GPI R&D

Trajectories & Successful Cases

2023

GPI SpA
I-38123 Trento (TN). Via Ragazzi del '99, n. 13
T +39 0461 381515 / F +39 0461 381599
info@gpi.it / PEC gpi@pec.gpi.it

R.I TN | C.F. | P.I. 01944260221
R.E.A. C.C.I.A.A. TN nr 189428
Cap. Soc. Euro 13.890.324,40 i.v.

Summary

Introduction

GPI is a group of companies controlled by the parent company GPI spa, which has been operating in the healthcare and social services market since 1988.

This alliance combines the range of skills of the companies involved with the aim of extending and improving the services offered.

GPI's business focuses on healthcare information systems, welfare and social services, home assistance, rehabilitation, rest homes, domotics, multi-services for healthcare and hospital organisations, telecommunications, social and sanitary call centres, online payment systems, mobile technologies, as well as consulting services on technology and organisation.

There are over 2400 customers (1500 in healthcare) that work with GPI Group solutions: institutions, healthcare bodies, public and private hospitals, day-care centres communities, residential healthcare centres, nursing homes, childcare centres, credit institutions, organised distribution retailers, small and medium sized enterprises and more.

At the end of 2021 the number of employees rose to almost 7.000. To learn more, visit www.gpi.it.

Among them, the GPI Research Centre was founded in order to achieve innovative objectives. The R&D department focuses on:

- Virtual Care
- Digital twins
- Patient Empowerment
- Artificial Intelligence
- Digital Therapeutics
- Population Health Management
- Augmented Telemedicine
- Augmented Epidemiology
- IoT devices
- Health Analytics

GPI's ability to go beyond the limits is proved by a big number of successful use cases.

POHEMA, end-to-end Digital Care platform

The aim of transforming the actual waiting medicine into an initiative and proactive one is at the heart of our vision. Today, most of the healthcare systems wait for traumas, pathologies and symptomatology to occur before intervening.

GPI's Pohema solution instead, provides a modular architecture that encompasses all of the dimensions of the eHealth opportunities to support an organisation and a supply of services individually for each patient, with the scope of coordinating care by anticipating the health demand.

The system can be configured in endless ways: components can be progressively selected and implemented, based on the requirements of the specific region and/or pathology.

According to this, our R&D laboratories focused on developing solutions that can meet different needs expressed by territories and related organisations, with the aim of delivering a configurable platform à la carte. In fact, our solution Pohema is an architecture made of software, technologies and services structured on a composable care approach. The modular structure ensures flexibility and adaptability to the different environments where it is implemented.

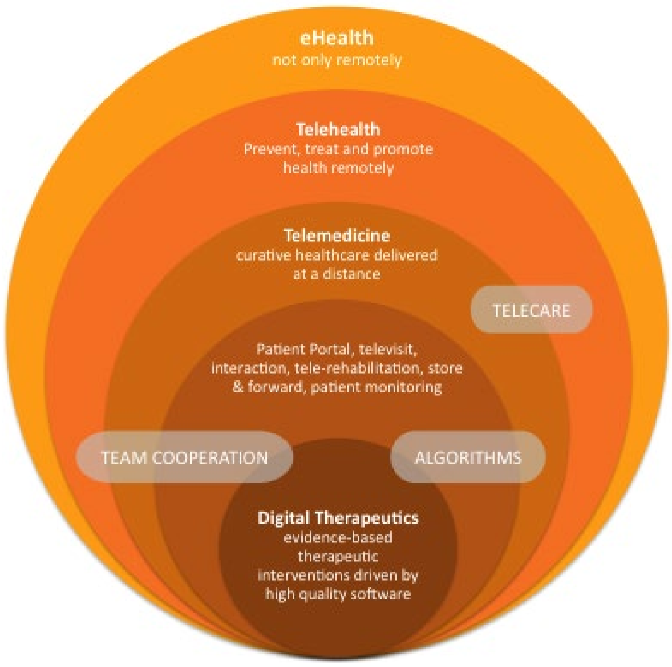
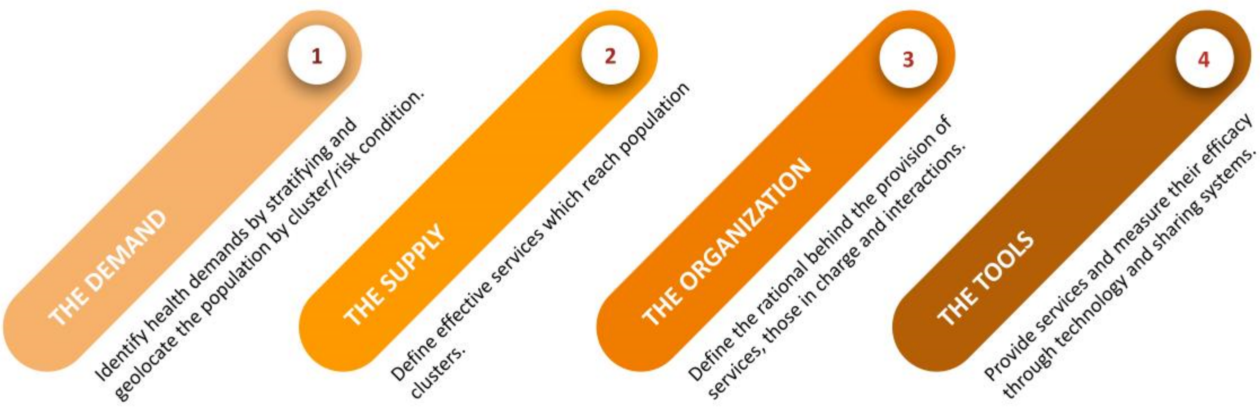


Figure 1 - Virtual Care

POHEMA'S operational model: the Population Health Management

The organisational structures of Population Health Management (PHM) provide an operational model intended to "simultaneously improve *Health*, *Costs* and *Experience* (Engagement and Empowerment) for a given group of people (Risk Stratification)" **- Gartner 2016.**

GPI has achieved the status of **global solutions provider** for PHM; it is in fact able to autonomously supply all the necessary services and technologies.



Our value proposition is based on the assumption described in the Population Health Management model that is intended to "*simultaneously improve Health, Costs and Experience for a given group of people*".

In fact we focused on the 4 dimensions of the model: identify and map the health demand, define services for baskets of patients sharing the same risk, redesign the organisation interactions and structure, provide tools & services to the involved actors for the supply of the healthcare services.

The intention of developing such a kind of end-to-end solution aims at ensuring specific benefits to the healthcare systems. We want to: improve the experiences of the actors involved in the processes (patient-centred approach), improve the adherence to therapies, improve the coordination of the care teams, improve access and sharing of the information, improve system efficiency and response time.

In order to do so, we had to focus on all of the different dimensions of Virtual Care: telemonitoring, televisit, engagement, interactions, imaging and team coordination. I want to specifically highlight the two dimensions which are really important for us: AI Algorithms and Digital Therapeutics.

The first topic regards the possibility of training algorithms towards disease prediction and clustering ability. Just to give an example, we have been able to develop an AI algorithm able to predict acute symptoms for patients characterised by cardiological pathologies.

Digital Therapeutics is important because of the concept of empowering the patient. In fact, our aim, together with Research Centres and University, is to demonstrate via trials that Digital Therapies have a positive and measurable impact on the patient's health status, such as a pharmaceutical active principle. Many projects regarding Chronicity management are following this scope.

Modular architecture, technology and methodologies that improve the digital experience for all stakeholders.

- **Patient Monitoring**
- **Patient Portal**
- **Televisit**
- **Patient Interaction**
- **Tele-rehabilitation**
- **Store & Forward**
- **A.I. Algorithms**
- **Care Team Cooperation**
- **Digital Therapeutics**

Figure 2 - POHEMA features

POHEMA: Augmented Telemedicine

Thanks to the use of IoT devices for the continuous detection of vital parameters, such as T1 Heart Monitor of Umana Technologies, it is possible to receive about 70k data per second for each patient monitored. This makes it possible to increase the information available by developing algorithms, including predictive ones. Below are two enabled algorithms:

- *COVID-19 Monitoring Algorithm*: Based on Respiratory Rate and Heart Rate, the algorithm recognises frequency peaks due to motor activity of the person being monitored, distinguishing them from positive covariance trends, typical of a trend towards respiratory failure, this allows early alarms to be triggered.

  In particular, the values of respiratory rate and heart rate, detected continuously, are pre-processed, a smoothing operation of the data is performed in order to reduce any peaks that are not due to emergency situations, but are instead caused by daily activities, such as climbing the stairs or doing physical activity and therefore should not produce any alert.

  From the results obtained, the normalised covariance or correlation coefficient is calculated. In our study, the focus was on a reference range between 0 and +1, as the aim is to identify a positive correlation between the number of heartbeats and time and between respiratory acts and time, in fact this type of correlation indicates that as time increases there has also been an increase in respiratory acts and/or heartbeats, therefore, if this situation continues it may be indicative of an emergency for which the doctor should be alerted.

- *Predictive Algorithm for Acute Stages of Heart Failure*: The clinical parameters used are Respiratory Rate, Heart Rate and Saturation. The algorithm performs a continuous assessment of the patient's physical activity using frequency peaks, analysing the distance between them over time. A thinning of the peaks indicates a lower propensity for movement in the patient and if a progressive decrease in saturation is observed on the peaks, this may indicate a change in the patient's clinical status. The clinic tells us that these changes in the heart failure patient begin gradually and even 10 or 15 days in advance of a life-threatening acute phase, which almost always requires hospitalisation, from which it can be deduced that a proactive approach can produce significant preventive effects. Alerts are provided to warn of the progression of the worsening clinical condition in order to prevent the onset of the acute phase of heart failure.

Talking About: Speech Emotion Recognition (SER)

In recent years, there is a growing interest in building technologies that are able to perform Speech Emotion Recognition (SER) from human speech. The project "Talking About" consists of a suite of Artificial Intelligence (AI) algorithms for building a reliable SER device from human speech.

Overall, SER technologies may be particularly useful for telehealth (e.g. video communication between patient and doctor). The audio stream from the patient can be relayed on the doctor's computer screen by which the doctor can see the patient's emotional state in real-time.

This modality should provide extra emotional pieces of information about the patient's current emotional state. Hopefully such SER feedback may help the doctor in listening more carefully to the patient's needs, therefore showing more empathy and checking the patient's compliance to therapy.

In conclusion, SER technologies are a promising field in AI which are in rapid expansion and whose implementation in healthcare will deliver several improvements for the coming years.

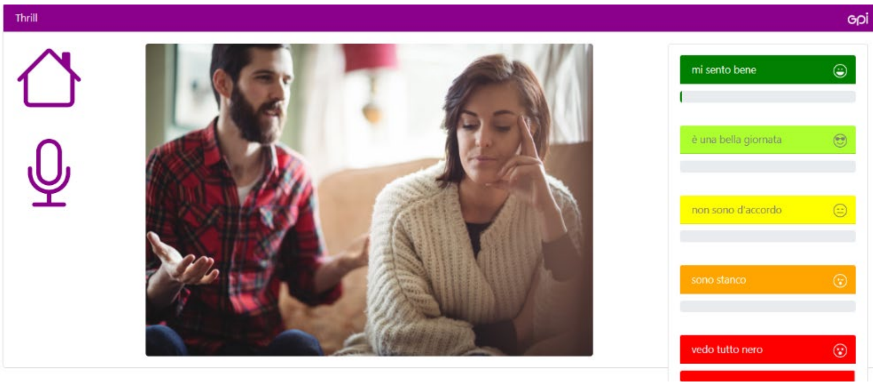


Figure 3 - Talking About front end

Augmented Epidemiology: One Health

"*While an unprecedented worldwide effort is under way to fight the COVID-19 pandemic, the persistent threats to the health of our planet call for urgent remedies too. Climate change, environmental pollution, biodiversity loss and an unsustainable use of natural resources pose multiple risks to human, animal and ecosystem health. They include infectious and noncommunicable diseases, antimicrobial resistance and water scarcity*"[1].

The disruptive nature of our initiative consists of the paradigm shift that we aim to, enabling the transition from an ex-post reporting model to a "predictive" paradigm and therefore ex-ante scenario approach, based on Pattern Recognition and AI Agent techniques enabling automated alerts critical to a more timely and effective Health Care and Environmental Policy making.

In fact, we strongly believe that Healthcare systems and operations can benefit from AI to help drive sustainability, efficiency and patient outcomes, starting from a new surveillance approach.

Since the European Commission adopted a recommendation on a European electronic health record exchange to unlock the flow of patient summary across borders, the setup of a Digital Health Infrastructure supporting evidence based decision making has become a priority for European Countries.

The pandemic has further reaffirmed the need to for Policy Makers to have new tools not only to timely track disease evolutions and epi conditions across the population, but also to inform faster intelligence, based on predictive techniques and scenario analysis, around the evolution of infectious diseases, and their correlation with social and environmental factors. A "One health" approach, in this regard, will be a further step forward ensuring tight coordination across human, environment and social dimensions, determinants and policy levers.

In fact, current epi detection practices, based on a long term, ex post monitoring model, have limitations, such as lack of standardisation, fragmentation, data capture and quality exist to varying degrees across Member States (MS), not only in the clinical space. Such variability, coupled with disease-specific surveillance objectives and priorities, necessitates to move towards a disease-by-disease and country-by-country, and even territory by territory approach.

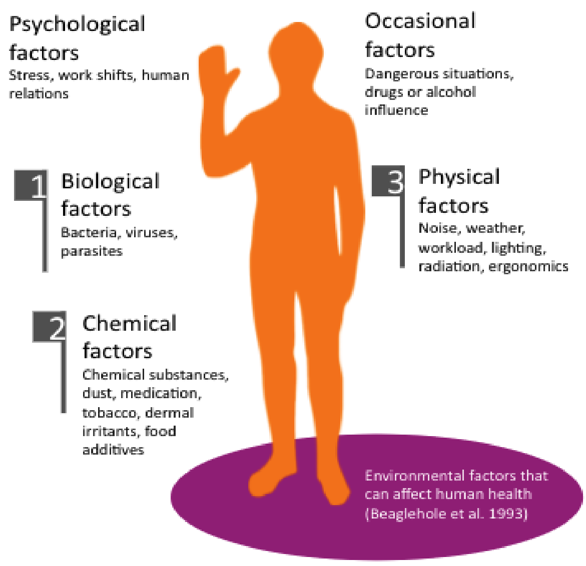


Figure 4 - One Health

With this initiative, GPI wants to invest jointly with IQVIA, with the objective of overcoming these limitations, looking for disease specific and enriched data sets available across Sentinel Areas, including laboratory, clinical and claims data matched with local environment and social vectors, providing the best proximity to the wider population. To this enriched locally relevant data sets, pseudo anonymized and anonymised, we will apply Pattern Recognition techniques, to identify non-evident patterns and consequently generate an alert system starting from the evidence that induces suspicion on the target phenomenon.

It is possible to measure a phenomenon through the observation of its effects, independently from its causes. It can be observable in relation to some qualitative and quantitative parameters. However, it is also possible to add another dimension: the observation of factors that can affect the phenomenon, or, in other words, making use of the determinants of exposure.

---

[1] COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS

This in particular allows to "increase" the number of variables that activate an effect and enables predictive processes. In others, the possibility is representing a phenomenon through the variables or patterns of variables that occur in the same phenomenon.

For example, age, sedentary life, high BMI (Body Mass Index) and waist circumference are very often observed in conjunction with type 2 diabetes. This type of approach allows epidemiological monitoring processes through the use of Pattern Recognition algorithms.

"*The field of Pattern Recognition concerns the automatic discovery of regularities in data through the use of computer algorithms and the use of regularities is to undertake actions such as the classification of data into different categories*" (C. M. Bishop 2006)

In other respects, the same approach may lead to the automatic discovery of little known, hidden or rather not evident phenomena upon observation of the facts as such.

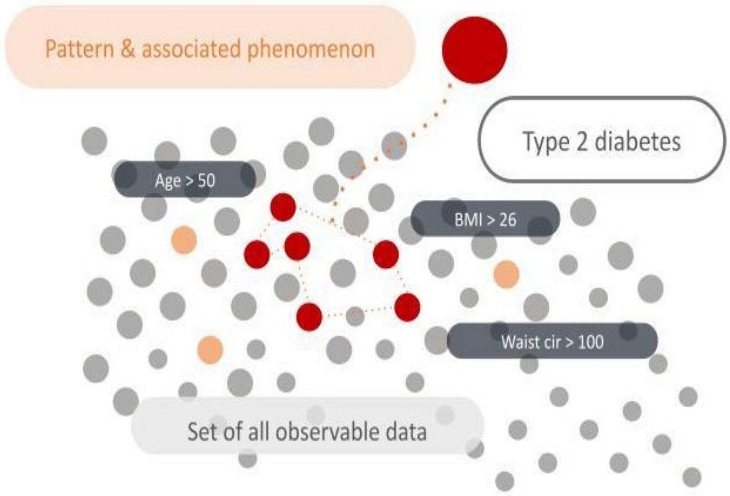


Figure 5 - Pattern recognition

Returning to the example of type 2 diabetes, it is obviously not certain that the occurrence of the conditions observed "very often concomitantly" implies the presence of the disease, rather it is very common for this to happen. Measuring the "very frequent" is equivalent to defining a level of risk of developing (or not treating) the disease, associated with the presence of concomitant factors. The number of these concomitant factors is a relevant aspect.

Pattern Recognition, in the presence of a sufficiently large set of observable data, is able to identify non-evident patterns and consequently generate an alert system starting from the evidence that induces suspicion on the target phenomenon.

The Project therefore aims to create an ICT platform capable of intercepting potential risks of the onset of pathological situations, not from ex-post analysis, but through algorithms capable of identifying, within the data managed, patterns that "trend" or "with great probability" can be considered precursors or concomitants of a public health phenomenon.

The innovative concept and tool we want to test and implement, also to overcome current limitation in Data Interoperability, is the "Hyper Individual", which we can define as: *"a limited territory, in which a population lives, which inherits the (anonymous) health data of each inhabitant from the latter, associating them with environmental data, so that it is possible to investigate and describe, also in an epidemiological sense, the health of a Hyper Individual who is in his territory as a person in his organism"*.

What we propose is to use anonymous environmental, veterinary, and climatic data, Electronic Health Record data and Clinical Pathology data referring to regional areas; by doing so, our Hyper Individual is represented over time by a Dataset defined by the boundaries of the territory for which the set of all data is incident.

These enriched datasets will allow through the adoption of algorithms to implement predictive objectives to introduce risk stratification models with forecasting components. At the same time, to introduce "Sentinel" techniques of observation of the territory in the continuous and immediate search for phenomena clinically relevant for the population.

RIPE: Risk Prediction in Elderly people

As part of the RI.P.E. Project, predictive models have been developed using Machine Learning techniques, that are able to provide indications about:

- glucose metabolism (prediction of average blood glucose values),
- renal insufficiency (prediction of average creatinine values),
- hypertension (prediction of average blood pressure values).

GPI in collaboration with the Galliera Hospital of Genoa was able to dispose of a dataset, Real World Data of about 8 million records (anonymous data). Starting from the Electronic Clinical Record of Diabetology with 20.000 enrolled patients and a depth of 10 years, it was possible to reconcile, for the same patients, the results of the clinical pathology examinations in the same period.

The dataset used for training the agent consists of features obtained from the patient's history (sex, age, BMI, waist circumference, systolic and diastolic blood pressure, smoking, years of disease) and from his laboratory tests (blood glucose, creatinine, triglycerides, cholesterol, albumin, plasma osmolarity, sodium, haemoglobin, glycated haemoglobin, microalbumin in urine).

After investigating several machine learning algorithms, the focus was on the Gradient Boosting Regressor (GBR). Boosting is a technique used to combine several simple estimators (e.g. decision trees) into a single complex model in order to obtain more robust predictors. For this reason boosting is known as the Ensemble Method. The term Gradient is derived from the fact that the algorithm uses gradient descent to minimise loss.

In order to verify the correct functioning of the algorithm, the value of the coefficient of determination (R2) was calculated. In addition, the mean squared error (MSE) values were calculated for the training and test phases. A further check was the application of cross-validation, which allows the algorithm score to be evaluated through cross-validation. The trained models performed well, with an accuracy of 97%.

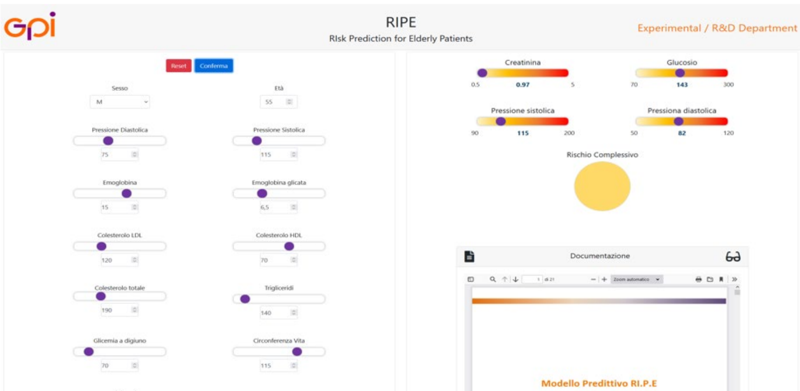


Figure 6 - Risk Prediction AI AGENT Front-end

In addition, it was possible to plot the histogram of the reported features in order of importance, i.e. the features that the model showed to be influential with respect to the Target variables (feature importance). This allowed the identification of regularities that are repeated within the dataset (Pattern Recognition). Specifically, it was interesting to recognise among the patient characteristics those that had a significant role in determining the pathological condition. For the models implemented, the characteristics of greatest diagnostic interest are:

- glucose metabolism: plasma osmolarity, albumin, microalbumin in urine;
- renal insufficiency: albumin, systolic blood pressure, plasma osmolarity;
- hypertension: plasma osmolarity, albumin, sodium.

What appears suggestive is that there is a Pattern of association that is preserved in the determination of the three targets, namely: increase in plasma osmolarity, decrease in albumin.

Risk Stratification: ML Algorithms for efficient vaccine plan definition

The underlying purpose of the study is to define Homogeneous Clusters according to the patient's health demand, on the basis of administrative data that can be marginalised from historical data and with respect to the services to be dedicated to each cluster in an offer framework territorial health services.

In this context, the study was guided by factors such as the number of patients affected by the disease and the resources absorbed by each patient, such as hospital admissions, emergency room access, outpatient services, drug consumption, home hospitalisation, etc.

The work privileged the aspects more typically referable to the categories of the data and its processing in a defined domain context; the activities did not have an epidemiological research purpose in the strict sense.

In order to describe the methodological approach implemented to obtain the stratification in risk cohorts, it is important to identify the individual phases that make it up:

- Input data collection
- Definition of prevalence
- Definition of risk cohorts

Once the subdivision of the reference population into the health risk clusters has been obtained, it is important to identify the context in which to use the information obtained and identify the objectives of the individual use cases.

As already mentioned, population stratification is a health planning tool, which defines the goal of providing a quantitative and qualitative characterization of health demand. The macro-objective that the stratification supports is Sustainability, understood as the aim of a multivariate approach in the strategies of the health offer according to an objective stratification: basket of services dedicated to homogeneous clusters of patients.

As previously indicated, one of the possible deployments of stratification in operational contexts is certainly related to the implementation of an efficient and effective Vaccination Plan from both a health and socio-economic point of view.

It is important to emphasise what are the three main objectives pursued through a vaccination plan:

- achievement of population immunity, and elimination of the risk of contagion
- effective and efficient allocation of resources used to cope with the disease, especially when characterised by acute symptoms
- improvement of the socio-economic living standards of the reference population, and implementation of lifestyle “normalisation” strategies

The importance of these objectives evidently lies in the need to protect the health of the target population and to allow the socio-economic fabric to have a prompt and safe restart.

Intuition: semantic inferential engine ontology based

Intuition[2] is a **semantic inferential engine ontology based**, designed to support **Public Job Centre** operators and **Company recruiters** in their search for ideal candidates for a specific job position or in the identification of job offers that are most similar to some specific curriculum vitae.

Intuition, thanks to a typing of the ontological profiles of the candidate and of the job offers, proposes a result that has a logical, hierarchical or proximity relationship with the one sought. In particular, Intuition, having received a job request as input (professional figure and associated attributes), returns a list, ordered by ranking, of the CVs that best respond to the request.

Intuition is an intuitive and easy-to-use application, able to obtain results from implicit information, through transversal connections of concepts and entities, connections that escape direct searches.

It uses an ontological model (OLM, Ontology of the Labour Market) describing the entities or classes living in the labour market and their existing or reciprocal relationships, expressed through natural language (informal) and frame based, i.e. based on concepts and properties. In addition, meaning is extracted from the text contained in CVs and applications from companies and job centres using NLP.

The development of Intuition took place in 3 phases:

1. In the first phase, the focus was on the construction of an ontology (a formal representation model of reality). Having identified the occupational profiles (ISTAT sources), the structure was elaborated by collecting relative attributes, synonyms and the resulting logical links. A total of 587 occupational profiles were elaborated, represented by 3827 attributes which generated about *4 million relations between attributes.*
2. The second phase allowed the creation of an inferential engine (based on Ontologies) to search, within the CVs, for the highest number of attributes related to a specific professional figure.
3. The third phase is the typification of the figures, a fundamental operation for the matching action implemented in the inferential engine. To this end, the attributes were processed (StopWords were removed, useful parts of the text were identified (PartOfSpeech) and the lemmatization operation was performed on the words). The lemmatization technique is a process of reducing a flexed form of a word to its canonical form, called “lemma”.

The project is currently in production since November 2021 in the Tuscany Region, integrated with the Job Portal, available to Employment Centres, citizens and companies looking for candidates.

[2] Intuition has been adopted by the Tuscany Region (Italy) since November 2021; more than 1000 Tuscan companies have already used it, producing over 2500 expressions of interest

PARENT: Monitoring and Integrating Neonates Behavioral and Physiological Parameters

PARENT is an EU-funded project aiming for early diagnosis of newborn motor/cognitive impairments. PARENT combines the efforts of a multidisciplinary network of leading European research groups, industry partners, paediatric hospitals and parents' associations to develop a technological infrastructure to conduct top-notch research in leading academic institutions, hospitals and industry R&D divisions. PARENT multidisciplinary approach and technologies will make a critical contribution towards an open neurodevelopmental disease diagnostic software infrastructure by interlinking disciplines from clinical data, neuroimaging collection and processing, biomarkers, data fusion, machine learning applied to clinical data, novel prediction algorithms. PARENT approach can be included in the more general paradigm of evidence-based medicine, precision medicine and patient centred-care, as well as decision support systems in the clinical field.

The study in the research of software architectures and IoT-based solutions to track neonates behavioural and physiological parameters in an incubator, as well as incubator settings (environments). The research will explore the ready-to-use solutions but also propose a novel and independent hardware-software solution to be easily integrated in any incubator to support wellbeing.

In this project GPI intends to achieve three goals:

- Standardised input and output data feeds, new AI algorithms for infant management.
- A ready-to-use software architecture and reusable models to describe the entire system too. This outcome can support software engineers and developers in the implementation of similar projects. The different software modules implemented to build the platforms or tools required. Novel IoT-based solutions to improve the neonates behavioural;
- A clinical decision support system for detecting signs of motor/cognitive impairments in premature infants based on various neurodiagnostic data sources.

More detail on Project website www.parenth2020.com.

ALCMEONE: Integrated planning and clinical management of the Headache patient

Industrial research and experimental development project, subsidies under the Fund for Sustainable Growth, Ministry of Economic Development, Horizon 2020 - PON 2014/2020

In line with the strategic objective of the Horizon 2020 Community Programme, new organisational models have been defined for the management of chronic conditions, often resulting from the increase in life expectancy, as well as the identification and development of innovative ITC systems to support the enhancement and control of care continuity paths.

A Machine Learning algorithm was developed to automatically recognise common data (Pattern Recognition) in a dataset, in order to provide one or more proposals for an Individual Care Plan (IAP) specific to each patient.

The peculiarity of this study is the high variability of the input data, in particular, in addition to the primary and secondary pathology, gender and age group, a list of services that make up the Care Pathway (PAI) was also considered and that are useful to "train" the algorithm to recognise the defined patterns.

The algorithm developed is based on the concepts of Machine Learning and Pattern Recognition. In particular, starting from a raw dataset containing the patient's anamnesis and the performances associated with it, it was necessary to identify patterns and define a proposal of PAI, consisting of a set of performances, specific for that patient who has particular characteristics.

The initial dataset was modified so as to obtain a matrix of size n x m, where n indicates the number of samples, in this case of patients (about 30,000), while m indicates the number of features, i.e. primary pathology, a list of secondary pathologies, sex, age group and list of services.

After having taken the first 4 fields from the dataset (primary pathology, secondary pathologies, sex and age group), for each of these different PAI proposals are identified, which can change in terms of quantity, type and periodicity with which to perform the services. Each proposed IAP is associated with a percentage, which gives more or less relevance to the care plan. This percentage is obtained from the number of times that IAP has been provided for that pathological situation, compared to the total number of IAPs formulated for that same condition.

In conclusion, the proposed algorithm, having received as input a set of information related to a patient, is able to provide proposals of IAPs, in order of relevance, specific for that pathological condition.