



Towards an Ethical Datamining of the Medical Data Sphere

Jérôme Béranger

Researcher (Ph.D) Ethics in e-health, 1 impasse Augustin Fresnel - ZA du Moulin New - BP 60227 - 44815 - SAINT-
HERBLAIN CEDEX, France

ABSTRACT: With current development of Big Data technologies, the algorithms that exploit data flow are now playing an increasingly decisive role in personal choices. Though it would be exaggerated to say that these algorithms control us, they do orientate a certain number of our decisions. Therefore, a reflection on the study of risks associated with ethical issues around personal health data is imposed on us. Our study is based on many field surveys, interviews targeted at different actors, as well as a literature search on the subject. This work led to the realization of an innovative method of risk analysis and determination of the requirements of the ethical goals of personal health data. The aim is to make sense of realization, implementation and use of personal data in order to better control it. We propose to align concepts of ontology of the risks to those of ontology of ethical objectives requirements on the basis of best practices in the field of data security. In addition to the guide mechanism we show in this paper, the approaches used for the construction and alignment of two ontologies. Our research wish to make the first elements of a moral framework contributing to significant changes in attitudes and practices of people involved in this data. Under these conditions, our ethical analysis is the foundation of a new ethical-technical approach, respectful citizen, oriented in the direction of thought, conscience and human responsibility.

KEYWORDS: New Information and Communication Technologies (NICT), Personal data, Health, Risks, Ethics, Ontology

I. INTRODUCTION

The way we live today is almost inconceivable without the data processing systems implemented in the computers, smartphones, pads or GPS that we use on a daily basis and to which we now add on other connected objects. We bathe in a controlled digital reality where a multitude of information flows converge. To the constant development of e-health, telemedicine and medical Big Data, legal changes that are accelerating, economic disruptions that amplify a feeling blur occurs around personal health data. Handling and use of medical information are then constantly balance between confidentiality and transparency. Now, these new information and communication technologies (NICT) substantially disrupt the medical microcosm up to challenge the paradigms of Medicine Hippocrates as we previously know.

Processing the data has become a sensitive issue inasmuch as they relate to our private sphere, to our intimacy. The personal control function is only partly effective since almost nobody knows how to implement it seriously. This is why experts have been discussing the opportunity of Big Data governance. How can we proceed? So how can we get back in control? One of the most interesting possibilities will consist of putting together and implementing an appropriate model to analyze, understand and process the large amounts of data involved. It is perfectly possible to imagine ethical modeling of complex data, for reasons that are integral to the way datamining operates. The inductive algorithms at the heart of Big Data processes are driven by logic amazingly close to the practical wisdom which is at the core of ethics.

The increasing digitization of medical data, the ability to store ever-increasing digital data, the accumulation of information of all kinds that follows, then help bring some fears and uncertainties due to its multiple uses (or complex) it's hard to measure impacts affecting very different populations, and its information sometimes out of control. The best way to address new ethical challenges posed by Big Data is probably to do it from an approach that treats as authentic and genuine all forms of existence and behavior, even those based on synthetic and artifacts from engineering. This "synthetic e-environmentalism" approach requires a change in our vision the relationship between physics (nature, reality) and technique



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

(the practice of science and its applications). Therefore, it becomes necessary to know and understand the use of NICT in the health sector in order to protect the privacy of every citizen.

In these circumstances, it becomes important to establish a reflection on personal health data via a Neo-Platonic systemic ethical prism (Ψ , G , Φ) to raise doubt and controlling uncertainties and devices uses of medical data in this new paperless ecosystem. For this, we have implemented a guiding approach to move, a study of the risks to the expression of ethical objectives around personal health data through the alignment of these two bridges ontologies. What we can say today risk business is that it is essential to identify these risks, prioritize, link them to processes, and implement an appropriate governance model to manage both from a financial performance perspective, compliance, continuity, image and protection of information, etc. Therefore, it is necessary to have a conscious and explicit approach to assessment and treatment of risks. It must then incorporate this approach into an information protection policy.

Today, the expression of needs is difficult: that protect and at what level?

How to conduct an ethical reflection on the current use of medical IS and Big Data healthy?

A pragmatic approach, progressive and risk mapping is a prerequisite both for a risk management approach, but also an efficient approach information protection.

II. MODELING NEO-PLATONIC SYSTEMIC ETHICS (Ψ , G , Φ) OF COMPLEX DATA

The appearance of Big Data allowed the emergence of a government powered algorithmic essentially of raw data, signals extra and intra-personal but quantifiable, operating by configuration anticipative possible rather than by standards of good conduct, and speaking to people only by warning devices independent of the human will. From this, our thoughts were quickly directed towards the development and implementation of a model adapted to analyze, understand and handle these large amounts of complex data.

The individual perceives "data", interprets "information", makes connections with other "knowledge" stored and is then able to act with "knowledge", he acquires the skills he is able to then to work repeatedly reflecting a "practical wisdom" [1]. Thus, this approach to epistemological and ethical time is part of the basic model for understanding the human dimension of the use of Information and communication technologies (ICT). Our interdisciplinary study course invites an epistemological critique focusing on "knowledge-process" that integrate Neo-Platonic thoughts of "doing to understand and understanding to do" by Leonardo da Vinci [2, 3].

To pass this move to simplification, we need to use the correct representation and modeling of substituting a description of processes at a state description. To be operational, the concept of information must be understood and used in a more systemic framework that connects it to the action via knowledge. Thus, knowledge is finalized in action and it reflects the shift from knowledge to know-used. Knowledge is information with respect to an end, she teaches on a goal, she taught for a given action. This is a real challenge to find out, theoretically and practically. Thus, this process of development of info-ethics to switch from one state (A) of complex knowledge, disorganized and unfocused to a state of (Ω) mere knowledge, structured and teleological. This passage of Big Data to info ethics is illustrated by the action of making sense of large scale. (Figure 1)

We have chosen to represent the state of simplified and complex knowledge, respectively, by the Greek letters Alpha (A) and Omega (Ω) with respect to their symbolic aspect.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

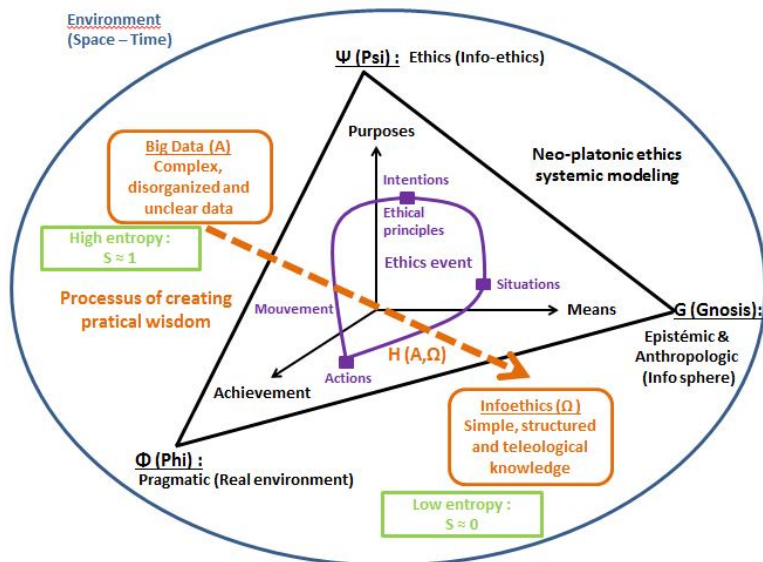


Figure 1: Transformation of the Alpha to Omega via the Neo-platonic prism systemic ethics

This transformation occurs through a systemic modeling ethical Neo-platonic (Ψ , G , Φ), which incorporates aspects of Ethics (Ψ : Psi), the Epistemic & Anthropology (G : Gnosis-Term Gnosis means knowledge in Latin) and Pragmatics (Φ Phi): The real environment).

This research infoethics helps strive for entropy (S : degree of disorder) very low, that is to say a degree of disorder almost zero. This process of legitimization of ethical knowledge of the infosphere (or Data sphere) associated with the pragmatism of the real environment, allows us to simplify our knowledge by approaching both a static and dynamic by its environment through its interactions. For [4], ethics is inseparable from the complex knowledge: [5] "Knowledge that connects, expressed and that is built as soon as one wants responsibility or solidarity citizen." This ethical reflection on the use of NICT is fueled by a triptych (action - competence - responsibility) that applies as much in IT in medicine.

The systems approach is relevant health decision where many elements from multiple players are interacting. The ethical decision is a fragile compromise between the actors involved in a given situation. This decision is unique and not reproductive as an environment and a given time. Always subject to questioning, it is in perpetual motion. Ethics appears as soon as there was a conflict because it must choose some values, while sacrificing others. Finally, this modeling ethical (Ψ , G , Φ) can be likened to what we called the "ethical data mining" implementing algorithms that are designed to bring out new information called info-ethics through analyzing a large amount of data. Reducing the amount of information is at the heart of an ethical dilemma. Should we lose or not volunteer information to regulate and operationalize this information (info-ethics)?

III. THE TECHNICAL AND ETHICAL GUIDANCE PROCESS

To our knowledge, no formal approach to achieve the interrelationships between risk analysis and the demands of digital medical personal data is established in the literature. The purpose of this section is to present an associative integration of these two processes. Specifically, we propose to design a guiding device in which the ethical goals of data derived from the associated risks. This is achieved through coherent and semantic associations between the ontology of these two processes. In addition, our guiding approach that allows you to switch from the study of risks to the data owners, to the expression of ethical requirements is data offers an improvement tool for a more controlled and supervised practice of personal health data. (Figure 2)

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

This alignment approach to concepts of ontology of the risks to those of ontology of ethical objectives requirements is set to change.

We can imagine such actions that:

- Enriching the two ontologies by introducing bridges of association between concepts that will, inter alia, to reveal potential conflicts between risks and / or ethical issues. This would provide better governance in the specification of the risks and would propose several scenarios to satisfy ethical issues.
- Automate the risk characterization step.

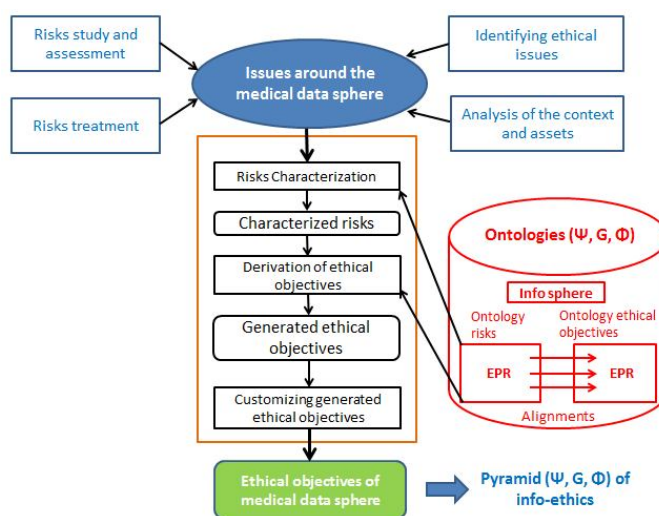


Figure 2: Architecture of the translation system of ethical objectives of medical data sphere

The primary objective is to model a body of knowledge in a given area, which can be real or imagined. In our case, the field is real and corresponds to the infosphere characterized by environmental parameters of the real (EPR) surrounding Big Data. From a literature search on the subject [6-10], we were able to distribute EPR in four specific areas order: Structural and Technical (ST), Strategic and Methodological (SM), Organizational and Regulatory (OR), Relational and Cultural (RC). These "keys to play" real multisectoral are articulated and allow them to give a practical application and value ethical framework.

The analysis about the context of medical IS and Big Data contributes to ask a series of questions that can be classified into two families of different nature:

A. Firstly, the field of IT:

- What computer models and mathematical choose?
- How to implement them?
- How to integrate and cooperate heterogeneous data?
- How to implement IS?
- How to make accessible Big Data? With what interfaces?
- How appropriate the new technologies of Big Data?
- How do you sort and find correlations between medical data, organizational, logistical and technical?

B. Secondly, the Social and Human Sciences Sector:

- What are the criteria for medical decision?
- What are the rules and limits of society?
- What are the business rules?

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

- How to find a consensus on the use of Big Data tool in daily practice?
- This technological tool is it used "on site"?
- How morally accept these NTIC among health professionals and health users?

C. Our governance approach is carried out in three stages:

- The potential risks identified following the definition of the problem of medical data sphere are characterized on the basis of the ontology of risk.
- The ethical issues associated with the risks identified and characterized, are derived.
- The list of ethical requirements obtained by derivation is customized based on the context of medical data sphere.

The approach used for the development of ontologies is based on an organization of concepts, beginning with the most general and continuing with more specific concepts. Each concept must have a unique and independent meaning, but each sub-concept represents a specificity corresponding concept. For example, organizational risks associated with decisions due to dysfunction [11, 12], or the risks "business", accidents and malignancies associated with internal / external dynamics of the business [13].

Furthermore, we built the ontology of ethical objectives inspiring us part of the analysis method of the ethical code for the use of Information and Communication Technology (ICT) proposed by D. Johnson [14]. Indeed, in his book "Computer Ethics" (1985), Johnson suggests that professional ethical codes should be examined through four main types of obligations, namely with regard to: the company, employer, customer and colleagues and other professional organizations. To this, we have combined two levels of action as:

- A reminder to the four principles of bioethics of [15], namely the principle: Beneficence, Autonomy, Non-maleficence, and Justice that must govern the personal data to avoid carelessness, the lack of foresight or negligence.
- A reference to which reference should be made to clarify the issues and duties of citizens and professionals of personal medical data.

Thus, building on numerous field surveys, targeted interviews with a hundred actors (information system managers, IS editors, health facility managers, health professionals, patient associations, users health, public institutions, etc.) and on a literature search on the subject, we could develop a model of ethical analysis of personal health data. Our ethics target consists of an analysis of: (Figure 3)

- The intrinsic value (design): Perspective of Organizational Learning.
- The management value (set-up): Perspective of internal processes.
- The operating value (use): Customer Perspective.

Our study tool is inspired by the "Balanced Scorecard" by Kaplan and Norton [16] incorporating the perspectives of their dashboard, except the financial perspective that we are dealing with an ethical valuation.

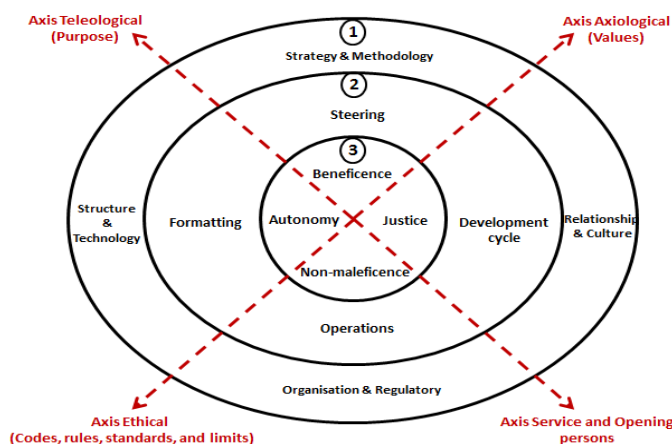


Figure 3: Target ethics of personal health data

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

IV. THE IDENTIFICATION OF RELATIONS BETWEEN THE TWO ONTOLOGIES

The linking of concepts of both ontologies reduces the direct involvement of users in the system matches the risks and ethical issues. This approach is carried out in five phases: (Figure 4)

- Recovery of all the databases provided by risk assessment methods.
- Identification of potential scenarios with their precise description and a list of associated actions.
- Study of each identified scenario that is related with the ontology of the risks associated with it associated risks.
- In parallel to this, we study the actions of each scenario attached to the ontology of ethical objectives.
- Implementation of associations related Risks / Scenarios with associations Scenarios / Ethical actions generates alignment bridges between the concepts of both ontologies.

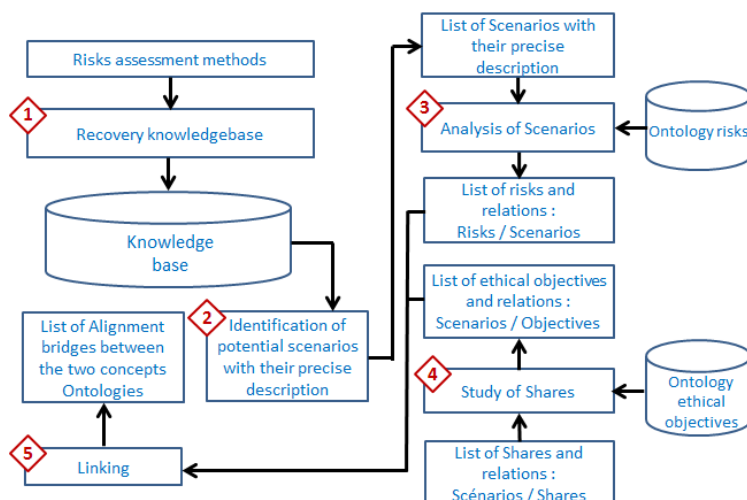


Figure 4: Identification System of alignment bridges

Note that the list of risks obtained through the guide device is more accurate. This precision is due to the fact that the bonds of heritage can reach all subtypes of risk. This accuracy is that by inheritance links can reach all subtypes of risk. This precision allowed us an accurate identification of ethical issues.

The following (Table 1) shows the alignment relationship between the risks and ethical objectives around the design, implementation and use of personal health data, obtained from our method.

Ontology risks	Ontologies ethical goals	Characteristics data sphere
	Assist in medical decision made by the health professional	
Business Risks	Assist the Ministry of Health to meet the expectations and taking care for the health user	Strategic and Methodological
	Establish with the patient legitimacy of law and information processing	Relational and Cultural



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

	Promoting quality, organization, management and planning of patient supported	
Organizational Risks	Divided fairly drawbacks and benefits of such a tool in the health professional workload	Organizational and Regulatory
	Assess performance and identify areas where action is required by listing the malfunction of points	
	Follow the statutory regulations of medical data	Organizational and Regulatory
Risks Threats	Working for the patient's well	Relational and Cultural
	Share a transparent and accessible information between the patient and the health professional	
Risks Errors (Management Information)	Ensure the quality and choice of the information given to patients	Structural and Technological
	Improve and enhance interactivity with actors outside the health facility	Relational and Cultural
Risks Errors (Usage - Transmission)	Improve continuity of care	
	Replace the patient at the center of the decision by providing a more complete and timely medical information: Better patient autonomy	Structural and Technological
	Develop and share accurate and appropriate information to the entire population	Strategic and Methodological
Errors risks (Application - Handling)	Minimize or eliminate the harm done to patients because of misinformation	
	Adapting technology to the knowledge and expertise of the health professional	Structural and Technological
	Establish a safety duty, integrity, traceability and protection of medical data	Relational and Cultural

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

Malignancies (Risks Disclosure Information - Risks Intentional Loss Information - Theft Risks/ Unauthorized Use Information - Risks Handling Information)	Ensure the technical relevance and human merits of the tool	Strategic and Methodological
	Allow an epidemiological or statistical analysis (SAE)	
Risks Failure Operation	Reduce unnecessary risks or miscalculated	Strategic and Methodological
	Enhance the availability of health professionals	Organizational and Regulatory
	Ensuring consent and the accession of the data owner	
Risks Vulnerabilities	Respect the privacy and the right to medical confidentiality and privacy	Relational and Cultural
	Respect the right to prior information, rectification and opposition to the describe the data owner	Organizational and Regulatory
Residual risks	Ensure the reliability of medical data and permanence	Structural and Technological
Risks Accidents	Respect the rules of storage, hosting and dissemination established by regulatory public institutions	Organizational and Regulatory

Table 1: Alignments between ontology risks and ontology of ethical objectives of data sphere

From these alignments ontological we built previously on the achievement, the establishment and operation of personal medical data, several measures and recommendations are binding on us to ensure the quality and protection of health data personal. (Table 2)

Identify the source of the data: raw, consolidated by different producers, constructed from third-party.
The personal health data can be collected and processed for a specific and legitimate purpose (See principle of finality).
Only the relevant and required data to operating objectives must be addressed. The analysis must be justified by the legitimate interest of the company (See principle of proportionality and relevance of data).
The medical data must flow guaranteeing ensure confidentiality and quality processes.
Clearly indicate the date of transfer or accommodation of data to ensure the good news and updated information.
Encrypt the data to ensure traceability and monitoring of the latter.
The information transmission tools must be valid.
Medical Data without constantly evolving and must be regularly updated.
The data is revalidated (revalidation capacity of diagnosis, therapeutic strategies, etc.).



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

Providing consistent information and based on medical evidence from medical sources and references (internal or external validating bodies) to ensure the credibility and well-founded data: percentage of validated data, validated method of selecting data, validation method (criteria), result (qualitative and quantitative) of validation, validation frequency.
Clearly indicate whether the information is based on scientific studies, expert consensus, or experience or professional or personal opinion.
Ensure that medical opinion or advice is given by qualified practitioners.
Describe the recruitment process (typology of contributors, contract supplier, method of remuneration, individual's consent).
Identify making modality, of incorporation or of the data processing.
Describe the customary use and conservation of its medical data by the supplier and that the conditions in which the user himself can access his data to ensure the integrity, security and confidentiality of data personal.
Personal health data must not be stored indefinitely and talk time is determined by reference to the purpose of processing.
Evaluate rigorously and fairly information, including that used to describe the product or service.
Respect the rights associated with intellectual protection: source licensing, rights of the producer / supplier of the data, license impairment due to the treatment of the data.
Ensure that suppliers comply with the generic and specific regulations related to the use of personal data.
Unless legal exceptions, the employer has a duty to ensure safety of personal data it operates so that they are not disclosed to unauthorized persons (See principle of security and confidentiality of data).
Consistency of data requires coordination time. [17]
Making information more accessible to patient or non-specialized healthcare professional so he can understand and use it in the best conditions. For this, products or services used should be described in plain language, easy to read and suitable for targeted users.

Table 2: Recommendations surrounding the development, implementation and operation of the health data personal

In light of these considerations, it seems essential to bring the heart of the establishment and operation of Big Data, sociology and ethics in charge of producing a framework of good practices of such data. Proper use of personal data based on organizational intelligence is therefore to eliminate, prioritize and sort the data available to give them meaning and coherence, not to accumulate. This means that a technology without ethics and without knowledge equivalent to a body without soul and without spirit.

Finally, the Code of Ethics and these recommendations should not be used as a rigid framework, but rather as a flexible structure to incorporate into the design, implementation and use of personal health data.

This model will be seen in three aspects interacting [18]:

- The first view, says ontological (or structural aspect), considering the IS in its structure. It is the vision of "being" of the device.
- The second point of view, to functional aspect, considers the IS in its function. It is the vision of "doing" of the device.
- The third view, says genetic (or dynamic aspect), considering the IS in its temporal evolution. It is the vision of "becoming" of the device.

Our reasoning around our ethical model is structured around four main themes that complement and fit together:

- The axiological axis: including human values;
- The teleological axis: describing the purposes of an action or fact;
- The ethical axis: explaining the rules, standards, laws, codes, standards and limits;



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

- The axis of the service and openness to people: translating the human relations between the different actors involved.

This cutting ethical target will highlight 12 areas that form what we call "Infoethics Management" (Charity, Justice, Non-maleficence, Autonomy, Strategy & Methodology, Organization & Regulatory Structure & Technology, Relation & Culture, Management, Formatting, Development cycle, and Operations) necessary for optimization and better control of Big Data in health.

V. ETHICAL DATA MINING

Faced with the problem of information overload illustrated by the emergence of Big Data, Data Mining technologies offer, thanks to Artificial intelligence (AI) devices to handle these large masses of data to extract relevant and critical information for an effective decision-making. The use of this intelligent NICT on personal data allows the profiling of individuals, predict their behavior and act accordingly. The idea of our "Ethical Data Mining" is to integrate all along the stages of the process of Data Mining, an ethical framework in order to bring more meaning and purpose based on the merits of the tool in question.

Thanks to AI devices it uses, Data Mining allows for the economy of time, resources and people to be able to understand and grasp these Big Data. The virtuous circle of our Data Mining is to convert unstructured data into massive info-ethics, the info-ethical in beneficial and consistent decisions. Data Mining is in the service of effective decision making, which consists of any process that will allow brewing, sort and analyzing terabytes of data in order to extract the right information. Often, data will report sources to various databases; their large number will require bringing them all under a single architecture called "data warehouse", before applying them the Data Mining technologies.

We talk about AI because this system bundles computer tools that will automatically operate research and study of data without the user having to enter or predetermine any assumption that the process would be expected to verify as accurate or not. Everything is automated, research such as analysis. Therefore, Data Mining is based on the ability of the machine to learn by itself, "the machine learning Ability" [19].

Our "Ethical Data Mining" tool integrates our Neo-systemic ethics modeling to brew, sorts the data to reveal the hidden relationships and correlations between the various data and to discover a pattern of behavior. The reliability of our model of treatment depends mainly on the validity and sourcing Big Data originally collected: these are not to be biased and flawed. Hence the interest of companies in the information including on the behavior of its customers, the utility for her detailed questionnaires on individuals from which the machine will simulate reasoning not only their behavior but the appreciation of the decision maker and produce effects to their respect. The intrinsic and qualitative value of "mega data" initial is essential and critical for optimal treatment of the latter.

We can summarize this process of Data Mining in six major steps: (Figure 5)

A. Knowledge of industry and trade: Any analytical meets specific needs according to the working environment and the field of activity. Therefore, to ensure the usefulness and sense of analytics, it essential to know the ethical issues and risks relating to the field of the company that will operate the Big Data.

B. Knowledge of raw data: The next phase is to assess the massive data available (or that the company may be obtained), to determine what they can contribute as information and understand how they apply to needs corporate clients, see the company itself if it concerns internal use.

C. The preparation of raw data: In this step, the data are prepared for their study. Data sources include combined relevant and consistent way. Integration is an important part of this device.

D. Ethical modeling: This phase requires our Neo-platonic ethics systemic modeling (Ψ , G , Φ) to monitor and supervise the most rigorous and fair as possible treatment of these Big Data.

E. Evaluation: It is then necessary to evaluate the model to check its correct functioning and operability. This causes most often adjustments and tune-ups, the possible addition of new data sources, the merger of these data, a modification of the model that has the advantage of being very flexible and evaluating a another hypothesis.

F. Deployment: When the model is validated and functional, the company can take the final step of the deployment.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

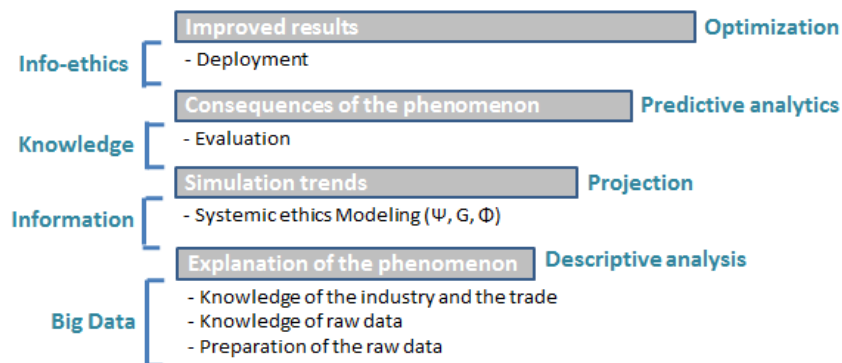


Figure 5: Stages of Ethical Data Mining

These steps constitute a complete cycle: once the ethics modeling on and the info-ethical obtained, the "Ethical Data Mining" provides new knowledge and craft new responses depending on the given context. At this point, the cycle repeats when the company wants to tune in response to changing environmental and market strategy around Big Data.

Ultimately, it is essential to make sense of the data. For this, it is essential to learn to "speak to the data", namely cross-check, intersect with other volumes of internal and external data, optimize operations either qualitatively and quantitatively. Such an approach can be made without an assessment tool to measure and monitor the ethical value of the raw data throughout its treatment: Design (intrinsic value) - Implementation (management value) - Use (value operating). According to the literature on the design of data quality, we can categorize and list three types of distinct approaches:

- First, empirical approaches that employ concrete experiences. We can take the example of Wang and Strong works [20] who interviewed directly to digital data users. The results of their interviews helped to highlight four dimensions around data quality aspects: the intrinsic quality, representation and the form, contextual, and accessibility.
- Second, the theoretical approaches that uses concepts of "real world" and "IS" for confronting with each other. Wang and Strong [20] also used this method in order to observe the differences between reality and technology. These differences translated as "defects" data quality revealed including three new categories namely: completeness, consistency and relevance, and accuracy.
- Third, intuitive approaches that use a ranking determined intuitively. Thus, in the work of Redman [21], data quality is categorized based on the value, size and nature.

The innovative aspect of our ethical analysis model is to have compiled and compared these three approaches to establish a Neo-platonic systems approach around Big Data. Under these conditions, we decided to build on a framework for study of three reading lines: actions, situations, and intentions that form an ethical event.

This modeling allows the simulation to better understand, support, foresight and evaluation of Big Data, especially in the medical sector. As we have seen, it is inspired in particular Shannon, Wiener, Morin and Floridi works. The challenge is to study both the epistemological elements, anthropological, philosophical and cultural related to communicational and informational arrangements that the processing of Big Data required.

Finally, our study model is neither a scholar of applied ethics instrument nor an empirical study to paint a picture of all the ethical issues raised. It is even less of a tool offering a ready answer to questions, also legitimate, professionals face daily in professional ethical for decision making around Big Data. This cross between empirical investigation and concepts proposed by theorists ethicists is to equip those involved in the use of data so that they can assume now unavoidable responsibility, a structured justification of the purposes and means of the underlying the large volumes of data. Therefore, our vision is

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

divided into two stages: firstly, the conceptual and methodological tools needed to build an "algorithmic ethics" of data, and secondly the description of the ethical issues concerning a specific situation.

From this thought process, our model can describe in detail statically, the environment of Big Data, and dynamically, activity flows, including existing interactions between these Big Data[22].

So, what ethical approach is best suited to equip professionals and decision makers involved in the given market in order to create healthy environments or strategies of behavioral changes around NICT?

It is in this context that we sought to lay the foundations of an ethic in Big Data of e-health. A 3.0 medicine based on an ethical analysis of modeling and improved performance through better dissemination of knowledge of the health professional. This study was made possible through the development of our ethical analysis modeling illustrated by our pyramid (Ψ , G , Φ) of info-ethics previously seen and that can make an encrypted weighting of all these results. From this technical and ethical guidance, the listing of ethical objectives on the design, implementation and use of personal health data, as well as criteria on quality and the management of this "data" in the info sphere, we could build an ethical evaluation of the data. This ethical valuation of personal data to a successful development of a radar graph, to visualize the axes of "Infoethics Management" (Beneficence, Justice, Non-maleficence, Autonomy, Strategy & Methodology, Organization & Regulatory, Structure & Technology, Relation & Culture, Management, Formatting, Development cycle, and Operations) and to identify areas for enhancement to improve personal health data. The objective is to establish an evaluation of the ethical value of health personal data. (Figure 6)

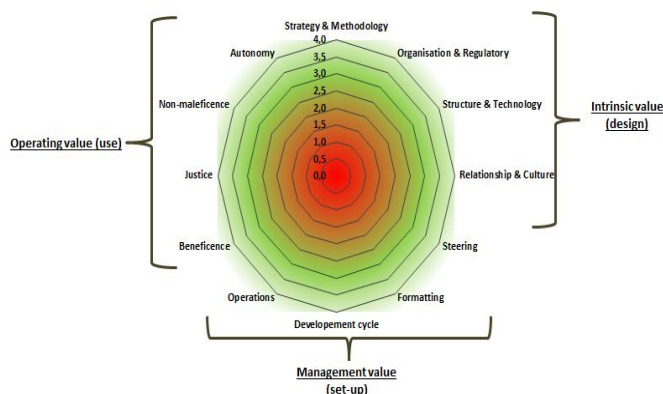


Figure 6: Ethical Value of Personal Data

The chart radar product at the end of assessment serves to position themselves on each of the 12 lines of "Infoethics Management" and identify areas to improve recovery to best use and master Big Data in Health. The 12 axes are distributed in three value categories for prospects of BSC Kaplan and Norton [16]. We described previously (Table 3).

- Intrinsic (design): Perspective of organizational learning.
- From management (set-up): Perspective of internal processes.
- Operating (use): Customer Perspective.

	Summary of average scores per axis	
Intrinsic value (design)	Strategy & Methodology	... / 4
	Organisation & Regulatory	... / 4
	Structure & Technology	... / 4
	Relation & Culture	... / 4
Management value (setup)	Management	... / 4
	Formatting	... / 4



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

	Development cycle	... / 4
	Operations	... / 4
Operating value (use)	Beneficence	... / 4
	Justice	... / 4
	Non-maleficence	... / 4
	Autonomy	... / 4

Table 3: Criteria for evaluating the value of a personal data

We apply the following scale of "scoring" of the 93 issues that are associated with indicators of intrinsic value, management, and operation of Big Data.

This scale is constructed as follows:

- One answer: "Yes totally" is 4 pts / 4.
- One answer: "Yes partially" equivalent to 3 pts / 4.
- One answer: "Neither yes nor no" is 2 pts / 4.
- One answer: "Not really," equivalent to 1 pt / 4.
- One answer: "Not at all" is 0 pt / 4.
- One answer: "I do not know" is X (that is to say, do not take this issue into account in the average score of the endpoint).

Therefore, from the results obtained, we can define and implement an "ethical management" device on the design, implementation and use of personal data from NICT. This pilot integrates ethics, among others, the conventional steps of methodology, such as "The definition of data quality, measurement of each of the measures, the analysis of the results of the evaluation, quality improvement, and monitoring of the effects of improvement" [23].

The techniques used for each of these phases depend on the type of processed data (structured, semi-structured, non-structured), the nature of the IS operating data (e.g. peer to peer, distributed, or warehouse) quantification and measurement strategies, continuous improvement techniques and quality dimensions to be considered.

Ultimately, our valuation tool ethics of Big Data is the cornerstone of what we have called the "algorithmic ethics".

VI. CONCLUSION

Big Data is not just a matter of volume and scale. NICT that can exploit, particularly in the health sector, are about a radical change in the medical paradigm in the doctor-patient by the processing of personal health data. Now, traditional medicine Hippocrates has given way to a 3.0 medical both personalized, participative, preventive, prognostic and predictive analysis based on digital data. The development of consumerism medical information then reveals a significant change in the legitimate demand of patients [24]. These naturally require more accountable for how they are cared for and how are used personal health data. Therefore, health professionals, producers and information providers must adapt their services and the way they work, taking into account this new situation.

Contemporary medicine provides a better understanding of chronic diseases, infectious diseases, new approaches to diagnosis and therapeutic treatment, of disease surveillance and health risk factors in public health, according to our habits and behavior by internet, health applications on smartphones and connected devices.

Digital performance concerns the volume of data studied, the diversity of sources and the search for a real-time response. This new approach led to new strategies for analysis of based data on probabilities, manipulation of complete distributions. This will translate into problems of automated decision processes and develop new visualizations from the new rules of interaction with the data. The Big Data release so our relationship to the data, we have to face it in a permanent exploratory mode. Therefore, the Big Data is a growing place within companies, marking a shift in strategy and organization of the latter.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

However, this data deluge raises some ethical concerns of their uses and deviations that may arise. IT management of personal health data calls for increased vigilance particularly for its protection and will be considered in compliance with certain conditions. This leads us to ask the question of the use and protection of health data.

Therefore, the quest for meaning which marks this digital shift necessarily involves a questioning of the values that guide our behavior and our actions both individual and collective [25]. This ethic, in the heart of Big Data, is all the more important that it is the health sector. Under these conditions, an awareness of the value of data from NICT becomes essential, as better supervision of the use of processing algorithms against individuals. Indeed, if we want to live harmoniously in this new world governed by the "algorithmic governance", we must intensify our efforts on the use of data that allows the "Machine Learning" to learn on their mastery by the user and on the transparency of their use.

It is from this observation that we have established a risk assessment resulting translation to ethical issues in medical data sphere based on ontologies. That successful in the development of a Neo-platonic ethics systemic modeling (Ψ , G , Φ), a technical and ethical guidance system, and a selective prioritization of health data. Now, each individual must reflect, sort and retrieve each data in a scale of rational value in the target can assume your choices and decisions on its use including health.

The realization of our "Ethical Data Mining" all along the stages of the data life cycle is illustrated by our tool of "Infoethics Management" that identifies and analyzes the ethical value of unstructured data. One of the challenges is to be clear about the rules of use of our personal health data in which the actors have, by facilitating traceability and security with respect for dignity and human freedom.

In our view, the proper use and protection of these data are essential elements of an ethical reflection on control procedures and supervision of the latter to retain a prominent place on privacy and trust among producers and providers of information, and contribute to a certain mastery of risks and deviations of these. Therefore anonymity by Big Data has become algorithmically impossible. Therefore, the "algorithmic ethics" becomes decisive so that professionals Big Data strike a balance between the rational and measured data processing, while respecting privacy and over-treatment that would be counterproductive especially in the doctor-patient relationship. This ethic will bring meaning to the treatment of these gigantic volumes of data in particular restoring causation correlations that are revealed by these farms.

Ultimately, we hope that our ethical vision will aim to strengthen public / private partnerships on sharing and the use of personal medical data, while preserving the security of personal data in trusted networks.

REFERENCES

1. Ricœur P, "Oneself as another threshold", pp.202, Paris, 1990.
2. Valery P, "The notebooks of Leonardo da Vinci, English Translated", Ed Gallimard, collection TEL, 1941.
3. Valery P, "Views, Ed the Round Table", pp. 228, Paris, 1948.
4. Morin E, "Method 6: Ethics threshold", Paris, 2004.
5. Morin E, "Method 4: Ideas threshold", Paris, 1991.
6. Dherse JL, Minguet DH, "Ethics or Chaos?", Renaissance Press, pp. 1-448, Paris, 1998.
7. Poncon G, "Information System: Towards an ethics of use", Hospital Review 531: 74-7, France, 2009.
8. Abbad J, "Organization and hospital management", Berger-Levrault, Paris, pp. 25-103, 2001.
9. Breton P, Proulx S, "The explosion of communication at the turn of the century", Editions Discovery, Paris, 2002.
10. Garro O, "Contribution to modeling the design of mechanical systems", Document for the Habilitation Searches, Grenoble I, Joseph Fourier University, pp.82, 1997.
11. Lambert SM, "Dynamic dashboards: An approach based on ontologies", European Univ., Dec, 2011.
12. Alberts C, Dorofee A, "Managing Information Security Risks: The OCTAVE (SM)", 2002.
13. Vasquez M, Lammari N, Wattiau IC, Akoka J, "From risk analysis to the expression of security requirements for information systems", 30th conference INFORSID (Informatics des Organization et Systems the Information et de Decision), Montpellier, pp. 337-362, May, 2012.
14. Johnson D, "Computer Ethics", First Edition, Englewood Cliffs, NJ: Prentice-Hall; Second Edition, Englewood Cliffs, NJ: Prentice-Hall, Third Edition Upper Saddle River, NJ: Prentice-Hall, 1985.
15. Beauchamp TL, Childress J, "Principles of Biomedical Ethics", Oxford University Press, 5th edition, New-York/Oxford, 2001.
16. Kaplan RS, Norton DP, "Using the Balanced Scorecard as a Strategic Management System. Harvard Business Review, January-February, pp. 75-85, 1996.
17. Fournier T, "Ethics and Cancer: Making Sense of care", ONCORA, 29 January 29, pp. 12, 2008.
18. Beranger J, "Health information systems and ethics: to Hippocrates", ISTE Editions, London, pp.1-418, 2015.



ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

19. May YC, "Web mining technology and academic librarianship: human-machine connections for the 21st century", Peer-Reviewed Journal on the Internet, 4: 6-7, June, 1999.
20. Wang RY, Strong DM, "Beyond accuracy: what data quality means to data consumers", Journal of Management Information Systems, 12: 5-33, 1996.
21. Redman TC, "The impact of poor data quality on the typical enterprise", Communications of the ACM, 41: 79-82, 1998.
22. Belkadi F, "Modeling approach of collaborative design situation", Digital document, 8: 93-106, 2008.
23. Joseph DF, Barnard W, "JURAN Institute's Six Sigma Breakthrough and Beyond – Quality Performance Breakthrough Methods", McGraw-Hill Professional, 2003.
24. Fainzang S, "The doctor- patient: Information and Lies", University Press of France, Paris, 2006.
25. Jean AS, "Ethics Information", University Press of Montreal, Quebec, pp.22, 2002.