

# THE EVOLUTION OF ARTIFICIAL INTELLIGENCE TOWARDS AUTONOMOUS SYSTEMS WITH PERSONALITY SIMULATION

Joël Colloc

Normandie Univ, UNIHAVRE, UNICAEN, UNIROUEN, CNRS IDEES

email: joel.colloc@univ-lehavre.fr

## KEYWORDS

Autonomous systems, Ethics, Robot personality, Emotion modeling, Responsibility, Forensics, Legality

## Abstract

Evolutions of artificial intelligence towards autonomous systems with personality-simulating devices, artificial emotions and languages will increasingly interact with human users and become partners in their everyday lives. This paper will examine the consequences of these interactions both at the individual level and on the place that these autonomous systems will take in the society. In February 2017, the European Parliament approved a report presented by an European deputy. This report analyses the legal consequences of the growing presence of robots in European daily life. A key measure proposed in the text is to give legal personality to robots. According to industry partners and stakeholders, this evolution of the law is made necessary by the precautionary principle and by the challenges of competitive and growth advantages. Are these systems useful to humans and acceptable to them? The presentation presents the main technological advances and feasibility of an artificial and emotional personality for robots, the ethical aspects of robots' interactions with humans, in particular the medico-legal notion of responsibility to individuals and the societal consequences of robotic deployment in society.

**Warning:** This material is protected by the Agence des dépôts numériques all rights reserved.

## Introduction

This article will focus on the evolution from memetic (MAI) to autonomous artificial intelligence (AAI) with the important question of machine that can be able to modify itself, to set up by itself its objectives and to decide alone how to accomplish them? Moreover, in the future some researchers trust that a machine could develop an artificial way of thinking making a robot to become an artificial person with legal personnality, responsible of the dammages caused by their acts to people and goods. A lot of stakeholders and researchers develop AI systems in medicine. Thus, we will take the healthcare

domain to illustrate the evolution of AI and their consequences for patients and caregivers. The article is structured in five sections: the first section describes the evolution from memetic AI to autonomous AI, the second section concerns the feasibility of autonomous systems and what features must be implemented to simulate artificial thinking capabilities. The third section presents the motivation to design and to develop autonomous robotics. The part four describes the impact on citizen life and particularly the social, legal consequences and a toolbox for digital ethics to take appropriate decisions. The next section try to forsee consequences of autonomous AI and robots endowed with consciouness and transhumanism.

## EVOLUTION OF ARTIFICIAL INTELLIGENCE: FROM MEMETIC TO AUTONOMOUS SYSTEMS

The AI is located at the crossroads of four families of disciplines: cognitive science that studies the nature of human intelligence: psychology, Linguistics, sociology, philosophy) neuroscience dealing with the operation of the brain, perception, emotions. Biology that describes the phenomena of living including genetics and social insects (Entomology), the behavior of the animals (ethology). Artificial learning (Machine learning) that uses mathematics and computer models to model learning, pattern recognition, knowledge modeling. Jean-Louis Le Moigne describes the evolution of information systems and a classification inspired from K. Boulding (1956). Like in epigenesis, the following level adds features to the previous one. By conciseness, only the levels 6 to 9 are presented Figure 1.

## Information hierarchy

On level 6, the system becomes able to memorize his decision O.I.D (Operating system (OS), Information system (IS), System of Decision (DS)). On level 7, the system coordinates numerous decisions of actions at all the time t, concerning its internal activity, regulation and the external information from and to its environment (SCS). On level 8, the system is endowed with a subsystem of imagination and design (SID). On level 9, the system is able to decide on its decision and to determine

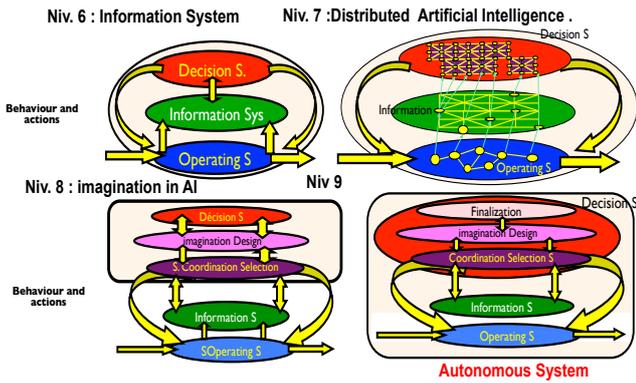


Figure 1: Complex system classification J-L. Le Moigne

the positive and negative aspects of its actions. This finalization of a complex system is close to the human thought (FS) which confers it an autonomy of decision allowing it to set its own goals. Autonomous multi-agents systems (AMAS) belongs to this kind Le Moigne (1990).

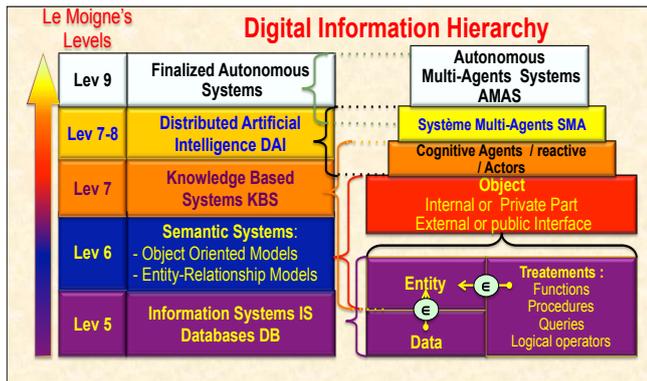


Figure 2: Digital Information Hierarchy

### Information hierarchy

Information relies on the triptych: data, information and knowledge Abiteboul (2012). But this classification must be completed with two important concepts: digital trace Galinon-Méléne (2011) and metaknowledge Pitrat (1990) Figure 2. AI is not just data but involves a hierarchy of information sources: On figure 3 from bottom to top: AI is not just data but involves a hierarchy of information sources: from bottom to top: -the digital trace is left intentionally or not on the net by the user: The clicks provide information on the preferences of users. They are time-stamped by the website. Data are characteristics or properties of a user

and the value belongs to an integrity set. Knowledge are Guidelines, heuristics are efficient methods to solve a problem in the same type of situation, they are independentes from individuals. Meta-knowledge determines what type of knowledge is necessary and how to use it to solve a specific type of problem. For example how to make a diagnosis. Obviously, meta-knowledge is also independent from individuals. Information is the inter-

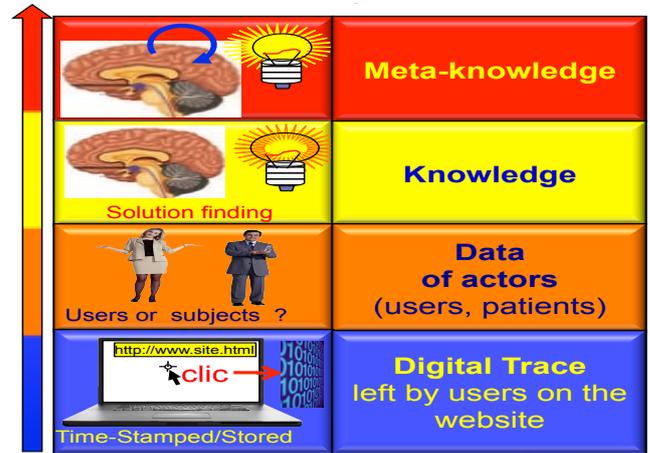


Figure 3: Hierarchy of Information sources from AI

pretation of one or more data associated with an object or actor (for example a patient in an hospital) whose treatment creates meaning in a context with arithmetic  $+$ ,  $-$ ,  $*$ ,  $/$ , logical {and,or,xor,not} or comparison operators  $\{=, <, \leq, >, \geq, \neq\}$ . The DateOfBirth attribute is used to compute the age of the patient with the function  $Age()$ . Th attriute Glycemia to assess the risk of diabetes by computing the  $DiabetesRisk()$  function when it exceeds the threshold of  $1.26\text{gram/liter}$  during two successive measurements.

**Knowledge:** In short, knowledge is the result of epidemiological studies in the population people and founded an aid to the diagnosis, prognosis, treatment. Databases index data collected and analyzed for each identified patient (medical record). The patient leaves traces of use on the Internet and social networks that allow a secondary re-identification and consolidate more data. For Alain Mille, "observation is a cognitive process to distinguish the imprint left in the environment (in the proper sense) as trace of something that may make sense... A paradox of security appears very quickly: access to the resources available in the computerized environment (not only the informatics environment so) very often requires procedures for identification of actors... Associated imprints are therefore very often specifically associated with the activity of its users and the observed traces reveal private behaviors." ...Mille (2013). The management of the property of computer traces is a nontrivial issue. Increasingly,

Modèles de l'IA	Applications
Very Large Data Base	Indexation by agents, Big Data, CRM, profiling of customers, CRM
Bayesian Networks	Decision support Systems combining statistics and Bayes Theorem
Decision Tree	Tree of decision: alternatives (If.. Then) that describes a decision process
Rule Base systems	Expert Systems using logical rules: deductive reasoning
Fuzzy Logic	Fuzzy induction reasoning, medicine, human sciences
Reactive multi-agents systems	Ants Algorithms, Transportation Simulation, Supply-chain management
Cognitive multi-agents Systems	Distributed AI, Cooperation of Decision support Systems, SHS, Emotion
Genetics Algorithms	Genetic encoding of solutions → Mutation → Competition → Selecting of best solutions
Evolutionary algorithm	Design, Artificial Invention, AI and Arts, music
Neural networks	Pattern recognitions, signals, situations, diagnosis
Case Based Reasoning	Analogic learning by computing distances with previous cases
Fuzzy Vectorial Spaces	Analogic Reasoning: with dynamic fuzzy vectorial spaces (time modeling of emotion)

Figure 4: Models used in machine learning

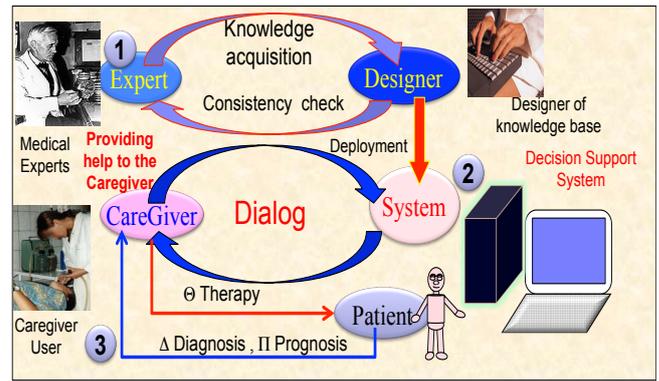


Figure 5: Memetic IA Decision support system

operators or companies who offer software or licensing systems, condition their utilization to the abandonment by the user of the property of its personal data and tacitly grant a right of communication to third parties for commercial purposes.

### Memetic AI, Knowledge base and decision support systems

AI today is a catch-all concept. However, it includes two main tracks: the Mimetic AI (the initial way) and autonomous AI (the new way). There are three steps in the Mimetic AI, the first step is the transfer of knowledge from experts to build a knowledge base in a given area, the second step is the designer of the knowledge base provides a support system to the decision Figure 3. The third step is the health care user (for example a medical student) dialogues with the system to solve clinical cases. In medicine, many systems have been developed which render services in aid to diagnosis and prescription for example assistance to users in areas or territories where few physicians are available. Convincing medical simulators are offered for students to place them in a situation of care without risk to the patient. The AI systems are implemented with models that are listed Figure 4 in an non exhaustive manner.

For example SIAMED offers a prescription from the diagnosis of infectious disease and its location in five successive stages which exploit the knowledge base of the antibiotics and clinical data of the patient's medical record Figure 6. The system implements the following clinical steps: the indications binded to the bacteria that could cause the disease and its localisation, the contraindications related to the State of the patient, drug-drug interactions with current treatments and takes into account the risks of side effects for each antibiotic. From the selected molecules, it offers the specialties of drug with a dosage and route of administration (Colloc 1985). Further works extend the capabil-

ities of the system to infectious disease diagnosis (Shen et al. 2018). Two inseparable aspects of medicine coex-

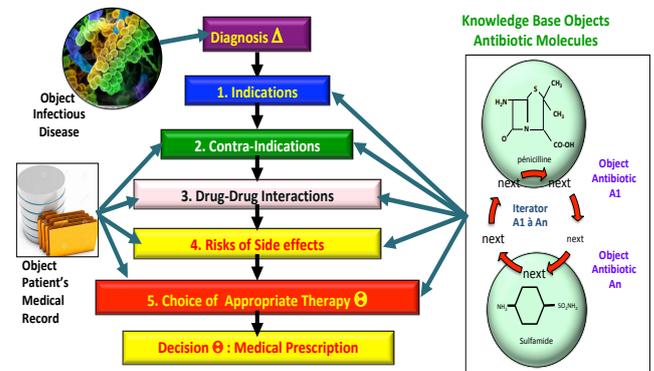


Figure 6: SIAMED: Antibiotherapy decision support system Colloc (1985)

ist, complement themselves: firstly, the clinical medicine in the service of the patient, each individual in the relationship (caregiver-patient) care with doctors on the left part of Figure 7. Secondly, the public health at the service of all of the population. The public health focuses on risk factors, to epidemics, to bad behaviour based on statistics in the population. It issues recommendations, hygiene measures and means of prevention on the right part of Figure 7. The objectives of public health and clinical medicine are sometimes contradictory. Care cannot be imposed on patients. In short, medical knowledge are the result of epidemiological studies in the population people and founded an aid to the diagnosis, prognosis, treatment. Knowledge comes also from medical publications that precisely describe the diseases and compare the treatment efficiency from the clinical experience of physician mainly in hospitals. Data are indexed in databases and are collected and analyzed for each iden-

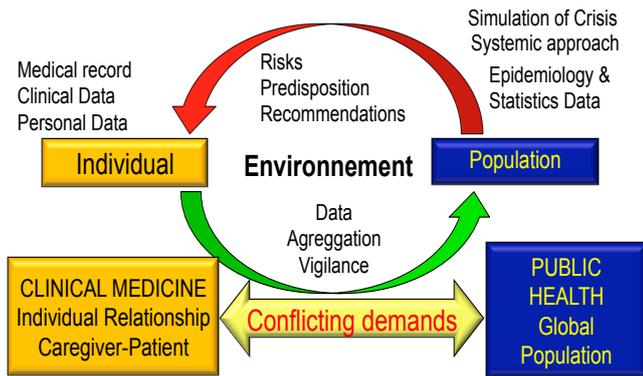


Figure 7: Clinical Medicine and Public Health

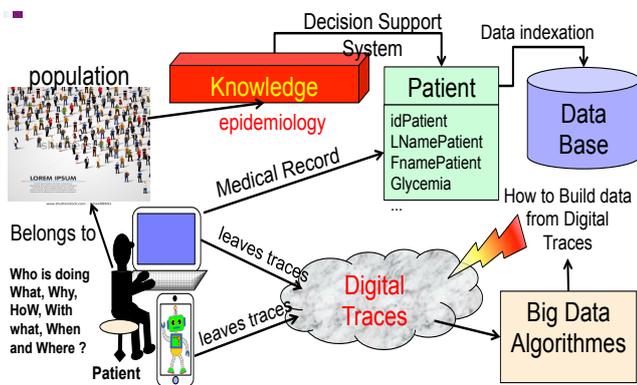


Figure 8: Reidentification of the Web surfer

tified patient (medical record). During the same time, the patient leaves traces of use on the Internet and social networks that allow a secondary re-identification and consolidate more data concerning the health and personal data of each person. For example, he will consult a medical website or a medical forum to understand the signification of clinical signs or to get information concerning the treatment of his/her disease Figure 8.

### Seven W to profile the Internet user

According to the method of Marcus Fabius Quintilianus (35-96 apr J.-C): who is doing what, why, how, with what, where and when. How Big Data turns digital traces of a hypothetical user (or patient) into data? The user profiling is determined by its repeated uses: Who : User identifier : IP Adress or email What : What website is visited by the surfer, Why: The traces leaved by the user (clicks and cookies) allow to know the motivations of the Internet Surfer for visiting the website. How: Clicks and keyword of research used by the Internet Surfer determines his/her interests in the visited

website With What: What Computer, Tablet, Pad or Smartphone is used. Where: The Website gets the geolocation of the machine When: The user action (query or click) is logged and timestamped. However, the actual user is uncertain, because this is the machine that is identified and several users (e.g. of the same family) may use it in turn.

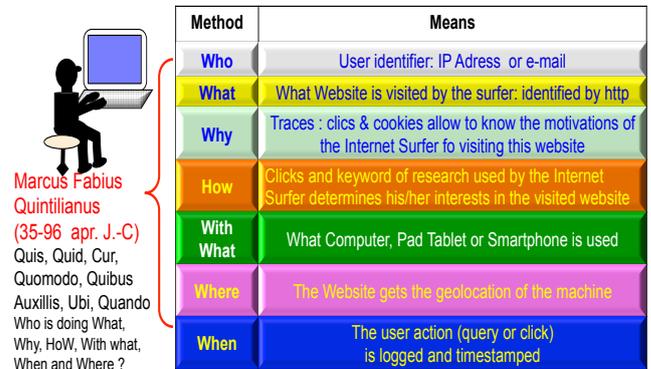


Figure 9: The Web surfer profiling

### Deep learning

The deep-learning aims to classify a large number of images or profiles of people from the net, in an autonomous way (Big-Data). He uses an artificial neural network with a large number of layers. In spite of his surprising performances, the main inconvenient of deep learning is that it doesn't implement meta-connaissance to explain the basis of the knowledge obtained and how and when to use this knowledge. John Searle's Chinese room: a man closed in the chamber is doing the matches between input questions with the output answers both written in chinese with the help of a manual of rules. The man could not pretend that he is able to master the chinese. In fact, to recognize does not mean to understand.

### HOW TO BUILD AUTONOMOUS SYSTEMS?

The recent progress of the cognitive sciences, neurosciences, computing and robotics boosts the project of autonomous artificial being. The deep learning methods allow to enhance the artificial perception (sight, hearing, touch, smell, proprioceptive sensitivity). However cognition especially reasoning and thinking remain a big challenge. The feasibility of autonomous artificial thinking systems needs to compare the way the human beings acquire their information and develops the thought with the current capacities of the autonomous information systems. We proposed a model based on four hierarchies: the Le Moigne's hierarchy of information systems supplies indicators of complexity and autonomy;

the cognitive hierarchy describes the sub-symbolic acquisition and the emergence of our personal experience, our knowledge, whereas the linguistic hierarchy builds the speech describing the knowledge acquired in terms of concrete and abstract objects on the environment and on oneself Figure 15. The digital information hierarchy relies on the necessary concepts, models, methods and tools to build autonomous information systems (Colloc 2016). The production of autonomous systems is rather different than memetic AI systems and the role of the end-user is totally changed Figure 10. An investor sells

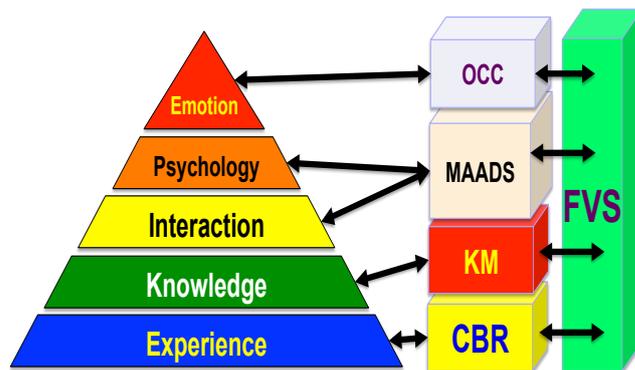


Figure 11: The EPICE Model

### The emotion layer E

The emotional layer E is composed of attributes that express balance and define the agent’s emotional states. It is based on a the Ortony Clore and Collins (OCC model) proposed in Ortony et al. (1988) and the Roseman’s model Roseman (1996). The layer E model defines a set of 11 emotional parameters that represent both positive and negative intensities Figure 12. Each emotion can change rapidly over time. The emotional parameters of the OCC model are represented by fuzzy functions and combined in a fuzzy vectorial space (Colloc 2017)

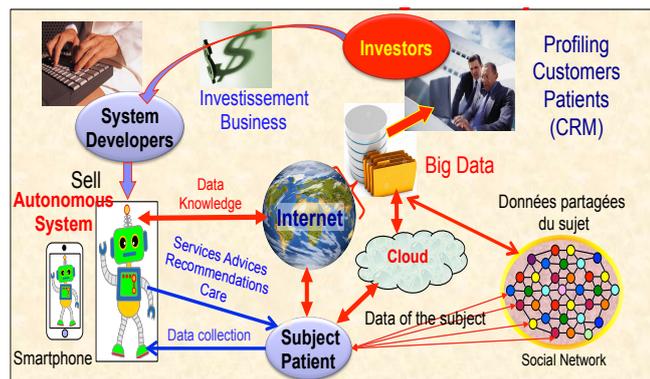


Figure 10: Autonomous AI context

robots companions or smartphone autonomous virtual assistants which provides advice, service, recommendations (in the future some cares) to subjects directly through a dialogue without the help of a caregiver. The autonomic system collects data of the subject and is able to get access and query data and knowledge on the Internet (Cloud, social networks, Big Data). He makes decisions in complex situations for the well-being of the subject. The collection of data allows to better understand the needs and expectations of the subject and how to answer or even to prevent his needs (may be too much).

### The EPICE model

EPICE is a french acronym for (Emotions, Personnalité, Interactions, Connaissances, Expériences) is a model that takes into account (Emotion, Personality, Interaction, Knowledge, Experience) and implements them in a decision support systems of medical ethics Colloc and Léry (2008). The implementation of the layers of the model is done with the fuzzy vectoriel space (FVS) Figure 11 which is described in our previous work Colloc and Summons (2017).

		+	-
Consequences of events	For others	Happy for	Resentment
		Gloating	Pity
	For self	Hope	Fear
		Joy	Distress
Actions of Agents	Self Agent	Pride	Shame
		Gratification	Remorse
	Other Agent	Gratitude	Anger
		Admiration	Reproach
Aspects of Objects		Gratification	Remorse
		Gratitude	Anger
		Love	Hate

Figure 12: The OCC parameters

### The personality layer P

The personality layer P describes the psychological and behavioural structure of an actor. These are generic prototypes describing typical psychological structures. These prototypes are specialized and thus adapted to the different actors in order to model the behaviours

actually observed. They correspond to decision-making patterns that are likely to evolve over an individual's lifetime. The personality layer is defined by a library of automatons modeling general behaviours and strategies. These automatons define metaknowledge in psychology. They model the decision-making modes corresponding to the evolution of an actor's emotional states. The personality layer is in close relation to the emotional layer which allows to represent sympathy, antipathy, aversion, self-esteem. The shape of automatons and the functions of transitions allow to represent particular psychological behaviors. For example, an obsessive actor will be represented by an automaton having a number of feedback loops due to his checking behavior before getting satisfaction.

### The interaction layer I

The interaction layer I describes the relationships and interactions that an actor has with other actors in the system. The complexity is linked to the diversity of the actors involved. Each actor can maintain relations with both legal persons and natural persons. The question that arises is what is the place of the various actors within a system of decision-making support in ethics in the context of humanitarian health. Catherine Fuchs in her book revisits the hypothesis formulated by Sapir and Whorf of linguistic relativism Fuchs (2003; 2004). For Edward Sapir, the language reflects a community's vision and interpretation Sapir (1930). His work is based on the classification of Native American languages. For Whorf there is no objective or universal reality, but only representations of it that would be determined by language. Catherine Fuchs explains that each language constructs a different "world view" because each language community selects distinct isolates of experiences and gives them shared meaning. Regarding the question of variants and invariants (one-sided): "the diversity of representations constructed through languages is a central question for linguistics: the theory of articulation concerning variations and invariants." We think that the triptych Language, Thought and Cognition is in action here: according to Jakobson, thought differs according to language. For Pinker and Fodor: there is no thought without language? Is there a thought and cognition without language? Do we think the same way as the Japanese or the Lapps? Louis Léry clearly showed that he was not. The culture of a people is organized around verbal and non-verbal modes of communication, beliefs, custom describes all the areas and constraints that will be involved in the decision. It also shows the hierarchy of decision-making criteria: law, ethics, charters, customs, beliefs and religion Léry and Colloc (2008) which is forward described in Figure 18. The brain is like a plant that grows in an environment: Our knowledge is linked to our culture and our environment. Can it be transferred to other peoples of different cultures? It is

doubtless possible, if one spends long years in a country, to acquire the customs and customs and to forget, unlearn one's previous culture, which will gradually fade away. This is just a hunch based on observation of a few cases. Our person can be fully and develop only in interrelation with others, hence the importance of the «relational flower». How to model interactions between legal and physical persons, groups of persons? A n-m relationship is necessary in order to represent all the interactions that may exist between many actors. The graph complexity increases dramatically ad  $(na.nr)^2$  where na is the number of actors and nr is the number of modeled relationships between them. Models from the work of linguists such as Quillian on semantics have been reused and adapted by computer scientists Quillian (1968) J.F. Le Ny shows the interest of semantic networks Ny (2005). The ontology models are intended for the axiomatisation of a field of knowledge Shen et al. (2015).

### The Knowledge layer K (Connaissance C)

The knowledge layer defines the sources of knowledge in which an actor has access because of his roles and of his skills. We define five categories of knowledge in a independent way chosen models of representation and domains of knowledge. Factual knowledge: it is about data describing an object of the world real and generally admitted by all. The observed facts are confidentially connected to the truth and classified according to their degree of certainty and precision. It is about a statistical argument: for example: the majority of the people have a similar perception of a characteristic of an object, for example its color is red. Knowledge heuristics: if a situation S is then observed we have knowledge which are relevant and valid in this situation S. It can be properties of objects or a usually applicable method successfully in this situation. The causes of the validity of a knowledge heuristics are not always available. Procedural knowledge: how acts on the world: know the chains of tasks to be made to reach an expected result. It is about all the procedures or the courses to follow expressed by a list of tasks to be realized to be effective in a given situation. Dynamic or behavioral knowledge: they concern the spontaneous variation of the facts, or a behaviour in time which is usually observed (for example, the earth rotates around its axis once every 24 hours). They are useful for the simulation of natural phenomena. It is about the perception of the various states spontaneously taken by objects during given period, of their interactions. The behavioral knowledge have a major importance in sociology, in economy, in botany, in medicine and in physics where we observe the behavior of a system. Deep learning is the new type of knowledge that can be now implemented in systems . System knowledge modelling relies on operators detecting the similarity of knowledge objects. Other impor-

tant aspects are time modelling and the reasoning operators: deduction, induction, abduction, sumsumption and analogy that allow to combine, compare knowledge objects. Deduction is based on the modus ponens or modus tollens. Induction tries to propagate a property observed in one object to all objects that belong to the same class and allow to split the class in two subclasses the one with the property and the other without the property. On the contrary, abduction tries to refute a property usually observed in the objects of a class. Sumsumption tries to generalise properties to more general concepts: Think of the individual under the general (an individual under a species, a species under a genus); consider a fact as understood under a law. General sumsumption could be applied to implement induction as proposed by Buntine in 1988 Buntine (1988). Analogy transposes the relationships and properties of objects from one universe to another one provided that these universes and object classes are sufficiently similar. The similarity is computed by a distance.

### The Experience layer E

The experience allows to enhance the knowledge in a specific area of the science. Case-based reasoning (CBR) implements a kind of analogy. CBR is a model of experience that allows to index and store cases in an object-oriented database and then, in next consultation, to use a distance to retrieve in the database the more similar cases to the new case in order to apply the most appropriate methods of the previously stored cases to solve the new problem Kolodner (1983)?Kolodner (1993)(Aamodt and Plaza 1994). CBR involves semantic distances developed by different approaches: algorithms of structural similarity Boulanger and Colloc (1992); statistical learning as proposed by Zighed D.A. and G. (1992); digital approaches from neural networks and fuzzy logic. The distances are using to implements the different stages of the CBR cycle as depicted on Figure 13 Colloc et al. (2007) The researches about semantic distances tend to combine symbolic and numerical aspects Shen et al. (2015).

### MOTIVATIONS FOR AUTONOMOUS ROBOTICS

#### The advantages of robotics

The advantages are the ones of the cybernetics or robotics which aims at the improvement of the human condition by unloading people of hard work and to assist them in the daily life as diligent and tireless companions in any sorts of activities of the life: education, supervision, safety of the elderly, the patients, the children. They are endowed with emotion and have a behavior which is going to look like more and more that of the human being, what makes the interaction with

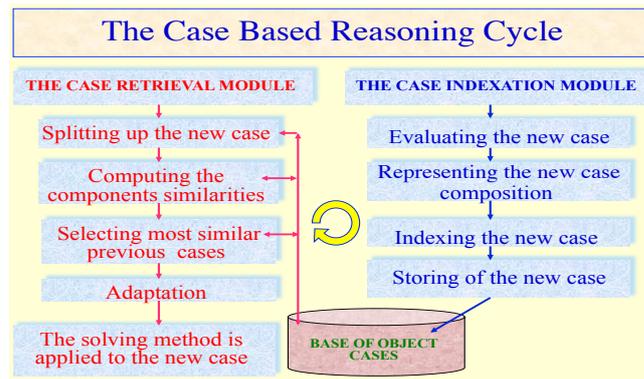


Figure 13: A Case Based Reasoning Cycle

them much more pleasant. The Conscious Autonomous Multi-agent Systems (CAMAS) has an internal memory to store facts of consciousness, but especially, they reach the global memory of the humanity from now on stored in Internet: not only in the knowledge but also in the data concerning each of us establishing our profile of way of life. The main question is addressed in the following subsection.

#### What is thinking ?

For the linguists as Pinker and his pupil Fodor, there is no thought without language, that is without words indicating the objects of the world. Jerry Fodor proposes an internal language named the "mentalese" which gives to each a reflexive thought, the capacity to speak to oneself and so to repeat for one of the information useful for his/her life. This reflexive thought is strongly bound to the consciousness of oneself and his/her body, to exist, to be an alive entity of the world with his history, souvenirs, feelings and projects. Even there, the self-awareness: "The I" appears late in the life of the individual, during the childhood and even the adolescence. The capacity to think is previous to the implementation of the language. When we speak to oneself, the internal verbalization stimulates the same intellectual cortical zones as during the expression by the word with request of the driving ways, but the motricity of apophyses aryténoïdes which tighten the vocal cords would be inhibited by the brain preventing the emission of sounds of the voice (Jacquet-Andrieu and Colloc 2014). Thus there it would be no important difference to speak and speak to himself (in the sense of the mentalais of Fodor), in both cases the areas of the language are requested Figure 14. We speak to ourself to strengthen our capacity of analysis and resolution of the problems but this activity is not necessary for the thought and do not even maybe fatal in a perception more lit of the world. Is to speak to oneself constitutes the only way of thinking?

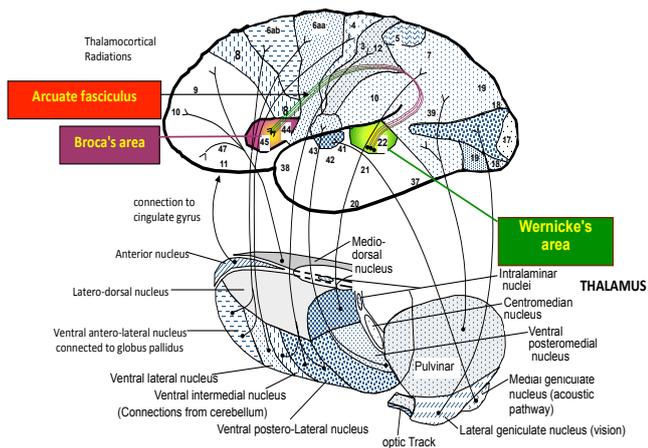


Figure 14: The brain, language area and thalamic connexions

### Autonomous systems: Could computer think ?

Computers could think in a very different way than human, as a plane flights differently but faster than a bird Pitrat (1993). The Zen is a total holism, the world cannot be absolutely divided into parts. The dilemma is that for every object of the world, according to master Zen Mummon: "we cannot express him with words and we cannot express him without the words". According to the Buddhism : to trust the words to reach the truth is equivalent of trusting an always incomplete formal system ?. For Jiddu Krishnamurti, our consciousness is common to all the humanity: All the human beings think that contribute to build it. He considers the individualism, the ego as an obstacle to understand the consciousness with rare moments of clarity ("insight") Krishnamurti (2005). "The thought is a movement in the time and the space. The thought is memory, memory of the past things. The thought is the activity of the knowledge, the knowledge which was gathered through millions of years and stored in the form of memory in the brain." Krishnamurti (2005). There are two forms of thinking: -The first is a reaction of the memory which contains the knowledge, the result of the experience from the beginning of humanity (phylogenesis) and since our birth (epigenesis) in a loop: experience knowledge memory thought action and so on, necessarily limited by the time, it is used every day, rational, individualistic, power-hungry and of progress submissive to the knowledge which accumulates, in the words which divide and this division is responsible for all the suffering, for all the troubles of the world. This first type of thinking relies on the symbolic treatment of the information and the linguistic hierarchy upper right part of Figure 15 with linguistic operators and able to implement the mentalesse as described by Jerry Fodor. -The

alternative thinking: the "active-attention" that occurs in rare occasions, we pay simply our attention to the world, without interpreting it, without naming anything, virgin of any prejudice, knowledge and especially spontaneously, by living this moment without thinking of it and without the will. For example: the direct perception of a wonderful landscape of mountain one morning with all our senses uses our complete attention where we forget ourselves and banish the use of the words" Krishnamurti (2005). Such a full perception requests only

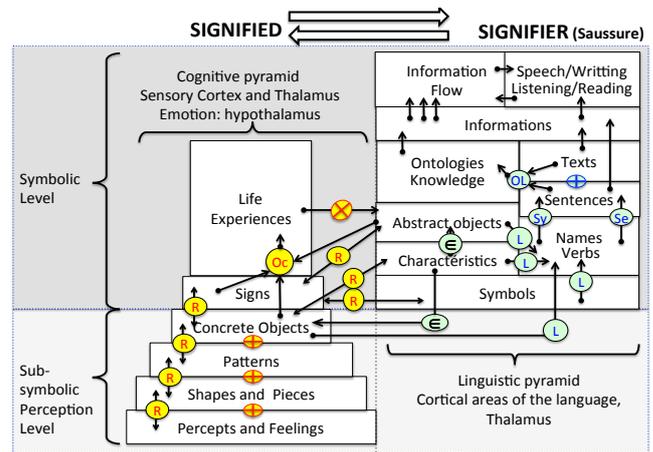


Figure 15: The cognitive pyramid

the left lower part: the sub-symbolic cognitive pyramid and not the linguistic pyramid Figure 15. The first way of thinking is a fatal vicious circle implemented soon in computers that will think better than human beings! And there nothing prevents that a computer invents a new religion at the origin of new sufferings for the humanity. The language during the evolution would have made lose to the man his spontaneity in the immediate perception of the world such as it is. The digital technology, which strengthens the symbolic nature of our relation to the world, is doubtless going to establish the peak of our ignorance. The realization of CAMAS able to think is now possible. Alain Cardon proposes a psychic system that relies on knowledge in psychology and capable of generating flows of thoughts that take place in the temporality under the shape of organized groups of processes to build the artificial psychic system and its interactions. This conscious autonomous multi-agent system (CAMAS) is able to choose its objectives and set its goals to achieve them (level 9 of the Le Moigne's classification of complex systems. Alain Cardon recently published a method to develop systems with psychological states and with the symbolic way of thinking inspired from Sigmund Freud Cardon (2016) Cardon (2018). The CAMAS can be reproduced in very large numbers and coordinate each other to reach complex goals fixed in common. The CAMAS is thus going to overtake the

human being in his capacity to be thought and to act in the world. Furthermore, it is endowed with mechanical, robotics architecture allowing it to move and to act in the world with more power and efficiency than human beings. It can be endowed with artificial senses like ordinary sight but also infrared vision, large spectral hearing, nano-cameras amplifying its perception far beyond our possibilities. The main advantage would be to be able to improve the capacities of investigation of our world including in hostile circles as the conquest of space by taking advantage of the concept of telepresence without risk Hofstadter and Dennett (1981). These systems have access to Internet and will quickly become more powerful than the human beings with disturbing consequences for the future of the humanity. We have shown that computers can think in the first way only. The main characteristic of humanity is to still be able to think also in the alternative way but how long before losing our essential faculty ?

## IMPACT ON CITIZEN LIFE

### Medicolegal aspects

Civil responsibility, which is based in France on article 1240 of the new civil code (ex 1382) is different for the mimetic AI and autonomous AI. For example in medicine, mimetic AI tools like a surgery robot is used under the responsibility of the surgeon who uses it. Consequently, the surgeon is responsible for any possible damage. In the context of an autonomous systems that makes alone a care act (example: drug delivery system). If this act causes injury to the patient. The company that produced the Robot is responsible for the damage caused by the system if and only if the link between the act and the damage is proved Figure 16.

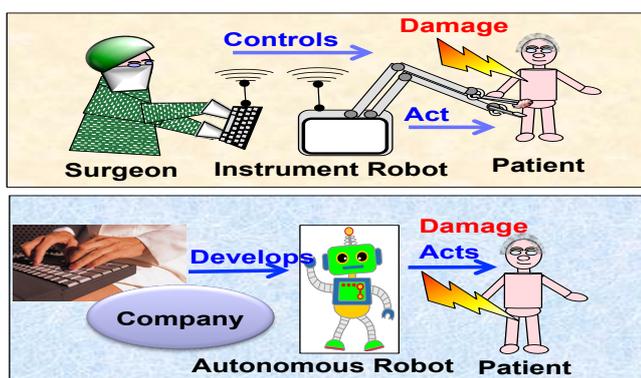


Figure 16: The civil liability of Robots'acts

## Ethics of artificial intelligence

There are two level of ethics: the first is the population and public health level where for example in France, the National Consultative Ethics Committee (CCNE) offers general recommendations useful to the population to address ethical decisions related to health and listening to citizens. Digital clinical ethics proposes a method of ethical decision-making about each patient the relationship of caregiver-patient care at the hospital or the practitioner's office. They try to foresee the impact of digital on the relationship of care and outcomes (benefits and risks) on the life of patients. The second is the clinical level that considers each individual in his environnement: Ethics begins where the ready-to-wear of morals is no longer sufficient. Ethics act is first a response to a situation limits and complex from the latin "responderere" with the double meaning: to warrant the acts i am doing to care this person, and to bring an answer to the problem; which is responsibility Léry and Colloc (2008). A Toolbox has been developed by Nicole and louis Léry with whom we formalized a metamodel of ethical decision published in 2008 ?.The method is pragmatic and is based on the Bergson's motto: thinking as a man of Action, act like a man of thought. Initially a question concerning the patient (this other myself). The question is necessarily intimate, unique, complex, because it concerns that person, his/her history. The question solicits his entourage (relational flower Figure 17) and need a reflection, then a decision that must lead to an action. This path leads us from the theoretical ethics (such as proposed by some Ethics Advisory Committees CCNE in France) to practical clinical ethics by successively visiting the following arguments: technical skills: what can be done with the professional skills in the medical art, knowledge, know-how, then, the risks/benefits/cost equation for assessing the achievable potential solutions. Then there are legal targets: law, jurisprudence and ethics for professionals, the recommendations and charters, moral, religious, beliefs the sociopsychological landmarks, cultures. All of the families, friends, caregivers who coexist the person participate in ethical reflection Figure 18. We are currently working to apply this ethics toolbox to digital ethics and especially to autonomous systems.

### Societal consequences

The recording information on the populations is former. Totalitarian regimes use it and have now powerful means to profile anybody in very precise way and with facial recognition they control the activity of whole population. The constitution of personal databases under the guise of public health needs have shocked the doctors: for example : the System GAMIN who organized the priori discrimination of socially disadvantaged children (french law of 15 July 1970). This law required

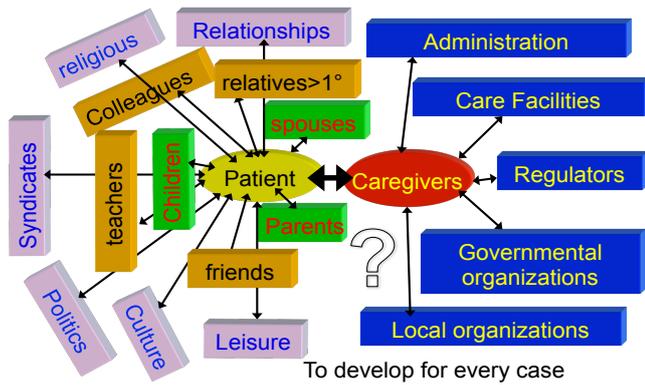


Figure 17: The relational flower of a patient

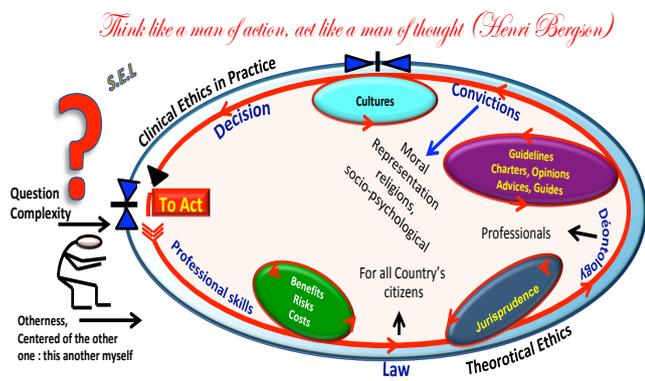


Figure 18: The ethics decision toolbox

that physicians declare disabled children, which constituted a priori discrimination of these children Colloc and Lery (1989), Vitalis (1981). We have seen that in technical terms, the concept of local administered and protected database whose purpose is known, declared and controlled in France by the *Commission Nationale Informatique et liberté* CNIL becomes obsolete due to transborder flows of personal data. This institution has therefore adapted and the french law "informatique et liberté" and has been enriched by the european General Data Protection Regulation (GDPR). But its means remain insufficient and especially its scope is limited to the national or the european perimeter while the Internet is a global network with different rules and laws. Outside european borders the controls and prosecutions remain difficult and most of the time impossible.

### Social risks of disclosure of medical data

The disclosure of patient health data may include a wide range of negative consequences. Firstly it can lead to a "loss of luck" related to the disclosure of disease (HIV

Infection, Hepatitis, Cancer, Addictions, state of depression) to a third party (insurer, employer, lender) who will refuse to grant a loan, employment or will increase the insurance fee: because he know the risk inherent to the disease. The exploitation of information against the person in a dispute (divorce, inheritance, childcare,...) can cause unpleasant results. The attacks on the reputation, brand image, ridicule, harassment related to the knowledge of a disease or an impairment... can cause very serious consequences with an exclusion from mainstream society, to be pushed aside in the company, job loss, rejection, marginalization, depression that can lead to suicide.

### Societal risks of autonomous systems

The main societal risk is that the CAMAS is put in the service of the ambitions and in the service of the power of an oligarchy of people who can acquire these very expensive sophisticated AI machines and on behalf of economic necessities, to impose their wills, their power on the rest of the humanity with the implementation of a totalitarianism based on the superhuman capacities (physical strength, speed of calculation, strategy, artificial thought) of such machines Figure 10. The study from the ethical point of view compares the arguments supplied by western and oriental philosophers on the nature of the thought and its role in the evolution of the humanity. The thought occurs at the sub-symbolic level in the cognitive pyramid. According to the Buddhist philosophy, but also the western logicians, giving a name to things is to classify and to classify it is to divide as defined by Lewis Carroll Carroll (1896). This division is source of the conflicts and prevents us from seeing the world directly. We are conditioned in an individualistic thought organized around the "me". The realization of systems conscious autonomous multi-agents endowed with the capacity to think is now possible. These systems will quickly become more powerful than the human beings with disturbing consequences for the future of the humanity. The AI mainly offers domination tools to people who have economical means, politic power to acquire them cognitive skills to use them. AI (eg. profiling, facial recognition, geolocalization...), could be used to enslave the poor and vulnerable populations, thus, main drawback of AI could be to enhance totalitarianism in the world.

### Transhumanisme

The transhumanisme is an elitist movement which aims at improving the human being by means of the use of the science and the technology to improve its physical characteristics (strength, longevity, disappearance of the suffering, the disease, the ageing and the death) but also mental by amplifying its capacities of perception, of memory and acquisition of knowledge and by delet-

ing the mental disorders. This movement joins in the lineage of the eugenics and constitutes certainly a very dangerous idea for the humanity. This idea confirms the apology of the individual thought and the glorification of the "Me" denounced by Krishnamurti who will bring more suffering to the humanity.

### Autonomous weapon

The development of drones and sophisticated autonomous armament systems makes us be afraid that this evolution is inevitable. There are many more weapons than hospitals in our world. An open letter with petition against the autonomous killing robots "Autonomous weapons: the year open letter from HAVE and robotics researchers" Future of Life Institute (2015) was recently displayed online by the researchers in artificial and robotics intelligence to warn the advent of such systems which could fall in the wrong hands and create the chaos at the world level. It seems to me reasonable to sign her and to make her known. The freedom, our lifestyle and even the future of the humanity is at stake. It is our duty of responsibility for warning the disasters bound to the threats which we perceive as seriousness Dupuy (2002) and to warn by throwing this alert. This imperious necessity will maybe make excuse the unusual length of this work

### Conclusion

We described the advantages and drawbacks of memetic and autonomous AI, the hierarchy of information, a classification of the AI systems and the different interactions and roles of human beings that provide and use them. We used as example the AI medical systems that are used in medicine and we compared the consequences of memetic AI and autonomous AI on the patient-caregivers relationship. The state of the art concerning the nature of the thought and the consciousness allowed to confront the arguments in favor and set against the actual feasibility of autonomous system able to act alone and in a second time to be endowed with the capacity of artificial thought towards an artificial consciousness necessarily different from the human consciousness but more ubiquitous and powerful because connected to Internet. The Krishnamurti's first way of thinking that relies on the experience knowledge memory accumulated from humanity history will be implemented soon with AI and Internet, leading to the emergence of an artificial robot personality. The second way of thinking is the individual insight "active-attention" that each of us can live in some occasion, could certainly not be implemented in robots, because it comes from the fact that we belongs to the humanity and share the same phylogenetics: we are all cousins. Robots have no history. We presented the medical ethics toolbox that was extended to study the digital ethics when autonomous

systems are implied in people care and the liability consequences in these situations. We also presented the social impact of the AI development at the population level and at the individual level. Today, most people are connected to Internet and are labelled with their e-reputation which becomes so important that each person who want to be in relation (professionally, financially, for business, personally) with you will have a look on your records on the social networks to get informations concerning your e-reputation. The impact on your life is so important that Big data and autonomous AI, deep learning are building the new digital world where everybody should try to exist. In one hand, AI provides new means to enhance people's life and in the other hand, the powerful AI systems provides means to control and enslave the populations and to destroy democracy. A reflexion on digital epistemology and ethics must be one of the most important challenge of the AI researchers today.

### REFERENCES

#### References

- Aamodt A. and Plaza E., 1994. *Case-Based Reasoning, Foundational Issues, Methodological Variations, and System Approaches*. *AI Communications*, 7(1), 39–59.
- Abiteboul S., 2012. *Sciences des données : de la logique du premier ordre à la Toile*. <http://books.openedition.org/cdf/529>.
- Boulangier D. and Colloc J., 1992. *Detecting Heterogeneity in a Multidatabase Environment through an O.O Model*. In *IFIP, DS5, International Conference on Semantics of Interoperable Database Systems*, Victoria, Australia.
- Buntine W., 1988. *Generalized Subsumption and Its Applications to Induction and Redundancy*. *Artificial Intelligence*, 36, 149–176.
- Cardon A., 2016. *L'usage des systèmes informatisés pour le contrôle de tout et de tous*. *Les cahiers du numérique*, 12, no. 1-2, 171–185.
- Cardon A., 2018. *Au-delà de l'intelligence artificielle de la conscience humaine à la conscience artificielle*. ISTE.
- Carroll L., 1896. *Symbolic Logic, Part I Elementary*. London Macmillan and Co.
- Colloc J., 1985. *Informatique Médicale (SIAMED) Application : Logiciel original d'antibiothérapie médicale*. Med. Phd, Fac. Medicine, University Lyon 1.

- Colloc J., 2016. *Perspectives and Ethics of the Autonomous Artificial Thinking Systems*. ACTA SYSTÉMICA, 16, no. 1, 31–36. ISSN 1813-4769. Meritorious Scholarly Contribution Award.
- Colloc J., 2017. *A Fuzzy Vectorial Space That Avoids to Defuzzify the Membership Functions*. In P.J. Gonçalves (Ed.), *ESM'2017 The 2017 European Simulation and Modelling Conference*. EUROSIS, 13–24.
- Colloc J.; Cuvelier M.H.; and Summons P., 2007. *A Clinical Decision Metaknowledge Model*. In *ISC'2007, EUROSIS, Delft, Nederland*. 23–29.
- Colloc J. and Lery N., 1989. *Secret Professionnel et Exploitation de Bases de Données Médicales Informatisées*. In *Association des Epidémiologistes de Langue Française. A.D.E.L.F. Bordeaux II*.
- Colloc J. and Léry L., 2008. *Un métamodèle d'aide à la décision en éthique médicale*. *Santé Décision Management*, 11(1-2), 255–274.
- Colloc J. and Summons P., 2017. *An approach of the Process of Addiction: A model of the experience*. In D.M. Dubois and G.E. Lasker (Eds.), *Proceedings of the Symposium on Reversible Time, Retardation and Anticipation in Quantum Physics, Biology and Cybernetics 29th Int Conference on Systems Research, Informatics and Cybernetics, Baden-Baden*. IIAS, Vol 1, 103–107.
- Dupuy J.P., 2002. *Pour un catastrophisme éclairé Quand l'impossible est certain*. 2-02-066046-6. Editions du Seuil.
- Fuchs C., 2003. *Les langues entre Universalisme et relativisme*. In *Le cerveau et la pensée la révolution des sciences cognitives*, Sciences Humaines éditions. 143–151.
- Fuchs C., 2004. *Pour Introduire à la linguistique cognitive*. In *La linguistique cognitive*, Ophrys MSH. 1–16.
- Future of Life Institute, 2015. *Autonomous weapons: an open letter from AI & robotics researchers*. URL <http://futureoflife.org/open-letter-autonomous-weapons/>.
- Galinon-Méléneq B., 2011. *L'homme trace Perspectives anthropologiques des traces contemporaines*. 410. CNRS EDITIONS.
- Hofstadter D. and Dennett D., 1981. *The Mind's I*. Basic Books Inc.
- Jacquet-Andrieu A. and Colloc J., 2014. *From Self-Awareness to the Consciousness of the "Speaking Subject"*. *International Journal of Computing Anticipatory Systems*, 28, 201–217.
- Kolodner J.L., 1983. *Maintaining Organization in a Dynamic Long-Term Memory*. *Cognitive Science*, 7, no. 4, 243–280.
- Kolodner J.L., 1993. *Case-Based Reasoning*. California: Morgan Kaufmann.
- Krishnamurti J., 2005. *La nature de la pensée*.
- Le Moigne J.L., 1990. *La modélisation des systèmes complexes*. DUNOD-BORDAS.
- Léry L. and Colloc J., 2008. *Prise de décision dans l'éthique au quotidien ? Comment décider le soin ? Santé Décision Management*, 11(1-2), 243–254.
- Mille A., 2013. *Des traces à l'ère du Web*. *Intellectica*, 59, 7–28.
- Ny J.F.L., 2005. *Comment l'esprit produit du sens*. Odile Jacob.
- Ortony A.; Clore G.; and Collins A., 1988. *The cognitive structure of emotions*. Cambridge University press.
- Pitrat J., 1990. *Métaconnaissance Futur de l'intelligence artificielle*. HERMES.
- Pitrat J., 1993. *L'intelligence artificielle : au-delà de l'intelligence humaine*. In D. Béchillon (Ed.), *Le cerveau : la machine-pensée*, L'Harmattan Paris, 2-7384-1818-X.
- Quillian M.R., 1968. *Semantic Memory*. In M. Minsky (Ed.), *Semantic information processing*, Cambridge, MA: MIT Press. 216–270.
- Roseman I.J., 1996. *Appraisal determinants of emotions: Constructing a more accurate and comprehensive theory*. *Cognition & Emotion*, 10, no. 3, 241–278.
- Sapir E., 1930. *The Southern Paiute language*. *Proceedings of the American Academy of Arts and Sciences*, 65, no. 1, 1–296.
- Shen Y.; Colloc J.; Jacquet-Andrieu A.; Guo Z.; and Liu Y., 2018. *Constructing Ontology-Based Cancer Treatment Decision Support System with Case-Based Reasoning*. *CoRR*, abs/1812.01891. URL <http://arxiv.org/abs/1812.01891>.
- Shen Y.; Colloc J.; Jacquet-Andrieu A.; and Lei K., 2015. *Emerging Medical Informatics with Case-Based Reasoning for Aiding Clinical Decision in Multi-Agent System*. *Journal of Biomedical Informatics*, 56, 307–317.
- Vitalis A., 1981. *Informatique, pouvoir et libertés*. Economica.
- Zighed D.A. A.J. and G. D., 1992. *SIPINA : Méthode et Logiciel*. Ed. Lacassagne.

## BIOGRAPHY

**Joël Colloc** earned his M.D. at the medical faculty of Lyon and a specialty degree of forensic medicine with a degree of clinical toxicology. He received a MSc. degree of IT from the Business School of Lyon (IAE) and a MSc. degree of computer sciences from the engineering school INSA of Lyon. He served as forensic physician at the Edouard Herriot Hospital in the neurological emergency department to cure drug addicted people, medical ethics and developed drug and addiction database. He went on to earn his Ph.D. in computer sciences at the INSA of Lyon. As Hospital assistant at the laboratory of medical computer science and he taught IT at the medical faculty. He was elected as associate professor in computer sciences at IAE of Lyon and he earned his accreditation to supervise researches in sciences at the Lyon 1. He is a Le Havre Normandy University professor in computer sciences since 2003. His main research topics concern e-health and particularly, fuzzy vectorial spaces (FVS), multi-agent clinical decision support systems (MADSS) and knowledge bases, Case Based Reasoning, ontologies, nervous system modeling and cognitive sciences and AI applications in medicine and human sciences. His human sciences researches try to conciliate the ethics of using Big Data in epidemiological studies, autonomous systems and robots and keeping ethics use of AI in order to improve clinical decision in medicine while preserving the patient-caregiver relationship, the privacy and the freewill choice of the patients.