Expanding Horizons
Business, Management and
Technology for Better Society
20–22 May 2020
Online Conference

make learn
TIIM

Management,
Knowledge and Learning
International Conference 2020
Technology, Innovation
and Industrial Management

Artificial Intelligence and Big Data in The Healthcare Sector. The Revolution of Traditional Medicine and The Birth of Precision Therapy. Analysis and Evolution of *Big Players* in Healthcare and New Start-Ups, Growth Assumptions, and Investment Prospects. National and European Legislative and Regulatory Limits.

Gino Fontana
Pegaso International Malta, Italy
sl.gf@libero.it

Abstract

The advent of artificial intelligence will change society as we know it, producing significant implications in different fields of human interaction, making some professions and entire traditional production processes obsolete. Given the trends of the new technology based on the computational calculation in a predictive key, a substantial reflection of the social model conceived to date is absolutely pivotal. In fact, traditional regulatory concepts i.e. privacy, the protection of personal data, intellectual property, laws and professional orders, appear incompatible with artificial intelligence and the big data chain in general and therefore it is considered necessary to modernize the existing technical-legal and jurisdictional framework in a way that is able to keep up with the times. The research analysed one of the possible uses of the A.I.: the healthcare field, oriented more and more towards the personalisation of the "precision medicine" cure and the prediction of the disease.

Keywords: Big Data, Healthcare, Privacy, collective intelligence, medicine law framework, intellectual property

INTRODUCTION

In healthcare, it is a common view to make a diagnosis of a disease resulting from the appearance of signs and symptoms, with the application of a standardized and therefore non-personalized therapy; this explains the occurrence of failure in some cases. Considering the current limits of the human mind of data processing and storage of information deriving from the experience as well as from research, we can only endorse the thought that brings *big data* to the centre of the debate in healthcare as in other fields of knowledge. In fact, the processing of unstructured macro-aggregated data such as genetics, genomics, clinical history, instrumental and laboratory data thanks to the big data computer technology leads us to say that the new frontiers of medicine are predictive, in other words the disease

can be predicted before its manifestation, such as cardiovascular diseases, stroke, heart attack and some metabolic diseases as diabetes. The application of the new technology would represent a leap forward so that many pathologies cured today at the manifestation of the symptomatology could be foreseen well in advance. (The application of big data to the local medicine will lead in the near future to the lengthening of the patient's life expectancy). There are therefore many opportunities to improve life expectancy with the possibility of managing health in a new key, revolutionising the work of doctors with a consequent improvement in the citizen's life quality on the one hand and in the professional performance on the other. Obviously, the risks connected to the reliability of the data (good data), entails now and in the immediate future a mandatory validation in terms of reliability. The ability to manage the boundless mass of data will change healthcare as we conceive it today and will revolutionise the way we do medicine. The challenge of healthcare 4.0 is the improvement of the patient citizen's standard and life quality but also the creation of business opportunities for the players.

In literature

Making an initial overview we can start from the assumption that there is no universal definition of big data but many possible ones. The first was coined by NASA about twenty years ago when the data volume of the age was not in the slightest close to the current ones. *Forbes* magazine reports there are at least twelve big data definitions, all more or less exhaustive. *Gartner* summarises much of the quality of big data given in *volume, variety,* and *speed*: "Volume" given the big dimension of data, in fact we are estimated to have four times more data than today in 2020; "Variety" considering the various non-standardized sources of data, this represents inter alia the main pitfall as their automatic procession is really challenging; in the end by "Speed" it is meant that the data are produced continuously, such as it happens with mobile phones, that even if switched off they don't stop working, e.g. in the case of the geolocation function.

Research objectives. In the near future

The research with a holistic approach aims to analyse the impact of artificial intelligence and big data on the health system and the professions of the health sector in general; also reporting the limits and the technical and legal implications connected to the use of data to privacy and the vulnerability of data with respect to their security. The reading key undertaken will bring to focus the inadequacies of the protections existing today. Both with respect to the world of the companies involved, such as patentability, and in terms of completeness of the study, market developments and investment opportunities will also be reported.

Through an algorithm that crosses an infinite amount of information, the computation calculations manage to operate in the medical field, deepening the diagnosis for patients and drastically reducing the probability of error.

The question under analysis relating to the use of big data and artificial intelligence in the health sector must answer the following question:

- How are they used?
- Predict diseases.
- Prevent diseases.

- Improve care efficiency.

- Public health.

Predicting diseases will be much simpler, - if we put the data together we will create predictive models and systems in health matters -; *preventing diseases* will be the second step "such as changing the patient's lifestyle, nutrition, etc. .."; *efficiency*, will mean improving the services and interventions aimed at the patient (e.g. waiting lists predicting risk); *public health* - through collective impact actions such as in the case of epidemiology, infectious diseases etc.

The two fundamental components of a near future in medicine and healthcare that we still do not see but that is already part of our reality "*By now, it's almost old news: big data will transform medicine*" are:

- *Patient generated data* (data with clinical relevance) data with information with clinical relevance.

- *Health record*, data already contained in hospitals and health companies.

In the near future we would have many other data available to clinicians that are definitely wider and more complete than today; in fact, in the health record, not yet present information will flow, such as genetic data.

The foundations of the new medicine are certainly based on the personalization of therapy through predictive approaches, the right care for the right patient at the right time. Through the use of Machine Learning and Artificial Intelligence, (for which we intend the teaching of intelligent machine behaviours through two functions: learning and problem solving), we will come to the development of therapeutic plans. Inside artificial intelligence there is *machine learning* that is when the machine is taught to do something without having it programmed to do so - they are therefore made through the analysis of statistical model data to solve and identify the clinical future of the patient. The artificial agent who receives the information outputs them for modelling.

The United States have been the forerunners in the introduction of the electronic patient record, as a set of different information and more structured data, and this already in 1972 with slowdowns caused by the legal implications related to the concept of privacy and protection of sensitive data; however, the greatest development took place in the 2000s with the implementation of ICT and the Internet and the lowering of costs related to the use of the new technology. Italy arrives much later than other European countries, and only today it is seriously approaching this phenomenon. In the USA, we started with an approach mixed with subsequent corrective measures, relating to the cleaning of the data as the same clinical concept can be defined with different words depending on the operator who intervenes and who enters the data into the system. By “unstructured” we mean text data.

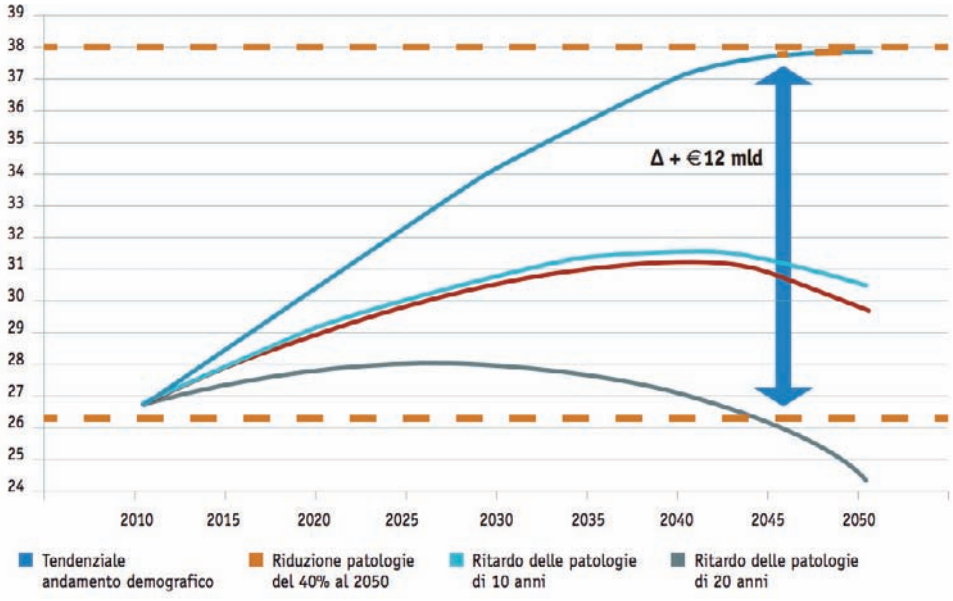
The electronic medical record (CCE) was introduced for the first time in Italy in 2012, with the amendment on simplifications, and represents the digital equivalent of the traditional medical record. It collects all the patient's clinical data (e.g. personal data, specialist visits, reports and test results). It also allows the management and instant sharing of all information, with the help of the IT system used by hospital companies. The information is managed in total respect of the patient's privacy and only

healthcare professionals, from whom the patient is being treated, can access it. The electronic health record (ESF), unlike the CCE, traces the entire clinical history of the patient. Through an electronic card, the citizen has the access credentials to share information with the attending physician and update their clinical data in total safety, while health professionals can consult it through an authentication system. Patient data can also be used for scientific and statistical research, provided that anonymity is guaranteed. In Italy, the data provided by the Digital Italy Agency maintain that the ESF is active only in six Italian regions (Emilia-Romagna, Lombardy, Tuscany, Sardinia, Valle d'Aosta, and the Autonomous Province of Trento), while in eleven others it is still developing. The regions that have not activated any procedure on this issue yet are Campania, Calabria and Sicily and the autonomous province of Bolzano. This data explains why the ESF, in 2017, was used only by 5% of citizens in these areas.

About the American experience that has now reached phases of momentary standardization and also of obligation with respect to the Anglo-Saxon regulatory system, we can affirm that the electronic record must contain the following elements:

- *ELECTRONIC HEALTH RECORDS*
- *MOST COMMON EHR DATA TYPICS*
  - Patient demographics
  - Clinical notes
  - Vital signs
  - Medical histories
  - Diagnose
  - Clinical images
  - Laboratory and test
  - Result

Picture 1: Increase of Delta over time comparing the demographic trend and the decrease in pathologies



Source: VII RBM Report -Italian Centre for Social Investment Studies on Public, Private and Intermediate Healthcare (2017)

In a context in which the demographic increase and the delay of the onset of chronic diseases increase over time, the investment in a correct and timely prevention activity in 2050 could generate savings of 12 billion euros, equal to 25% of the public spending on healthcare. Acting in this direction, in fact, many hospital admissions could be avoided, intervening in a minimally invasive way during the initial stages of a pathology. However, in order to put this innovative approach into practice, it is necessary to carry out a radical change at the basis of the incentive systems of those who provide the medical care service, modifying the payment system that currently is based on sales volumes, to move to a system that rewards the value and quality of the interventions carried out.

New investments, development opportunities in healthcare. Big players and innovative start-ups. New international trends. Protection of intellectual property and the impasse on patentability.

Not all players on the market today have the assets to be able to analyse big data and this is why the healthcare market is opening up to new big players such as IBM, Google, Sap, Siemens, Philips, Microsoft, 3M, all companies that we would not typically associate with the healthcare sector but which in recent times are making new and strong investments in such sector. The phenomenon occurs because competitors have technical data analysis skills. Among the others particularly remarkable is IBM. It has created a system called *Watson for oncology* that has the aim of correlating and interrelating the clinical data of patients to study new and personalized treatments for cancer on the one hand, while on the other it aims to index research and scientific articles in order to optimize the doctors' time "also by drawing on self-diagnosis formulas on the Internet for future patients" - that is, people who research on their symptoms. Noteworthy is also SAP, a giant known for the management systems of healthcare companies through software and analytics systems. Taking advantage of its platform it has suggested new clinical tests for specific diseases. Furthermore, Microsoft has launched a new division worldwide *Elche*, the system that aims to use data in healthcare through the use of artificial intelligence and machine learning, a real ecosystem that goes from the consumer to the optimization "for example of shifts in the hospital or patient flows".

Healthcare start-ups are particularly lively and active, in fact 20% of the world's most funded start-ups operate in the healthcare sector and compared to big players they are much more focused on analysing a specific problem. Healthcare start-ups have a greater ability to attract funding with very heterogeneous teams not necessarily from the same area e.g. doctor, technology expert, fundraising expert, administration expert, and so on. At an international level, health start-ups are growing very rapidly. This does not apply to Italy, where unfortunately space has been left for the public and not for private companies. Internationally, some successful start-ups are reported, such as *American Flatiron* focused on cancer data which collects data from all cancer centres and puts them together in order to allow clinicians to create increasingly personalized therapies. This company allows the clinician to also analyse their own data, aimed at their sample; another success case is ADA that works using artificial intelligence able to communicate with the user in natural language with possible suggestion cases "e.g. I feel a fever - how do you feel? - take an aspirin". Infervitions is a Chinese start-up focused on the recognition of images whose mission is to exclude surely healthy images to alert about those that present pathologies. Image recognition is an emerging trend with enormous uses (e.g. I am taken a picture and from the colour of my skin they can understand if there is a possible disease) to support the doctor in the diagnosis. In the future, doctors will no longer have to deal with traditional companies that they are used to knowing. By contrast, they will increasingly have to deal with technological giants. Technology is an enabler from which information must be correctly drawn which

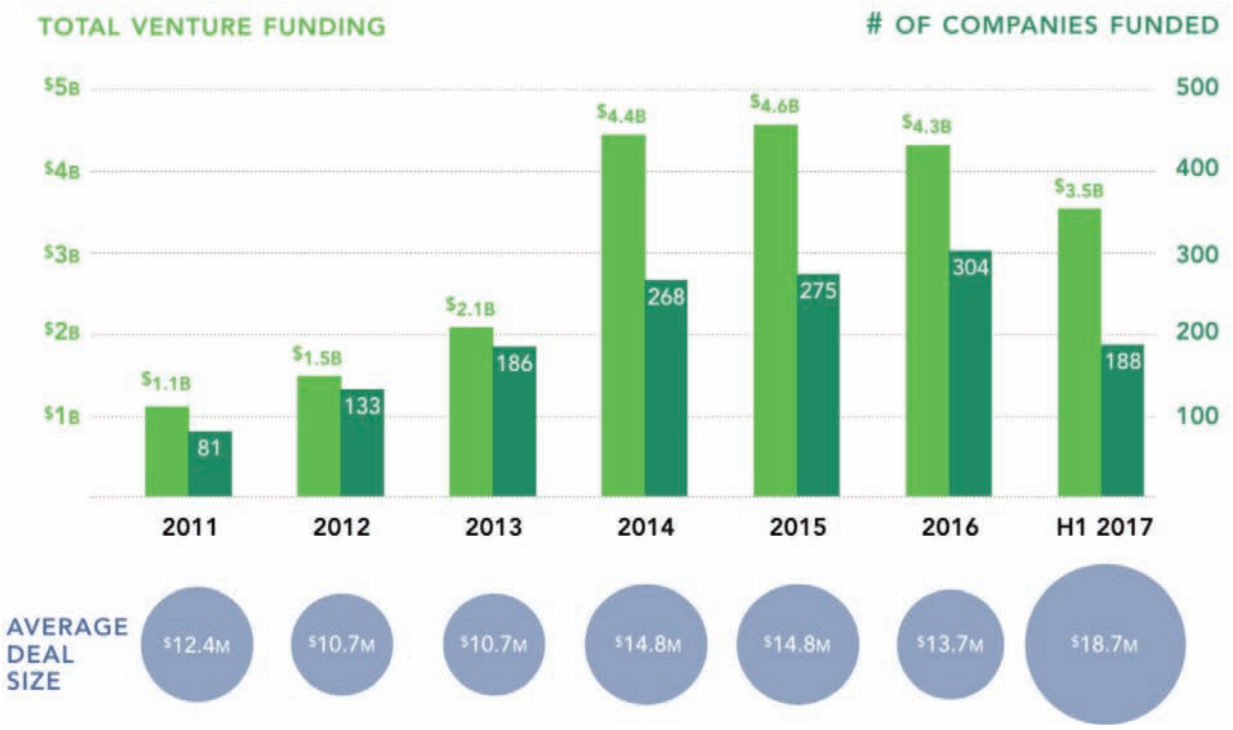
will lead, through appropriate analysis, more and more to a personalization of care. Basically, the new companies, even if small, must interact with the patient in order to focus the needs by proposing technical solutions, also avoiding expensive exams. The stereotype of having a cure for a disease that is good for everyone will be dispelled through the use of this technology.

There are therefore new and great opportunities for the new players determined by innovation and targeting in terms of prevention, in fact the quality of the health system affects only 10% of the patient's life while the further 90% is determined by the individual lifestyle, genetics and environmental factors. New business formulas such as telemedicine, and / or linked to *the Internet of Things* (IoT) could also be successful in business intelligence. It goes without saying that many innovations carried out by start-ups such as the use of Smartwatches and other wearable devices, virtual assistance, implantable electronic sensors, digital wearables meet the legislative limits of a national and European regulation not yet ready to protect the interests involved related to patient data.

The giants market, as it is obvious, is full of pitfalls linked to the initial stages concerning the launch of new "smart business" businesses, for new start-ups; in fact, the big players provide the huge initial capital to be used to start a new business, this determines that the most innovative start-ups that were previously financed by the technological giants are often acquired by these big players, in a complex market logic tense to the maximum employment of the market segments - for example google finances many innovative start-ups in order to acquire them if they go well and a new brand launched on the market is successful, therefore in the positive case the result is internalized and these companies are also acquired with the capital increase technique and / or with deadbolt contracts or methods for controlling the majority of the boards. We reserve a broader reasoning on patentability in general and artificial intelligence products in particular elsewhere.

In any case, the healthcare sector is a business in strong growth and the technological creativity applied to the context will increasingly represent an excellent investment sector in the various declinations for the next few years, such as insurance, for example, whose insurance premium is linked the customer's health conditions that are monitored through a smartwatch ". In support of this reasoning, the index of seniority of the population in Italy and in Europe must also be taken into consideration; according to ISTAT, in fact, the number of people who will pass the 65-year-old threshold will be 39.5% of the total population by 2050, with an average estimated life of around 82.5 years (80.6 for men and 85.1 for women). To this another relevant phenomenon must be added: in recent years, in fact, chronic degenerative diseases are increasingly common, such as heart failure, obstructive bronchial respiratory disorders; therefore, the new markets will be oriented towards the production of smart devices such as: heart rate monitors, smart bands, reminds, alerts, wireless systems that communicate via Bluetooth and GPS with other devices etc.

Picture 2: Total venture funding from 2011 till 2017



Source: Rock Health Funding Database (2017)

The European Union framework program for research and innovation Horizon 2020 contains the section entitled "*Societal challenges*" and the subsection entitled "*Health, demographic change and well-being*" in which all private initiatives find a possible space, which aim at the improvement of health and wellness throughout the life of the citizen (*Long Term care*) with an allocated budget of 323.5 million euros. We come to an interesting point regarding the world of business relating to the patentability of artificial intelligence systems, that is, the adequacy of current intellectual property regulations also in relation to AI. This question is the subject of debate both at national and European level. AI technology is mainly made up of software and / or source code which is protected by legislation, however this protection expands only to the elements that are the result of the author's creativity; in other words, only the source code that was created by the author is protected by law. Instead the machine language based on algorithms and computational predictive capabilities is unprotected. In Italy art. 45 of Legislative Decree 30/2005 establishes that software as such and mathematical methods cannot be considered as inventions and therefore cannot be patented. Therefore, an algorithm being mathematical in nature cannot be patented; however, a method that uses a mathematical method to solve a technical problem is patentable. By issuing guidelines, the European Patent Office admits patentability in the presence of certain conditions: a) the invention must have a technical character; b) it must be the product of an inventive activity; c) it must be clear and sufficiently described. The largest number of patent applications is presented by companies, such as IBM (8,290 inventions) and Microsoft (5,930) in the lead, followed by Toshiba (5,223), Samsung (5,102) and NEC (4,406).

Privacy and data vulnerability in healthcare. Legislation and Regulations

The information concerning a person's health status, the pathologies, and treatments to which they have undergone, constitute the personal history of the patient to whom the maximum confidentiality of the information must be guaranteed. The misappropriation or disclosure of this type of data can constitute, for instance, a discriminating factor in the professional area, as well as a serious moral damage for the person involved. However, personal information also attracts the interest of insurance companies, or large pharmaceutical companies, for their immense value in defining a market strategy. The health sector is the one most exposed to the risk of violations of sensitive data, with a percentage that exceeds 70% of cases. The study published within the 2017 *Data Breach Investigation Report* of Verizon, a leading telecommunications company, reported that 458 accidents would have occurred in the Healthcare sector, including 296 with data theft.

The violations were committed in 80% of cases due to errors or data breach and would have been committed:

- 68% by subjects inside the healthcare facilities,
- 32% by external subjects,
- 6% by close family members.

The motive behind the theft of sensitive data is driven by economic interests for 64%, while 23% for personal revenge. The data were reported:

- 69% in the healthcare sector,
- 33% in the personal sphere,
- 4% related to payments.

In the healthcare sector, in fact, confidentiality does not only concern information that can be extracted from patients' medical records but focuses more on safeguarding the patient's health state. To protect the integrity of the data, a recent regulation was issued within the Law n.24 of March 2017 containing "*Provisions on the safety care and of the assisted person, as well as on the professional liability of health professionals*". The new European regulation on the protection of personal data, which entered into force on 24th May 2016 and definitively applicable starting from 25th May 2018, has therefore introduced numerous and substantial innovations regarding the protection of personal data, surpassing Directive 95/46/ EC. The role of the subject's consent changes; in fact, art. 6 of the GDPR lists six different cases in which the data processing "*principle of accountability"* is legitimate, placing possible alternatives to consent. The Regulation introduces for the first time a definition of "health-related data", according to which these data consist of *"personal data relating to the physical or mental health of a natural person, including the provision of health care services, which reveal information related to their health*". Article. 4 no. 11 of the GDPR defines the consent of the interested party as "*any manifestation of free, specific, informed and unequivocal will of the interested party*". The new configuration of the Regulation is based on the combined provisions of art. 6 and art. 9 dedicated to the "*Treatment of particular categories of personal data*". The letter h) refers to the hypothesis in which "*the treatment is necessary for the purpose of preventive medicine or*

occupational medicine, assessment of the employee's working capacity, diagnosis, assistance or health or social therapy or management of health and social systems and services under Union's or Member States' law or in accordance with the contract with a healthcare professional, except for the conditions and guarantees referred to in paragraph 3". Paragraph 3 specifies that personal data must be processed by or under the responsibility of a professional. The letter i) instead refers to the hypothesis in which "*the treatment is necessary for reasons of public interest in the public health sector, such as the protection from serious threats to cross-border healthcare or the guarantee of high quality and safety parameters of the healthcare and medicines and medical devices, based on Union's or Member States' law which provides for appropriate and specific measures to protect the rights and freedoms of the interested subject, in particular professional secrecy*". In essence, the health professional is not required to acquire the patient's consent when the treatment is necessary; however, it is still necessary for all data treatments that are connected to the health service in the other sense, therefore, by virtue of this principle, the patient can withdraw consent to consult the data contained in the electronic health record at any time. Namely, the electronic health record is a tool for the collection, in electronic form, of data relating to all clinical events concerning a particular patient. The difference compared to the DSE consists in the fact that while the latter collects the data generated on the occasion of clinical events that took place only within the healthcare facility where the dossier was created, the ESF is managed at national or regional level and, therefore, it can contain data relating to clinical events that have taken place in any healthcare facility in the area (national or regional) and can be powered by each doctor operating in each of these facilities.

CONCLUSIONS

Artificial intelligence in a few years will transform society as we know it and make a myriad of professions and working activities obsolete. We are still unable to perceive all the application reverberations of the new technology, in fact health is one of a thousand sectors, but, surely we can say that we will not be in the presence of a simple technological innovation generating adaptive formations; simply those professional figures, such as the lawyer, radiologist, journalist, etc. will no longer exist. The task of research together with the world of business and institutions is the profiling of new professional figures on the one hand and the protection of weak rights on the other, without neglecting a modern and avant-garde regulatory framework.

REFERENCES

Harvard Business Review (2017): "*Healthcare needs real competition*", Leemore S. Dafny e Thomas H. Lee
Rapporto ISTAT (2017): Indicatori demografici
Rapporto RBM - Censis (2017): Sanità Pubblica, Privata ed Intermediata
Rapporto WHO (2017): stime per l'anno 2017
Regolamento europeo 2016/n.679
Report (2017): *Brand Finance Glabal 500*
Report Commissione Europea (2010): "La strategia Europa 2020"
Report dell'Osservatorio Innovazione Digitale in Sanità (2016) "*Sanità digitale: non più miraggio, non ancora realtà*", School of Management del Politecnico di Milano
Report ISTAT (2017): serie storiche del PIL e della spesa sanitaria
Report OECD (2017): Health Statistics
Report Ponemon (2016)*:* Cost of Data Breach Study
Report Research2guidance (2017): mHealth App Developer Economics
Report Verizon (2017): Data Breach Investigation

Report World Economic Forum (2016)*:* Digital Transformation Healthcare.
Articolo online: Digital Health Italia (2016) "*Monitoraggio biometrico di tipo medicale"*
Articolo online: Digital Health Italia (2016) "*L'intelligenza Artificiale al servizio della medicina nella diagnosi di disturbi mentali*", di Sara Scarpinati (*www.digitalhealth italia.com*)
Articolo online: Digital Health Italia (2016). "*Realtà Virtuale: da fenomeno geek a nuove possibilità di impiego in campo medico*", di Sara Scarpinati (*www.digitalhealthitalia.com*).
Articolo: *ilsole24ore (2014) "Professione: scienziato del dato*."
Portale online: Advocate Health Care Center Chicago, Illinois
Portale online: Agenda Digitale 2017 "*Big Data per la Salute*", di Eugenio Santoro
Portale online: Agenzia per l'Italia Digitale (AgID), Presidenza del Consiglio dei Ministri
Portale online: American Health Information Management Association (AHIMA)
Portale online: *Amiko.io*Health Outcomes