Journal of Information Assurance and Security.
ISSN 1554-1010 Volume 15 (2020) pp. 053-064

Building trust over intelligence for autonomous systems

Marwa Brichni[1], **Said el Gattoufi**[2]

[1]Higher Institute of Management, SMART laboratory
University of Tunis
Bardo 2000, Tunis, Tunisia
marwabrichni@gmail.com

[2]Higher Institute of Management, SMART laboratory
University of Tunis
Bardo 2000, Tunis, Tunisia
algattoufi@yahoo.com

Abstract: **Trust building follows a dynamic process from initial trust defined by the propensity to trust in automation to trust development which is dependent on the trustworthiness of automation. Based on prior studies, many factors intervene in propensity formation and trustworthiness constitution, nevertheless their influence on trust is not detailed enough to order them according to their standing. In this survey, we highlight the importance of intelligence among the factors influencing trust and we suggest an uncertainty derived approach to build trust**
Keywords: Autonomous systems, Uncertainty, Machine intelligence, Trust management

I. Introduction

Trust, like many aesthetic concepts that are clustered among the category of ambiguous concepts, suffers from the absence of an agreed definition and objective concrete measures.Efforts were gathered to elucidate this matter, but trust results did not show obvious improvement in what concerns human acceptance towards automation.

Deloitte reports about customers' tendency to trust autonomous systems regularly. According to a study they have carried [1], authors have reported interesting statistics about autonomous vehicle's acceptance which reflects their trust towards these systems. The study shows that fully self-driving cars will not be safe percentage decreased from 72% in 2017 to 47% in 2019 which is a huge transformation within human beliefs about automation. Although the statistics are promising, the adoption of these autonomous systems in real life still is not considered as only 7% have tried riding autonomous vehicles and only 26% who are very interested in.

Trust in computer science has been tied to HMI "Human-machine interaction" and HRI "Human-robot interaction", principally through robots than other autonomous systems such as self-driving cars and autopilots. Numerous works are interested in quantifying trust by understanding the concept behind the interaction between humans and machines. Usually, the resulted trust is related more or less to three pillars, the human trust in the machine, the machine trustworthiness, and the environment influence in trust The ironic fact about trust is as the level of autonomy increases, trust decreases and this has been shown by prior studies. Thus, it is important to rely on features that are unquestionably reliable such as intelligence. Nevertheless during this century an abuse of the term intelligence was extremely obvious even though the fact of portraying intelligence is still challenging. The act of misusing intelligence, as a marketing term, is selling illusion to trust buyers.

II. Automation

A. Definition of automation

Autonomy, automation and autonomous are all appellations of the same coin but with a slight difference. They are accordingly used to describe an aspect, a process, a system of tasks usually executed by a human. Automation is defined as one process by which a machine executes a function previously accomplished by a human [2]. In a more detailed explanation, Lee [3] defines automation as one multitask technology that collects and filtrates data, transforms information and makes decisions.
In the same direction, automation according to Sheridan [4] is considered as the involvement of environmental variables sensing and processing information with the corresponding mechanical action. Autonomy, under what reported Bradshaw and fellows[5]: "idealized characterization of observed or anticipated interactions between the machine, the work to be accomplished, and the situation". As for autonomous systems, they are a sort of advanced automation able to learn, to evolve to change functional capacities effectively[6].

B. Trust importance in the use of automation

Lee's definition of trust in automation was "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" [3].
The importance of examining trust in the use of automation is clear from the increasing complexity of autonomous systems that might cause severe consequences concerning their predictability. In another work [7], Lee has concluded when to use automation and when to not rely on it, by verifying if the trust in automation exceeds an operator's ability to perform the same task that has to be automated. The vice versa is true as if an operator's self-confidence exceeds the trust in automation than automation is not.
Lee [8] treated other aspects of trust when it comes to how to deal as a human operator in front of an autonomous system. He defined the over-reliance on trust as the full dependence on automation even when it is faulty which can result in the misuse of automation. The second aspect is under-reliance which resumes in the full dependence on the human operator when the autonomous system is able to perform the same tasks and it leads to disuse.
One of the main challenges for the integration of autonomous systems in the human lifestyle is the trust allocated to them. Lack of trust hardens the involvement of advanced autonomous systems and over trust increases invalid expectations yielding to distrust. Thus, any wrong quantification of trust might beget a sort of misuse, disuse, abuse of the autonomous system.
Prasuraman and Riley [2] have identified differences between expectations, requirements while dealing with automation and they classified them into three categories as disuse, misuse, and abuse. Disuse of automation refers to failing while using automation when it is supposed to enhance performance. Misuse of automation is the over-reliance on automation. Automation abuse happens when it is implemented without regard for a human operator.
Hoff and Bashir [9] have introduced a solution to reduce the frequency of misuse and disuse of automation by means of appropriate levels of trust in automation. Thus, it is mandatory to quantify trust to calibrate automation use.

III. Trust management in automation

According to a recent survey adopted by [10], autonomous systems are a fascinating technology. Nevertheless, 43% of the survey participants state that they are hesitant about driving in an autonomous car. Hence, an important percentage of participants are worried about riding any autonomous system for their questioned reliability.
Authors in [11] described building trust as crucial in the development and the acceptance of artificial intelligence. We believe that building trust is also a responsibility based on Siau's citation: "Like any type of trust, trust in AI takes time to build, seconds to break, and forever to repair once it is broken"[11].

A. Definition of trust/distrust

Trust plays an important role in many contexts. Authors [12] have identified different views about trust by meeting expectations socially, learning from experience psychologically and taking the risk based on a moral relationship.
Most of the trust-based solutions are used mainly for e-commerce applications and cloud computing for security reasons. The adoption of trust has been invoked two decades ago for intelligent systems by the pioneer Marsh [13]. He argued that artificial agents should be able to make decisions about subjects of trustworthiness.
Rotter [14] characterizes trust as an attitude of reliance over an expectation made by the trustor towards the trustee's behavior. According to Mayer and fellows [15], trust is considered as an intention based on the expectations that the other will perform the exact action.
Fishbein and Ajzen's [16] presented the theory of Reasoned Action to settle the conflicting definitions about trust by a framework to define the relationship between attitude, intention, and behavior. Trust acts as a decision heuristic applied when a total understanding is absent[17].
According to [18] the majority of trust definitions include two sides, the trustor to give trust and the trustee to receive it and behavior or action to be achieved as a trust simulator. Muir [19] described trust as a "hypothetical latent construct" and that means constructs that cannot be observed or measured.
As for Lewicki and fellows [20] they presume that the huge available literature about trust since a long time ago only makes the distinction between the provided definitions, aspects, types more difficult than how it should normally do.
Trust as other aesthetic concepts such as intelligence and awareness suffers from inconsistencies between their definitions and this is due to how each scholar sees it, some consider trust as an attitude, some as an intention or behavior. Muir [21] has introduced the concept of mistrust as the error of trusting incompetent or distrusting a competent trustee.
It is important to consider carefully user's opinions about autonomous intelligent systems that draw their fearful perceptions from mistrust caused by lack of experience first and anticipated emotions towards artificial intelligence.
According to Rotter [22], Jian [23]and schoorman [24], distrust and trust are opposite ends of the same construct which is the attitude towards the trustee. Slovic [25] insists on the fact that they are not opposite since they are conceptually different. He is supported by Mc Allister's opinion [26], as he asserted that trust and distrust are distinct concepts by defining both as the possibility of either desirable or undesirable outcome. Authors [27] [28] from their own reviews have concluded to the fact that understanding and defining trust is always of an ongoing debate such as the other aesthetic subjects and in many fields and especially in automation.

B. Trust implementation in automation on the basis of trust between humans

Sheridan and fellows [29] [30] in some very old statements have presented trust as the means mediating between human and automation as it is mediating between humans. The similarities between the human-human interaction and the human-autonomous system are also highlighted by Nass [31].
Lewandowsky [32] argued that the similarities detected between both interactions also involve trust in both sides of in-

teractions based on the ongoing dynamics during a task completion in uncertain environments.
Since many works have been commonly stating the similarities between the trust in human-human interaction and human-automation interaction, a review was made by Madhavan [33] to detect the differences in the sense to know how to generalize trust in people to trust in automation with minimalist deficiencies.

C. Concepts and Factors of influence on trust

1) Concepts:

The theoretical basis of trust formation as a dynamic attitude is established on many dimensions such as predictability because it depends on performance stability over time, dependability and faith according to Rempel [34]. Thus, trust as a construct was used interchangeably with predictability, confidence and many other concepts which leads to confusion. The relationship between these concepts is deemed ambiguous according to Mayer [15].
In an old statement [35], Deutsch considered that trust and prediction are means of reducing uncertainty. While Deutch stressed on trust, Muir examined predictability as it leads to initial trust. He presented a framework that matches Rempel's theoretical dimensions [19].
The assessment of predictability in this framework is performed based on three elements. These elements are the actual predictability of the expected behavior, the ability of the trustor to assess the predictability and finally the environment stability.
Rempel's work was the initial proposition to understand and to model trust. Mayer [15] suggested a theoretical basis also in terms of ability, benevolence, and integrity in the context of organizational trust.
He evoked another misunderstanding about trust and confidence and stated that the relationship between trust and confidence is imprecise. Confidence was defined in Siergrist's work [36] as a belief that is based on facts, the expectations will occur as intended. Thus, confidence is different from the trust since it does not represent an attitude.
Trust is also confused with reliance for some scholars. Lee claimed That reliance is a behavior consequent to the attitude of trust[3]. Over-trust conduct to over-reliance and it is explained by a faulty calibration of trust, and under trust, which was called initially distrust, conduct to under-reliance [3].
In a recent survey, Gaudiello and fellows [37] have made an argument about the difference between trust and reliance. While trust is an attitude, reliance is a behavior resultant to trust. An attitude involves the mind's predisposition to certain aspects of life, while behavior relates to the actual expression of feelings into action.
The presented concepts are mostly used instead of trust by abuse of language. Some of them are a sort of beliefs or behaviors and others are factors affecting trust as an attitude directly. We sought to make a sort of a representation of these concepts and factors to unveil the relationship that leads to confusion and needs clarification.
We depicted in figure 1 all the concepts that were used interchangeably to describe trust. The ambiguous difference between these concepts led us to define each of them separately and connect it somehow to trust since confusing concepts means that a strong relationship already exists.
The first issue in the literature is related to what is trust exactly. Opinions vary from deeming trust as a belief, an attitude or behavior. The most common definitions agreed on considering trust as an attitude. Thus, all the derived aspects like over trust, under trust and distrust are kind of an attitude towards automation. Faith and confidence are considered as beliefs that represent the prerequisites to trust. Faith is the result of feelings influenced by the operator's traits as factors of impact on the propensity to trust. These factors affect directly the key aspects of faith such as predictability and dependability.
On the other hand, confidence also as a belief is more objective than faith, it is constructed over analytic from past experience and knowledge mainly which are influenced by automation and environment factors such reliability and culture. The confidence side calibrates trustworthiness towards automation, together with faith they define the attitude of trust.
One important point to clarify is that trustworthiness and propensity to trust influence each other, which explains the different levels of trust from one person to another and for the same person for certain automation to another. The goal of the coming sections is to show the impact of intelligence as an automation factor on trustworthiness and the propensity to trust. It is considered among the factors that could only influence positively to trust automation, otherwise, many factors cannot be influenced too but they are neutral in front of other factors.
Behavior is the result of an attitude. The aim of a trust is to push towards using automation. Thus, the use of automation may vary according to how much the operator trusts automation: use, misuse or abuse of automation which represents the behavior of reliance and its derivatives from over-reliance to under reliance.

2) Factors:

In the last decade, efforts have been concentrated on improving trust models. These models rely on trust as the key attitude and the concepts related to it in addition to factors influencing trust in automation. The factors are devised into three categories: factors related accordingly to the operator, to automation and environment[38].
The operator factors are classified into traits and states. Traits detain static attribution of values. They are the source of estimation about the propensity to trust, they can also affect one's beliefs. As for operator states, they affect directly to trust and change adequately to circumstances. These factors are also defined as characteristics and abilities [39].
Among traits factors that have a considerable impact within this classification are age, ethnicity, gender, and culture. The workload is a state factor that plays an interesting role in moderating between trust and reliance proposing to rely on autonomous systems when the workload is high even though trust levels are low. Biros [40] studied the influence of task load and automation trust on deception detection which highlights this factor.
Automation factors according to Lewis [41], in a recent review about the role of trust in human-robot interaction, are reliability and the system fault's effect. He introduced pre-

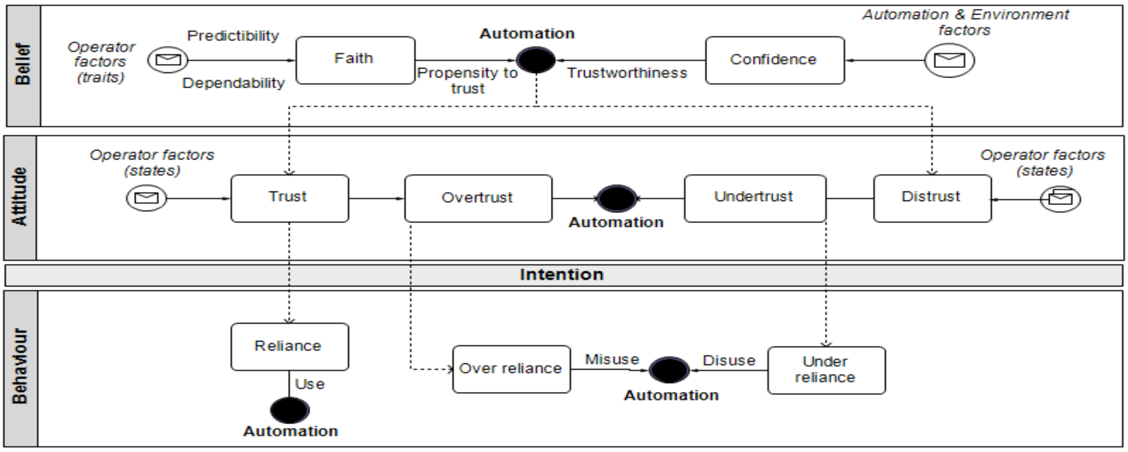

Figure. 1: Theoretical framework of trust

dictability as an important factor too as it informs about failures, thus, it reduces the level of uncertainty and risk which is considered as an essential trust request.

Some other key automation factors that influence users' trust upon intelligent systems such as predictability, control, and transparency were detailed in Holliday's work[42]. The effects of these concepts on trust may decrease it. For example, due to control, the system may change the predicted output without authorization from the user such that the predictability aspect is also affected.

The last category is specific to the environment factors. The effect of environmental characteristics on trust in automation is affecting automation use that indirectly affects trust. Risks related to uncertainties and task types along with task complexity are the key factors on which existing researches have been focusing [15].

In figure 2, we aimed to synthesize the main factors with a strong influence on trust in automation. Plus, we brought special interest on the direct impact of them on each other. The factor's wheel presents human factors divided into traits and states which helps to estimate the operator's propensity to trust without external influence. On the other hand, machine factors are split into capabilities and features to define the machine trustworthiness without an external influence neither. The environmental factors are also depicted as the outer factors that could influence the operator and the automation factors at the same time. All these factors influencing trust in automation can be devised into two parts, factor of internal influence which belong to the same division and factors of external influence that belong to different divisions.

Internal influence:

An operator is characterized by states and traits, states are a dynamic property like fatigue, stress, workload, etc. On the other side, traits are more or less static and they represent the identity of the operator like age, gender, personality, etc. These latter constitute the propensity to trust automation and each category of people might have their own propensity and each person has his individual difference of trust tendency.

Many studies were conducted to provide statistics about culture, age, and other traits to inform about their influence on the propensity to trust automation such as Chien's survey for how personality traits for US, Turkish, and Taiwanese participants make a difference in trust levels [43]. In our opinion, the propensity to trust score issued from surveys relatively based on human traits is influenced by human states, but no studies have been made on this basis to the best of our knowledge. We explain this internal influenced by the fact that stress can decrease the initial trust by influencing the propensity to trust a human being of a given age, gender, culture, personality, etc..

As for automation, any machine is presented by a set of features and a set of capabilities. Features are considered the identity of the machine, such as personality, level of automation and intelligence. Capabilities are the set of the behavioral aspects of the machine and they are influenced by key features, which means that any variation in features values would affect the machine capabilities.

As an example, we mention the LOA's role to predict the behavior and quantify reliability in an internal influence relationship. Machine trustworthiness is defined by machines' features and capabilities. The aggregated value of the capabilities influenced by the machine's features as explained before corresponds to the value of trustworthiness that affects trust in automation.

External influence:

The external influence is depicted in the relationship between the factor categories, operation, automation, and environment. There exists one side effect from environment factors towards automation and operator factors and both sides' effect for operator and automation. As described by the arrows in the above figure, operator factors define the propensity to trust and this variable is influenced by machine trustworthiness and vice versa.

To explain this fact, we recall that we might have different levels of trust from one person to another which means for the same trustworthiness value each propensity is the judge on adjusting trust in automation. Similarly, one person could have various trust levels for different machines and that is due to their trustworthiness effect on his propensity to trust.

Even though studying automation factors is out of our scope, we wanted through this section to shed light on the influence

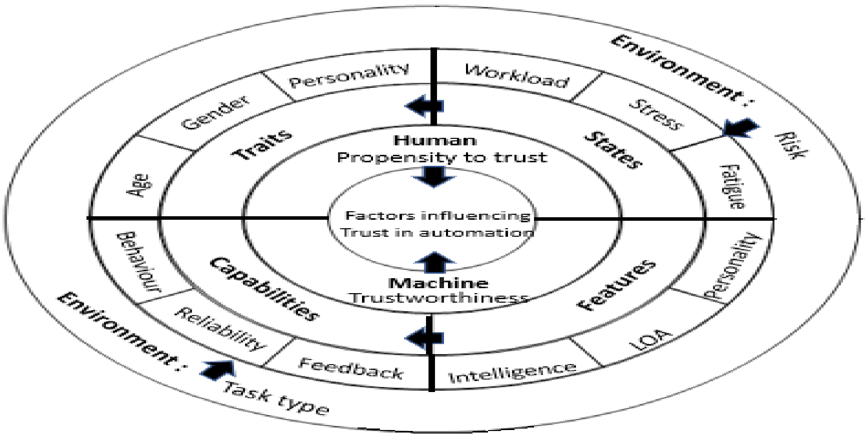

Figure. 2: Factors relationship to influence trust in automation

relationship between factors of automation because, to our knowledge, this has not been discussed before and it helps to introduce about intelligence's influence as a feature factor on trust in automation.

IV. Trust evaluation in automation

Authors in [44] accentuates the lack of clarity of defining and measuring trust, although the existing definitions, models, and measures of trust. The authors propose measuring reliance and compliance behaviors instead of using subjective trust measures. A new vision adopted by the authors[44] is to be based on a unified model for intelligent systems with four key abilities, namely, to acquire, master, create, and feedback knowledge. Intelligence measurement according to their belief relies on evaluating the cited abilities at the testing time.
Authors in [45] have developed a framework using an estimation method and a statistical model checking approach during a verification process to assign trust based on safety and security aspects.
Jensen and fellows [46] have dug into a cognitive lens by using emotions as a trust calibration factor. The reason for leveling trust is the consequence of an inappropriate calibration given the reliability of the system.
Holliday and all [42] reported an approach employing quantitative and qualitative measures to examine how to trust change over time and whether the reason is that the system offered explanations or not. In this research paper [10], we have found a new formalization of the explanation approach defined in other works.
The authors propose a driver interface that visualizes the car's interpretation of its current situation and adequate actions. To this end, the authors proposed different types of visualizations where the survey results show that the world in miniature visualization increased the trust the most compared to a chauffeur avatar or a display of the car's indicators.
The aim of the study of Garcia [47] is to predict user adoption of autonomous vehicles. The authors developed a scale to quantify trust related to the levels of vehicle autonomy by a factor analysis of the collected data.
In [48], an intelligent agent approach was implemented using human traits to increase trust since trust is often given to a human, other existing approaches adopt this analysis by implementing the aspect of anthropomorphism to shape appearances, while the proposed approach uses anthropomorphism to shape interaction.

A. Trust models

The implementation of trust models is considered a recent implication from researchers to objectively assess aesthetical concepts like trust. Based on the factors discussed above, several models were implemented as an updated presentation of the famous theoretical model of Muir [19] which has been used as a reference model of trust.
The choice of factors, with an impact on trust, depends on the goal of modeling. These models facilitate the development of measurements of trust or the implementation of trust in systems. In this work, we will introduce some of the most referenced models for trust in general automation and trust in specific automation. We present in the following table **??** some of them according to their date of appearance within the literature.

B. Trust measures

We should resolve some related issues that are blocking from giving trust entirely to the trustee and to quantify trust, Billings and fellows have suggested addressing the issues by defining what, how and when to measure to trust autonomous systems [54].
Morris and fellows [55] have focused on the possibility of trusting autonomous vehicles. Some of the main issues they detected about trust are related to safety, user trust, acceptance, ethics, and Legal issues and hacking appeal and communication. They conducted a survey to trace human confidence into an autonomous car by investigating driver behavior in a highly automated driving environment.
Based on surveys about adopting autonomous vehicles, authors admit that people are captivated by autonomous driving as an idea, but they are always hesitant to give control to a vehicle and lack of trust might be the main reason for their retention. Current measures of trust are split into three cat-

Table 1: Trust models across literature

Finding	Source
A general model within which trust according to them only affects behaviour as an attitude rather than all the given definitions of intention, belief or a simple behavior. Their model is a description of the dynamic process of trust and how it affects reliance on automation	Lee and See [3]
A model that aims to compare the human development trust into autonomous systems against the same trust development between humans. He shed light on the generalization from trust in human into trust in automation	Madhavan [33]
On the base of Lee's model, Hancock developed a model dividing trust factors into separate factors that are related accordingly to human, robots and environment categories	Hancock [38]
As a proposition of a specific type of automation, this model separates the characteristics of the trustor and the trustee which is an online system plus the integration of influences of effective and cognitive trust	Skarlatidou [49]
Chien also proposed a model that is defined from a set of cultural factors that are influencing trust in autonomous systems as a means for the development of a subjective psychometric measure for trust	Chien [50]
Hoffmann's model is built over the antecedents of performance, process and purpose which are key composites of trust formation. Hoff and Bashir's model specificity is that it could be applied to a range of automated systems, it is founded on the same categories defined in Hanckok's work.	Hoffmann [9][51]
Bindewald and fellows have developed a model that describes trust behavior and situations where the autonomous system acts on behalf of the human operator.	Bindewald [52]
Ekman proposed a framework for trust in vehicle systems composed of three sections each one is specific in a particular level from the trust cycle formation as a generic section to the driving event and the factor affecting trust upon these events in the second and third more detailed sections	Ekman [53]

egories: subjective measures like report scale and objective measures including the behavioral, physiological and neural measures. Self-report scales are the most used measure of trust but they are subject to problems according to [56] since questionnaires are not able to capture real-time trust changes in addition to the difficulty of conducting them out of an experimental context. They are using indicators for specific items to rate from low to a high level of trust. In 2000, Jian and his fellows [23] have developed a 7point scale as an odd scale to affect a neutral trust level. Their work has been used as a reference for many researchers, it was also recently discussed in details in a sort of examination by [57] and [58]. For the last decade, many scales have been developed to give more trust in automation. Some of them are general and culturally derived [50] [59] and some are related to a specific automation type such as [60] [47]. The physiological and neural measures are today's most investigated approaches for measuring trust by enhancing situational awareness [61]. Research works in this direction were focusing on observing gaze behavior through eye-tracking [56], measuring anticipatory stress by heartbeat rating [62] and examining brain activation [63] as obvious or measurable indicators to evaluate trust.

V. Intelligence's influence on trust: Uncertainty oriented evaluation

A. Definition of intelligence

Intelligence concept has been profoundly discussed for over 100 years and led out to several definitions that made researchers arguing about but still never well defined. All the existing definitions collected about intelligence were merely about learning and knowledge [64]. As follow we extracted general, psychological and AI researcher's definitions.

1. "The capacity to acquire and apply knowledge" The American Heritage Dictionary, fourth edition, 2000
2. "The ability to learn, understand and make judgments or have opinions that are based on reason" Cambridge Advance Learner's Dictionary, 2006
3. "..A person possesses intelligence insofar as he has learned, or can learn, to adjust himself to his environment." S. S. Colvin quoted in [65]:
4. ". . . that facet of mind underlying our capacity to think, to solve novel problems, to reason and to have knowledge of the world." M. Anderson [66]
5. ". . . the essential, domain-independent skills necessary for acquiring a wide range of domain-specific knowledge – the ability to learn anything. P. Voss [67].

With relevance to the collected postulations in [64], intelligence is an abstract notion detected while applying knowledge. Two approach types of knowledge are distinguished as rationalist and empiricist[68]. The rational approach proposes the distinction between the sensitive and the rational knowledge acquired respectively from senses and reason. The empiricist approach is related to experience.

Numerous viewpoints about intelligence were handled by scholars. Sternberg [69] believes that intelligence is an ambiguous concept that has been laid by psychologists at a first place to explain a certain process of learning from experiences and of adapting to every new situation to deal with it in a certain way. Intelligence, as commonly described by numerous authors is a combination of learning, in a formal or informal way, posing problems and finally solving them [70].

Warwick [71] defines intelligence as a set of information processing processes enabling an individual autonomous survival. The real nature of intelligence is associated with acting like a human to represent the weak AI or thinking like a human and steward the strong AI.

James Albus [72] confesses that intelligence is a controversial concept as it has not a commonly accepted definition. The provided definitions in his work are focusing on performance as the key measured feature for intelligence and they can be used for both biological and machine embodiment. Albus views the concept of intelligence as a multi-level concept with three main parameters such as hardware capacities (memory, power), data processing (world modeling, behavior generation) and information quality/quantity.

Meystel [73] has conducted his research with logic as he stated that stupidity and ignorance are not synonyms, such so intelligence does not substitute possessing knowledge. Thus, being informed cannot be admitted as a base to be judged intelligent. According to Meystel, intelligence depends on processing information and adapting to uncertainty, unlike claims relating it to knowledge. Wallace [74] believes that intelligence is a system internal property but not a behavior that can be determined by tests.

Intelligence is an intriguing topic across many fields. Biology led to explain that intelligence is the adaptation capacity to environmental condition and it is setting the organism's balance. In other terms it is the behavioral strategy that permits to maximize success in uncertain environment[72].

From the robotic point of view, intelligence is the ability to act autonomously in uncertain conditions. Albus [72] stated the viewpoint of control theory from the machine embodiment side to define intelligence as a knowledgeable "helmsman of behavior" and a result of integrating knowledge and feedback into a control system. Intelligent control is usually implemented to build intelligent machines. Also, it is linked to the machine intelligence concept since it is observed in the reasoning process performed by the machine [75]. Hence, adopting intelligent control definition helps to clarify machine intelligence purpose.

The presence of intelligence is defined by the presence of a control relation. Furthermore, good control is interpreted as a negative feedback loop. Explicitly, unexpected situations are counteracted by compensating actions to make the measured state as close to the desired state [76]. Thus, an intelligent control system is supposed to handle the sensor information about its own state and the state of the environment to make strong reasoning under unexpected situations.

As a synthetic definition, intelligence is considered as a control tool resulted from the evaluation consequent from the rewarding system after mission success under uncertainty.

B. Human-like intelligence impact on trust

One of the common reasons that human lacks trust in autonomous systems is the technological challenges like uncertainty preventing advances in terms of decision making. Thus, the key to assure autonomous systems's adoption henceforth is to address the trust issue related to uncertainty. Intelligence is confirmed when uncertainties are defeated. The thrust to measure autonomous systems' intelligence has been enticing to gain users' trust, so, quantifying trust under uncertainties is tightly related to measuring intelligence.

Another definition of trust in automation can be recalled to shed light on the importance of intelligence since this latter intervenes when uncertainty occurs, is: "trust is seen as a feature of entities that can be calculated as the probability of reliable behavior in the presence of externally induced uncertainty" [12].

We believe that this definition of trust in automation is interesting since it induces a new vision by involving uncertainty which is a key feature of intelligence. Thus, we think that intelligence as a concept resumes the other fields' trust definitions. Measuring intelligence by rendering uncertainty, expectations, experience variables, to concrete measures will inform about trust in a very accurate way.

Behavioral aspects drawn from human intelligence are required for automated systems to be trusted. We believe also that human-like intelligence is more important than human-like appearance. As explained, intelligence is the ability to reflect a logical understanding then show an accepted behavior. The point is that although the authors [48] have highlighted the importance of the human-like intelligence to increase trust in machines, the approach they have adopted does not imply specific features of intelligence into consideration, we believe that our approach has the same aim, that is to emphasize on the role of anthropomorphism to increase trust, but it introduces key variables to keep it effective. In [77], authors have thoroughly surveyed the existing literature of trust in autonomous systems. The authors examined four categories of trust for autonomous systems; trust in humans-robots interaction, human-machine interaction, self-driving cars, and autopilot systems. The literature review opens up about approaches dedicated to manage trust or to improve it for autonomous systems. Least of attention was given to trust on autopilot systems since the authors have only targeted a few technical works. We assume that few attempts on assessing trust in autopilots were taken since the survey is the newest in this research field and we took advantage of this situation to propose a new approach based on the intelligence of autonomous systems.

According to [44], Sheridan has pointed through his works that trust between humans and automation is similar to trust between humans and humans, thus, gaining trust could be set on a human basis and improved by retrieving important aspects of human involved in maintaining these relationships.

C. Uncertainty oriented evaluation for intelligence

A major defect of considering machine's intelligence as a result of a successful task or successful ability and trust it is that it does not take into account that it might be a chance driven result or by dint of redundancy and that is not intelligence. Thus, one of the main issues to our opinion is that scientists neglect the most important keyword in the definition of intelligence, namely uncertainties. Experts have overlooked machine intelligence evaluation when it is enduring uncertainties. Hence, we have noticed that they neglected an area where intelligence stands on autonomy and performance, that refers accordingly to abilities and tasks, to overcome disturbances.

The most intuitive representation of how uncertainty intervenes to confuse intelligence is the information loss game. This game returns the analytical intelligence capacity of the machine with obscure input due to undergoing uncertainties which correspond to the entropy theory process used already

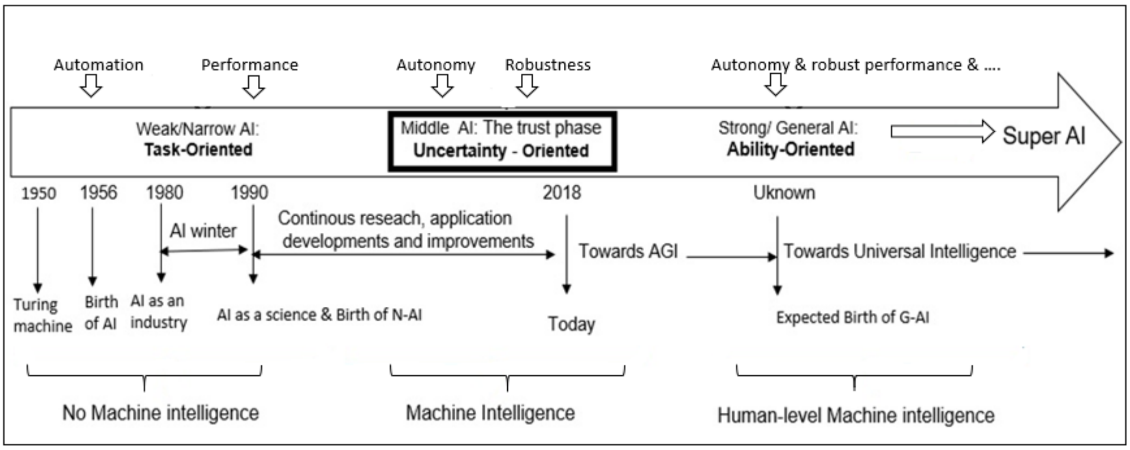

Figure. 3: AI timeline: An uncertainty oriented evaluation for the middle AI

as a measurement approach[78].

We believe that today's inventions do neither belong to the narrow AI nor to the general AI, they are more than simple task executors and less than a human being in terms of abilities.Thus, our main goal is to add a new evaluation that fills the middle AI.

There is still also considerable controversy surrounding splitting task oriented from ability oriented evaluation methods for nowadays' machines because they lead help from each other to exhibit a more apparent intelligent behavior. For example, Turing tests were attributed mainly to task-oriented evaluation. They also took a significant place within ability oriented evaluation by dint of the anthropocentric side that they hold[79].

We have raised disagreement with regard using these evaluations types interchangeably in an ambiguous way. In light of these concerns, there is now considerable interest in designing a new evaluation type for the middle AI current inventions.

Concerns call into question the validity of our proposition but since we feel supported by Iantovics [80] for the encouragement of presenting new measurements methods for intelligence, by daugherty [81] for the possibility of defining an intermediate artificial intelligence type, we took the assumption of the need of a new evaluation type for the middle AI. However, there still a need for scientific support evaluating machines on the basis of uncertainties in extension to performing tasks autonomously using task and ability.
As endorsed by Hernandez-Orallo in [79] "There is a hierarchical continuum between task-oriented evaluation and ability-oriented evaluation." It is permitted to define intermediate evaluation relying basically on tasks and evolving to introduce ability throughout the way but to not be mistaken with considering specific purpose systems or general purpose systems are indicators of AI progress. The main objective from the present paper is to reveal the gap that was caused by artificial intelligence evaluation milestones depicting the AI progress within AI lifetime.

Several endeavors are destined to organize artificial intelligence evaluations. In [82] Goertzel categorizes measures and tests as either helping for evaluating the achievement of human-level AGI like Turing test or the partial progress toward human-level AGI. The study conducted in [79] falls back to the postulate of the human reference to classify systems designed for help from those for human replacement. We realize that AI is doing much more than helping human due to integrated autonomy and robust performance. Nevertheless, it is not yet reaching the phase of replacement because the ability of AI is not yet developed to reach the human ability level.

The questions we raise are: to what evaluation kind is the middle AI oriented? Are today's inventions evaluated on the basis of tasks or abilities or maybe both? If so, a new evaluation type seems interesting. We depicted in figure 3 the middle AI trust phase and the proposed uncertainty oriented evaluation that rests on basic tasks' presence along with primordial abilities such as the decisional process which relies on advanced learning techniques to deal correctly with uncertainties.

The towards human-level machine intelligence involves a bunch of concepts according to the machine situation. A task-oriented evaluation can be helpful when the task is defined and the result is granted. When we deal with adaptive systems, uncertainty is unavoidable due to the changing context, which cannot be handled with the mentioned evaluation.

VI. Conclusion

When man safety is called into question, autonomous systems are unable to grant sufficient trust to be used similarly as any other type of automation. Anthropocentric approaches have shown the successful impact of increasing trust levels. These statements inform about intelligence importance as a human property to settle trust into automation and as the only

factor dealing properly with uncertainties to face situation randomness.
Statistics revealed a significant tendency to trust advances in technology but this tendency is limited when the advances are related to autonomous systems. We explain this retention with the study of the influencing factors which rely on specific factors more than others such as on performance more than autonomy, or reliability more than dependability, etc. Regarding autonomous systems, intelligence is a global factor that results from a set of situational conditions gathering automation, operator and environment factors. Thus, we sought to quantify trust based on intelligence as the influencing factor that impacts the human tendency to trust automation.

References

[1] 2019 deloitte global automotive consumer study, accessed: 2019-12-06.

[2] R. Parasuraman, V. Riley, Humans and automation: Use, misuse, disuse, abuse, Human factors 39 (2) (1997) 230–253.

[3] J. D. Lee, K. A. See, Trust in automation: Designing for appropriate reliance, Human factors 46 (1) (2004) 50–80.

[4] T. B. Sheridan, Humans and automation: System design and research issues., Human Factors and Ergonomics Society, 2002.

[5] J. M. Bradshaw, R. R. Hoffman, D. D. Woods, M. Johnson, The seven deadly myths of"" autonomous systems"", IEEE Intelligent Systems 28 (3) (2013) 54–61.

[6] P. A. Hancock, Imposing limits on autonomous systems, Ergonomics 60 (2) (2017) 284–291.

[7] J. D. Lee, N. Moray, Trust, self-confidence, and operators' adaptation to automation, International journal of human-computer studies 40 (1) (1994) 153–184.

[8] J. Lee, N. Moray, Trust, control strategies and allocation of function in human-machine systems, Ergonomics 35 (10) (1992) 1243–1270.

[9] K. A. Hoff, M. Bashir, Trust in automation: Integrating empirical evidence on factors that influence trust, Human factors 57 (3) (2015) 407–434.

[10] R. Häuslschmid, M. von Buelow, B. Pfleging, A. Butz, Supporting trust in autonomous driving, in: Proceedings of the 22nd international conference on intelligent user interfaces, ACM, 2017, pp. 319–329.

[11] K. Siau, W. Wang, Building trust in artificial intelligence, machine learning, and robotics, Cutter Business Technology Journal 31 (2) (2018) 47–53.

[12] P. Andras, L. Esterle, M. Guckert, T. A. Han, P. R. Lewis, K. Milanovic, T. Payne, C. Perret, J. Pitt, S. T. Powers, et al., Trusting intelligent machines: Deepening trust within socio-technical systems, IEEE Technology and Society Magazine 37 (4) (2018) 76–83.

[13] S. P. Marsh, Formalising trust as a computational concept (1994).

[14] J. B. Rotter, A new scale for the measurement of interpersonal trust 1, Journal of personality 35 (4) (1967) 651–665.

[15] R. C. Mayer, J. H. Davis, F. D. Schoorman, An integrative model of organizational trust, Academy of management review 20 (3) (1995) 709–734.

[16] M. Fishbein, I. Ajzen, Belief, attitude, intention, and behavior: An introduction to theory and research (1977).

[17] R. M. Kramer, Trust and distrust in organizations: Emerging perspectives, enduring questions, Annual review of psychology 50 (1) (1999) 569–598.

[18] C. Castelfranchi, R. Falcone, Trust is much more than subjective probability: Mental components and sources of trust, in: Proceedings of the 33rd annual Hawaii international conference on system sciences, IEEE, 2000, pp. 10–pp.

[19] B. M. Muir, Trust in automation: Part i. theoretical issues in the study of trust and human intervention in automated systems, Ergonomics 37 (11) (1994) 1905–1922.

[20] R. J. Lewicki, E. C. Tomlinson, N. Gillespie, Models of interpersonal trust development: Theoretical approaches, empirical evidence, and future directions, Journal of management 32 (6) (2006) 991–1022.

[21] B. M. Muir, Trust between humans and machines, and the design of decision aids, International journal of man-machine studies 27 (5-6) (1987) 527–539.

[22] J. B. Rotter, Generalized expectancies for interpersonal trust., American psychologist 26 (5) (1971) 443.

[23] J.-Y. Jian, A. M. Bisantz, C. G. Drury, Foundations for an empirically determined scale of trust in automated systems, International Journal of Cognitive Ergonomics 4 (1) (2000) 53–71.

[24] F. D. Schoorman, R. C. Mayer, J. H. Davis, An integrative model of organizational trust: Past, present, and future (2007).

[25] P. Slovic, Perceived risk, trust, and democracy, Risk analysis 13 (6) (1993) 675–682.

[26] D. J. McAllister, Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations, Academy of management journal 38 (1) (1995) 24–59.

[27] A. C. Costa, C. A. Fulmer, N. R. Anderson, Trust in work teams: An integrative review, multilevel model, and future directions, Journal of Organizational Behavior 39 (2) (2018) 169–184.

[28] T. B. Eikeland, T. Saevi, Beyond rational order: Shifting the meaning of trust in organizational research, Human Studies 40 (4) (2017) 603–636.

[29] T. B. Sheridan, Considerations in modeling the human supervisory controller, IFAC Proceedings Volumes 8 (1) (1975) 223–228.

[30] T. B. Sheridan, W. R. Ferrell, Man-machine systems; Information, control, and decision models of human performance., the MIT press, 1974.

[31] C. Nass, Y. Moon, P. Carney, Are people polite to computers? responses to computer-based interviewing systems 1, Journal of Applied Social Psychology 29 (5) (1999) 1093–1109.

[32] S. Lewandowsky, M. Mundy, G. Tan, The dynamics of trust: Comparing humans to automation., Journal of Experimental Psychology: Applied 6 (2) (2000) 104.

[33] P. Madhavan, D. A. Wiegmann, Similarities and differences between human–human and human–automation trust: an integrative review, Theoretical Issues in Ergonomics Science 8 (4) (2007) 277–301.

[34] J. K. Rempel, J. G. Holmes, M. P. Zanna, Trust in close relationships., Journal of personality and social psychology 49 (1) (1985) 95.

[35] M. Deutsch, Trust and suspicion, Journal of conflict resolution 2 (4) (1958) 265–279.

[36] M. Siegrist, T. C. Earle, H. Gutscher, Test of a trust and confidence model in the applied context of electromagnetic field (emf) risks, Risk Analysis: An International Journal 23 (4) (2003) 705–716.

[37] I. Gaudiello, E. Zibetti, S. Lefort, M. Chetouani, S. Ivaldi, Trust as indicator of robot functional and social acceptance. an experimental study on user conformation to icub answers, Computers in Human Behavior 61 (2016) 633–655.

[38] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. Chen, E. J. De Visser, R. Parasuraman, A meta-analysis of factors affecting trust in human-robot interaction, Human factors 53 (5) (2011) 517–527.

[39] K. E. Schaefer, J. Y. Chen, J. L. Szalma, P. A. Hancock, A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems, Human factors 58 (3) (2016) 377–400.

[40] D. P. Biros, M. Daly, G. Gunsch, The influence of task load and automation trust on deception detection, Group Decision and Negotiation 13 (2) (2004) 173–189.

[41] M. Lewis, K. Sycara, P. Walker, The role of trust in human-robot interaction, in: Foundations of trusted autonomy, Springer, Cham, 2018, pp. 135–159.

[42] D. Holliday, S. Wilson, S. Stumpf, User trust in intelligent systems: A journey over time, in: Proceedings of the 21st International Conference on Intelligent User Interfaces, ACM, 2016, pp. 164–168.

[43] S.-Y. Chien, K. Sycara, J.-S. Liu, A. Kumru, Relation between trust attitudes toward automation, hofstede's cultural dimensions, and big five personality traits, in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 60, SAGE Publications Sage CA: Los Angeles, CA, 2016, pp. 841–845.

[44] B. French, A. Duenser, A. Heathcote, Trust in automation—a literature review (2018).

[45] R. G. Dutta, X. Guo, Y. Jin, Quantifying trust in autonomous system under uncertainties, in: 2016 29th IEEE International System-on-Chip Conference (SOCC), IEEE, 2016, pp. 362–367.

[46] T. Jensen, M. M. H. Khan, Y. Albayram, M. A. A. Fahim, R. Buck, E. Coman, Anticipated emotions in initial trust evaluations of a drone system based on performance and process information, International Journal of Human–Computer Interaction (2019) 1–10.

[47] D. Garcia, C. Kreutzer, K. Badillo-Urquiola, M. Mouloua, Measuring trust of autonomous vehicles: a development and validation study, in: International Conference on Human-Computer Interaction, Springer, 2015, pp. 610–615.

[48] P. Ruijten, J. Terken, S. Chandramouli, Enhancing trust in autonomous vehicles through intelligent user interfaces that mimic human behavior, Multimodal Technologies and Interaction 2 (4) (2018) 62.

[49] A. Skarlatidou, T. Cheng, M. Haklay, Guidelines for trust interface design for public engagement web gis, International Journal of Geographical Information Science 27 (8) (2013) 1668–1687.

[50] S.-Y. Chien, M. Lewis, Z. Semnani-Azad, K. Sycara, An empirical model of cultural factors on trust in automation, in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 58, SAGE Publications Sage CA: Los Angeles, CA, 2014, pp. 859–863.

[51] H. Hoffmann, M. S"ollner, Incorporating behavioral trust theory into system development for ubiquitous applications, Personal and ubiquitous computing 18 (1) (2014) 117–128.

[52] J. M. Bindewald, C. F. Rusnock, M. E. Miller, Measuring human trust behavior in human-machine teams, in: International Conference on Applied Human Factors and Ergonomics, Springer, 2017, pp. 47–58.

[53] F. Ekman, M. Johansson, J. Sochor, Creating appropriate trust in automated vehicle systems: A framework for hmi design, IEEE Transactions on Human-Machine Systems 48 (1) (2017) 95–101.

[54] D. R. Billings, K. E. Schaefer, J. Y. Chen, P. A. Hancock, Human-robot interaction: developing trust in robots, in: Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction, ACM, 2012, pp. 109–110.

[55] D. R. Large, G. Burnett, A. Morris, A. Muthumani, R. Matthias, A longitudinal simulator study to explore drivers' behaviour during highly-automated driving, in: International Conference on Applied Human Factors and Ergonomics, Springer, 2017, pp. 583–594.

[56] F. Walker, W. Verwey, M. Martens, Gaze behaviour as a measure of trust in automated vehicles, in: Proceedings of the the 6th Humanist Conference (June 2018), 2018.

[57] R. S. Gutzwiller, E. K. Chiou, S. D. Craig, C. M. Lewis, G. J. Lematta, C.-P. Hsiung, Positive bias in the 'trust in automated systems survey'? an examination of the jian et al.(2000) scale, in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 63, SAGE Publications Sage CA: Los Angeles, CA, 2019, pp. 217–221.

[58] L. Buckley, S.-A. Kaye, A. K. Pradhan, Psychosocial factors associated with intended use of automated vehicles: A simulated driving study, Accident Analysis & Prevention 115 (2018) 202–208.

[59] S.-Y. Chien, Z. Semnani-Azad, M. Lewis, K. Sycara, Towards the development of an inter-cultural scale to measure trust in automation, in: International conference on cross-cultural design, Springer, 2014, pp. 35–46.

[60] R. E. Yagoda, D. J. Gillan, You want me to trust a robot? the development of a human–robot interaction trust scale, International Journal of Social Robotics 4 (3) (2012) 235–248.

[61] L. Petersen, L. Robert, J. Yang, D. Tilbury, Situational awareness, driver's trust in automated driving systems and secondary task performance, SAE International Journal of Connected and Autonomous Vehicles, Forthcoming (2019).

[62] B. Reimer, B. Mehler, J. F. Coughlin, Reductions in self-reported stress and anticipatory heart rate with the use of a semi-automated parallel parking system, Applied ergonomics 52 (2016) 120–127.

[63] K. Goodyear, R. Parasuraman, S. Chernyak, E. de Visser, P. Madhavan, G. Deshpande, F. Krueger, An fmri and effective connectivity study investigating miss errors during advice utilization from human and machine agents, Social neuroscience 12 (5) (2017) 570–581.

[64] S. Legg, M. Hutter, et al., A collection of definitions of intelligence, Frontiers in Artificial Intelligence and applications 157 (2007) 17.

[65] R. J. Sternberg, Components of human intelligence, Cognition 15 (1-3) (1983) 1–48.

[66] R. Anderson, Body intelligence scale: defining and measuring the intelligence of the body., The Humanistic Psychologist 34 (4) (2006) 357.

[67] P. Voss, Essentials of general intelligence: The direct path to artificial general intelligence, in: Artificial general intelligence, Springer, 2007, pp. 131–157.

[68] J.-M. Gallina, Les représentations mentales, Dunod, 2006.

[69] R. J. Sternberg, Handbook of intelligence, Cambridge University Press, 2000.

[70] D. G. Moursund, Project-based learning using information technology, International society for technology in education Eugene, OR, 1999.

[71] K. Warwick, S. J. Nasuto, Historical and current machine intelligence, IEEE instrumentation & measurement magazine 9 (6) (2006) 20–26.

[72] J. S. Albus, Features of intelligence required by unmanned ground vehicles, NIST SPECIAL PUBLICATION SP (2001) 243–252.

[73] A. Meystel, Evolution of intelligent systems architectures: What should be measured?, NIST SPECIAL PUBLICATION SP (2001) 361–382.

[74] S. Wallace, S. A. Wallace, J. E. Laird, Intelligence and behavioral boundaries (2002).

[75] E. Grant, G. K. Lee, Properties of learning knowledge-based controllers for robotic workcells and dynamic systems, in: Systems, Man, and Cybernetics, 2000 IEEE International Conference on, Vol. 2, IEEE, 2000, pp. 1216–1221.

[76] C. Joslyn, Towards measures of intelligence based on semiotic control, Tech. rep., Los Alamos National Lab., NM (US) (2000).

[77] S. Shahrdar, L. Menezes, M. Nojoumian, A survey on trust in autonomous systems, in: Science and Information Conference, Springer, 2018, pp. 368–386.

[78] G. Saridis, Definition and measurement of machine intelligence, NIST SPECIAL PUBLICATION SP (2001) 441–452.

[79] J. Hernández-Orallo, Evaluation in artificial intelligence: from task-oriented to ability-oriented measurement, Artificial Intelligence Review 48 (3) (2017) 397–447.

[80] L. B. Iantovics, A. Gligor, M. A. Niazi, A. I. Biro, S. M. Szilagyi, D. Tokody, Review of recent trends in measuring the computing systems intelligence, BRAIN. Broad Research in Artificial Intelligence and Neuroscience 9 (2) (2018) 77–94.

[81] P. R. Daugherty, H. J. Wilson, Human+ machine: reimagining work in the age of AI, Harvard Business Press, 2018.

[82] B. Goertzel, Artificial general intelligence: concept, state of the art, and future prospects, Journal of Artificial General Intelligence 5 (1) (2014) 1–48.

Author Biographies

Marwa Brichni obtained her bachelor's degree from the higher institute of management of tunis, University of tunis in 2013. She is currently a Ph. D. student under the supervision of Prof. Said Gattoufi. Her research is centered on evaluating and quantifying machine's intelligence using a human based measure for natural intelligence following the new paradigm for human level machine intelligence.

Said El Gattoufi got his bachelor's degree of Science from Faculty of Science de Tunis, Tunisia, 1985. His Master's degree of Science from Institute Superieur de Gestion, Tunisia, 1992. He is a Doctor of Philosophy in Operations Management from Sabanci University in Turkey ,2002. He heads the Department of Operations Management and Business Statistics in Muscat, Oman from 2007 to 2011 and joined Laboratoire SOIE, Institut Superieur de Gestion de Tunis, Tunisie right after and till these days as a professor of quantitative methods