EXPERT SYSTEM
SEMANTIC INTELLIGENCE

EMPOWERING PERSONALIZED MEDICINE WITH SEMANTIC TECHNOLOGY

INTRODUCTION

While long available in healthcare, big data tools and technologies are still largely unable to fully implement effective systems for decision support and for what is currently its most revolutionary challenge: Personalized Medicine.

Where traditional evidence-based medicine relies largely on general, de-personalized statistics, today's movement toward personalized medicine relies heavily on large amounts of unique, individual data in the form of unstructured information, otherwise known as big data.

The opportunity presented by big data and personalized medicine—improved quality of diagnosis, treatment and care—can only be realized with a cognitive technology that, by reading and understanding text, can unlock the full value of unstructured information.

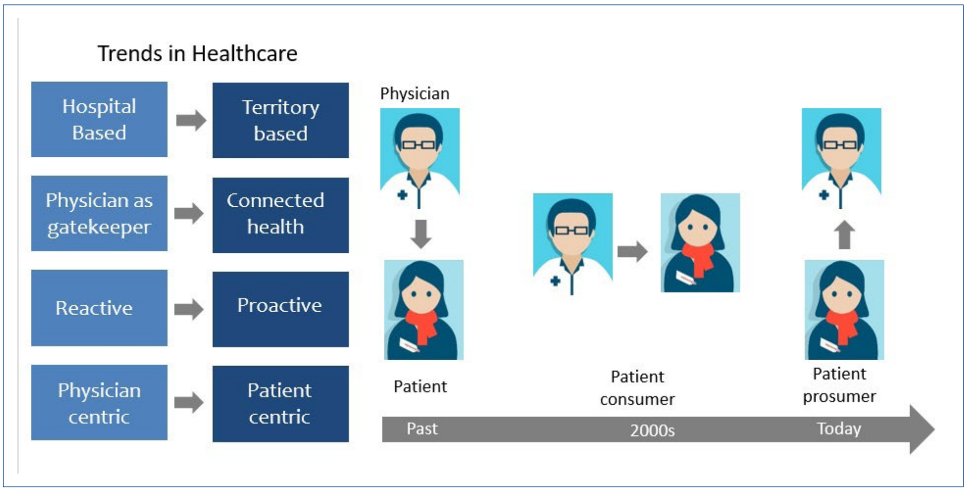
In this document, we will introduce a scalable and expandable data-driven infrastructure powered by cognitive technology that provides access to knowledge for the individual patient, for new diagnostic tests, treatment protocols and solutions that ultimately reduce healthcare costs for the long term.

Today's Healthcare System is Fueled by Data

The digitization of clinical health content is driving a new frontier in healthcare worldwide. New sources of information—data produced by traditional devices and sensors, streaming data produced by new wearable devices—are producing interesting data from the point of view of both medicine and analytics. The speed, volume and variety of constantly growing health data is driving creation of a diverse knowledge base that supports and empowers the use of big data analysis techniques.

What is Personalized Medicine?

Rather than basing diagnosis and treatment solely on general information and guidelines from population-based studies and clinical trials, Personalized Medicine[1] (PM) employs an approach that is tailored to the individual. Also referred to as Precision Medicine, PM emphasizes customized medical treatment for patients based on their genetic makeup and other personal characteristics (e.g., clinical co-morbidities), essentially matching the right treatment and dosage to the right patient at the right time. Its overarching goal is to create a framework that leverages patient electronic health records (EHR), OMICS (primarily genomics) data and crowdsourced information to facilitate clinical decision support, making it predictive, personalized, preventive and participatory (P4 Medicine).



Trends in healthcare

To understand the volume of data we're talking about and the speed at which it's becoming available, consider these statistics. According to the Institute for Health Technology Transformation, medical electronic archives are growing 20 to 40% each year in the United States and around the world. By 2013, U.S. heathcare organizations were said to have generated 150 exabytes of health care data. According to the results of a study by the Office of the National Coordinator for Health Information Technology, deployment of EHR systems increased eight-fold between 2008 to 2014.

New sources of information

Electronic health record data is just a fraction of the information available to healthcare professionals. Approximately 80% of available healthcare data is contained in unstructured formats such as text and images. It is here, in content that is unstructured or semi-structured, that the true opportunity for PM exists. Let's take a closer look at these emerging sources of information.

1 Excerpt from Panahiazar, Taslimitehrani, Jadhav, Pathak.- Empowering Personalized Medicine with Big Data and Semantic Web Technology: Promises, Challenges, and Use Cases 2014 Oct. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4333680/pdf/nihms656338.pdf

Patient narratives

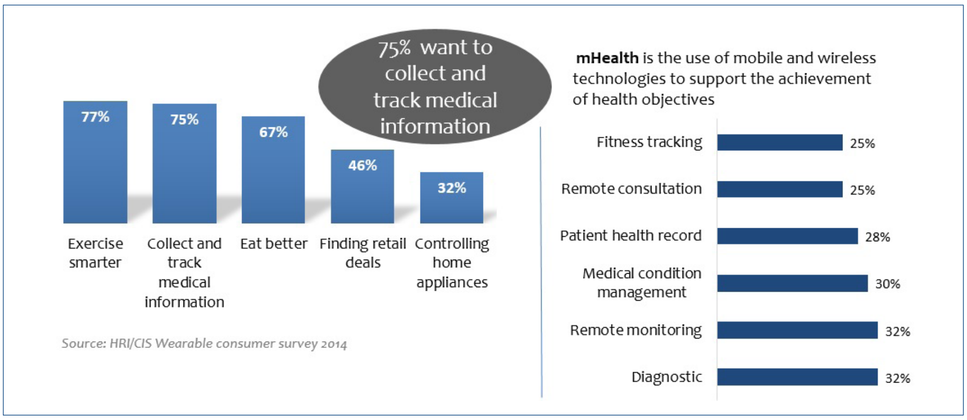
Thanks to the widespread deployment of EHR, medical records sharing between patient and doctor is becoming the new reality worldwide.[2] Digital medical records often include detailed patient information that goes beyond the traditional medical record. Patient diaries or transcripts of doctor-patient chat applications offer a deeper, more complex narrative about a patient's condition, such as their current medicines and treatment, family history and lifestyle and habits (according to the guidelines of Narrative Medicine).

Crowdsourced data

The web is an extensive source of narrative-based health related content, from forums and discussion boards on medical association websites (e.g. www.aism.it for multiple sclerosis, www.rareconnect.org for rare diseases), to websites like www.patientslikeme.com and social media (Facebook, Twitter or WhatsApp). Despite the positive aspects of allowing users to share information, these supervised sites may encourage self diagnosis over seeing a doctor, and may perpetuate scams and even harmful treatments not backed by medicine or science.

Wearable devices

Wearable devices are providing extremely large volumes of health- and activity-related data. While not strictly used for medical purposes, this data—tracking everything from sleep, to the number of steps taken, to one's heart rate and other lifestyle habits—could eventually be incorporated into a patient's existing medical profile data, only adding to the big data mix. In fact, 75% of consumers who use these devices cite the collection of medical data as one of the main reasons for purchasing them. The impact of data from wearables may be even greater if we consider that 68% of consumers are willing to anonymously submit health data in exchange for insurance premium discounts or freebies offered by providers. This presents a revenue-generating opportunity for the wearables market from the analysis and sharing of consumer data.

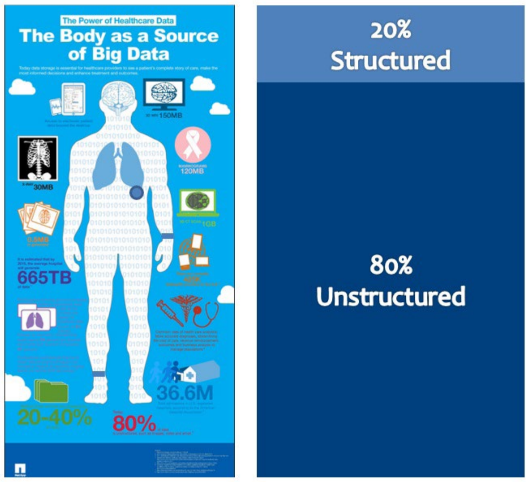


Most consumers want health information from their wearable devices the domains where they are used

2 For example: TreC – FBK in Trento, Italy: http://www.fbk.eu/spotlights/trec-la-cartella-clinica-del-cittadino

The Big Data Opportunity for Decision Support

Traditional data analysis systems have typically relied on just 20% of healthcare knowledge—the part that is already structured—for decision support. Despite the fact that it offers a limited view of the patient's health and condition, structured information (such as vital signs, lab results, allergies, immunizations, medications, insurance codes and other standard medical record data) is readily available for analysis processes. This is one of the factors that has prevented wider implementation of effective decision support systems or PM.



The body as a source of Big Data (Infographic from NetApp)

Analysis of the other 80% of information—unstructured and semi-structured sources—can help medical professionals with decision making support by accessing and analyzing data in a variety of types and from multiple sources (for example, various patient sources but also large databases from multiple institutions). This is where cognitive technology can help.

Introducing Cognitive Technology for PM

The wealth of diverse, detailed information available today requires technologies that are able to extract and exploit knowledge present in unstructured content, to understand natural language (NLP), extract entities and manage ontologies. It requires cognitive technology.

For unstructured content that is made primarily of text—words, phrases and sentences—approaches that treat text as a sequence of characters, count word frequency or require that you train a matching system are ineffective not only because they are time consuming and costly, but more importantly because they overlook the granular detail of content. Why? Because they don't understand language. We do.

Our cognitive software, Cogito, leverages both Artificial Intelligence based on semantics and deep learning to understand text in a way that mimics human comprehension of information.

Semantic technology at the core of Cogito is differentiated by our semantic engine, which performs deep analysis of text and our concept-rich knowledge graph that resolves any ambiguity in text to understand the precise, accurate meaning of any word in its proper context.

By understanding the meaning of words and all their nuanced expressions in your unstructured content and big data, your systems and processes that depend on information can leverage cognitive intelligence to capture all of the information present in content and make the knowledge found there actionable.

Cogito contains health domain-specific knowledge and uses deep learning and machine learning algorithms to acquire new knowledge from any source. This allows it to make strategic connections across millions of documents that would be invisible to traditional systems.

Technology that speaks your language

Codes and numbers alone cannot describe the most individual differences that only an improved ability to listen can reveal, both in the course of doctor-patient dialogue and in the deep analysis of new types of narrative-based content. This is precisely where cognitive technology can make a difference.

The language of health

The English language distinguishes between the three fundamental aspects of the concept of illness: Disease is the pathology of a condition in a medical context and how it develops in the body; Illness describes one's personal experience of a condition and how they feel about it; Sickness refers to how society and culture regards a condition, whether it is accepted or not and any societal / behavioral norms around it. Personalized Medicine looks not only at the disease itself, but also at its more individual and unique aspects.

Bridging the gap between EBM and PM

New technologies and the availability of new health-related information place the traditional Evidence Based Medicine (EBM) approach (based on statistics and meta analysis of the average patient), in difficulty. While EBM has yielded great results in the past, it struggles to respond to the individual and personalized needs of today's e-patients. This is especially true for the elderly, who often suffer from multiple morbidities and complex illnesses that are not easily identified or attributable to a universal code.

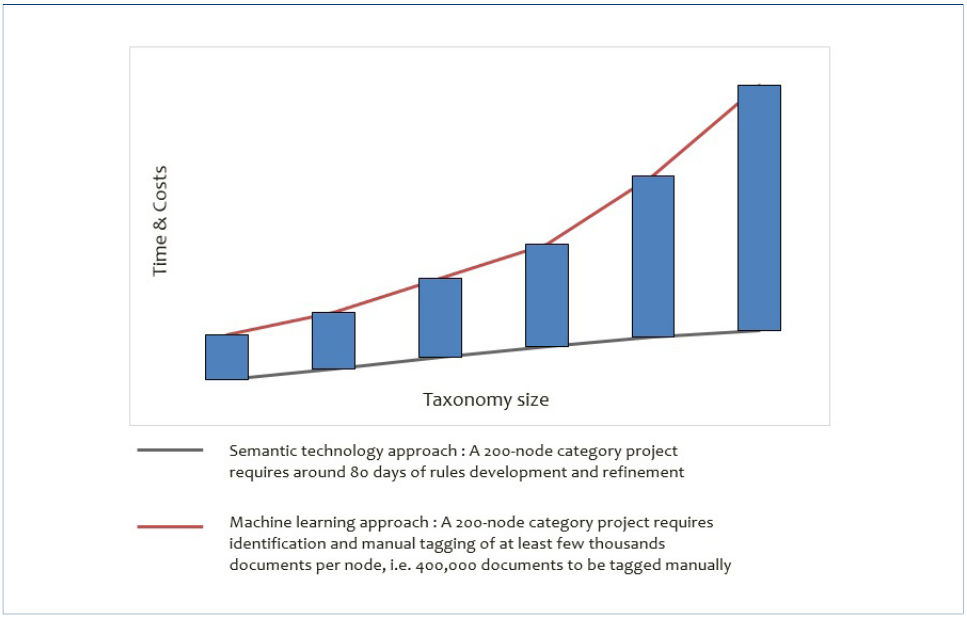
There are various strategies that may supplement or even substitute EBM, such as Network Physiology[3] (NP), a new approach for medicine and diagnostics based on the connections between events and how our different organ systems work together and change over time. For both NP and PM, cognitive technology is fundamental for extracting the valuable relationships between different elements in unstructured data.

Knowledge discovery in data

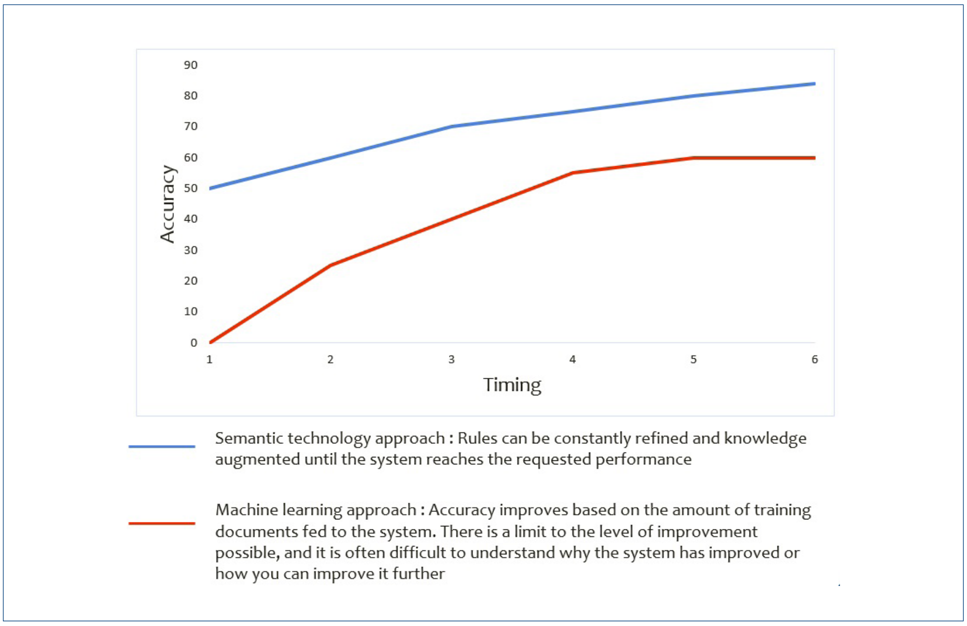
Places, people, drugs, prognoses and organs are examples of entities present in text that can be extracted to enrich structured data in analysis. The ability of deep linguistic analysis to identify mood in content, such as whether fear or anger prevail in a text, can also be valuable for patient care and in the evaluation of different treatment or therapies.

3 Network Physiology is a rapidly emerging discipline at the intersection of statistical physics, biomedical engineering, applied physiology and clinical medicine. This field refers to the development of theoretical frameworks and system-wide analysis approaches to understand how several different organ systems, each with its own regulatory mechanisms, dynamically interact with one other to preserve the physiological function. NP advances a new vision about human neural regulation, and ultimately provides a novel paradigm of system medicine, which exploits the big data generated by the human organism to define complex phenotypes, orient clinical decisions and provide personalized patient care.

Although automation can help make the coding process more efficient, the real challenge is not finding what is missing, but discovering the connections and relationships in complex information. Looking for information in unstructured data that can enrich the knowledge of a patient's unique features, drawing from their history (if possible) and from their personal narratives[4] is of great value. To perform this with consistency and accuracy requires technology that is not subject to human weaknesses of limited time, distractions, error and emotion. In this capacity, text analysis technologies based on a cognitive approach can work with great precision while reducing the costs of managing such processes.



Increased efficiency as a result of semantic technology



Greater accuracy as a result of semantic technology

4 Narrative Medicine (also known as Narrative Based Medicine) attempts to include the biographical (and not just the biotechnical), in accordance with Personalized Medicine.

By reading as a person does, Cogito can identify the most relevant concepts in a document; if these concepts are related to diseases, chemistry or pharma, it can immediately assign the document to these domains. Customization or training is limited to augmenting the knowledge base in the event that new or previously unknown terms become part of the standard lexicon.

Technologies that employ a machine learning approach do not have embedded knowledge but require a set of documents to train the system (the larger the set of documents, the better). Generally, this training set must be manually pre-tagged by experts. By keeping track of the keywords and sequences that it finds in the course of processing a large number of documents, at some point the system will be able to recognize when a document is talking about certain topics by matching known patterns. However, any additions or changes to the knowledge will require that training start over from the beginning.

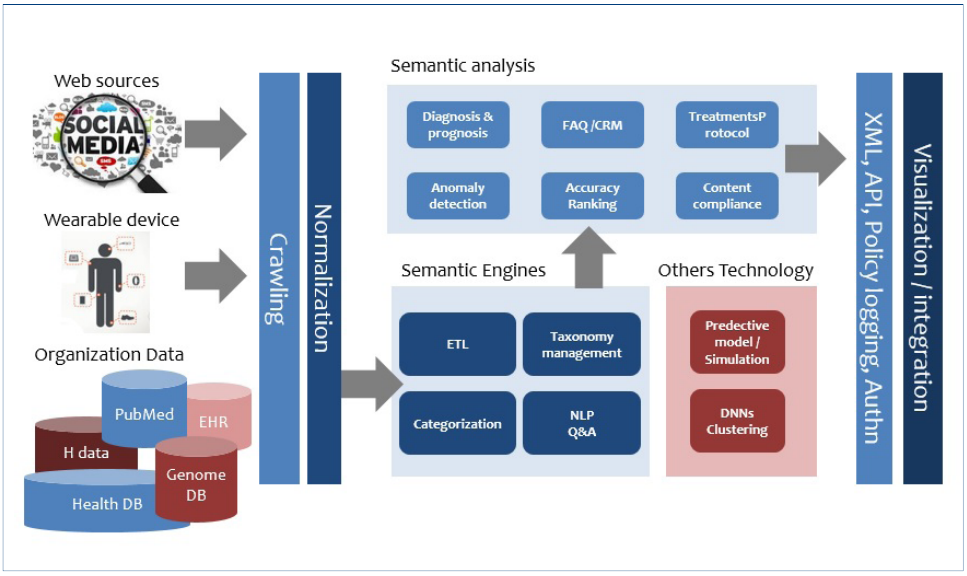
Key Takeaways: The Value of Cognitive Technology for PM

The explosion of health and medical information that we're currently experiencing will only continue to grow. With worldwide healthcare data expected to grow eight-fold to 25,000 petabytes by 2020, according to IDC Healthcare Insights, companies need to be honing their ability to properly capture, analyze and make use of this big data today.

To make PM a reality requires transforming all of the available information into intelligence that can be used for actionable decision support across the organization.

When meaningfully used, these new sources of information can provide significant insight and avoid unnecessary treatments, minimize adverse drug events, maximize overall safety, and eventually lead to a more effective and efficient healthcare system. It can also support the objectives of personalized medicine.

The following diagram shows the possible architecture of a Computational Health Platform oriented to Personalized Medicine.



Architecture of a personalized medicine platform

These are just a few examples of how cognitive technology can benefit Personalized Medicine:

- Ease the transition from ICD9[5] to ICD10 coding through a deep understanding of content, not with statistical algorithms that try to match codes.
- Improve doctor-patient communication and understanding on digital channels such as self-help and Q&A mechanisms and decision support processes using natural language processing.
- Extract important information about doctors, drugs, products, hospitals, lab values, etc. from different types of text (emails, papers, medical reports, news, etc.) to enhance structured information that the medical community typically rely on.
- Support predictive analytics and "what if" scenarios to identify innovations and trends in even the most complex healthcare information flows.
- Ensure the validity of crowdsourced, user-generated information to guide users to trustworthy and credible information and resources.
- Verify the contents of official documents for consistency or compliance with legislation, standards, etc. (for example, does the discharge letter contains everything that is needed?).
- Identify anomalies or other interesting content by crossing machine- or device-generated content with "less relevant" information such as log files of a complicated process or routine checks with negative results.

5 https://en.wikipedia.org/wiki/List_of_ICD-9_codes

About Expert System

Expert System is a leading provider of cognitive computing and text analytics software based on the proprietary, patented, multilingual semantic technology of Cogito. Using Expert System's products, enterprise companies and government agencies can go beyond traditional keyword approaches for the rapid sense-making of their structured and unstructured data. Expert System technology serves some of the world's largest industries including Banking and Insurance, Life Sciences and Pharmaceuticals, Oil and Gas, Media and Publishing, and Government.

www.expertsystem.com
info@expertsystem.com

EXPERT SYSTEM
SEMANTIC INTELLIGENCE