corresponds to large companies specialised in R&D in healthcare that adopt federated learning in one or more of the services proposed in their catalog (*e.g. BC Platforms, Celerium, C3 AI Suite, TriNetX, DataFleets, Iqvia*), and pharmaceutical companies that develop software associated with this modality to different medical specialties (*e.g. Merck or AstraZeneca*).

Definitely, in this category strategic alliances continuously flourish along with technology transfer agreements, being considered as one of the possible scenarios for the MIP startup, with special attention to certain multinationals recognised in related fields. Options may include agreements subject to terms or services of interest to each party and therefore not necessarily be restricted to the platform. Thus, although the alliances suggest an opportunity for growth given the experience of multinationals, some possible advantages and associated risks for the MIP startup are summarised:

ADVANTAGES

- Expansion to new markets and greater visibility reaching new lead users, which would hardly be found through the current sources.
- Introduction of more algorithms highly specialised to specific brain diseases.
- Reduced costs in the internationalisation process in terms of the networks already stablished inside and outside Europe.
- Expert feedback to renew training programs, manuals and consulting services, adapting the platform to the IT systems and needs of these new users.

[16] Experts in innovation management show that service life-cycles become shorter. (D' Alvano and Hidalgo, 2011)

[17] The bargaining power of suppliers is a concept originally developed by Michael Porter in the 80's, as one of the known Porter's Five Forces of industry competitive analysis. (Porter, 1985)

- Speed in the exchange of knowledge in medicine, avoiding repetition of studies and generation of diverse high-precision AI models in neurology.
- Acceleration in learning technical processes and collaboration in unforeseen technical difficulties (greater risk management).
- Access to new knowledge and human resources increasing the innovation skills.
- Option to support datacenter outside Europe, sharing costs or distributing the associated high energy consumption (1).

RISKS

- Depending on the technology transfer agreement, less control of the software, may involve the partial or total assignment of the license and the trademark (2). It would therefore be convenient to commercialise its use without distributing copies.
- Transfer of valuable information or metadata of neurological diseases associated to European hospitals. In case of offering only the federated version, given the limitation of available algorithms and its current UX, it would be less competitive as there are platforms with similar characteristics already incorporated to their catalogue.
- Possible conflict of interest arising from agreements reached with other platforms, increasing the bargaining power of these organisations, would also be reflected towards a higher standard in the industry when considering new business proposals. In other words, a disadvantage given the time it is introduced to the market vs emerging projects.
- Lack of experience in knowledge integration mechanisms inherent in alliances, increasing costs in personnel, financial resources and time. Two main ways are identified in the literature (3) *direction* or *organisational routines*, which involve the establishment of rules and procedures that contribute effectively to the development of technology and associated services.
- Risk the main objective of alliances and their evolution due to the interest in acquiring specialised knowledge rather than its application, favouring a competition scenario rather collaboration. It is also reflected in differences between partners in terms of mutual learning (convergence) or acquisition and specialisation of knowledge (divergence) (3).

Main advantages and risks of strategic alliances and technology transfer of the MIP.

- Consortium and public-private partnerships (PPPs): As described in the previous section, they are projects in development composed of pharmaceutical companies, R&D centres and federated AI firms seeking as much as data as possible from third parties to train their models, as part of its trials for drug development and precision medicine. From specific participation projects coordinated by a

(1) Specialised studies calculate that data centres consume an estimated of 200 terawatt hours (TWh) annually, more than the consumption in countries such as Iran (Jones, 2018); (2) Hidalgo et al, 2009; (3) Grant and Baden-Fuller, 2004; Mowery et al, 1996.

leading participant that provides the federated technological structure (e.g. Mellody), federated data repositories available to other organisations (e.g GAAIN), non-for-profit organisations (e.g. Pistoia Alliance), to more ambitious projects that seek to incorporate organisations and data under this modality (e.g Pistoia Alliance, EHDEN by IMI), the latter also recommended for the MIP under collaboration scheme, with currently 41 selected data partners (EHDEN, 2021).

- ***Startups and unicorns:*** Undoubtedly they are looking for positioning in the industry, eager in most cases to increase the number of customers who provide them with data (*e.g. Owkin, Apheris AI),* along with an accelerated growth in the market, expanding the possibilities of support from VC investors (*e.g. Rhino Health, Edgify*). In this category we also include startups based on privacy-preserving schemes on blockchain, a framework that has gained popularity in recent years, also incorporating collaborative schemes that attract programmers and experts in data analysis, furthermore, they also promote dynamic virtual ecosystems, strengthening their competitiveness in their AI solutions, inside and outside the healthcare field. They have the technical capacity to provide services to hospitals and research centres in data analysis, so in the long term those that manage to maintain their growth strategies have the technological capacity to become first group companies and venture into this market, therefore, increasing the possibility of being part of the future competition for the MIP startup. Unicorn startup such as *Innovacer* in the USA also have attractive characteristics for the medical community, with a workflow powered by EHR, laboratory, etc. with analytical solutions that, although they do not start from a federated structure, they have partnerships with strong organisations there, so it could be considered a collaborative initiative towards this attractive market.

- ***Companies of upcoming introduction or emerging technologies:*** From the perspective of alliances and collaborations, these are teams enriched by networking promoted by incubators, entrepreneurship programs, technology centres that allow them to develop their technology to the maximum. Some provide guidance in terms of protecting their intellectual property, promoting a climate of trust, innovation projects and potential investors. Naturally, given the scenario of uncertainty to which these groups are exposed, they cannot be satisfied with a single strategy to enter the industry, some

of them opt for differentiation generic strategies (Porter, 1985), others by innovation strategies through validated learning based on methods such as Lean Startup (Ries, 2009), improving the value proposition in the Minimum Viable Product (MVP), strengthening their competitive advantage. Others have managed to gain position in complex scenarios such as Rhino Health in USA, which grew rapidly given the urgency of COVID and the need to work data from hospitals, strengthening its differentiation and cost strategy -free of charge-, collaborating with 20 hospitals worldwide with a striking growth in less than a year.

In summary, the generic strategies that characterise the federating learning industry together with the strategy that we suggest to be followed for the MIP to its introduction, is shown below:

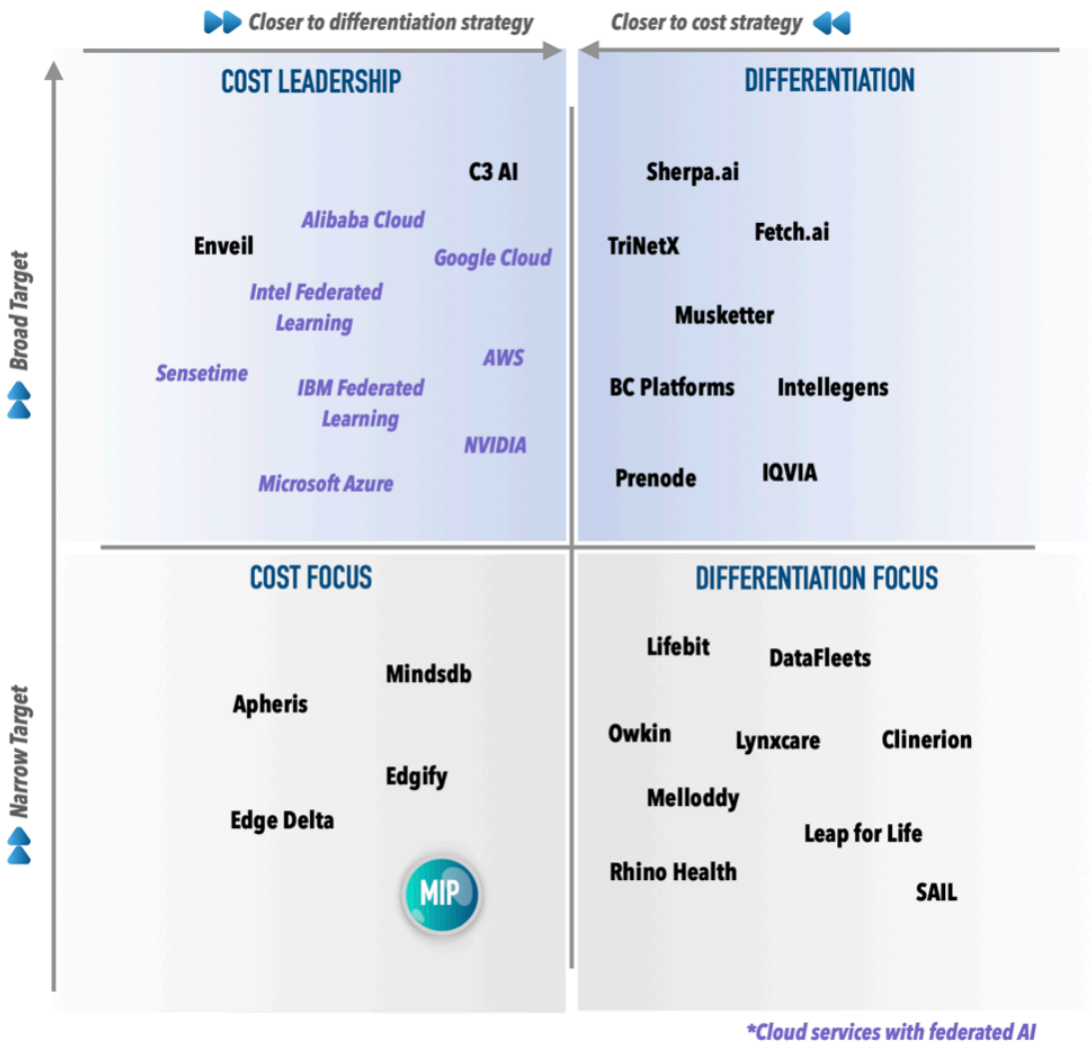


Figure 11. Leading companies in the ecosystem of federated learning and the recommended generic strategy for the MIP market introduction. Adapted from Porter (1985)

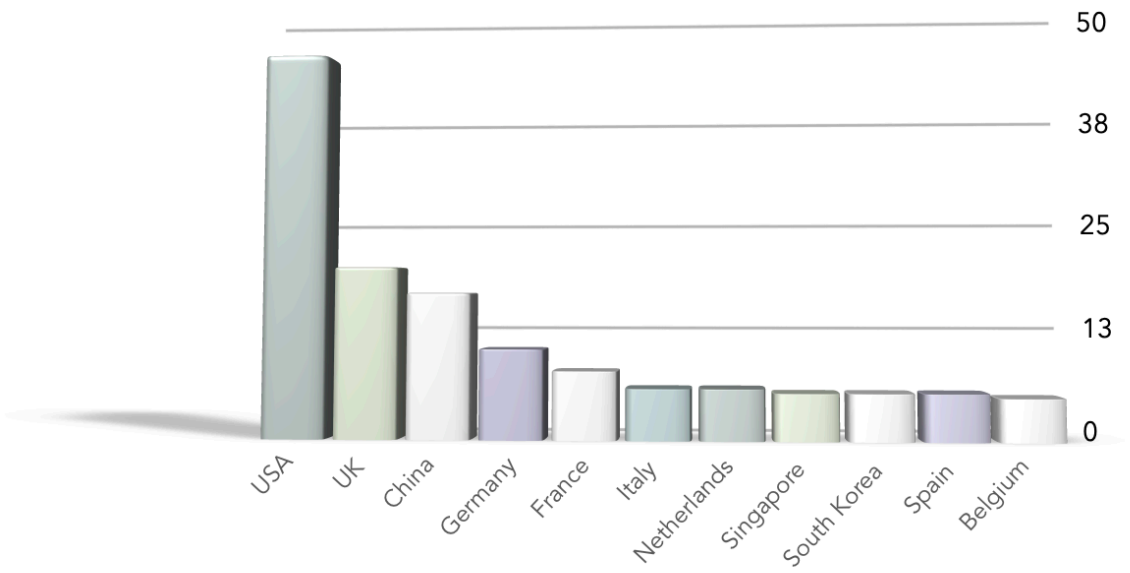


Figure 12. Main countries with the greater number of publications in federated learning within the healthcare sector. Data: Web of Science

- Additional considerations of the federated learning ecosystem:

Since our interest has been precisely to find as many projects and initiatives as possible that can be in line with the future startup, worldwide publications related to the health sector based on federated learning denote a significantly lower proportion (9,2%) compared to all industrial areas (1,599), being the STEM areas the most prominent in global terms[18]. This only suggests an approach to identify the regions and possible technological clusters that are evolving in this field, in addition to identifying the main organisations involved. The most relevant factors are simplified:

It is an increasingly accessible technology to medical teams: Understanding that federated learning allows to train AI models without accessing source data combined with the integration of an architecture that invites cross-organisational collaboration, different open-source frameworks are offered on platforms as GitHub (e.g. FATE, Federated AI Technology Enabler[19]). Encouraging in this sense the search for quality services that can be marketed and increasing the competitive advantages among companies. It is clear that the business model for the MIP should be based on Software As Medical Device (SaMD) scheme, with an emphasis on offering AI models, training programs adapted to medical teams and clinical researchers,

[18] It is inferred that there could be more, at least for those publications that do not include the term as the main method or key words, limiting searches and the current capacity of the most specialised portals.

[19] Created by Webank's AI Department in China, more at: https://fate.fedai.org

driven by sales and post-sales service, expanding job opportunities in specific areas.

Simplifying the main areas of medical research in which this technology stands out: Medical Informatics (21%), Health Care Sciences Services (16%), Radiology Nuclear Medicine Medical Imaging (14%), General Internal Medicine (10%), Neurosciences Neurology (8%), Public Environmental Occupational Health (8%), Biochemistry Molecular Biology (7%), Mathematical Computational Biology (7%), Oncology (7%), Pharmacology Pharmacy (6%), Psychology (5%). An opportunity for the future company to be alert to the obstacles that arise in the studies carried out by these teams with similar platforms, as well as to formulate and encourage the quality and quantity of new publications. In neurology, the leading AI corporations that predominate include organisations such as Intel or IBM.

Undoubtedly, a major part of the funding for these projects in Europe comes from public bodies, mainly from the European Commission, internationally recognised institutions collaborate in their areas of specialty such as NIH, specialised research agencies and foundations (e.g, National Institute For Health Research, CGIAR, Uk Research Innovation, Wellcome Trust). This is also reflected in more ambitious European projects such as Melloddy (IMI, 2021) that although it involves the pharmaceutical sector as partners, the funds comes mostly from IMI, universities and the AI companies that provide the technology. Thus, the future MIP startup could incorporate as part of its activities the promotion and validation of studies in the development of drugs and therapies with federated learning, encouraging new collaborations with this relevant sector.

Main sales strategy: Subscription-based models with possibility of demo version or free trial up to 30 days, then subscription plan (usually monthly), cloud service provider (e.g. AWS, Microsoft Azure, NVIDIA), in this regard, the MIP would have the advantage of having FENIX service[20]. Likewise, the product or service which it specialises should be considered, the competitive advantage of the startup: From AI models, access to advanced studies, train algorithms services, cohorts, as well as the entire post-sales service scheme (fast support from engineers and technicians, maintenance, tutorials and training).

[20] https://www.humanbrainproject.eu/en/ebrains-tools-2020/

PART II.

MARKET SEGMENTATION & PERSPECTIVES

A strategic scheme has been defined to evaluate actions consistent with the objectives of the MIP together with current market characteristics and their needs, in accordance with publications and statistical data from official bodies, research centres, hospitals, specialised sources at the scientific level, published opinions from industry experts, as well as real-world cases of related business and academic resources. The aim is to increase support for the platform as much as possible, in a personalised manner in terms of its development and growth as an upcoming startup.

From a practical approach, this scheme is developed into different perspectives, from a more general to a more specific scope: The first corresponds to a more detailed framework in terms of the number of hospitals and their characteristics, by region and income level -HIC, LMIC-. **The Table "Potential Opportunities and Possible Business Models" is taken as a reference, also combining with international classifications such as the Systems Health Accounts (SHA) and National Health Accounts (US NHA) described below**, to obtain a specific structure and propose actions aligned with the variety of diseases that the MIP offers to experiment on its platform. This perspective also assesses country prevention programmes, which are another factor in delivering it with potential utility. Not only by helping to mitigate the risk to the advent of new conditions, as explained above, but also by the associated costs that would reduce hospitals, including pharmaceuticals in accelerating specific studies for drug development.

Naturally, each medical specialty comprises certain needs, demands and priorities, in this sense the second perspective focuses on defining actions at a strategic and practical level in these markets in limited specialties of the MIP given its scope, as well as possible links to other specialties that have priority related to the platform. This section also describes examples of some related companies by sector, their investments capacity, health technology assessment frameworks (HTA), partnerships, merger and acquisition strategies (M&A), as well as closures of hospitals and centres. Finally, the last perspective defines the main legal frames affecting technology, and related property (IP) protection issues. With this knowledge the exploitation plan is strengthened, allowing to foresee appropriate actions for its deployment.

REMARKABLE INITIATIVES TAKEN BY THE MIP TEAM AS BASIS FOR THE EXPLOITATION PLAN

The following image (N°13) published in EBRAINS[21] shows the status of the software installed in major hospitals -currently more than 30- in most part of Europe, as well as other nearby countries such as Israel and Russia. In addition, according to the MIP team, almost 25 clinical research centres outside the HBP have the platform installed. This constitutes a valuable effort by the team to be able to consolidate and coordinate in the medium term the dynamics of work with the centres and hospitals already associated, in which the requests for personalised cases by medical teams and their feedbacks should be also incorporated in a regular way, including a continuous update of the application. Also, this distribution in Europe can be considered an advantage over the rest of the federated platforms that are starting in the industry -some of them explained in the previous section, to establish comparisons in the sector-, and it is also an advantage to count with the support of the infrastructure FENIX within EBRAINS, offering greater confidence to existing customers and attracting new users.

Figure 13. The MIP Network

In addition, the MIP team also explains that their action plan includes presentations and collaborations with key associations such as the European Academy of Neurology (EAN) and the European Stroke Organisation, arguing that these efforts allow them to compare quality indicators in health conditions that may differ between different regions in Europe. Likewise, they explain that other no less relevant networks such as EpiCare or European Reference Network (ERN) also offer them a frame of reference to comply with these European policies, organisations with solid experience and contacts with a large number of hospitals networks and specialised centres, in which the platform has also been supported for its deployment. Another aspect that the team points out is the limitation of funds to continue with a more ambitious exploitation plan and that also contemplates the expansion of its human resource, of utmost importance to be able to

[21] https://ebrains.eu/service/medical-informatics-platform/

efficiently distribute the activities for the spin-off. Some alternatives proposed include the possibility of joining projects such as the European Health Data Space (EHDS) and the Brain Health Data Space (BHDS). With this action plan, we offer also other recommendations that could be considered for future deployment in a fast-growing and learning industry in the next segment.

POTENTIAL OPPORTUNITIES AND POSSIBLE BUSINESS MODELS

A table of the industry organisations is prepared below, following initially the two groups of classification for the health sector of the Global Industry Classification Standard GICS (2018): 1) Health Care Equipment & Services and 2) Pharmaceuticals & Biotechnology & Life Sciences. However, the description of the subgroups is broad so we have identified individually the institutions in which we believe that the platform can develop its technological potential: The left column shows the *Lead users* or those organisations that would be willing to pay for its use, the second column explains the purpose, while the third indicates the priority that the technology would have for these companies or the level of affinity according to their capabilities. Although obviously its progress will depend on multiple factors inherent to the product and service, design, speed to market, among other aspects; **represent possibilities of introduction as a business model in the medium or long term and invite to further develop applications to satisfy these segments** (some of the market segments and lead users are expanded in the following section):

Table 11: Potential Opportunities and Possible Business Models

HBP Current Technology: MIP		
LEAD USERS (Would pay for specific services)	**OBJECTIVE:**	**POSSIBILITY OF ESTABLISHING A RELATED SPIN-OFF IN MEDIUM-TERM (*):**
Two groups according to the Global Industry Classification Standard:		
1) HEALTHCARE EQUIPMENT & SERVICES (Health Care Equipment & Supplies, Health Care Providers & Services, Health Care Technology):		
IoT Companies (To Healthcare)	Interconnected system in hospitals and clinics. If this platform connects to smart devices then we will have better quality and interconnectivity of the precise healthcare system from the patient's home to their journey to the clinic as well as within the clinics, also with pharmacies and medical suppliers.	High

AR-VR Companies	Clinical intervention and simulation in procedures and evolution of the disease. Only if it has the possibility that some results can be displayed virtually, in case it is useful for certain pathologies. Also, digital twins could be a possibility in the near future.	Medium (depends on the speciality)
Other Healthcare Technology Startups	Data entry, model validation.	(To define separately, due to the variable number of related startups)
Rare Diseases Platform Databases	Provide Specialised Datasets (i.e, Orphadata).	High
Plasma Fractionation Companies	Databases for their drug development (including rare diseases).	Low (heavily regulated)
Market Research Organizations	Data availability for market research and new reports.	Medium (more useful for CRO)
Contract Research Organizations (CRO)	Data availability for health research, clinical trials protocols and new reports.	High
Company Software of Laboratory Information Systems (LIS)	They distribute the software to private laboratories and manage the sample and analysis data (i.e technidata medical software).	High
EHR and EMR Software Platforms	Digitization and management of data and medical records, facilitating the integration of data into the platform.	
2) PHARMACEUTICALS, BIOTECHNOLOGY & LIFE SCIENCES (Biotechnology, Pharmaceuticals, Life Sciences Tools & Services)		
Hospitals & Clinics	To provide technological solutions for diagnosis, prevention and therapy based on data storage and management, AI, etc.	High (Only as an important local and federated platform for patient data management for diagnosis, prevention and treatment).
Medical Research Foundations		
Clinical and Translational Research Centres		
Blood Banks (samples to hospitals and clinics)	Digital control of blood donations and transfusions, integrated with the rest of the platform.	Medium
Private/Commercial Laboratories (Pathology, microbiology, biochemistry)	Digital control of pathology, microbiology, biochemistry samples integrated with the rest of the platform.	
Biobanks (samples to research)	Digital control and management of biological samples (stem cells, serum hematology, blood, plasma, tissue, etc.)	High
Pharmaceutics	Supporting pharmaceutical industry research on drug administration and condition evolution, taking into consideration other variables already integrated in the MIP.	Medium (only as supporting platform to pharmaceutics research and trials)

Clinical Trials	Databases for support Clinical Trials research.	High (As supporting platform to research)
Other Research Institutes	To provide technological solutions for diagnosis, prevention and therapy based on data storage and management, AI, etc.	To evaluate
Universities	As part of a training process prior to medical practice.	High (only as supporting platform to training and research)

HBP Current Technology: MIP		
FINAL CUSTOMERS (Beneficiaries)	**CHALLENGES FACED BY THE MIP FOR FINAL CUSTOMERS:**	**POSSIBLE SOLUTIONS**
1) HEALTHCARE WORKERS:		
Physicians	Integrate the platform as a key reference platform in diagnostics and facilitate its learning. Facilitate the integration of EHR when these usually are incomplete, with typos, or inconsistent. Harmonisation of medical terminology: Structure a more precise and consistent ontology in its definitions compared to the current ones.	Incorporate it as part of the educational programme in medical careers (at least as an optional subject) and educational books. Promote a regular use in medical teams, demonstrating greater accuracy in diagnosis with personalised algorithms.
Surgeons		
Radiologists		
Pharmacists		
Medicine Students		
Hospitals Managers & Coordinators	Facilitating and simplifying the process of incorporating heterogeneous data.	Tutorials and in situ cases with certified development staff and training for medical teams.
Data Administrators		Quality control tools for data transfer, automatic data cleaning
2) RESEARCH WORKERS		
Clinical Investigators	Integrate the platform as a key reference platform in diagnostics, clinical trials and facilitate training.	Promote their participation: Feedbacks, short but precise meetings, tutorials, and regular interaction with engineers and technicians.
Other Research Specialists	Reduce the gap between related researchers and the community.	Promote their participation, feedbacks and learning.
Engineers	To compose specialised and certified teams for installation, consulting and maintenance support.	Provide professional certificates of recognition. Offer timely support.
Programmers and other Specialised IT workers	Need for more data specialists. Achieve smoother and continuous integration between the knowledge of data scientists and the medical workers.	Provide professional certificates of recognition, periodic interactions between physicians, engineers and technicians.

3) OTHERS:		
Patients Associations	To get new feedback and perspectives from patient association representatives that can be useful for integrating the platform and improving diagnostics (1)	A business model that includes corporate social responsibility, collaborating with these organisations which also promote activities in support of research and dissemination of knowledge, relevant for the MIP.
Governments	Provide up-to-date, reliable and accurate data on global health conditions according to specific regulations.	Join European Parliament hearings and meetings (as a spin-off).
International Healthcare Organisations	Reach a strategic meeting point with public organisations in the healthcare sector that can facilitate reaching new and potential markets. Also, to benefit the medical sector that resides in vulnerable regions or with low digitisation of their services, providing them with valuable tools and new knowledge.	Create initiatives together with healthcare organisations, strategic partnerships with companies to reach foreign segments.
Foundations	Introduce the platform as tool for training doctors and specialised researchers affiliated to these centres, in order to capture the interest and promote its use.	Partnering with local foundations diversified into specific activities (2)
Business ecosystems	Growth alternatives through the right networks. Explore new markets and contacts. Gain trust and brand visibility.	Participation in corporate events, forums, think tanks, special summit initiatives (as a spin-off).

(1) Other authors suggest that although the patient's perspective does not contribute to a technical aspect, the social component provides a broader view on the development of diseases. (Wehling et al, 2015)

(2) In Spain, for example, the scope of activities managed by the foundations are aimed at education and research 21,57% and health 5,44%, respectively. (AEF, 2020)

Based on the previous structuring about the members of the industry coupled with the challenges faced by the technology to reach them, **we continue with a more detailed analysis on its initial segmentation.** This will make it possible to formulate actions plans in line with the functionalities of the software vis-à-vis hospitals (its main users) not only by the fact it has been installed in several of these centres, but also by the growing demand to treat the burden of neurological diseases which, among other aspects, include regional differences in terms of accessibility, medical equipment, limited availability of clinicians or training. A next segmentation at this stage derives in the market of clinical trials in neurology, being an activity that brings together lead users of interest for the platform -both hospitals and pharmaceuticals, CRO, sponsors- and that, despite being organisations with different ambitions they need diversified tools that help process the enormous amount of data generated without being infringed, accelerating their R&D processes, and ultimately converging as common objectives. The MIP since its creation has been evolving in a constant federated learning and must continue to deploy analytical models for its different lead users, so market segmentations must undoubtedly be aligned with the acquisition of specific data. Starting with the main ones:

MARKET SEGMENTATION I: THE HOSPITAL SECTOR AND AVAILABILITY OF MENTAL HEALTH SPECIALISTS WORLDWIDE

Two standards were developed: The System of Health Accounts (SHA) which includes the International Classification of Health Accounts (ICHA), by OECD in the year 2000; and the Guide to Producing National Health Accounts or NHA Producers Guide, between the WHO, World Bank and USAID. OECD/ Eurostat/WHO (2017)

The types of providers in health care have been extensively studied in recent years by renowned international organisations such as OECD, WHO, Eurostat (2017), for example, creating remarkable standards for the administration and control of financial flows related to the health system. As far as this section is concerned, within the System of Health Accounts (SHA) can be found the Classification of health care providers (ICHA-HP) since 2011, which comprises an organised and detailed description of the broad range of good and services in health offered in different countries, also related to their financial classification. Based in this description, figure 14 shows the number of hospitals active in 34 countries, initially following the latter classification (HP.1.0) and simplifying some of the ramifications reflected by the OECD: Hospitals -which also include general hospitals (HP.1.1), mental health hospitals (HP.1.2), specialised hospitals (HP.1.3)-

and For-profit privately owned hospitals (Fig. 15), constituting the broadest groups. **The variation in 2012 and 2018, respectively, is reflected in more recent data publicly, including all the categories mentioned:**

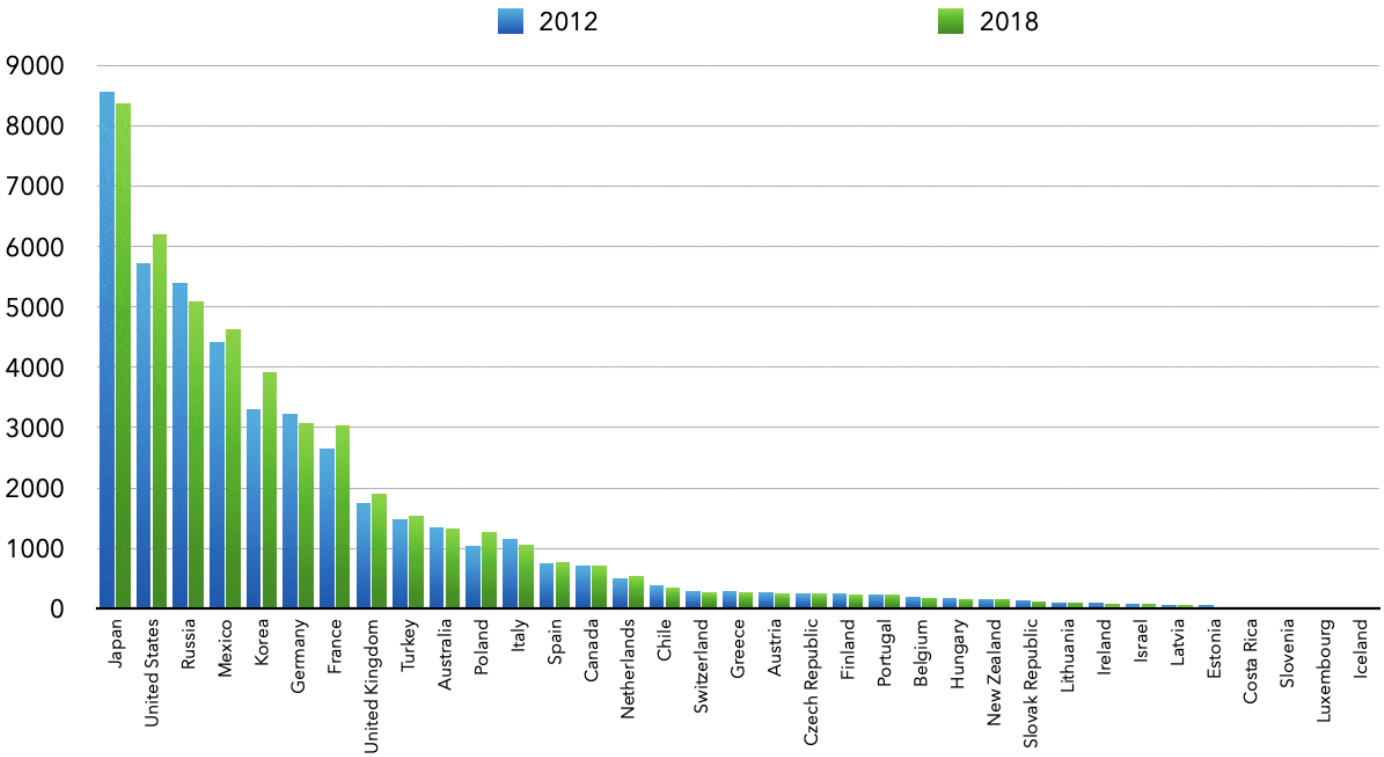


Figure 14. Total number of public hospitals in 34 countries worldwide (2012-2018). Data from: OECD 2021. *Some data had to be obtained outside OECD, such as Russia, the most recent shown there are approximately 5,4 thousand hospitals. *The Russian Federal State Statistics Service (2021)*

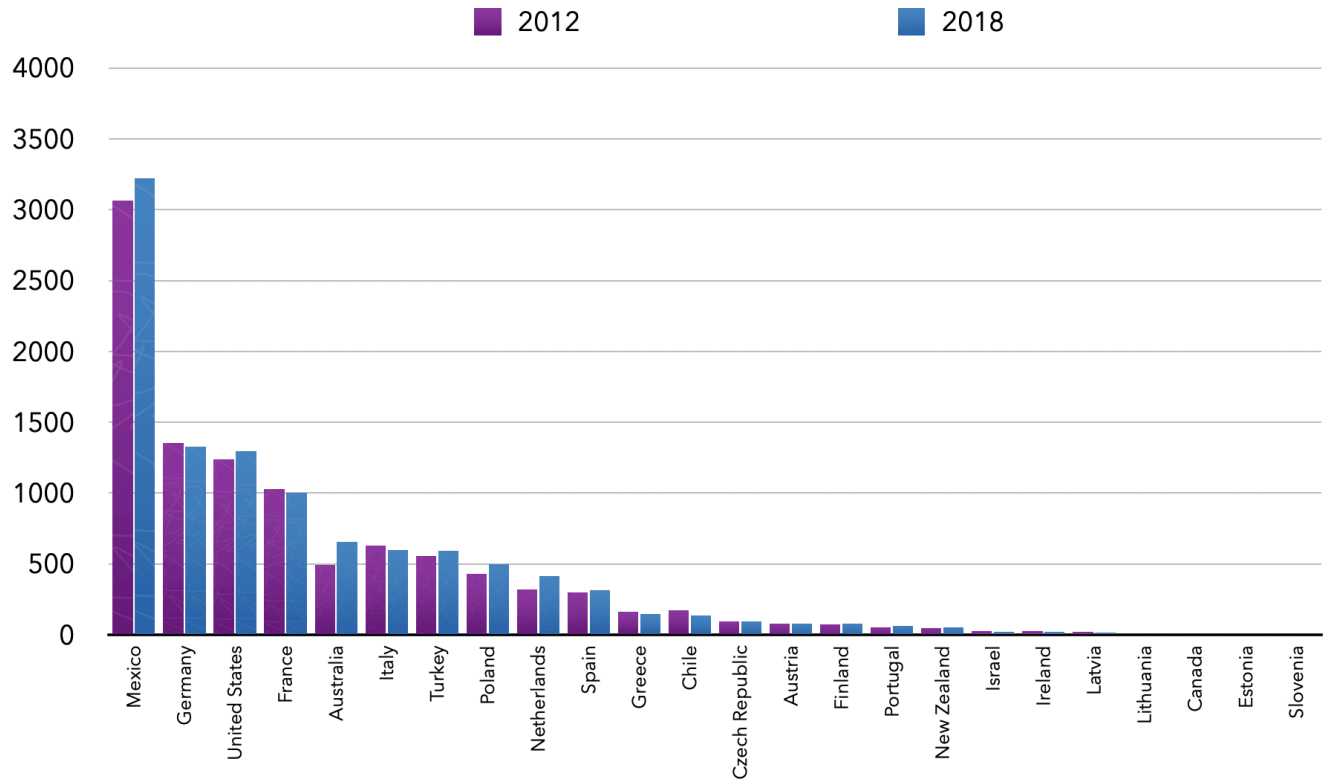


Figure 15. For-profit privately owned hospitals in 24 countries worldwide (2012 and 2018). Data from: OECD 2021

The lack of neurologists is a problem worldwide, and in the HICs concrete actions are being implemented to increase their availability. Most of these professionals experience burnout given the high workload in healthcare facilities today (Burton, 2018).

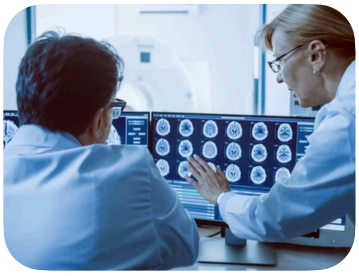


Image: Adobe Stock

Although the hospitals in which the platform has been installed have neurology services with a high volume of patients and a wide operability of its medical teams, **we also recommend its installation in the centres furthers from the big cities or with low population density**. This could greatly benefit those medical teams and researchers limited in their research experience and with fewer specialists, given the available centres reflected in Figures 14 and 15, respectively. Starting mostly form the European market, the following EAN table simplifies the recent scenario on the number of neurologists available per patient, in which about 85,000 specialists would be qualified in EAN/WHO regions of Europe, equivalent to 10,000 patients per neurologists, emphasising that these differences are further accentuated between regions: "**There are huge disparities across European countries, ranging from 2,500 patients to 46,000 patients per neurologist**" (EAN, 2019).

These differences are also explained by the effort of some European countries previously to establish changes in their structural policies, to optimise the costs dedicated to the health sector, which has influenced this disparity in specialists in brain diseases. Sweden is one of many examples in adoption of these strategies to control healthcare costs, which has led to a withdrawal of health professionals -and these specialists are not exception- "They started to reducing the size of health settings and the decreasing the number of healthcare professionals in 2000" (Ross et al, 2000 mentioned by Matin et al, 2019). Despite this, these changes have favoured the greater demand for technologies that help compensate for these shortcomings and **try to reduce the pressure on the sector in the face of the next challenges in mental health worldwide**, an example to demonstrate this will soon be described with teleneurology, a service that can be complemented by the MIP platform.

	Population	Prevalence of neurological diseases	Deaths due to neurological disease	Number of neurologists
European Union (EU28)	512,355,000	307,859,199	1,116,038	43,306
Western Europe	**432,969,000**	**260,827,756**	**892,162**	**36,009**
Central Europe	**114,803,000**	**67,368,506**	**300,317**	**9,031**
Eastern Europe	**210,199,000**	**130,372,328**	**604,144**	**29,016**
WHO Region Europe	925,631,000	542,935,521	1,981,463	83,397

Table 12: The burden of neurological disease in Europe. Source: European Academy of Neurology (EAN, 2019). **Bolded** regions where the MIP is deployed according to this data.

Furthermore, other similar analyses share that the global median of neurological workforce is 3.1 per 100,000 population, while in Europe slightly more optimistic data put the average at 9 per 100,000 inhabitants (WHO, 2017), also reflecting concerns about the profound differences in the number of specialists per patient. Clearly, considering the increase in neurodegenerative diseases associated with the platform, the limited availability of neurologists, also related to an increased workload for the medical and research service with limited resources; specialised operating centres will require with greater justification the use of software (including the lead users of the table 13).

DEVELOPING A MARKET STRATEGY

Considering the initial deployment in some hospitals mostly in Western Europe and Central Europe, it would be appropriate for the team to delve into the centres where the availability of neurologist equipment is lower and help to reduce these gaps. The following tables expand the number of specialists per patient adapted to the current MIP deployment (selected from 155 countries), to support the future spin-off as part of its lead users, the key regions for its introduction, both in Europe and beyond its borders, including the MENA countries which have shown significant growth in the hospital sector and pharmaceutical industry (see also *Clinical trials in neurology and the MIP*).

Table 13. Availability of Mental Health Specialists Worldwide, Neurosurgical Centres (including university hospitals) and Psychiatric Hospitals & Other Mental Health Centres.

		Final Users			Lead Users		
	Country	**Psychiatrists working in mental health sector** (per 100,000 population) (2)	**Nurses working in mental health sector** (per 100,000 population) (2)	**Psychologists working in mental health sector** (per 100,000 population) (2)	**Total Neurosurgical Centres** (Including University Hospitals) (3)	**Total Psychiatric, Specialist Hospitals & Other Mental Health Centres** (Latest data available) (4)	**Main References**
		Western Europe (1)					
★	Belgium (HIC)	18	125.69	10.46	43	53 mental hospitals, 63 psychiatric units in general hospitals, 128 residential care, 2 forensic (inpatient). 101 mental health centres (outpatient)	(2) WHO (2020), Civio (2021); (3) EANS (2021); (4) Baig and Delvenne (2008), Mental Health Europe (2017)
★★	France (HIC)	20.907	98.017	48.704	47	182 departments in general hospitals, 90 in specialists hospitals, 63 acute specialist hospitals (not for profit), 150 acute specialist hospitals, (private). 9 UHSAs created for inmates/ forensic*	(2) WHO (2020); (3) EANS (2021); (4) Mental Health Europe (2017); Insee (2020). *The forensic psychiatry is not yet an academic field in France (Fovet et al, 2020)

	Final Users			Lead Users		
Country	**Psychiatrists working in mental health sector** (per 100,000 population) (2)	**Nurses working in mental health sector** (per 100,000 population) (2)	**Psychologists working in mental health sector** (per 100,000 population) (2)	**Total Neurosurgical Centres** (Including University Hospitals) (3)	**Total Psychiatric, Specialist Hospitals & Other Mental Health Centres** (Latest data available) (4)	**Main References**
Ireland (HIC)	6.1	-	21.17	3	66 (28 psychiatric hospitals, 22 general hospital psychiatric, 1 forensic inpatient, others)	(2) WHO (2020), Civio (2021); (3) EANS (2021); (4) Daly and Craig (2019)
Germany (HIC)	13.202	-	49.555	145 centres. 184 neurosurgery departments, 451 neurology departments.	407 psychiatry - psychotherapy, 262 psychotherapy - psychosomatic medicine, 147 child - adolescent psychiatry (departments associated to hospitals), 77 forensic	(2) WHO (2020); (3) EANS (2021), Destatis (2019); (4) Destatis (2019)
Monaco (HIC)	31.326	83.536	53.515	2	2 psychiatric units in general hospitals	(2) WHO (2020), WHO (2011); (3) EANS (2021)
Netherlands (HIC)	20.870	-	123.464	21	114 mental health institutions, 30 of them inpatient and outpatient, 2 general mental hospitals, 479 nursing homes	(2) WHO (2020); (3) EANS (2021); Kroneman et al, (2016)
Switzerland (HIC)	43.956	92.659	84.137	18	51 psychiatric hospitals, 56 rehabilitation clinics	(2) WHO (2020); (3) EANS (2021); (4) Office fédéral de la statistique OFS (2021)
United Kingdom (HIC)	8 England, 10 Scotland, 6 Wales, 8 Northern Ireland	-	24.61	38	54 NHS, 221 independent mental health services. *Inpatient*- NHS: 53 psychiatric units, 30 child - adolescent psychiatry, 46 rehab, 53 wards for older people and 37 with disability; Independent: 33 psychiatric units, 24 child - adolescent psychiatry, 88 rehab, 40 wards for older people, and 41 with disability.	(2) WHO (2020),Civio (2021), RCPSYCH (2017); (3) EANS (2021);(4) CQC (2017)
	Northern Europe (1)					
Estonia (HIC)	16.188	23.484	6.460	2	9 psychiatric units in general hospitals, 2 mental health hospitals, 46 special care institutions, 3 child - adolescent psychiatry, 1 forensic. (Inpatient)	(2) WHO (2020); (3) EANS (2021); (4)) WHO (2020)
Denmark (HIC)	18.9	-	54.31	4	30 psychiatric units in general hospitals (Inpatient)	(2) WHO (2020),Civio (2021), Eurostat (2018); (3) EANS (2021); (4) Mental Health Europe (2017)
Finland (HIC)	23.586	51.970	109.486	5	41 psychiatric units in general hospitals, 168 residential care facilities, 26 child - adolescent psychiatry, 3 forensic. (Inpatient)	(2) WHO (2020); (3) EANS (2021); (4) WHO (2020)
Latvia (HIC)	9.987	22.834	18.72	6	6 mental health hospitals, 3 psychiatric units in general hospitals, 3 child - adolescent psychiatry, 1 forensic. (Inpatient)	(2) WHO (2020),Civio (2021); (4)) WHO (2020)

	Final Users			Lead Users		
Country	Psychiatrists working in mental health sector (per 100,000 population) (2)	Nurses working in mental health sector (per 100,000 population) (2)	Psychologists working in mental health sector (per 100,000 population) (2)	Total Neurosurgical Centres (Including University Hospitals) (3)	Total Psychiatric, Specialist Hospitals & Other Mental Health Centres (Latest data available) (4)	Main References
Lithuania (HIC)	18.452	49.763	15.860	7	4 mental health hospitals, 20 psychiatric units in general hospitals, 1 forensic. (Inpatient). 115 mental health centres (outpatient)	(2) WHO (2020), (3) EANS (2021); (4) WHO (2020)
Norway (HIC)	48.040	-	73.522	6	30 mental health hospitals, 25 psychiatric units in general hospitals, 24 child - adolescent psychiatry, 29 forensic. (Inpatient). 36 mental health centres (outpatient)	(2) WHO (2020), (3) EANS (2021); (4) WHO (2020)
Sweden (HIC)	20.863	50.566	58.32	7	(Psychiatric units associated to general hospitals: 4,800). 20 forensic	(2) WHO (2020),Civio (2021); (3) EANS (2021); (4) WHO (2020)
	Southern Europe (1)					
Cyprus (HIC)	11.6	34.970	-	6	1 mental health hospitals, 2 psychiatric units in general hospitals, 1 child - adolescent psychiatry (Inpatient).	(2) WHO (2020); (3) EANS (2021), Eurostat (2018); (4) WHO (2020)
Greece (HIC)	5.803	12.748	7.54	41	3 mental health hospitals, 37 psychiatric units in general hospitals (public) and 36 (private), 7 child - adolescent psychiatry, 2 forensic (Inpatient). 63 mental health centres (outpatient)	(2) WHO (2020),Civio (2021); (3) EANS (2021); (4) Mental Health Europe (2017), WHO (2020)
Italy (HIC)	5.978	23.492	8.62	148	(Psychiatric units associated to general hospitals). 354 psychiatric units in general hospitals, 25 child - adolescent psychiatry.	(2) WHO (2020),Civio (2021); (3) EANS (2021); (4) WHO (2020)
Portugal (HIC)	13.4	-	9.60	14	3 mental health hospitals, 30 psychiatric units in general hospitals, 2 child - adolescent psychiatry, 1 forensic (Inpatient). 3 mental health centres (outpatient)	(2) WHO (2020),Civio (2021); Eurostat (2018), (3) EANS (2021); (4) WHO (2020)
Spain (HIC)	9.694	2.873	5.60	256 (93 public, 164 private)	93 mental health hospitals, 477 psychiatric units in general hospitals, 2 child - adolescent psychiatry (Inpatient). 358 mental health centres (outpatient)	(2) WHO (2020),Civio (2021); (3) EANS (2021); (4) Ministerio de Sanidad (2016), WHO (2020)
	Central and Eastern Europe (1)					
Austria (HIC)	16.14	-	-	11	8 mental health hospitals, 23 psychiatric units in general hospitals, 12 child - adolescent psychiatry, 2 forensic (Inpatient). 31 mental health centres (outpatient)	(2) WHO (2020); (3) EANS (2021); (4) Mental Health Europe (2017), WHO (2020)

	Final Users			Lead Users		
Country	Psychiatrists working in mental health sector (per 100,000 population) (2)	Nurses working in mental health sector (per 100,000 population) (2)	Psychologists working in mental health sector (per 100,000 population) (2)	Total Neurosurgical Centres (Including University Hospitals) (3)	Total Psychiatric, Specialist Hospitals & Other Mental Health Centres (Latest data available) (4)	Main References
Bosnia and Herzegovina (UMIC)	8.032	22.144	1.527	6	3 mental health hospitals, 19 psychiatric units in general hospitals, 1 forensic (Inpatient). 10 mental health centres (outpatient)	(2) WHO (2020); (3) EANS (2021); (4) Mental Health Europe (2017), WHO (2020)
Bulgaria (UMIC)	7.203	17.736	1.867	18	54 residential institutions,12 mental health hospitals, 22 psychiatric units in general hospitals, 2 child - adolescent psychiatry, 2 forensic (Inpatient). 3 mental health centres (outpatient)	(2) WHO (2020); (3) EANS (2021); (4) Mental Health Europe (2017), WHO (2020)
Czechia (HIC)	12.364	30.348	2.518	15	18 mental health hospitals, 30 psychiatric units in general hospitals, 14 child - adolescent psychiatry, 5 forensic (Inpatient).	(2) WHO (2020) (3) EANS (2021); (4) Mental Health Europe (2017), WHO (2020)
Hungary (HIC)	11.079	33.739	2.484	15	259 psychiatric units in general hospitals, 19 child - adolescent psychiatry, 1 forensic (Inpatient), 91 community care for mental patients. 467 mental health centres (outpatient)	(2) WHO (2020); (3) EANS (2021); (4) WHO (2020), Mental Health Europe (2017), Ács et al (2019)
Poland (HIC)	24.176	30.916	9.60	54	48 mental health hospitals, 100 psychiatric units in general hospitals, 32 child - adolescent psychiatry, 41 forensic (Inpatient). 132 mental health centres (outpatient)	(2) WHO (2020), Civio (2021); (3) EANS (2021); (4) WHO (2020)
Romania (UMIC)	5.665	18.710	1.479	53	34 mental health hospitals, 78 psychiatric units in general hospitals, 22 child - adolescent psychiatry, 4 forensic (Inpatient). 116 mental health centres, 2 Crisis Centre, 67 neuropsychiatric recovery and rehab (outpatient)	(2) WHO (2020); (3) EANS (2021); (4) WHO (2020), Mental Health Europe (2017)
Slovenia (HIC)	11.953	36.727	22.00	2	5 mental health hospitals, 1 psychiatric units in general hospitals, 2 child - adolescent psychiatry, 1 forensic (Inpatient). 10 mental health centres (outpatient)	(2) WHO (2020),Civio (2021); (3) EANS (2021); (4) WHO (2020), The Association of Health Institutions of Slovenia (2021)
Ukraine (LMIC)	6.930	-	-	56	57 mental health hospitals, 23 psychiatric units in general hospitals, 2 child - adolescent psychiatry, 1 forensic (Inpatient). 22 mental health centres (outpatient)	(2) WHO (2020); (3) EANS (2021); (4) WHO (2020)

Final Users | **Lead Users**

Country	Psychiatrists working in mental health sector (per 100,000 population) (2)	Nurses working in mental health sector (per 100,000 population) (2)	Psychologists working in mental health sector (per 100,000 population) (2)	Total Neurosurgical Centres (Including University Hospitals) (3)	Total Psychiatric, Specialist Hospitals & Other Mental Health Centres (Latest data available) (4)	Main References
Russia (UMIC)	8.479	-	2.404	334 neurosurgical departments	195 mental health hospitals (Inpatient). 3,356 mental health centres (outpatient)	(2) WHO (2020), Krylov et al, 2017; (3) EANS (2021); (4) WHO (2020)
	North America (1)					
Canada (HIC)	14.679	68.659	48.740	14 academic centres, unknown total	35 psychiatric units in general hospitals (Ontario). 4 psychiatric hospitals, 6 specialty psychiatric tertiary hospitals, 200 general hospitals with mental health beds, 350 community agencies	(2) WHO (2020); (3) Staudt (2018); (4) Annual Report of the Office of the Auditor General of Ontario (2016)
Mexico (UMIC)	0.21	0.22	3.46	-	33 mental hospitals, 8 psychiatric units in general hospitals, 3 child - adolescent psychiatry, 2 forensic. 60 mental health centres (outpatient). 863 psychiatry offices	(2) WHO (2020); (3) ; (4) Díaz-Castro (2020); WHO (2020)
United States of America (HIC)	10.542	4.283	29.864	5,700 hospitals (1,600 trauma care, 1,000 stroke, 200 children's hospital)	605 mental hospitals, 1117 psychiatric units in general hospitals, 67 child - adolescent psychiatry (inpatient)	(2) WHO (2020); (3) AANS (2012); (4) WHO (2020)

Country	Psychiatrists working in mental health sector (per 100,000 population)	Nurses working in mental health sector (per 100,000 population)	Psychologists working in mental health sector (per 100,000 population)
	South America (1)		
Argentina (HIC)	21.705	-	222.572
Brazil (UMIC)	3.165	34.948	12.368
Chile (HIC)	6.970	1.880	-
Costa Rica (UMIC)	3.931	5.699	142.018
Dominican Republic (UMIC)	2.280	-	7.598
Panama (UMIC)	4.031	3.452	7.331
Peru (UMIC)	2.948	5.995	9.507
	Asia (1)		
Armenia (LMIC)	3.840	11.245	788
Azerbaijan (UMIC)	3.452	6.717	1.165
China (UMIC)	2.199	5.423	-
India (LMIC)	292	796	69

Country	Psychiatrists working in mental health sector (per 100,000 population)	Nurses working in mental health sector (per 100,000 population)	Psychologists working in mental health sector (per 100,000 population)
Indonesia (LMIC)	310	2.518	175
Japan (HIC)	11.867	83.805	3.037
Republic of Korea (HIC)	5.793	13.662	1.591
Malaysia (UMIC)	1.048	6.838	1.029
Thailand (UMIC)	721	7.418	1.749
	Oceania (1)		
Australia (HIC)	13.525	90.582	103.036
New Zealand (HIC)	28.540	75.132	-
	Middle East/North Africa (MENA)		
Bahrain (HIC)	5.467	27.918	1.239
Egypt (LMIC)	1.600	4.799	256
Iran (Islamic Republic of) (UMIC)	2.016	9.451	5.166
Iraq (UMIC)	343	1.218	111
Israel (HIC)	9.870	-	88.089
Jordan (LMIC)	1.125	3.297	1.266
Lebanon (UMIC)	1.213	3.145	3.298
Morocco (LMIC)	839	2.451	575
Oman (HIC)	1.738	3.000	786
Qatar (HIC)	2.712	9.933	1.413
Saudi Arabia (HIC)	1.321	10.660	2.034
Syrian Arab Republic (LMIC)	368	1.068	1.068
Tunisia (LMIC)	-	195	9
United Arab Emirates (HIC)	1.649	4.370	765
Yemen (LMIC)	201	323	409
Other African countries			
South Africa (UMIC)	1.521	-	-

Acronyms:
- HIC= High Income country
- UMIC= Upper middle-income country
- LMIC= Low middle-income country

Market to be explored

Countries with MIP installed

Further deployment in centers with lowest specialists is required

Additional notes:
- (1) Geographical sub-regions following the EuroVoc classification.
- (4) We are excluding residential care facilities and other community-based residential arrangements to focus on hospital and inpatient care, relevant for startups of this kind.

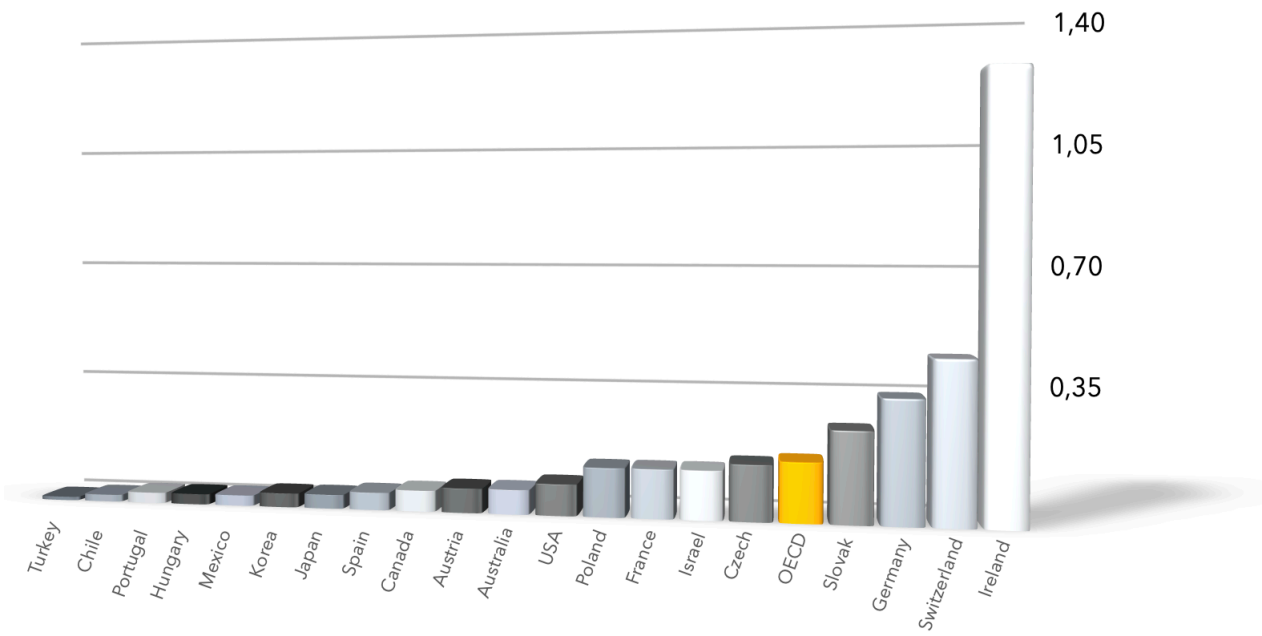
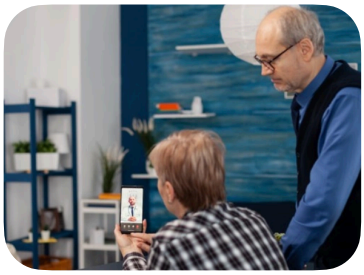


Figure 16. Public psychiatric hospitals per 100 000 population in selected OECD countries, 2010. Data: OECD (2014)

In order to formulate optimal and precise strategies for the internationalisation of the startup in the European market initially, we have considered some of the most relevant related factors. From the burden of neurological diseases, geographical proximity to the operational centre, similarity in IT systems in the hospital structure, less barriers to entry, lead users -number of public and private hospitals, specialised neurology centres, university hospitals-, availability of final customers, neurologists, psychiatrists, mainly-, up to the analytical income classification per country according to the international references of the World Bank and, although it is a broader indicator, the healthcare expenditure relative to GDP per country has also been previously evaluated. In **Europe**, it could be considered to formulate positioning strategies in regions such as:

Western Europe with greater advantage in sectors where it is installed: Germany, France, Belgium, Austria, Switzerland (head-quarters). Also, explore strategies for its deployment next in Netherlands, United Kingdom, Ireland. **Central and Eastern Europe** starting from the installation points, countries such as Bulgaria, Czechia or Russia, as well as exploring nearby markets in Hungary, Poland, Romania or Slovenia. In **Southern Europe** highlights Spain, Italy, Portugal, Greece, while in **Northern Europe** markets in Denmark, Sweden, Finland or Lithuania as starting points.

Image: Freepik



In India teleneurology has excelled in recent years in randomised controlled trials (RCT). In the United States this service increased dramatically during the pandemic by 2,000%, while in some European countries such as Italy heavily affected by COVID-19 -teleneurology offered by multidisciplinary teams for Parkinson's patients demonstrated its potential and a diversified organisational scheme. (Patterson, 2021)

Finally, other aspects that can be extracted from this industry related to diagnostics services in neurology include the increase in digital healthcare, especially teleneurology in hospitals in different corners of the world, with an emphasis on Europe, North America, or Asia; overcoming legal barriers in countries such as India, which were maintained before the pandemic, as it was not clear whether or not these services were a legal practice (Patterson, 2021). While, in other geographical points of Latin America, as in Brazil (Barros Domingues et al, 2000), for example, telemedicine[22] shows an important evolution, although with it, a greater intensity in the regulation of these services.

In this sense, the digital transformation driven by **adverse conditions has led researchers and neurologists to test their skills from multiple** computer applications and portable devices, so the possibility of incorporating the MIP to this type of services in these markets would also be a strategy and a differentiating component in an industry that continues to expand. To reinforce these considerations with recent data, according to a report by Grandviewresearch (2021), the market of Hospital Information Systems (HIS) approached USD 13.26 billion in 2015, projecting a growth of 11.4% annual; while in the neurotechnology market, industry reports share a CAGR of 12% over the period 2017-2023, with significant leadership from the United States, Europe and Asia Pacific (Mishra, 2021). Finally, other recent data such as those from CB Insights (Figure 17) also reinforce these perspectives.

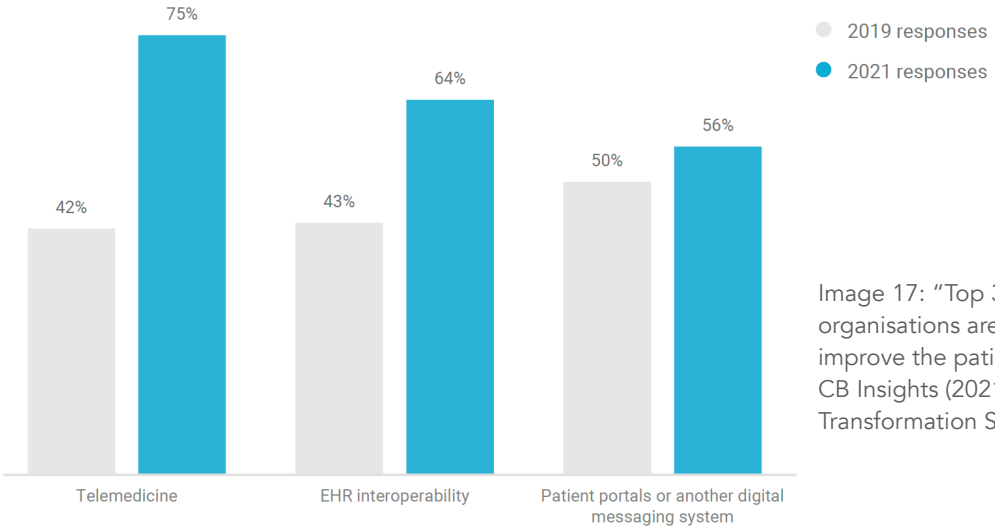


Image 17: "Top 3 areas where healthcare organisations are currently investing to improve the patient experience" Source: CB Insights (2021) Healthcare Digital Transformation Survey.

[22] Telemedicine in Brazil is understood as "The practice of medicine mediated by technologies for the purpose of assistance, research, prevention of illness and injuries and the health promotion". Law N° 13.989, of 15 April 2020. http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2020/Lei/L13989.htm

Other element to consider is the variation of hospitals infrastructures in Europe, which include lead users in the markets of interest. Some examples such as Portugal, that ten years ago had about 127 hospitals have almost doubled in number, around 238 for 2019 (Instituto Nacional de Estadística, 2019). In Russia, (where the platform is also deployed), although publications on the number of centres are limited, one of its official bodies suggests that for 2016 there were about 5,4 while for 2019 the outpatient clinics for example, stood at 5,8; respectively.

MARKET SEGMENTATION II: THE CLINICAL TRIAL MARKET IN NEUROLOGY AND THE MIP

The Global Burden of Disease (GBD) is a method that quantifies the degree of health loss caused by diseases, injuries and risk factors such as age, sex and geography at specific periods of time. The World Bank (2010)

An important consideration lies in the evolution of clinical trials in neurology worldwide, as well as the clinical factors that affect their growth and results in the field of **neurodegenerative diseases, psychiatric disorders and neurodevelopmental conditions**. Indeed, the global need to continue investing resources and research efforts in healthcare is growing, together with the testing and efficiency responsible for new treatments through clinical trials, understanding the differences between regions, for example, developing countries have an increasing risk for disease burden and where less than 20% of the studies are carried out (Alemayehu, 2018). **The MIP could support doctors and research groups in these countries in terms of data for new studies,** offering its structure for experiments in a federated way.

However, its application at the local level will depend on the electronic prior registration of the medical record and the activities inherent in the study to be subsequently incorporated into the platform. Nonetheless, many of these organisations have poor electronic records or do not have a software or IT systems yet to do so -of course this varies between Low Income Countries LIC or Low Middle Income Countries LMIC-, and therefore a great barrier to processing data on the platform. Probably for those clinical trials of smaller populations and depending on some conditions of the Data Catalogue, could be a benefit to these countries and incorporate

potential or new Common Data Elements. Though, these reflections will remain to be confirmed.

Another concern arises about the increase and **premature completion of clinical trials**, contrasted in multiple perspectives by the evident rise of industry-sponsored trials, 4% from 1975 to 57% in 2004 (Booth et al, 2008; mentioned by Buergy et al, 2020). Naturally, failure to complete them also does not exempt from the investment of human and economic resources that are not recovered, together with the process of selecting and recruiting patients for studies[23] (Rimel, 2016); which raises deeper concerns such as the opportunity cost of not accepting clinical trials that anticipate high financial investment and limited recruitment (Vickers, 2015). This latest study proposes some considerable solutions to try to address these problems, for example, simplifying the patient eligibility criterion, integration trial data with electronic medical records (as explained in diagram 2). Other authors assert the usefulness in determining the factors that can predict low accrual in trials before starting them (Bennette et al, 2015), that is in its designing process, which could help to select which trials have a higher priority, directing as best as possible the efforts and resources established.

FEDERATED LEARNING AS PART OF THE SOLUTION

Image: Freepik - Rawpixel

Between 15-50% of completed clinical studies are never published (Zou et al, 2018).

Considering that in some cases the data generated from randomised clinical trials is stored separately or not integrated with the patient's history in standard practice and is not exactly an advantage for physicians (Vickers, 2015). **The MIP could become as an invaluable platform to take advantage of the information generated by these clinical trials, as well as for those where the final results are not available**. It might be inferred that the data generated during the development phases, not only could contain some useful information for the doctor in his diagnoses, associated comorbidities or for future reference, but it would also help other medical groups to re-evaluate the hypotheses from new trials or put that information to better use. There is a possibility of integrating through the MIP, if **an alternative can be enabled to store this data of unfinished trials with no results and provide a useful perspective for future research.**

[23] In the United States, recruiting a single patient for a clinical trial can achieve 50,000$, about 42,000€. (Vickers, 2015)

In this context, there are various explanations in the literature for the designated data repository for clinical trials, and furthermore, they cannot necessarily be integrated with EMR/EHR, which as Vickers explained, is a problem for medical workers, who should be aware of the researcher's findings. Other studies explain that the nature of designated repositories can be diverse, for example, those devoted exclusively to clinical research data, specialised in a specific disease or repositories established by the institution itself, without recommendations on this, but on the processes (Ohmann et al, 2017). For this report it is important to consider it because in additional to the completed clinical trials whose data is recorded in the EHR and therefore in the MIP -under the scheme i2b2 explained in previous section-, those clinical trials sponsored and that ceased prematurely, could contain information not negligible for future research and enhance the MIP value, which also manages the data anonymously and federated.

Returning to the example of COVID-19, the gap between the number of clinical trials recorded and those that published their results in public repositories, makes it evident that there is a need for more communication on these data, in the face of such critical and severe situations for the human being. Rodgers et al (2021), reflect of the 3,754 completed trials, less than half (40,4%) did not publish results, among other ideas comment that this lack of data results in treatments based on incomplete information, potentially greater damage, in other words, asserts that doctors need greater support from trial sponsors regarding the publication of lost results, argue (Ibid, 2021): "We therefore recommend caution in experimental drug use for non-severe disease and urge trial sponsors to report missing results retrospectively".

Could the MIP help in this context? Or perhaps, could the MIP support the clinical trials in this definition of factors involved in its design, affecting the initiation of such trials in industry?

The platform could exert a favourable influence on the formulation and design of clinical trials related to the diseases in which specialises, accelerating the analyses of the variables associated with the Metadata, allowing to anticipate relationships between them and discard other premises. This would save a considerable cost in those trials with higher priority and uncertainty in their hypotheses, offering different perspectives in experiments, in this sense, facilitates greater

effectiveness in its planning, also complementing new and better predictions in diagnostics.

In a recent comparative study, Redolfi et al (2020) about Alzheimer Dementia (AD) using the platform, physicians explain that the information provided by it promoted a change in the original hypothesis of its three evaluations conducted. In addition to the increase in the level of medical prediction with two algorithms applied (CCC and GB) and other observations, point out that the MIP significantly affected the initial considerations in 15% of the cases studied. This type of analysis not only stands out for the surprising level of prediction of the ML algorithms executed (85-95% for GB, evidently with some level of variation and selectivity of cohorts), these could also place the MIP as a useful tool to support the advancement of trials in terms of their design and prediction of possible outcomes.

Naturally, the clinical trial market will continue to grow, and its methods evolve with them. Accelerated digital devices and AI solutions facilitate their definition, helping to manage limited resources and promoting its globalisation. However, it is urgent to find solutions to shorten time in its stages without sacrificing its quality, also linked to greater flexibility in protocols by regulators and, on the other hand, providing facilities for patients in terms of transportation to the centres when necessary, promoting better quality in regular doctor-patient communications and their recruitment through telemedicine, which also reaches disadvantaged populations.

Studies such as Kostis et al (2019), for example, sharpen judgment on the usefulness of trials in relation to the time taken from the design, analysis and publication of the results; which are often eclipsed by new findings in biological and epidemiological fields, as well as the emergence of new drugs. Once again, the challenges faced by medical teams and researchers could be minimised by AI solutions, such as design, detecting protocol failures, delay in patient recruitment, cost estimation, unexpected events, among others.

In the case of digital clinical trials, greater agility is offered in the process of patient recruitment, consent and data collection that falls on applications and communications via email, telephone, etc., and complementary analyses that are performed directly by visiting local nurses or laboratories (Cummings, 2021). Therefore, they are formats that could gain greater prominence over traditional essays or be

combined between physical and virtual format and will thus require powerful medical research and data storage platforms that can facilitate these processes and, at the same time, veracity in their results and possible predictions.

Knowing that the overall probability of success of a clinical trial for drug development is 13,8%[24] and that, in addition, a large part of studies completed do not publish their results, for a scenario in which this market remains booming, **probably the quantity of studies prematurely stopped would decrease by encouraging the use of platforms such as the MIP together with AI**, towards a better definition and efficient design of clinical trials. In other cases, it would avoid wasting time and resources in starting studies with inadequate or repetitive methods in similar experiments, direct them with different approaches and more collaborative; dynamics that in the past was almost impossible to reach. The integration of data at the federated level of dozens of hospitals, in addition to promoting the sharing of results or prediction of possible outcomes, is part of an evolution of digital medicine and that is what the MIP is getting prepared for.

On the other hand, a possible spillover effect for High Income Countries (HIC) and Low Middle Income Countries (LMIC) is the promotion of new jobs for the execution of clinical trials, in other words, the integration of research teams from multiple specialities including profiles of data scientists, engineers, as well as new digital skills for physicians that operate these technologies (Inan et al, 2020). For patients from LMIC, the impact would probably fall on not getting medicines and novelty treatments that they could normally obtain through pharmacological trials -in case its effectiveness can be proven without traditional testing-.

[24] In a study based on the data from 185,994 trials between 2000 and 2015. (Chi et al, 2018)

DEVELOPING A MARKET STRATEGY

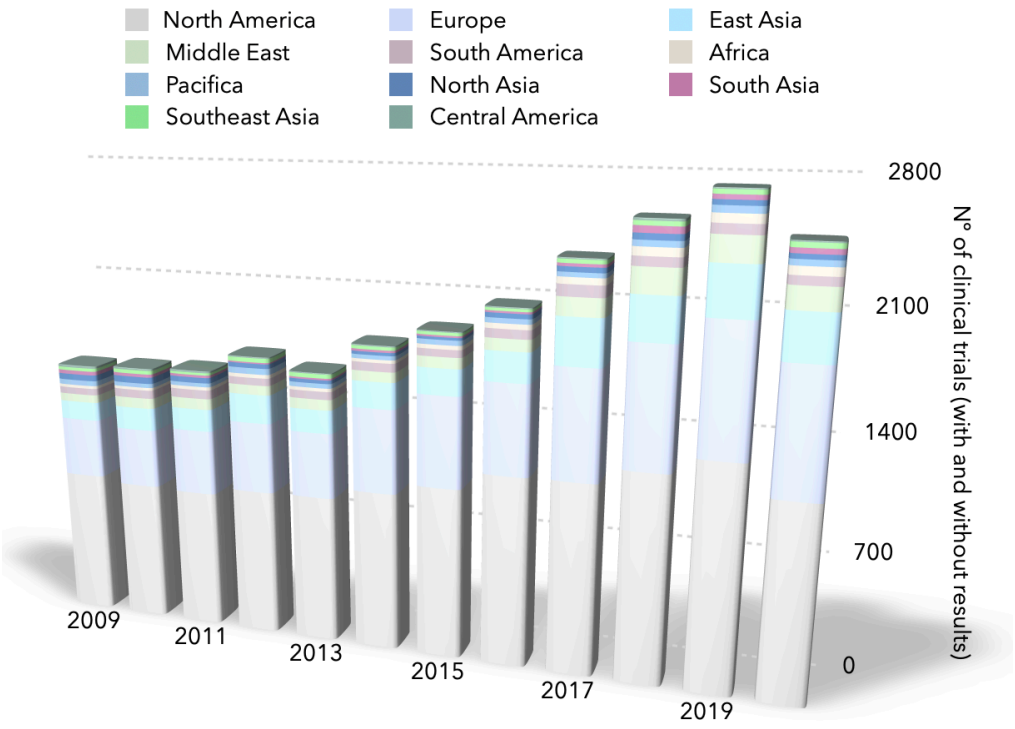


Figure 18. Total number of clinical trials registered in neurodegenerative diseases, psychiatric disorders and neurodevelopmental conditions (2009-2020). Data from: clinicaltrials.gov

The figure 18 shows the growth of clinical trials in recent years within the spectrum of neurodegenerative diseases, psychiatric and neurodevelopmental conditions by region. Among the blocs with the **highest participation are North America, followed by Europe and East Asia, while the rest of the regions show a slower growth and lower participation, except for Middle East** which in recent years has shown an interesting progressive increase. However, the average annual growth of the most prominent one along with the global (5,83%) is shown below: Although North America has a greater role in the number of records in progressive neurological diseases, has a more conservative average growth (2,81%) between 2009-2020, Europe denotes a higher rate (7,03%) and East Asia being one of the most pronounced (9,70%). It is understood that Middle East has lower volume of registers compared to leaders, however, its average growth rate is (10,85%) a sample of the attention in these countries for the deployment of this type of trials. Another aspect to highlight is the contrast between the growth of records worldwide -including 2020 when the pandemic caused by COVID-19 begins- and a noticeable decline in clinical trials in neurology.

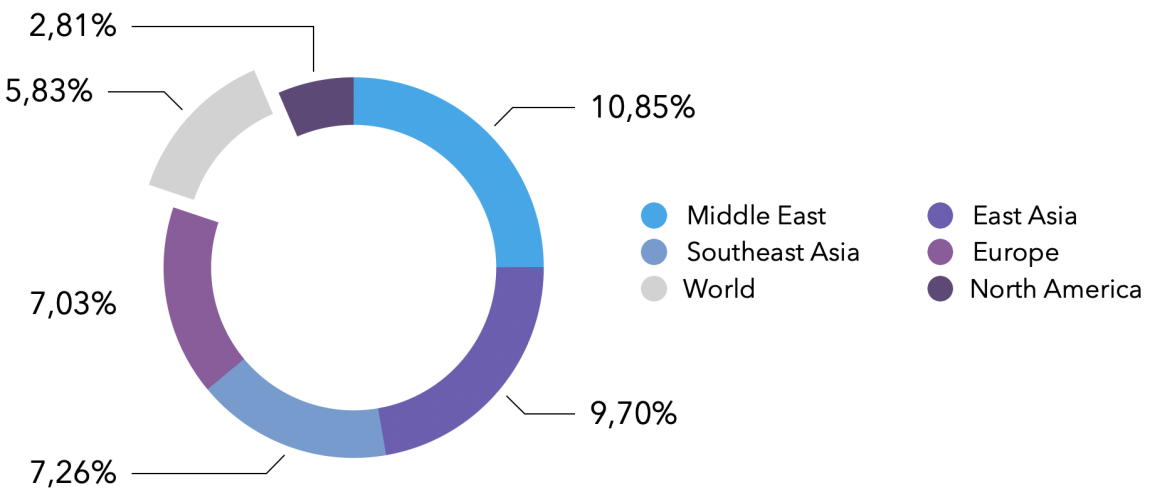


Figure 19. Average annual growth rate of neurodegenerative, psychiatric and neurodevelopmental conditions trials by major region.

These perspectives allow us to delve into the possible causes that influence the clinical trials market in neurology, on the one hand, emphasising those regions where their records are concentrated, therefore greater EHR or EDC (**Electronic Data Records for clinical trials and CTMS** for the informatics systems that manage them, *see figure 2 in MIP Introduction and the major Hospital IT Systems*), with a view to allowing the MIP to adapt to them, maintaining a continuous evolution in its AI methods. While, on the other hand, for those regions in LMIC and LIC, in whose hospitals and centres are not EHR or EDC available, may benefit from being offered as specialised training software both at the methodological level and the analysis for these conditions. In addition, there is also a possibility in these regions in relation to hospitals close to increasing their technological potential, e.g. having their records in electronic form, increasing the chances that the software can be considered an ideal complement to the analysis of trial data as part of their local structure.

In the internal structure, a periodic evaluation should be promoted to make the knowledge derived from these experiments more efficient in terms of data and practical to deploy in all its phases, considering the current barriers that affect the beginning of new clinical trials in neurology. In this sense, in markets with high volume such as the United States, updating studies in clinical research in neurology (Hall et al, 2018) explain the main reasons why this activity has remained stable over the last 10 years at least in that country, through a survey of more than 250 members of the American Academy of Neurology (AAN), and identifying the underlying causes of those specialists who have not conducted research in the previous year: **time, initial**

expenses and formal training for this activity are the most frequent causes.

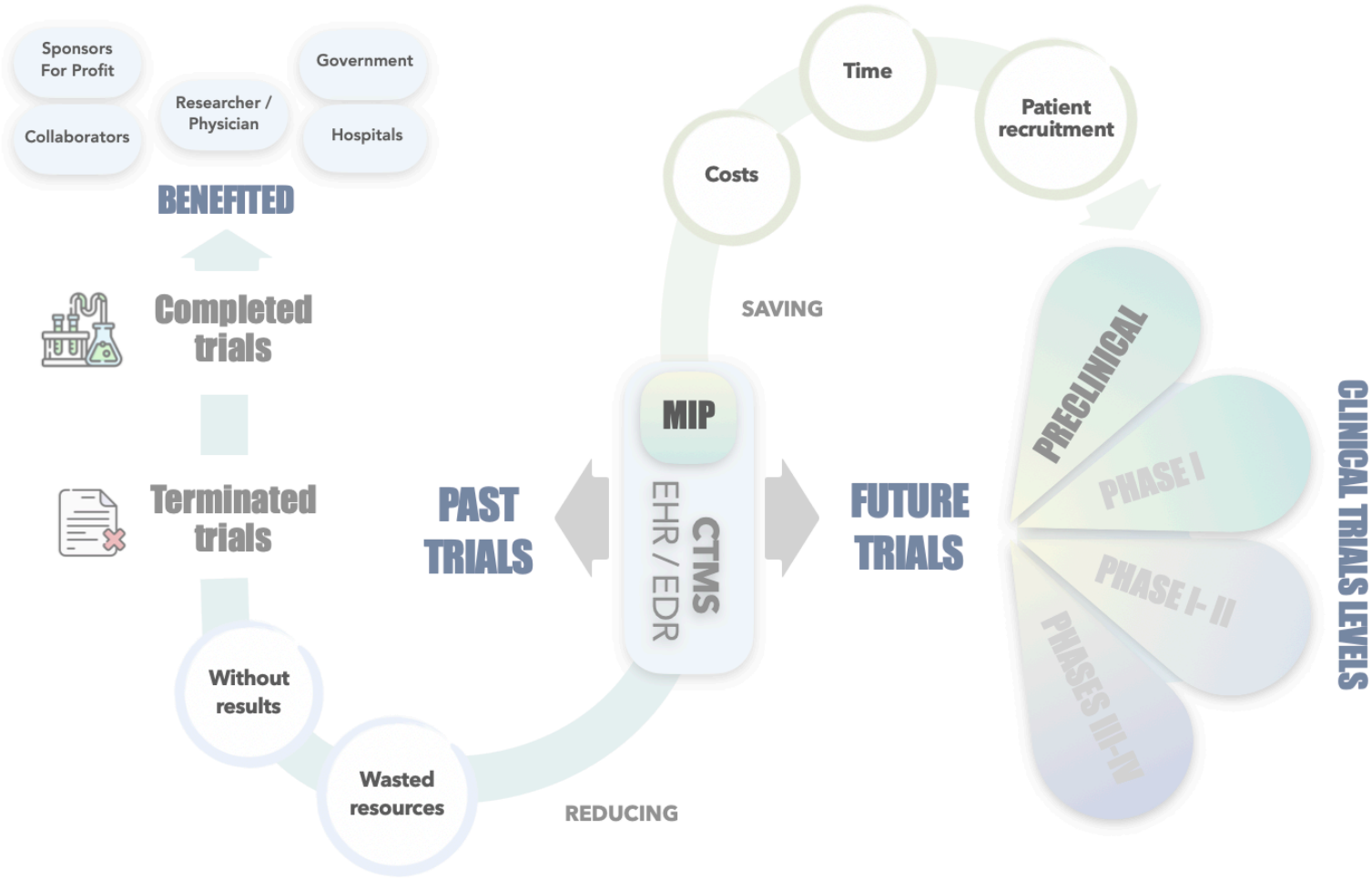


Figure 20. Dual perspective on the influence of software on clinical trials data (past and future trials)

In Europe, the dynamics are variable between countries, however, the most advanced and those with the highest health spending share such critical factors e.g. access to funding, fewer physicians dedicated exclusively to clinical research or not all of them have completed a standard GCP certification[25]. In Germany, for example, a survey conducted in university hospitals, non-university hospitals and medical practices suggests that in recent years have been conducted more phase III and IV trials in neurology in contrast with those in phases I and II, confirming the decrease in the latter given the difficulty in accessing funds (Lohmann et al, 2021), furthermore, they are mostly concentrated in university hospitals, while non-interventional trials

[25] It is an international certification of scientific and ethical quality for the development of clinical trials, demanded in large part by industrial sponsors. (Fougerou-Leurent et al, 2020)

(NIS)[26] were developed largely by centres of medical practices. In such markets of interest, **a dual perspective is considered** (designed in figure 20) which can contemplate historical data -both finished and completed trials- partly covering the software's need in acquiring more relevant and concrete data from centres; while the following seeks to influence those trials in early stages and mostly conducted at university hospitals, in which the platform would have a greater utility.

Indeed, it is complex to measure the future impact of platforms such the MIP in the development of clinical trials in neurology considering the limitations currently shared by the agents involved. For-profit sponsors (pharmaceuticals, devices manufacturers, medical software companies, research partners) and those who invest significant financial resources in the trials, continue to prioritise the protection of their knowledge, reducing failures, time and costs. From physicians as mentioned above, factors such as lack of time, costs formal training or access to sponsors. **Regarding the public sector, there is also a decrease in health spending, resulting in less funds to start new studies in these areas.**

However, the platform has certain technological advantages that support the prior definition and design of these studies especially in initial phases, which may include a key methodological training required by the medical sector, promoting the mastery of AI techniques and avoiding the repetition of studies. It also has the potential to communicate with other computer systems which may include the Clinical Trials Management Systems (CTMS), offering their data, and to help researchers plan more appropriate patient recruitment according to the regions in which it is deployed, influencing the *terminated* clinical trials or those that have been concluded prematurely.

[26] Trials where the drug is prescribed and the patient's evolution is observed. (Aronson, 2004)

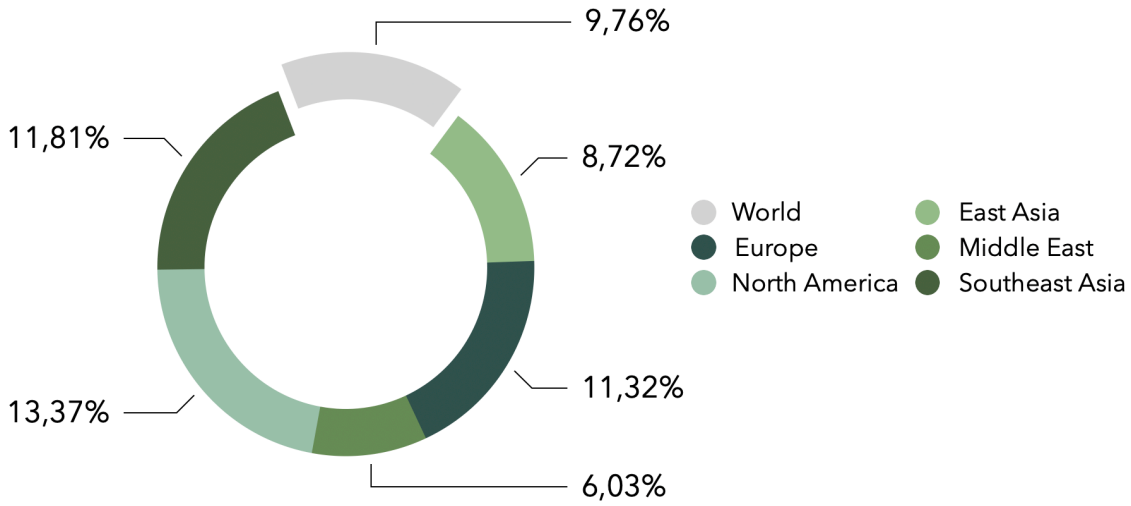
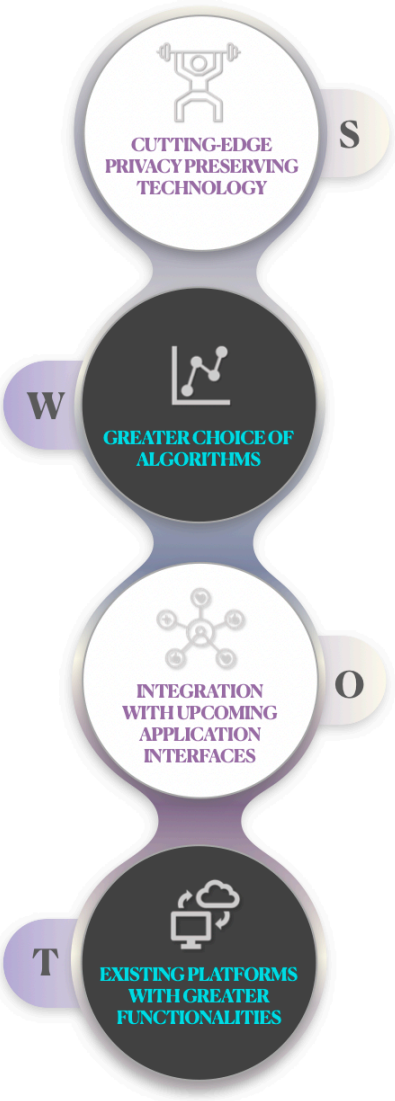


Figure 21. Percentage of terminated trials as a proportion of completed in neurodegenerative diseases, psychiatric disorders and neurodevelopmental conditions (2009-2020). Data from clinicaltrials.gov (2021)

It is understood that a lack of proper recruitment greatly influences an early termination of the trials, coupled with loss of resources and unanswered research questions (Kasenda et al, 2020), therefore, the valuable information collected by **the platform could help in an optimal planning in the aspect of recruitment of required patients.** Therefore, it would be appropriate to define a formal training in clinical research specific in neurology associated to MIP, as the possibility of being included as a support tool in the GCP certification modules. On the one hand, inviting neurologists to alternate their activities, for example, in periods in which performance may be compromised due to cognitive overload or burnout and, on the other, providing solutions for the recruitment of patients with specific characteristics within the virtual experimentation, taking advantage of the communication it has with other hospitals systems in different geographical points, influencing the reduction of those trials that fail due to the lack of optimal planning in recruitment stages.

SWOT ANALYSIS



After the above analysis and adding the valuable contributions between the MIP development team and the UPM Innovation team, we could establish some opportunities, weaknesses, threats and strengths; simplified in the next figures. There are many strengths of the platform that characterise it as an emerging medical research tool, whose attractiveness lies mainly on the deployment of AI algorithms based in machine learning on metadata and synthetic data of thousands of patients (pseudonymised locally and anonymised at the federal level as explained before), also this box describes the main virtues which will be continuously updated and aligned to the TRL.

Evidently, beyond the functionalities adhered to the technology, its long-term growth will depend in large part on combining its technical capacity with the issues that face the key users and rather facilitates the diagnostics and research processes (some of them as mentioned at the beginning of the report, decreasing the expenses related to wrong diagnosis, new clinical trials, greater precision in the diagnostics, to mention a few). In addition, some of them constitutes significant opportunities for a next deployment, e.g. to continuously evaluate that the data are accurate and sufficient for the explored conditions, measure the platform performance, as well as the possibility of identify findings and advert coming conditions or possible appearance of phenomena -in case of pandemics, also certain increases of specific conditions-, thanks to its use (for example through published articles about the MIP, however, there are uses that are more difficult to control.

Main MIP Strengths

- ***Powerful medical research platform integrated with cutting-edge privacy preserving technology that support clinical decision-making with patient data analysis*** - this is achieved by the informatics structure that supports it, allowing to experiment with AI tools the variables associated with the metadata supported in the Data Catalogue, where multiple brain characteristics and diseases can be displayed.

- ***Complex and personalised algorithms are created upon request by medical experts*** - an advantage for medical teams needing further specific analytics models adapted to

groups with similar conditions, while incorporated new data from dozens of hospitals, its results help to reduce future misdiagnosis and delayed diagnosis *(Could be offered as freemium feature).*

- ***Rare and specific diseases could be explored with the support of certain medical MIP teams and validate the experiments within the platform*** - it begins to be a reality from its partnership with ERN EpiCare, with clinical care teams distributed in 28 healthcare centres with extensive experience in rare and complex diseases (Epicare, 2021).

- ***It could be used both in online and offline mode*** - useful in remote sites or centres with less facilities in IT infrastructure, in addition to cost savings for clinical researchers.

- ***It does not require any download and does not use any storage space*** - knowing that the average data generated by a single patient is equivalent to about 80 megabytes each year[27], together with the increasing energy consumption of the data centres that support the systems of health centres globally, this type of solution can stand out in the long term among the wide range of medical exploration software on the offer today.

- ***It can compare diagnostic procedures and CDEs between countries*** - the MIP allows to take maximum advantage of the information from patient data stored in hospitals, understanding that each condition is composed of specific datasets worked from these centres, thus offering the possibility of comparing previous procedures and therapies.

- ***It can reduce the gap in the number of specialists per country or within the EU*** - if and only if the information on the number of affected persons per facility and per country between hospitals (while keeping the data anonymous) is transparent, through the exchange of specialists and online consultations. The differences in the number of specialists available between countries are notorious, especially in neurology, shortcomings that are increasingly reflected in the long waiting times that patients have to deal with to receive personalised medical care. These types of platforms should help speed up early diagnoses, as well as support in the online training of this limited availability of specialised human resources.

- ***Deployed in more than 31 hospitals*** and an ongoing effort to integrate more members. It has data from approximately 30,000 patients and regularly add up new records through collaborations with other organisations, in addition to EpiCare, such as the European Stroke Organisation. Also has to be considered an action plan that integrates in a federated way new data directly from hospitals, psychiatric centres and

[27] Huesch and Mosher, 2017.

neurology in those regions of Europe where they are most present, as seen in the previous section.

- ***Support of talented clinicians, engineers and researchers, as well as financial resources from the Human Brain Project and EBRAINS -*** Currently led by Dr. Ferath Kherif in Lausanne, Switzerland at the Centre Hospitalier Universitaire Vaudois (CHUV). The contribution of researchers has played an important role in its development from world-renowned universities, being a project with the capacity to attract new health professionals given the breadth of diseases it explores, along with other innovative disciplines such as engineering or computer science.

- ***Strict GLDP monitoring, it doesn't allow copying or transferring data, increasing the patient protection and privacy -*** Its structure makes it possible to visualise the results of multiple neuropsychological evaluations that clinician uses to delimit brain diseases, including large image datasets and EHR, without compromising personal data. Reinforces security level from the very beginning with a complete anonymisation of the local data, annulling the possibility of reconstructing the original information.

- ***Correlate variables that normally are difficult to calculate due to lack of more comprehensive data or data accuracy*** - With the catalog of ML algorithms in continuous evolution and a sustained growth of data (EHR, genetic information, scan images, laboratory, etc.) it allows the exploration of broad cohorts, through the variables previously standarised in its federated network and more than 260 embedded ontologies[28]. Also, the opportunity to explore the correct biomarkers that intervene in diseases, which are increasingly relevant in their prevention and therapy[29]. In this sense, the nature of the MIP data is more varied and offer quality in content, essential for more precise results in their analytics.

- ***Creation of CDEs or key factors of patient conditions (so far 168) allowing both their harmonisation process and support for the federated hospital/centre studies.*** Another of its advantages is to offer the researcher the possibility of creating models based on variables that compose the CDEs at the federated level. For some experts, part of the current CDEs are not always common, but are adapted a particular study or database[30], therefore, federated platforms such as the MIP have the capacity to overcome these limitations by being strongly structured following Findable, Accessible, Interoperable and Reusable (FAIR) principles and current standards with data from large

[28] Venetis and Vasalos, 2015.

[29] Baum, 2016.

[30] Kush et al, 2020.

cohorts. Also, this process includes an accompaniment during the registration of data and the submission of specific formats through continuous revisions.

- ***They help hospitals and specialised data harmonisation centres to implement AI algorithms*** - Starting from a double perspective: On the one hand, applying analytical tools based on local CDEs, therefore, in more restrictive groups, and on the other hand, AI at the federated level following the rules of the CDEs shared by all its members. In this sense, a practical and intuitive platform is provided to the centres that allows the selection of variable and co-variables of interest to be executed in the variety of algorithms available during the experiments. In addition, new hypotheses and research questions can be formulated by offering the possibility of comparing results at the local level with other demographic groups, with the exception that this type of data is partially limited in the MIP to ensure privacy protection.

- ***Data infrastructure support through Swiss National Supercomputing Centre (CSCS) as part of the recognised FENIX consortium*** - The MIP is directly related to a technological infrastructure integrated by five European supercomputing centres (Alam et al, 2020), surpassing the proposal of cloud services commonly offered by other organisations -both locally and federated-. In short, the generation of data in hospitals grows exponentially which demands diversified and high-performance systems to generate ML models efficiently, being FENIX the provider of excellence of EBRAINS in this area.

STRENGTHS

1. Powerful platform of medical research integrated with the latest technology for patient data analysis preserving their privacy.
2. Complex and personalised algorithms are created upon request by medical experts (could be applied a freemium mode in the near future).
3. Rare and specific diseases could be explored with the support of certain medical MIP teams and validate the experiments within the platform (if freemium mode would be applicable, this option could be considered).
4. It could be used both in online mode and offline.
5. It does not require any downloads and does not use any storage space.
6. It can compare diagnostic procedures and CDEs between countries.
7. It can reduce the gap in the number of specialists per country or within the EU, if and only if the information on the number of affected persons per facility and per country between hospitals (while keeping the data anonymous) is transparent, through the exchange of specialists and online consultations.
8. Deployed in more than 31 hospitals.
9. Support of talented clinicians, engineers and researchers, as well as financial resources from the Human Brain Project and EBRAINS.
10. Strict GLDP monitoring, it doesn't allow copying or transferring data, increasing the patient protection and privacy.
11. Correlate variables that normally are difficult to calculate due to lack of more comprehensive data or data accuracy.
12. Creation of CDEs or key factors of patient conditions (so far 168) allowing both their harmonisation process and support for the federated hospital/center studies.
13. They help hospitals and specialized data harmonisation centers to implement AI algorithms.
14. Data infrastructure support through Swiss National Supercomputing Center (CSCS) as part of the recognised FENIX consortium.

Main MIP Weaknesses

- ***Absence of visualisation in 3D models of the brain, it could be considered to make a link with TVB for certain disease conditions*** - Being the brain such a complex and relevant organ, the MIP competes with other software of high spatial resolution, as well as visualisation of other parts of the body. Powerful integrated AI applications guide clinicians in preoperative stages, detects anomalies, predict outcomes, in addition to automatic reporting and altogether significantly shorten the time usually invested in analysis.

- ***It must achieve the highest possible accuracy in a minimum of medical specialties*** - Despite the current advantages of AI, it continues to have limitations in accuracy diagnosis of multiple neurological conditions, increasing the need of reliable data and solving problems related to the absence of standarised protocols in its procedures. Covering a variety of diseases without a degree of depth and precision, it may risk losing a competitive advantage over other platforms.

- ***Applicable to High Income Countries (HIC) rather than Low-Middle Income Countries (LMIC) it is necessary to analyse how to help them*** and it would be very useful to have more precise data from these countries as well. The MIP should make a medium-term effort to reach sectors where the burden of neurological conditions is a major challenge, motivating their clinicians to achieve accessible and improved training to meet domestic demand.

- ***Improve Front-End design and visualisation parameters are repeated and others should be more specific, at least in the federal version*** - Clinicians are continuously evaluating and analysing information, therefore, making the visualisation and exploration of their variables an intuitive and reliable space would motivate the use on a regular basis. It would also encourage the training of relevant support staff such as nurses.

- ***Continuous improvement is needed in agnostic engine allowing it to be adapted to other formats and programs*** - Despite offering increasingly intuitive tools for data transfer from the centres, in order to achieve greater competitiveness and accelerate the acquisition of data, it would be useful to adapt its internal processes for the reception of data from new sources such as PGHD, as well as HIS frequently used in other specialties.

- ***The data catalogue is separate from the front-end of federated MIP*** - For new specialists who are becoming familiar with the platform, it would be useful to integrate links to other relevant resources into its main distribution, facilitating a fast learning.

- ***Its main algorithms are shared in a public repository on GitHub, forcing the creation of others that are much more personalised and a differentiated value proposition that makes it profitable in its business model*** - It is understood that at the federated level most of its algorithms are known by clinicians and researchers in this field, also offered by recognised cloud services in their licenses. Considering this relevant advantage of the industry leaders, at the spin-off level is evident to specify the modality and the price requested for the AI models adapted to the local datasets, together with the on-demand training services, taking advantage of the networks of dozens of hospitals in Europe and their alliances with regional organisations. A choose must be made between the time of use, number of registered computers, payment for punctual services, periodic subscription or a combination of these.

WEAKNESSES

1. Absence of visualization in 3D models of the brain it could be considered to make a link with TVB for certain disease conditions.
2. It must achieve the highest possible accuracy in a minimum of medical specialties (placing greater emphasis on quality rather than variety).
3. The Data Catalogue is not functional in: Firefox, Safari, Brave.
4. Applicable to High Income Countries (HIC) rather than Low-Middle Income Countries (LMIC) it is necessary to analyse how to help them, and it would be very useful to have more precise data from these countries as well.
5. Improve Front-End design and visualization (parameters are repeated and others should be more specific, at least in the federated version).
6. Continuous improvement is needed in agnostic engine: allowing it to be adapted to other formats and programs.
7. The demo version to attract new users has dispersed information and incomplete descriptions.
8. Small details that can make a difference: Federated platform does not open in different devices at the same time (i.e tablets).
9. Update and align manuals according to new versions and data catalogue (medical conditions).
10. The data catalogue is separate from the front-end of federated MIP.
11. Sometimes, when processing the variables in the platform, the data are not loaded, partly due to a lack of information.
12. Its main algorithms are shared in a public repository on GitHub, forcing the creation of others that are much more personalised and a differentiated value proposition that makes it profitable in its business model.

Main MIP Opportunities

- ***Ability to prevent future consequences due to mis-supply of drugs through the amount of data acquired*** - Medication errors may occur both at the time of prescription, as well as in dispensing, administration or control of the drug, which obviously involve variable damage and an associated annual costs. Therefore, promoting more concrete studies in this field are efforts that could help to save lives,

improve future diagnoses and should also be directed in support of the medical sector which faces daily pressure to improve its interventions.

- ***Certain countries may reflect weakness in their EHR systems by region, the MIP would enhance its use for consistent and massive digitisation***, under a common framework that can be moved to other hospitals in different communities. For those centres that do not have complete data of their patients, or in which their systems lack precise tools for their analysis, the platform offers a reliable alternative to help clinicians with limited technical resources.

- ***Integration with upcoming application interfaces in the medical industry: Digital twins, avatars*** - The most recent advances in digital medicine use the massive amount of data from multiple sources that, combined with AI algorithms, define the advanced precision medicine and with it the possibility of creating predictive models based on multiscale data integration scheme[31]. The MIP has the possibility of accelerating this technique, which is still in an early stage in neurology, however, the team should consider -among other aspects- the integration of real data (it maintains synthetic data) and a greater effort in the standarisation of these from origin.

- ***In the near future, consider reflecting conditions and therapies in other parts of the body with VR and AR*** - In addition to the previous point, if a spatial scale of augmented reality is integrated, it would offer an attractive alternative to study in the same workflow brain regions with higher precision, delving into the different brain levels of organisations and its variations[32], thereby integrating cutting-edge EBRAINS technologies such as NEST and TVB. If data from other related systems, such as Oncology Information Systems (*see section I on possible influence of MIP*), for example, are added, then the possibility of exploring other affected organs.

- ***Reduce future costs through the prevention and knowledge of incoming and existing data*** - The constant pressure on hospitals and mental health services, coupled with long waiting times in this specialty, intensify the demand for lower cost and reliable methods in their predictions before the supply of drugs and new therapies. Thus, it can be used as a support instrument in the planning of public health prevention programs.

- ***It would help reduce time in diagnosis and medical care, providing alternatives and greater accuracy in differential diagnoses*** - Delayed diagnoses are a critical problem in neurological diseases, undoubtedly there are multiple underlying causes. Only in dementia, for example, factors may be related with education or knowledge of those affected about the disease to seek support, lack of screen protocols or standarised

[31] Voigt et al, 2021.

[32] Pillai and Jirsa, 2017.

routines in evaluations, experience if the attending clinician, limited communication between the patient and their caregivers[33]. Other drawbacks no less relevant as the long waiting times to receive medical consultations with specialists, the proximity of specialised centres, the effectiveness of current treatments that depend on the progress of several of these diseases. Definitely, there is an urgent need for technological solutions that complement the efforts to reduce late diagnoses, reduce uncertainty and timely communicate the risk of developing a disease.

- ***Develop applications that can provide patients with health recommendations and advice according to their clinical data and habits, at the preventive level and/or for their prompt recovery*** - The possibility of including applications that relate data directly from patients or PGHD in a federated scheme opens new opportunities of analysis for real-time analyses of their reactions to medications or other therapies, knowing aspects ins routines and behaviours that can be of practical utility for the study of new variables on the platform. Technical capabilities would be expanded by obtaining data from multiple sources and different perspectives that invite the analysis of more accurate data or specific parameters that are not contemplated in the usual EHRs.

- ***Possibility of becoming a powerful structure of validation, diagnosis and prevention of future pandemics and/or unforeseen condition***s - Its federated structure records demographic data in dozens of its members hospitals, valuable information for decision-making by the agencies responsible for public health, create or optimise prevention programs, while helping towards an effective distribution of inputs and financial resources in the most affected localities.

- ***Integration into other types of programs is being considered, facilitating better future data reception*** **-** Hospital and medical device software are becoming increasingly sophisticated worldwide, recording information from millions of patients every minute. An agnostic engine approach is evaluated to strengthen the competitive advantage of the MIP, avoiding inconveniences of use restricted to the processing of data in specific formats, relevant to ensure greater loyalty or preference in its use, given the ease and time saved for researchers in continually consulting alternative sources of data provided by other companies. Being aware of the real problems of medical equipment in diagnostics and their feedback regarding the use of this platform plays a key factor for a better design and adaptability to frequently used devices, which are simple to handle for clinicians who do not have a long experience in the applications of AI analytics.

[33] Bradford et al, 2009; Woods et al, 2018.

- ***New business models related (described in table "HBP Current Technology: MIP") -*** Multiple lead users belonging to the Healthcare and equipment providers sector, as well as at the pharmaceutical and biotechnology level, were identified individually. Likewise, end users or those who will benefit for use the tool and other related organisations were also described. This structure allowed to identify future opportunities, the ways in which these organisations can benefit from the platform considering its characteristics and the priority it would have to target that segment.

- ***Collaborations with large-scale associations as EpiCare, ENIGMA, and others with a large outreach to key segments*** - New possibilities to acquire more specialised data of interest to the platform are generated (neuroimaging, genetics, EHR), in addition to venturing into new sectors with the collaborations of research groups dedicated exclusively to certain diseases. The ENIGMA[34] consortium is one of these organisations that brings together more than 50 active groups offering a flexible framework with the possibility of sharing and transferring data when joining as a member, while EpiCare stands out in both data and the breadth of research groups in epilepsy and rare diseases, in 300 institutions worldwide.

OPPORTUNITIES

1. Ability to prevent future consequences due to mis-supply of drugs through the amount of data acquired.
2. Certain countries may reflect weakness in their EHR systems by region, the MIP would enhance its use for consistent and massive digitization, under a common framework that can be moved to other hospitals in different communities.
3. It has the potential to capture data that relates misdiagnosis with a possible degree of patient harm (This could be done, for example, comparing patient with the same diagnostics, similar conditions or characteristics with the treatments provided. However, there is also an obstacle to different responses in treatments of similar diseases given the genetic variations between patients).
4. Integration with upcoming application interfaces in the medical industry: Digital twins, avatars.
5. In the near future, consider reflecting conditions and therapies in other parts of the body with VR and AR.
6. Reduce future costs through the prevention and knowledge of incoming and existing data.
7. It would help reduce time in diagnosis and medical care, providing alternatives and greater accuracy in differential diagnoses.
8. Develop applications that can provide patients with health recommendations and advice according to their clinical data and habits, at the preventive level and/or for their prompt recovery.
9. Possibility of becoming a powerful structure of validation, diagnosis and prevention of future pandemics and/or unforeseen conditions.
10. Integration into other types of programs is being considered, facilitating better future data reception.
11. New business models related (described in table "HBP Current Technology: MIP").
12. Collaborations with large-scale associations as EpiCare, ENIGMA, and others with a large outreach to key segments.

[34] http://enigma.ini.usc.edu

Main MIP Threats

- ***Scepticism among doctors and specialists about the effectiveness of the new tool*** - An increasing variety of companies are offering AI solutions for brain diseases, competing in speed, useful exploration tools and intuitive interface. However, its practical application will depend on the will of clinicians, who mostly operate with various healthcare systems and applications in their daily routines, therefore, the added value of the platform must respond to the need for these teams and to the final diagnostics, justified by facilitating the current workload instead of increasing it.

- ***It covers many diseases and lacks depths in the parameters and variables,*** there is a risk that it could be taken as a secondary or unreliable reference - Clearly the platform has the technical capacity to explore data from multiple diseases in a federated way, however, it is still in the early phase in hospitals, which have the final word on the current offering of data, the variables that can be explored and the ease of its initial processing.

- ***In the initial schema mapping using MIPMap to transfer data is vulnerable to errors*** because it depends of different medical specialties and its users, duplicating efforts in the following processes to assess the quality of the data. Definitely, the original data that is processed follows a terminology associated with each medical specialty and will depend mainly on it, therefore, those responsible for executing these procedures -including computing experts- have the challenge to analysing the quality of information to be anonymised.

- ***In the initial process and also to pass local data to Federated data*** -so the platform becomes more competitive- depends on the willingness and expertise of its users. The decision from hospitals and other partner centres to be part of the federated platform is key to delve into concrete data by cohorts across multiple hospitals, as well as provide new criteria for the creation of CDEs, among the most relevant factors. In this sense, there is a minimal risk of its participants who do not show interest to move to a federated level, it is essential to determine the causes and correct the deficiencies, improving the quality of the service offered by the software.

- ***New or existing platforms with greater functionalities, specialisation and user privileges.*** As explained in the analysis of competitors, new participants in healthcare continue to expand with federated learning technology, startups, consortiums, companies with years of experience in software and services to hospitals that have also chosen to offer these functions, including the recognised multinationals with cloud computing services and programs to promote similar projects.

- ***Blockchain technology with future potential in the field*** - With great success in fields such as finance, energy, security, etc. is also developed in the healthcare industry, with similar advantages in authentication and data preservation, allowing the generation of models with AI. New projects are combining the strengths of both federated learning and blockchain techniques, opening the way to a new kind of decentralised structures.

- ***Different IT systems between hospitals and centres in technical and data aspects, slowing down the process of harmonisation and integration*** - Knowing that medical records derive from different sources and formats, the need to integrate information, to document it digitally without requiring a greater effort from the medical team, to analyse unstructured data and ensure quality (Menasalvas et al, 2013). A good start to increase the attractiveness and versatility of the software requires at least identifying the basic dynamics to which clinicians are exposed in their daily work in terms of the most used software, coupled with the possibility of relating device data by specialty. The strong competition led by large corporation such as Apple[35], Google o Microsoft represent a significant advantage given the time they have been integrating their systems in hospitals through solid distribution strategies, allowing users to share information with their doctors from the same device they commercialise.

- ***Less financial resources, influencing future improvements and releases*** - Until the differences concerning cohesion funds between Switzerland and the European Commission are resolved, Swiss researchers and organisations will be limited in their opportunities to access funding in three of the Horizon Europe programmes[36]. In this area, the alternative is to ally with an organization within the EU or a member country, circumstances that undoubtedly slow down the learning of new technologies and future benefits to final users, pressing for seeking private funding.

THREATS

1. Skepticism among doctors and specialists about the effectiveness of the new tool.
2. It covers many diseases and lacks depths in the parameters and variables, there is a risk that it could be taken as a secondary or unreliable reference.
3. Mistrust of the tool by hospitals and users with limited use of the platform and its virtues, generally due to lack of information or knowledge of the platform's internal processes.
4. In the initial schema mapping using MIPMap to transfer data is vulnerable to errors because it depends of different medical specialties and its users, duplicating efforts in the following processes to assess the quality of the data.
5. In the initial process and also to pass local data to Federated data -so the platform becomes more competitive- depends on the willingness and expertise of its users.
6. New or existing platforms with greater functionalities, specialisation and user privileges.
7. Lack of harmonisation in the legal level and between countries.
8. Lack of clarity in priorities in terms of design or adaptation of all the registered users.
9. Blockchain technology with future potential in the field.
10. Different IT systems between hospitals and centres in technical and data aspects, slowing down the process of harmonization and integration.
11. Less financial resources, influencing future improvements and releases.

35 Apple has integrated with 700 hospitals and clinics. Reuter and Vaidya (2021)

36 European Research Council (ERC), Marie Skłodowska-Curie Actions (MCSA) and the European Innovation Council (EIC). Zubașcu, 2021, ScienceBusiness.

CONCLUSIONS

The MIP is an outstanding technology for training and analysis for the research clinician of tomorrow. It can be stated that is the convergence between the neuroscientific capacity, medical practice, and accelerated computer development, upshot of the effort of a multidisciplinary human resource. Integrating the exorbitant amount of data in health and organising it in a federated scheme with AI, has been an arduous and extraordinary task, currently **presented as one of the EBRAINS research tools with technical potential to grow in the market**, the main message of this report. Although it may seem like a truism to mention that technologies do not evolve on their own, it will require a sustained effort to identify the objectives of the future spin-off and adapt to current demands in practical scenarios, starting with its competitive advantages and what it can improve, instead of limiting its lines of action only to its competitors.

On the other hand, it should be reflected that this platform has not had a straight path since its inception, both because the degree of uncertainty in a market that was beginning to expand, and due the amount and heterogeneity of data scattered in hospitals to be recorded, essential resource for software. It was ultimately a large-scale project and a technology that did not have specific business models in the health field as a reference for more than a decade and that can now be taken as lessons learned to continue evolving in this exciting field of federated learning and AI.

In this regard, the reader will remember countless cases in the scientific field, whose leaders were unaware of the real impact of a new project and its derived results, technologies that are largely reborn when they are transformed sometime later, leaving an invaluable legacy in terms of specialised knowledge and behind it, the generous possibility of changing lives; which is why the MIP fits as one of these innovations. **If it is now a reality that it is possible to improve the hypotheses in current medical studies** (Redolfi et al, 2020), it will require duplicate efforts to reach new medical researchers, in order to continue the legacy of HBP and EBRAINS.

Throughout this report, the main outputs have been identified in which it stand out above other platforms, from the personalised model generation with AI, reduction of costs in clinical trials, collaborations with groups specialised in brain and rare diseases, decrease of misdiagnosis and late diagnosis, advance of an evidence-based precision medicine, to the generation of prevention programs with a more critical approach by region, as well as an efficient distribution of its resources; as some of the most noteworthy aspects. Although the current hospital systems do not technically depend on this platform, those who could benefit directly or indirectly from its explorations were identified, strengthening even more if manages to integrate data from other sources, evidently maintaining its privacy-preserving scheme. In this way, **time and effort could be saved**

for medical teams in looking for alternative sources of validation, on a framework that allows alternating between the study and application of AI in particular variables or a group of them, combined with the ability to receive information from other angles that are usually difficult to reach for the attending medical service. Another aspect to highlight is the current AI analytical tools, with useful algorithms and frequently used by the clinician. It is understood that the expert team of the MIP either through a startup or as part of the EBRAINS infrastructure will continue to deploy new algorithms and specific models according to the requirements of hospitals and research centres, and in this sense to balance itself with the offering of its competitors or other software commonly used in this medical specialty.

The current and forthcoming initiatives were described in a comparative scheme that allowed to define the limitations of other projects and combinations with competitive technologies such as blockchain. Issues are described on how the MIP could improve in the field of business, marketing, finance, alliances, collaborations, technical aspects, maintenance, technological surveillance aspects and other relevant issues in terms of revenues. **This led to identify in a practical way the participants in the industry,** study the advantages and disadvantages of their possible associations, also to consider strategies towards their first stage as a startup, final decisions that will obviously depend on the leaders of the technology.

The business opportunities presented and their possibilities of introduction represent exciting initiatives to be developed. The practical scheme deployed helped to delve in more detail into the main segments such as specialised hospitals, neurology centres, in addition to the clinical trials market, with a view to Europe mostly considering its installation points and other specific sectors according to the number of specialised centres and practicing specialists. **We close with a SWOT analysis gathering the most relevant observations,** emphasising its strengths, correcting the weaknesses aligned with its lead users, taking advantage of the opportunities in which it can excel while preparing for the possible threats described. In addition, some suggestions are offered in terms of business, product improvement along with actions that EBRAINS can take to support the MIP and other of its related projects.

References:

ACRDSI (2021). FDA Cleared AI Algorithms. https://models.acrdsi.org

Ács, A., Molnár, E., Molnár, G. Y., & Balogh, Z. (2019). The care of people living with mental illness in the Hungarian social care system: the process of deinstitutionalization and the phenomenon of stigmatization. Developments in Health Sciences, 2(1), 1-8.

Adamos, D. A., Kosmidis, E. K., & Theophilidis, G. (2008). Performance evaluation of PCA-based spike sorting algorithms. Computer methods and programs in biomedicine, 91(3), 232-244.

Agency for Healthcare Research and Quality, AHRQ (s.f). Computerized Provider Order Entry. https://digital.ahrq.gov/key-topics/computerized-provider-order-entry

Ahmed B, Brodley CE, Blackmon KE, Kuzniecky R, Barash G, Carlson C, Quinn BT, Doyle W, French J, Devinsky O, Thesen T. Cortical feature analysis and machine learning improves detection of "MRI-negative" focal cortical dysplasia. Epilepsy Behav. 2015 Jul;48:21-8. doi: 10.1016/j.yebeh.2015.04.055. Epub 2015 May 31. PMID: 26037845; PMCID: PMC4500682.

Alashwal, H., El Halaby, M., Crouse, J. J., Abdalla, A., & Moustafa, A. A. (2019). The application of unsupervised clustering methods to Alzheimer's disease. Frontiers in computational neuroscience, 13, 31.

Aldosari, B. (2017). Patients' safety in the era of EMR/EHR automation. Informatics in Medicine Unlocked. Volume 9, 2017, Pages 230-233. https://doi.org/10.1016/j.imu.2017.10.001

Alemayehu, C., Mitchell, G., & Nikles, J. (2018). Barriers for conducting clinical trials in developing countries-a systematic review. International journal for equity in health, 17(1), 1-11.

Amunts K, Ebell C, Muller J, Telefont M, Knoll A, Lippert T. The Human Brain Project: Creating a European Research Infrastructure to Decode the Human Brain. Neuron. 2016 Nov 2;92(3):574-581. doi: 10.1016/j.neuron.2016.10.046. PMID: 27809997.

Analytics Vidhya (2018). A Simple Introduction to ANOVA (with applications in Excel). https://www.analyticsvidhya.com/blog/2018/01/anova-analysis-of-variance/

Ancker, J.S., Edwards, A., Nosal, S. et al. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. BMC Med Inform Decis Mak 17, 36 (2017). https://doi.org/10.1186/s12911-017-0430-8

Auditor General of Ontario. (2011). 2011 Annual Report of the Office of the Auditor General of Ontario. LCBO: New Product Procurement. https://www.auditor.on.ca/en/content/annualreports/arreports/en16/v1_312en16.pdf

Argaw, S. T., Bempong, N. E., Eshaya-Chauvin, B., & Flahault, A. (2019). The state of research on cyberattacks against hospitals and available best practice recommendations: a scoping review. BMC medical informatics and decision making, 19(1), 10. https://doi.org/10.1186/s12911-018-0724-5

Aronson J. K. (2004). What is a clinical trial?. British journal of clinical pharmacology, 58(1), 1–3. https://doi.org/10.1111/j.1365-2125.2004.02184.x

Bahado-Singh, R. O., Vishweswaraiah, S., Er, A., Aydas, B., Turkoglu, O., Taskin, B. D., ... & Radhakrishna, U. (2020). Artificial intelligence and the detection of pediatric concussion using epigenomic analysis. Brain research, 1726, 146510.

Baig, B. J., & Delvenne, V. (2008). Psychiatry in Belgium. International psychiatry : bulletin of the Board of International Affairs of the Royal College of Psychiatrists, 5(4), 88–90.

Barker, J. W., Rosso, A. L., Sparto, P. J., & Huppert, T. J. (2016). Correction of motion artifacts and serial correlations for real-time functional near-infrared spectroscopy. Neurophotonics, 3(3), 031410.

Bayes, T. (1763). LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, FRS communicated by Mr. Price, in a letter to John Canton, AMFR S. Philosophical transactions of the Royal Society of London, (53), 370-418.

Bartholomew, D. J. (2010). Principal components analysis.

Benjamens, S., Dhunnoo, P., & Meskó, B. (2020). The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *NPJ digital medicine*, *3*(1), 1-8.

Bennette, C. S., Ramsey, S. D., McDermott, C. L., Carlson, J. J., Basu, A., & Veenstra, D. L. (2016). Predicting low accrual in the National Cancer Institute's Cooperative Group clinical trials. JNCI: Journal of the National Cancer Institute, 108(2).

Bielza, C., & Larrañaga, P. (2020). Data-Driven Computational Neuroscience: Machine Learning and Statistical Models. Cambridge: Cambridge University Press. doi:10.1017/9781108642989

Booth CM, Cescon DW, Wang L, Tannock IF, Krzyzanowska MK. Evolution of the randomized controlled trial in oncology over three decades. J Clin Oncol. 2008; 26: 5458- 5464.

Bourke, A., Dixon, W. G., Roddam, A., Lin, K. J., Hall, G. C., Curtis, J. R., ... & Bartels, D. B. (2020). Incorporating patient generated health data into pharmacoepidemiological research. Pharmacoepidemiology and Drug Safety, 29(12), 1540-1549.

Braud, A., Fromentoux, G., Radier, B., and Le Grand, O. "The Road to European Digital Sovereignty with Gaia-X and IDSA," in IEEE Network, vol. 35, no. 2, pp. 4-5, March/April 2021, doi: 10.1109/MNET.2021.9387709.

Burton, A. (2018). How do we fix the shortage of neurologists?. The Lancet Neurology, 17(6), 502-503.

C. Chen, C. Hsu, H. Chiu and H. Rau, "Prediction of survival in patients with liver cancer using artificial neural networks and classification and regression trees," 2011 Seventh International Conference on Natural Computation, 2011, pp. 811-815, doi: 10.1109/ICNC.2011.6022187.

Center for Electronics and Information Studies, Chinese Academy of Engineering (2020). The Development of Deep Learning Technologies: Research on the Development of Electronic Information Engineering Technology in China, 13-45. https://doi.org/10.1007/978-981-15-4584-9

Chang, C. H., Lin, C. H., & Lane, H. Y. (2021). Machine Learning and Novel Biomarkers for the Diagnosis of Alzheimer's Disease. International Journal of Molecular Sciences, 22(5), 2761.

Chapelle, O., Scholkopf, B., & Zien, A. (2009). Semi-supervised learning (chapelle, o. et al., eds.; 2006)[book reviews]. *IEEE Transactions on Neural Networks*, *20*(3), 542-542.

Chi Heem Wong, Kien Wei Siah, Andrew W Lo, Estimation of clinical trial success rates and related parameters, Biostatistics, Volume 20, Issue 2, April 2019, Pages 273–286, https://doi.org/10.1093/biostatistics/kxx069

Centre Hospitalier Universitaire Vaudois, Lausanne University hospital, CHUV (2020). Human Brain Project. https://www.chuv.ch/en/neurosciences/dnc-home/research/human-brain-project

Cleophas, T. J., Zwinderman, A. H., & Cleophas-Allers, H. I. (2013). Machine learning in medicine (Vol. 9). Dordrecht, The Netherlands:: Springer.

Cleophas, T. J., & Zwinderman, A. H. (2015). Machine learning in medicine-a complete overview (pp. 17-24). Springer International Publishing.

Clinicaltrials.gov (2021). Trends, Charts, and Maps. NIH, National Library of Medicine

Cohen, S., Fehder, D. C., Hochberg, Y. V., & Murray, F. (2019). The design of startup accelerators. *Research Policy*, *48*(7), 1781-1797.

Combi, C., & Pozzi, G. (2019). Clinical Information Systems and Artificial Intelligence: Recent Research Trends. Yearbook of medical informatics, 28(1), 83–94. https://doi.org/10.1055/s-0039-1677915

Council of Europe (2020). Bulgarian psychiatric and social care institutions: ill-treatment, acute understaffing, mechanical restraint misuse. https://www.coe.int/en/web/portal/-/bulgarian-psychiatric-and-social-care-institutions-ill-treatment-acute-understaffing-mechanical-restraint-misuse

Care Quality Commission (2017). The state of care in mental health services 2014-2017. State of Care. www.cqc.org.uk

Cummings SR. Clinical Trials Without Clinical Sites. JAMA Intern Med. Published online March 01, 2021. doi:10.1001/jamainternmed.2020.9223

Cybersecurity and Infrastructure Security Agency, CISA (2020). Alert (AA20-302A). Ransomware Activity Targeting the Healthcare and Public Health Sector. National Cyber Awareness System. https://us-cert.cisa.gov/ncas/alerts/aa20-302a

Daly, A., & Craig, S. (2019). Irish Psychiatric Units and Hospitals Census 2019 Main Findings. Health Research Board.

D'Alvano, L., & Hidalgo, A. (2012). Innovation management techniques and development degree of innovation process in service organisations. R&D Management, 42(1), 60-70.

Degoulet, P. And Fieschi, M. (2012). Introduction to Clinical Informatics. Springer Science & Business Media.

Demiraj, A., Spartalis, I., Karozos, K., and Vassalos, V. (2019). Meta-Data Management and Quality Control for the Medical Informatics Platform. ACM Digital Library. Proceedings of the 23rd International Database Applications & Engineering Article No.: 10 Pages 1–9https://doi.org/10.1145/3331076.3331088

Demiris, G., Iribarren, S., Sward, K., Lee, S. and Yang, R. (2019). Patient generated health data use in clinical practice: A systematic review. Nursing Outlook. Volume 67, Issue 4, July–August 2019, Pages 311-330. https://doi.org/10.1016/j.outlook.2019.04.005

Destatis (2019). Statistisches Bundesamt. https://www.destatis.de/

Díaz-Castro, L., Cabello-Rangel, H., Medina-Mora, M. E., Berenzon-Gorn, S., Robles-García, R., & Madrigal-de León, E. Á. (2020). Necesidades de atención en salud mental y uso de servicios en población Mexicana con trastornos mentales graves. *salud pública de méxico*, *62*(1, ene-feb), 72-79.

Diaz, R; Navia, A., González, F., Medina, J. (2019). Machine Learning to Augment Shared Knowledge in Federated Privacy-Preserving Scenarios (MUSKETEER). https://musketeer.eu/wp-content/uploads/2019/10/MUSKETEER_D4.1_final.pdf

Ding, Y., Qu, J., Yu, X., & Wang, S. (2014). The mediating effects of burnout on the relationship between anxiety symptoms and occupational stress among community healthcare workers in China: a cross-sectional study. PloS one, 9(9), e107130.

Domingues, Renan Barros et al. Telemedicine in neurology: current evidence. Arquivos de Neuro-Psiquiatria [online]. 2020, v. 78, n. 12 [Accessed 27 July 2021] , pp. 818-826. Available from: <https://doi.org/10.1590/0004-282X20200131>. Epub 07 Dec 2020. ISSN 1678-4227. https://doi.org/10.1590/0004-282X20200131.

Duarte, S. L., & Garcia, M. L. T. (2013). Reforma psiquiátrica: trajetória de redução dos leitos psiquiátricos no Brasil. Emancipação, 13(1), 39-54.

European Academy of Neurology (2019). FACT SHEET: The burden of neurological diseases in Europe. https://www.ean.org/research/resources/the-burden-of-neurological-diseases

Eraslan, G., Avsec, Ž., Gagneur, J. et al. Deep learning: new computational modelling techniques for genomics. Nat Rev Genet 20, 389–403 (2019). https://doi.org/10.1038/s41576-019-0122-6

European Union Agency for Cybersecurity (ENISA) (2020). Ransomware. https://www.enisa.europa.eu/publications/ransomware/

Eurostat (2018). Mental health and related issues statistics. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Main_Page

Eurostat (2021). Data Browser. https://ec.europa.eu/eurostat/data/database

Essinger, J. (2014). Ada's algorithm: How Lord Byron's daughter Ada Lovelace launched the digital age. Melville House.

Finazzi S, Poole D, Luciani D, Cogo PE, Bertolini G (2011) Calibration Belt for Quality-of-Care Assessment Based on Dichotomous Outcomes. PLoS ONE 6(2): e16110. https://doi.org/10.1371/journal.pone.0016110

Fleming, T. R., & Lin, D. Y. (2000). Survival analysis in clinical trials: past developments and future directions. Biometrics, 56(4), 971-983.

Florian Privé, Keurcien Luu, Michael G B Blum, John J McGrath, Bjarni J Vilhjálmsson, Efficient toolkit implementing best practices for principal component analysis of population genetic data, Bioinformatics, Volume 36, Issue 16, 15 August 2020, Pages 4449–4457, https://doi.org/10.1093/bioinformatics/btaa520

Fougerou-Leurent C, Chesnais J, Nekmouche S, Veislinger A, Le Saux M, Joumard C, Lorre V, Bellot C, Alleton N, Labourdette E, Marie C, Fin L, Bellissant E, Laviolle B. Enquête sur la certification aux bonnes pratiques cliniques en recherche académique [A survey on French hospital physicians'certification to the good clinical practices]. Therapie. 2020 Nov-Dec;75(6):537-542. French. doi: 10.1016/j.therap.2020.04.002. Epub 2020 Apr 19. PMID: 32409207.

Fradkov, A. L. (2020). Early history of machine learning. IFAC-PapersOnLine, 53(2), 1385-1390.

Frackowiak, R. (2017). Future Medicine: Modern Informatics. TEDxYouth@Zurich. https://youtu.be/_jzpwaiU6Lo

Friedman, C., & Rigby, M. (2013). Conceptualising and creating a global learning health system. International journal of medical informatics, 82(4), e63-e71.

Fovet, T., Amad, A., Horn, M., Thomas, P., & Chan-Chee, C. (2020). Utilization of hospital-level mental health care services for inmates in France: a transversal study. Psychiatric Services, 71(8), 824-828.

Galili, T., Mitelpunkt, A., Shachar, N., Marcus-Kalish, M., & Benjamini, Y. (2014, October). Categorize, cluster, and classify: a 3-c strategy for scientific discovery in the medical informatics platform of the human brain project. In International Conference on Discovery Science (pp. 73-86). Springer, Cham.

Galtier, M. N., & Marini, C. (2019). Substra: a framework for privacy-preserving, traceable and collaborative machine learning. arXiv preprint arXiv:1910.11567.

General Data Protection Regulation, GDPR (2016). Regulation (Eu) 2016/679 of the European Parliament and of the Council of 27 April 2016. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

Grant, R. M., & Baden-Fuller, C. (2004). A knowledge accessing theory of strategic alliances. Journal of management studies, 41(1), 61-84.

Hall, D. A., Ramos, A. R., Gelfand, J. M., Videnovic, A., Benatar, M., Cahill, C., ... & Goldman, M. (2018). The state of clinical research in neurology. Neurology, 90(15), e1347-e1354.

Haux, R., Ammenwerth, E., Winter, A., & Brigl, B. (2004). Strategic information management in hospitals: an introduction to hospital information systems. Springer Science & Business Media.

HealthIT (2018). What is computerized provider order entry?. https://www.healthit.gov/faq/what-computerized-provider-order-entry

Hicks JK, Dunnenberger HM, Gumpper KF, Haidar CE, Hoffman JM. Integrating pharmacogenomics into electronic health records with clinical decision support. Am J Health Syst Pharm. 2016 Dec 1;73(23):1967-1976. doi: 10.2146/ajhp160030. PMID: 27864204; PMCID: PMC5117634.

Hidalgo, A., Villanueva, M., Saturio, P. (2009). Derechos de propiedad industrial e intelectual en proyectos tecnológicos de cooperación público-privada. CDTI. Ministerio de Ciencia e Innovación. España.

Holwerda, S.W., Holland, M.T., Green, A.L. et al. Dissociation between reduced pain and arterial blood pressure following epidural spinal cord stimulation in patients with chronic pain: A retrospective study. Clin Auton Res 31, 303–316 (2021). https://doi.org/10.1007/s10286-020-00690-5
Human Brain Project (2021). "HOW DOES MIP WORK?. https://www.humanbrainproject.eu/en/medicine/medical-informatics-platform/how-does-mip-works/

Huesch, M. D., & Mosher, T. J. (2017). Using it or losing it? The case for data scientists inside health care. *Nejm Catalyst*, *3*(3).

IBM SPSS Statistics. Manual Del Usuario Del Sistema Basico de IBM SPSS Statistics 20.

IBM (2021). Natural Language Processing (NLP).

ICPE All Access conference abstracts (2020). Pharmacoepidemiol Drug Saf, 29: 3-634. https://doi.org/10.1002/pds.5114

Inan, O.T., Tenaerts, P., Prindiville, S.A. et al. Digitizing clinical trials. npj Digit. Med. 3, 101 (2020). https://doi.org/10.1038/s41746-020-0302-y

Insee (2020). Personnels et équipements de santé. https://www.insee.fr/fr/statistiques/4277748?sommaire=4318291#titre-bloc-3

INTERPOL (2020). Cybercrime, COVID-19 Impact. France. www.interpol.int

Intelligence, S. P. G. M. (2018). GICS® Global Industry Classification Standard.

Islam, M. M., Poly, T. N., & Li, Y. J. (2018). Recent Advancement of Clinical Information Systems: Opportunities and Challenges. Yearbook of medical informatics, 27(1), 83–90. https://doi.org/10.1055/s-0038-1667075

Institute of Medicine (US) (2009). Roundtable on Evidence-Based Medicine. Leadership Commitments to Improve Value in Healthcare: Finding Common Ground: Workshop Summary. Washington (DC): National Academies Press (US). Institute of Medicine: Roundtable on Evidence-Based Medicine. Available from: https://www.ncbi.nlm.nih.gov/books/NBK52847/

Instituto Nacional de Estatística - Estatísticas da Saúde : 2019. Lisboa : INE, 2021. Disponível na www: <url:https://www.ine.pt/xurl/pub/257483090>. ISSN 2183-1637. ISBN 978-989-25-0560-2

Irish Hospital Consultants Association (IHCA). Opening Statement to the Meeting of the Joint Committee on Health on Workforce Planning, 21st October 2020. https://data.oireachtas.ie/ie/oireachtas/committee/dail/33/joint_committee_on_health/submissions/2020/2020-10-21_opening-statement-professor-alan-irvine-president-irish-hospital-consultants-association_en.pdf

Jamal, S., Goyal, S., Shanker, A. et al. Integrating network, sequence and functional features using machine learning approaches towards identification of novel Alzheimer genes. BMC Genomics 17, 807 (2016). https://doi.org/10.1186/s12864-016-3108-1

Jones, N. (2018). How to stop data centres from gobbling up the world's electricity. Nature, 561(7722), 163-167.

Jung, B.C., Choi, S.I., Du, A.X. et al. Principal Component Analysis of Cerebellar Shape on MRI Separates SCA Types 2 and 6 into Two Archetypal Modes of Degeneration. Cerebellum 11, 887–895 (2012). https://doi.org/10.1007/s12311-011-0334-6

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., ... & Zhao, S. (2019). Advances and open problems in federated learning. arXiv preprint arXiv:1912.04977.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. Journal of the American statistical association, 53(282), 457-481.

Kasenda, B., Liu, J., Jiang, Y. et al. Prediction of RECRUITment In randomized clinical Trials (RECRUIT-IT)—rationale and design for an international collaborative study. Trials 21, 731 (2020). https://doi.org/10.1186/s13063-020-04666-8

Kashambwa, R. (2020). PND119 ASSOCIATION BETWEEN LIPID LEVELS AND HEMORRHAGIC STROKE AMONG WOMEN. Value in Health, 23, S282.

Kazancigil, M. A. (2019, July). Innovations and convergence in mobile medical applications and cloud-based hospital information systems for the real-time monitoring of patients and early warning of diseases. In *2019 IEEE World Congress on Services (SERVICES)* (Vol. 2642, pp. 301-306). IEEE.

Kesselheim AS, Cresswell K, Phansalkar S, Bates DW, Sheikh A. Clinical decision support systems could be modified to reduce 'alert fatigue' while still minimizing the risk of litigation. Health Aff (Millwood). 2011 Dec;30(12):2310-7. doi: 10.1377/hlthaff.2010.1111. PMID: 22147858.

Kim, M., Snowden, S., Suvitaival, T., Ali, A., Merkler, D. J., Ahmad, T., ... & Legido-Quigley, C. (2019). Primary fatty amides in plasma associated with brain amyloid burden, hippocampal volume, and memory in the European Medical Information Framework for Alzheimer's Disease biomarker discovery cohort. Alzheimer's & Dementia, 15(6), 817-827.

Kroneman, M., Boerma, W., van den Berg, M., Groenewegen, P., de Jong, J., van Ginneken, E., & World Health Organization. (2016). Netherlands: health system review.

Krylov, V. V., Konovalov, A. N., Dash'yan, V. G., Kondakov, E. N., Tanyashin, S. V., Gorelyshev, S. K., ... & Sytnik, A. V. (2017). Neurosurgery in Russian Federation. Zhurnal Voprosy Neyrokhirurgii imeni NN Burdenko, 1, 5-11.

Kononenko, I. (2001). Machine learning for medical diagnosis: history, state of the art and perspective. Artificial Intelligence in medicine, 23(1), 89-109.

Kostis, W. J., Dobrzynski, J. M., & Kostis, J. B. (2019). Advantages and Disadvantages in Clinical Trials. In Management of Hypertension (pp. 215-238). Springer, Cham.

Kramer, S., Bradfield, J.C. A general definition of malware. J Comput Virol 6, 105–114 (2010). https://doi.org/10.1007/s11416-009-0137-1

Krebs, K., Milani, L. Translating pharmacogenomics into clinical decisions: do not let the perfect be the enemy of the good. Hum Genomics 13, 39 (2019). https://doi.org/10.1186/s40246-019-0229-z

Kümmerle, C., Verdun, C. M., & Stöger, D. (2020). Iteratively Reweighted Least Squares for $\ell_1 $-minimization with Global Linear Convergence Rate. arXiv preprint arXiv:2012.12250.

Kush, R. D., Warzel, D., Kush, M. A., Sherman, A., Navarro, E. A., Fitzmartin, R., ... & Hudson, L. (2020). FAIR data sharing: the roles of common data elements and harmonization. Journal of biomedical informatics, 107, 103421.

Lampert-Grassi, María Pilar (2018). Modelos Comunitarios en Salud Mental: Chile, España e Inglaterra. Biblioteca del Congreso Nacional de Chile. Nº SUP: 116251

Larobina, M. and Murino L. (2014). Medical image file formats. Springer. Journal of Digital Imaging. 27(2): 200–206. 10.1007/s10278-013-9657-9

León, G. (2020). Tecnologías estratégicas relacionadas con COVID-19, impacto y dependencia para España. Fundación Rogelio Segovia para el Desarrollo de las Telecomunicaciones. Universidad Politécnica de Madrid. España.

Li, J. S., Yu, H. Y., & Zhang, X. G. (2011). Data mining in hospital information system. New Fundamental Technologies in Data Mining, 1st ed.; Funatsu, K., Ed, 143-172.

Lin E, Lin CH, Lane HY. Precision Psychiatry Applications with Pharmacogenomics: Artificial Intelligence and Machine Learning Approaches. Int J Mol Sci. 2020 Feb 1;21(3):969. doi: 10.3390/ijms21030969. PMID: 32024055; PMCID: PMC7037937.

Lohmann, L., Lammerskitten, A., Korsen, M., Dodel, R., Gaul, C., Hamer, H. M., ... & Meuth, S. G. (2021). Status of clinical research in neurology in Germany—A national survey. European Journal of Neurology, 28(5), 1446-1452.

Maciukiewicz M., Marshe V.S., Hauschild A.C., Foster J.A., Rotzinger S., Kennedy J.L., Kennedy S.H., Muller D.J., Geraci J. GWAS-based machine learning approach to predict duloxetine response in major depressive disorder. J. Psychiatr. Res. 2018;99:62–68. doi: 10.1016/j.jpsychires.2017.12.009.

Mangesius S, Mariotto S, Ferrari S, Pereverzyev S Jr, Lerchner H, Haider L, Gizewski ER, Wenning G, Seppi K, Reindl M, Poewe W. Novel decision algorithm to discriminate parkinsonism with combined blood and imaging biomarkers.

Microsoft Azure (2021). https://azure.microsoft.com/en-us/

Markram, H., Meier, K., Lippert, T., Grillner, S., Frackowiak, R., Dehaene, S., ... & Saria, A. (2011). Introducing the human brain project. *Procedia Computer Science*, *7*, 39-42.

Masurkar, P.P., Chatterjee, S., Sherer, J.T. et al. Antimuscarinic Cascade Across Individual Cholinesterase Inhibitors in Older Adults with Dementia. Drugs Aging 38, 593–602 (2021). https://doi.org/10.1007/s40266-021-00863-5

Matin, B. K., Karyani, A. K., Rezaei, S., Soofi, M., & Soltani, S. (2019). Do countries with higher GDP spend more on disabilities? New evidence in OECD countries. Medical journal of the Islamic Republic of Iran, 33, 122.

MedCalc (1993-2017). MedCalc Software bvba. Software manual

Menasalvas, E., Gonzalo, C., & RODRÍGUEZ-GONZÁLEZ, A. (2017). Big Data en Salud: retos y oportunidades. *Economía Industrial (Ejemplar dedicado a: Nuevas tecnologías digitales)*, *405*, 87-97.

Menear, M., Blanchette, MA., Demers-Payette, O. et al. A framework for value-creating learning health systems. Health Res Policy Sys 17, 79 (2019). https://doi.org/10.1186/s12961-019-0477-3

Miao, R., Xia, LY., Chen, HH. et al. Improved Classification of Blood-Brain-Barrier Drugs Using Deep Learning. Sci Rep 9, 8802 (2019). https://doi.org/10.1038/s41598-019-44773-4

Ministerio de Sanidad, Consumo y Bienestar Social. Secretaría General Técnica (2016). Estadísticas de Centros Sanitarios de Atención Especializada. Hospitales y Centros sin Internamiento. Centro de Publicaciones. España

MIP Deployment and Support Team (2020). Medical Informatics Platform (MIP). MIP DATA FACTORY. Human Brain Project SP8.2. Data Processing User Guide.

Mirsky, Y., Mahler, T., Shelef, I., & Elovici, Y. (2019). CT-GAN: Malicious tampering of 3D medical imagery using deep learning. In 28th {USENIX} Security Symposium ({USENIX} Security 19) (pp. 461-478).

Mitelpunkt, A., Galili, T., Kozlovski, T. et al. Novel Alzheimer's disease subtypes identified using a data and knowledge driven strategy. Sci Rep 10, 1327 (2020). https://doi.org/10.1038/s41598-020-57785-2

MONAI (2021). https://monai.io

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. Decentralized Business Review, 21260.

Nattino, G., Lemeshow, S., Phillips, G., Finazzi, S., & Bertolini, G. (2017). Assessing the calibration of dichotomous outcome models with the calibration belt. The Stata Journal, 17(4), 1003-1014.

Nunes, A., Ardau, R., Berghöfer, A., Bocchetta, A., Chillotti, C., Deiana, V., ... & Alda, M. (2020). Prediction of lithium response using clinical data. Acta Psychiatrica Scandinavica, 141(2), 131-141.

OECD/Eurostat/WHO (2017), *A System of Health Accounts 2011: Revised edition*, OECD Publishing, Paris, https://doi.org/10.1787/9789264270985-en.

OECD (2014), "Public psychiatric hospitals per 100 000 population in selected OECD countries, 2010", in Making Mental Health Count: The Social and Economic Costs of Neglecting Mental Health Care, OECD Health Policy Studies, OECD Publishing, Paris, https://doi.org/10.1787/9789264208445-graph22-en.

Office Fédéral de la statistique (2016). Statistiques des hôpitaux 2016. Département fédéral de l'intérieur DFI Confédération suisse

Ohmann C, Banzi R, Canham S, et al Sharing and reuse of individual participant data from clinical trials: principles and recommendations BMJ Open 2017;7:e018647. doi: 10.1136/bmjopen-2017-018647

Ohsfeldt, R. L., Ward, M. M., Schneider, J. E., Jaana, M., Miller, T. R., Lei, Y., & Wakefield, D. S. (2005). Implementation of hospital computerized physician order entry systems in a rural state: feasibility and financial impact. Journal of the American Medical Informatics Association : JAMIA, 12(1), 20–27. https://doi.org/10.1197/jamia.M1553

Patel, S., Maxwell, C., & Gliebus, G. (2019). MCI and OSA (P4. 1-022). Neurology, 92, P4.1-022 (AAN 71st Annual Meeting). https://n.neurology.org/content/92/15_Supplement/P4.1-022.abstract

Patterson, V. Neurological telemedicine in the COVID-19 era. Nat Rev Neurol 17, 73–74 (2021). https://doi.org/10.1038/s41582-020-00438-9

Pillai AS, Jirsa VK. Symmetry Breaking in Space-Time Hierarchies Shapes Brain Dynamics and Behavior. Neuron. 2017 Jun 7;94(5):1010-1026. doi: 10.1016/j.neuron.2017.05.013. PMID: 28595045.

Rashidi, H. H., Tran, N. K., Betts, E. V., Howell, L. P., & Green, R. (2019). Artificial intelligence and machine learning in pathology: the present landscape of supervised methods. Academic pathology, 6, 2374289519873088.

Ravelli DP. Deinstitutionalisation of mental health care in The Netherlands: towards an integrative approach. Int J Integr Care. 2006;6:e04. doi: 10.5334/ijic.146. Epub 2006 Mar 15. PMID: 16896384; PMCID: PMC1480375.

Redolfi A, De Francesco S, Palesi F, et al. Medical Informatics Platform (MIP): A Pilot Study Across Clinical Italian Cohorts. Front Neurol. 2020;11:1021. Published 2020 Sep 23. doi:10.3389/fneur.2020.01021

Reuter, E. and Vaidya, A. (2021). After Google, Microsoft struck out, can Apple get people to use its health records feature?. MedCityNews. https://medcitynews.com/2021/09/546989/

Rimel, B. J. (2016). Clinical trial accrual: obstacles and opportunities. Frontiers in oncology, 6, 103.

Rodgers, F., Pepperrell, T., Keestra, S., & Pilkington, V. (2021). Missing clinical trial data: the evidence gap in primary data for potential COVID-19 drugs. Trials, 22(1), 1-10.

Rosenfeld, B. (2015). History of Statistics 8. Vermont Mathematics Initiative. https://higherlogicdownload.s3.amazonaws.com/AMSTAT/1484431b-3202-461e-b7e6-ebce10ca8bcd/UploadedImages/Classroom_Activities/HS_8__FISHER_and_Design_of_experiments.pdf

Royal College of Psychiatrists (2017). Postcode lottery for psychiatric care. https://www.rcpsych.ac.uk/news-and-features/latest-news/detail/2017/09/11/postcode-lottery-for-psychiatric-care

Ryan, P., Rao, G., Feeney, K., Sena, A., Klebanov, G. OHDSI Tools Ecosystem. https://www.ohdsi.org/wp-content/uploads/2018/10/OHDSI-Ecosystem-tutorial-Ryan-Rao-Feeney-Sena-Klebanov-11Oct2018.pdf

Russian Federal State Statistics Service (2021). https://eng.rosstat.gov.ru

Salm, M., & Wübker, A. (2020). Sources of regional variation in healthcare utilization in Germany. Journal of health economics, 69, 102271.

Salto-Tellez M, Maxwell P, Hamilton P. Artificial intelligence-the third revolution in pathology. Histopathology. 2019 Feb;74(3):372-376. doi: 10.1111/his.13760. PMID: 30270453.

Saxena, A., Pare, S., Meena, M. S., Gupta, D., Gupta, A., Razzak, I., ... & Prasad, M. (2020). A Two-Phase Approach for Semi-Supervised Feature Selection. Algorithms, 13(9), 215.

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. IBM Journal of research and development, 3(3), 210-229.

Sánchez, I. (2021). Europa busca su independencia digital en los servicios de datos en la nube, a través del proyecto europeo GAIA-X para contrarrestar como en otros ámbitos el liderazgo norteamericano y chino. https://www.mundiario.com/articulo/economia/democracia-digital-europea-depende-solo-nube/20210502101912218292.html

Sewani, H., & Kashef, R. (2020). An Autoencoder-Based Deep Learning Classifier for Efficient Diagnosis of Autism. Children (Basel, Switzerland), 7(10), 182. https://doi.org/10.3390/children7100182

Shalev-Shwartz, S., & Ben-David, S. (2014). Understanding machine learning: From theory to algorithms. Cambridge university press.

Sittig, D. F., & Singh, H. (2016). A Socio-Technical Approach to Preventing, Mitigating, and Recovering from Ransomware Attacks. Applied clinical informatics, 7(2), 624–632. https://doi.org/10.4338/ACI-2016-04-SOA-0064

Stacey, J. L. (2020, July). TNF-α INHIBITORS IN PREVENTING MILD COGNITIVE IMPAIRMENT (MCI), DEMENTIA, AND ALZHEIMER'S DISEASE (AD): AN UPDATED RETROSPECTIVE REVIEW USING BOTH EMR AND CLAIMS DATA. In 2020 Alzheimer's Association International Conference. ALZ.

Staudt, M. D. (2019). Neurosurgical Training in Canada: Changing Landscape with Ongoing Challenges. *World neurosurgery*, *132*, 26-28.

Stern N., Freeman H. (1988) Who Invented the First Electronic Digital Computer?. In: Weiss E.A. (eds) A Computer Science Reader. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-8726-6_6

Stanford Medicine (2017). Harnessing the Power of Data in Health. Health Trends Report. https://med.stanford.edu/content/dam/sm/sm-news/documents/StanfordMedicineHealthTrendsWhitePaper2017.pdf

Stanton, J. M. (2001). Galton, Pearson, and the peas: A brief history of linear regression for statistics instructors. Journal of Statistics Education, 9(3).

Statista (2017) Mental hospitals in the biggest cities of France 2017.

Tavazzi L. (2019). Big data: is clinical practice changing?. European heart journal supplements : journal of the European Society of Cardiology, 21(Suppl B), B98–B102. https://doi.org/10.1093/eurheartj/suz034

Taquet, M., Geddes, J. R., Husain, M., Luciano, S., & Harrison, P. J. (2021). 6-month neurological and psychiatric outcomes in 236 379 survivors of COVID-19: a retrospective cohort study using electronic health records. The Lancet Psychiatry, 8(5), 416-427.

The Association of Health Institutions of Slovenia (2021). https://www.zdrzz.si/f/docs/zlozenka-en/THE_ASSOCIATION_OF_HEALTH_INSTITUTIONS_OF_SLOVENIA.pdf

Tiwari, N., Rai, A. K., Sahani, A., & Maurya, A. K. Survey Paper On Hospital Management System (HMS).

Toga, A. (2018). Data Sharing in Alzheimer's Disease Research. US Neurology. 2018;14(2):68–70 DOI: https://doi.org/10.17925/USN.2018.14.2.68

Tolosa, E., Wenning, G., & Poewe, W. (2006). The diagnosis of Parkinson's disease. The Lancet Neurology, 5(1), 75-86.

U.S. Food and Drug Administration (FDA). https://www.fda.gov

Venetis, T., Ailamaki, A., Heinis, T., Karpathiotakis, M., Kherif, F., Mitelpunkt, A., & Vassalos, V. (2015, August). Towards the identification of disease signatures. In International Conference on Brain Informatics and Health (pp. 145-155). Springer, Cham.

Venetis, T., & Vassalos, V. (2015, July). Data integration in the human brain project. In *International Conference on Data Integration* in the Life Sciences (pp. 28-36). Springer, Cham.

Vickers AJ. Clinical trials in crisis: Four simple methodologic fixes. Clin Trials. 2014 Dec;11(6):615-21. doi: 10.1177/1740774514553681. Epub 2014 Oct 1. PMID: 25278228; PMCID: PMC4229450.

Voigt, I., Inojosa, H., Dillenseger, A., Haase, R., Akgün, K., & Ziemssen, T. (2021). Digital Twins for Multiple Sclerosis. *Frontiers in immunology*, *12*, 669811. https://doi.org/10.3389/fimmu.2021.669811

Weber-Jahnke, J. H. (2011, May). A preliminary study of apparent causes and outcomes of reported failures with patient management software. In Proceedings of the 3rd Workshop on Software Engineering in Health Care (pp. 5-8).

Wehling, P., Viehöver, W., & Koenen, S. (Eds.). (2014). The public shaping of medical research: patient associations, health movements and biomedicine. Routledge.

WHO (2018). Ageing and health. https://www.who.int/news-room/fact-sheets/detail/ageing-and-health

WHO (2017). Depression and other common mental disorders: global health estimates (No. WHO/MSD/MER/2017.2). World Health Organization.

WHO (2020). The Global Health Observatory (GHO).

Winter, A., Haux, R., Ammenwerth, E., Brigl, B., Hellrung, N., & Jahn, F. (2010). Health information systems. In Health Information Systems (pp. 33-42). Springer, London

Wright, A., Sittig, D. F., Ash, J. S., Sharma, S., Pang, J. E., & Middleton, B. (2009). Clinical decision support capabilities of commercially-available clinical information systems. *Journal of the American Medical Informatics Association*, *16*(5), 637-644.

Zhang, Q., Schultz, J. L., Aldridge, G. M., Simmering, J. E., & Narayanan, N. S. (2020). Coronavirus Disease 2019 Case Fatality and Parkinson's Disease. Movement disorders : official journal of the Movement Disorder Society, 35(11), 1914–1915. https://doi.org/10.1002/mds.28325

Zwolenski, M., & Weatherill, L. (2014). The digital universe: Rich data and the increasing value of the internet of things. Journal of Telecommunications and the Digital Economy, 2(3), 47-1.