Fusion of Information and Analytics: A Discussion on Potential Methods to Cope with Uncertainty in Complex Environments (Big Data and IoT)

Eloi Bosse, Basel Solaiman

Fusion of Information and Analytics: A Discussion on Potential Methods to Cope with Uncertainty in Complex Environments (Big Data and IoT)

Éloi Bossé*

Expertises Parafuse &
Electrical and Computer Engineering, McMaster University,
1280 Main St W, Hamilton, ON L8S 4L8, Canada
Email: ebosse861@gmail.com
*Corresponding author

Basel Solaiman

Image & Information Processing Department (iTi), IMT-Atlantique,
Technopôle Brest Iroise CS 83818, 29238 Brest Cedex France
Email : basel.solaiman@imt-atlantique.fr

Abstract: Information overload and complexity are core problems to most organizations of today. The advances in networking capabilities have created the conditions of complexity by enabling richer, real-time interactions between and among individuals, objects, systems and organizations. Fusion of Information and Analytics Technologies (FIAT) are key enablers for the design of current and future decision support systems to support prognosis, diagnosis, and prescriptive tasks in such complex environments. Hundreds of methods and technologies exist, and several books have been dedicated to either analytics or information fusion so far. However, very few have discussed the methodological aspects and the need of integrating frameworks for these techniques coming from multiple disciplines. This paper presents a discussion of potential integrating frameworks as well as the development of a computational model to evolve FIAT-based systems capable of meeting the challenges of complex environments such as in Big Data and Internet of Things (IoT).

Keywords: Information fusion; analytics; decision support; situation analysis; complex systems: Big Data: IoT.

Reference to this paper should be made as follows: Author. (xxxx) 'Title', *Int. J. xxxxxxxxxx xxxxxxxxxx*,

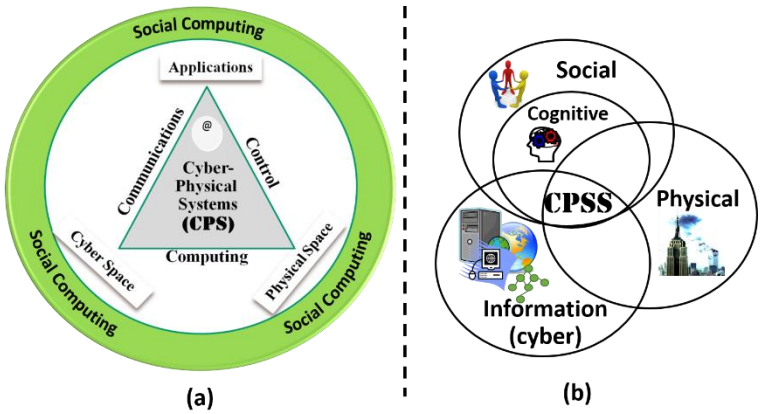
Biographical notes: Éloi Bossé, Ph.D**.** degree from Université Laval, QC, in Electrical Engineering. He worked at the Communications Research Centre, and Defence Research Establishments (DRDC) Ottawa & Valcartier, Canada, where he worked on radar signal processing, sensor fusion and decision support (1981-2011). He has published over 200 papers and technical reports. He retired from DRDC in Sept. 2011. Since then, he conducted some research activities under NATO Peace and Security Programme, as researcher at Polytechnic of Montreal, and at McMaster University, as associate researcher at IMT-Atlantique and since 2015, as president of Expertise Parafuse Inc., a consultant firm on Analytics and Information Fusion technologies.

Basel Solaiman received the Ph.D.(88) from Université de Rennes-I, France. In 1983 he joined the Laboratoire d'Electronique de Philips, Paris, where he worked on Image Compression standards. In 1988 he joined the Institute of Industrial Informatics, Brest, as a R&D Engineer in Chief. In 1992, he joined Telecom Bretagne, Brest, where he is actually a Full Professor and the Head of the Image and Information Processing department. He has published over 180 journal papers, several book chapters, over 230 conference proceedings and technical reports. Pr. Solaiman's research activities range from civilian into military ones: medical, remote sensing, underwater imaging and knowledge mining.

1 Introduction

All organizations have a tendency of generating more and more information that obviously challenges decision-makers and analysts with a deluge of potentially valuable data. The data deluge [1] is a result from a multitude of advanced sensors and sources capable of generating a diversity and a volume of data never envisioned before. Decision-makers and analysts cannot cope with that flow of data, without any impact on decisions quality and on actions efficiency. On the other hand, decision-makers can foresee value in that data deluge which can translate to technological opportunities, for instance, the Internet of Things [2-8] and the Big Data [9-12]. More and more, IoT and Big data are perceived as two sides of the same coin where Big Data would be a subset of IoT [13-15]. Big Data is evidently contextual to Cyber-Physical and Social Systems (CPSS) [16-20]. CPSS emerge from the interrelation of social, cognitive, information/cyber and physical worlds as pictured in figure 1. Social and cognitive dimensions interface with the physical world through the cyber world.

Figure 1 CPSS versus CPS (Source: [12])



Considering the interface between the information/cyber and physical worlds, one can imagine CPSS as a network of networks of humans and machines (Cyber-Physical Systems (CPS)). In CPS [21], computations are integrated with physical processes. This is either called Cyber-Physical Systems (CPS) or Internet of Things (IoT) depending on the community of origin (engineering or computer sciences) [22]. 'Embedded Systems' are considered as predecessors of CPS. CPS is not being discussed is this paper since it concerns more multisensory data fusion techniques and there exist excellent books [23] and surveys [24], [25], [26], [27] on that. Note that CPS and IoT do not necessarily imply complexity while CPSS can be considered as a typical complex system.

The discussion here concerns potential methods to benefit from Fusion of Information and Analytics Technologies (FIAT) to cope with complexity. CPSS demand multidisciplinary contributions from human and social sciences, physics, engineering and computer sciences to meet system overall challenges such as: 24/7 availability, 100% connectivity, predictability and repeatability, real-time, etc. The advances in Information and Communications Technologies (ICT) in particular smart ICT, to which FIAT belong, although providing a lot of benefits to improve dependability, efficiency and trustworthiness in systems, have also increased tremendously the networking capabilities so creating the conditions of complexity by enabling richer, **real-time interactions** between and among entities that compose CPSS. As a result, events that may once have had isolated consequences can now generate a cascade of consequences that affect badly system dependability and trustworthiness. The question is how to assemble a set of FIAT that would support decision makers in such complex systems?

The development of FIAT can facilitate decision-makers to get ways to communicate and understand complex insights and take efficient action. FIAT can be assembled into dynamic, real-time and near real-time, decision support systems (DSS) that are capable of supporting prognosis, diagnosis, and prescriptive tasks. Several applications involving data analytics and information fusion have already demonstrated that in current complicated systems. The question is then: what

theories, concepts, models and tools can help to achieve the same kind of objectives but within complex systems? Note that from a FIAT-based decision support system perspective, Big Data and IoT are problems to solve but from the decision-makers perspective, they represent opportunities and they are called Big Data and IoT technologies.

FIAT-based DSS meant to improve the quality of information (QoI) that in turn will improve the quality of decisions, which is directly linked with the capacity to ensure dependability and obtain trustworthiness in systems. By system dependability, we mean the main aspects described throughout the literature [28],[29]: reliability, safety, availability, maintainability and security. The objectives for the development of FIAT-based decision support tools (or a network of tools) could be illustrated as in figure 2 that represents our frame of thought throughout the whole discussion.

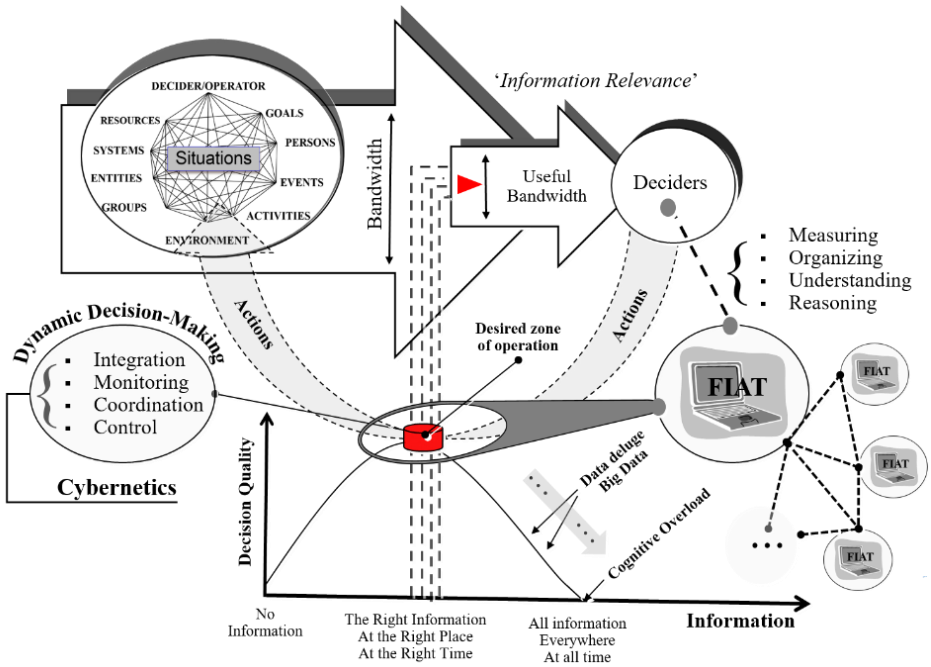
Situation Awareness (SAW) is somehow related to the quantity and the quality of information available to an individual say, a decision-maker: no information should result in poor SAW so very low decision quality as indicated on the lower left portion of figure 2. Providing "*all information, everywhere, at all time*" does not necessarily mean better SAW since too much information may exceed the human information processing capabilities, resulting in a cognitive overload (lower right portion of figure 2). Second, a portion of the data may be seen as noise to decision-maker because it is irrelevant. The decision-maker must detect and use only a relevant fraction of the information, called 'useful bandwidth' in figure 2. Such considerations bring the concept of "*the right information, at the right place, at the right time*", in turn leading to the notion of information relevance to provide actionable knowledge to deciders (desired zone of operation in figure 2).

FIAT are represented by the analytics and information fusion processes to support measuring, organizing, understanding, and reasoning with data-information-knowledge (DIK). Using an analogy brought by the signal processing community and pictured in figure 2, one can imagine a useful 'information' bandwidth where the overall goal would be to provide only useful information to deciders. That conceptual multidimensional '*useful bandwidth*' could be defined by assembling appropriate FIAT to lead to intelligent filtering and with metadata approaches [30]. In fact, in the presence of a deluge of data, the emphasis is upon classification and prioritization of information required to execute a given task [31]: "*As data become abundant, the main problem is no longer finding the information as such but laying one's hands on the relevant bits easily and quickly*."

Dependability has been addressed extensively in physical systems and computer engineering [29], [32], [33]. However, to ensure dependability in CPS [34] is very immature due to our poor understanding of the cyber-physical interface, mainly due to the consequences of the cyber-objects interactions. One can imagine the complications in CPSS with the addition of a social computing dimension. The dependability assessments of distributed physical systems have been particularly focused on the probabilistic modeling of random behaviors, given sufficient informative data. In future CPSS, traditional ways to provide DS for the four general cybernetic functions (integration, monitoring, coordination, control)

must be revisited, adapted or even more often reinvented. That reinvention requires an obliged passage through FIAT.

Figure 2 FIAT supporting deciders/operators to be in a desired zone of operation: (Adapted from Source [12] p.52)



This paper is in fact a transversal cut through the 8 chapters of the authors' recent book [12] published by Artech House entitled "*Fusion of Information and Analytics for Big Data and IoT".* The transversal cut provides an overview of potential methods that look promising for the design of FIAT-based decision support systems in complex dynamic environments as CPSS. Details of algorithms, technics, methods, examples and more complete literature surveys can be found in the book [12].

The paper is organized as follow. Section 2 describes the high-level characteristics of complex systems that challenge any FIAT design intervention. Section 3 discusses FIAT potential integrating frameworks. Section 4 introduces the essential elements toward the construct of a computational FIAT model. Section 5 terminates with the concluding remarks.

2. Characteristics of Cyber-Physical and Social Systems (CPSS)

Complexity is a consequence of interactions of elements within a system and between a system and its environment. Complexity is present in several natural systems but now, with the evolution of networking, human-made systems (e.g.

parallel and distributed computing systems, artificial intelligence systems, etc) exhibit complex behaviors. These behaviors emerge as a result of nonlinear spatio-temporal interactions among a large number of system components at different levels of the organization. Cyber-Physical and Social Systems (CPSS), or Complex Adaptive Systems or simply Complex Systems is the way they are called. A complex adaptive system (CAS) is being defined succinctly in [35] as "*a collection of semi-autonomous agents that are free to interact in unpredictable ways, and whose interactions generate system-wide patterns. Over time, those emergent patterns influence the subsequent behaviors of the participating agents.*" Table 1 below presents their main characteristics.

Table 1 – Main characteristics of complex systems.

Characteristics	**Description**
Emergence	Out of the interactions between the individual elements in the systems, behaviour emerges at the level of the system as a whole. - the overall behavior usually cannot be explained merely as the sum of individual parts.
Non-linearity	Non-linear dynamics: - may suddenly change behaviour or move to another regime. - relatively small changes may lead to large effects.
Limited predictability	Behaviour cannot be well predicted. Small changes in initial conditions can lead to very different dynamics over time.
Evolutionary dynamics	Systems are shaped by evolutionary dynamics: - selection of elements that are fit to cope with variations causing new variation. - cycles of variation-selection-multiplication-variation of elements. - elements in a system can change based on their interactions.
Self-organisation	Systems operate without central control: - often characterised by a certain order. - distributed control. - organise themselves from the bottom-up. - interrelationships between elements of the system produce coherence.
Fundamental Uncertainty	Complex systems are extremely hard to predict: - future states are fundamentally uncertain. - three pattern dynamics: organized, self-organized, non-organized - organized: close to certainty so easier to plan and control. - self-organized: hardly predictable so tension between stability and surprise. - non-organized: seek patterns to push towards self-organized.

What is the difference between a complex and a complicated system? A complicated system is the one where the number of parts, and their relationships are hidden from view. To understand such a system, a reductionist method can be used: the parts are separated from each other and the relationships clearly defined. A complicated system can be understood in terms of its parts. If the whole of the system is different from the sum of its parts, then it is complex. Each part is massively entwined with others, and the emergent pattern cannot be discerned from its components. Complicated and complex systems require different methods of analysis. With complicated systems, methods based upon repetition, replication, predictability, and functional decompositions are being used while in complex systems, we have methods based on pattern description, contextualization, and dynamic evolution. In CPSS, we are dealing with a spectrum of pattern dynamics from: organized, self-organized and non-organized. **Interactions** between elements of a complex system is the most important '*aspect*' to model the phenomenon since they partly determine the future states of the system [36, 37]. Agent-based technologies and network sciences are extremely important to model and simulate complex systems phenomena [38].

2.1 System of systems engineering principles: fundamental remarks

The realization of the complex interface between cyber and physical worlds is challenged by the concurrent nature and laws of physics governing our world, as opposed to the discrete and asynchronous nature of the cyber world. Adding the social dimension through the cyber world (by what is labeled social computing in figure 2) evolves CPS to CPSS sometime alternatively referred as 'complex networks', 'system-of-systems', 'network centric systems', 'socio-technical systems'. Some excerpts below reflect challenges encountered facing this complexity:

- from [39], "*CPS changes the notion of the physical systems (e.g., aircraft, vehicle) to include human, infrastructure, and platform in a system-of-systems, creating a uniquely large scope and context in which the system behavior must be predictable and provable. The resulting systems-of-systems are highly networked and dynamic in nature, with complexity, e.g., software size, growing at an exponential rate, with increasing time-critical interactions between purely physical elements and highly intangible cyber elements* (e.g. social computing)";
- from [40], "*System integration is the elephant in the China store of large-scale Cyber–Physical Systems (CPS) design. It would be hard to find any other technology that is more undervalued scientifically and at the same time has bigger impact on the presence and future of engineered systems. The unique challenges in CPS integration emerge from the heterogeneity of components and interactions. This heterogeneity drives the need for modeling and analyzing cross-domain interactions among physical and*

computational/networking domains and demands deep understanding of the effects of heterogeneous abstraction layers in the design flow."

System of systems engineering methodologies and Enterprise Architecture (EA) [41] are emerging approaches to deal with the development and maintenance of CPSS. Architectures help in managing complexity through techniques such as modularization and abstraction at various levels, and facilitate changes by providing a documentation of the components, relationships, and constraints of the systems or processes. It is worth considering, or perhaps, anticipating the general concepts or principles from which such complex systems might be constructed in the future. The following two general concepts and principles are emerging from literature:

1) the concept of composition, defined in figure 3 as composability and compositionality, that is very important to deal with complexity through system scalability, modularity, abstraction and to exploit homomorphism in constructing layered systems and services;

2) the principles of a transcendence/emergence pair: also defined in figure 3, where self- organization concepts such as the autonomic principles brought by IBM [42, 43] are guided by a transcendent concept 'task to be performed': the goal here is to realize software systems and applications that can manage themselves under high-level supervision from humans.

In CPSS, the combined scale, complexity, heterogeneous and dynamics of networks, systems and applications impose ever increasing constraints on information infrastructure that alternate paradigms and strategies need to be used to ensure system dependability and trustworthiness. Autonomic systems [44] are characterized by their self-properties: "self-configuration, self-healing, self-optimization and self-protection". The principles of systems transcendence and emergence impact greatly on the principle of composition for complex systems. Bernard-Weir [45] discusses the principle of transcendence and proposed points about the system's evolution that could be directed by "the concept of a task to be performed" (a sort of "blueprint") and stated that random causality and unpredictable unfolding may not be sufficient to explain such evolution. He states about the concepts of Self-Organization (SO), emergence, and the concepts of hetero organization and immergence. *"SO does not (self-) organize at random. … Self Organisation (SO) is a very important concept, but it is not sufficient in itself. Too much stress has been placed on the unpredictability and the novelty of the new states or levels that would emerge by virtue of SO."* Bernard-Weir [45] makes reference to Agonistic Antagonistic System Sciences (AASS) for the transcendence versus immanence pair that proved its efficiency in various domains from biomedical to social sciences. He presents a series of examples from the biological field where the recurrence of the same general mechanisms may be observed at every level of evolution (a kind of life model): "*actions of pairs such as stimulation/inhibition, gene expression/repression, trophotrope/ergotrope regulation, information/catalysis, phenotype/genotype …*".

Figure 3 General principles of complex systems engineering

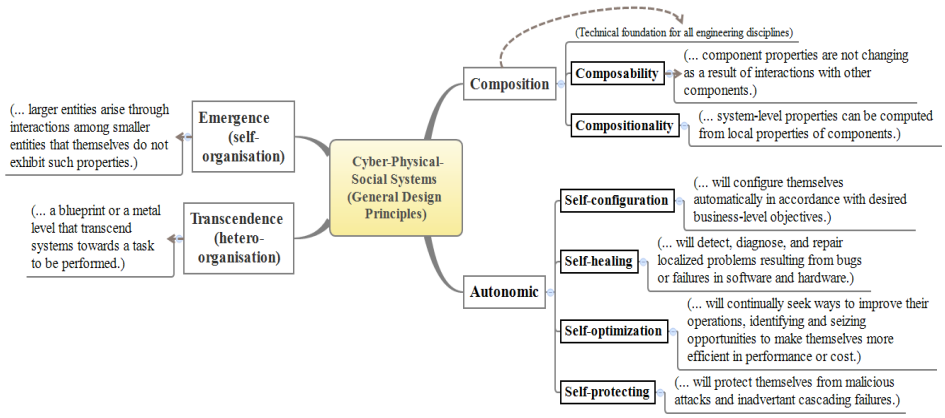
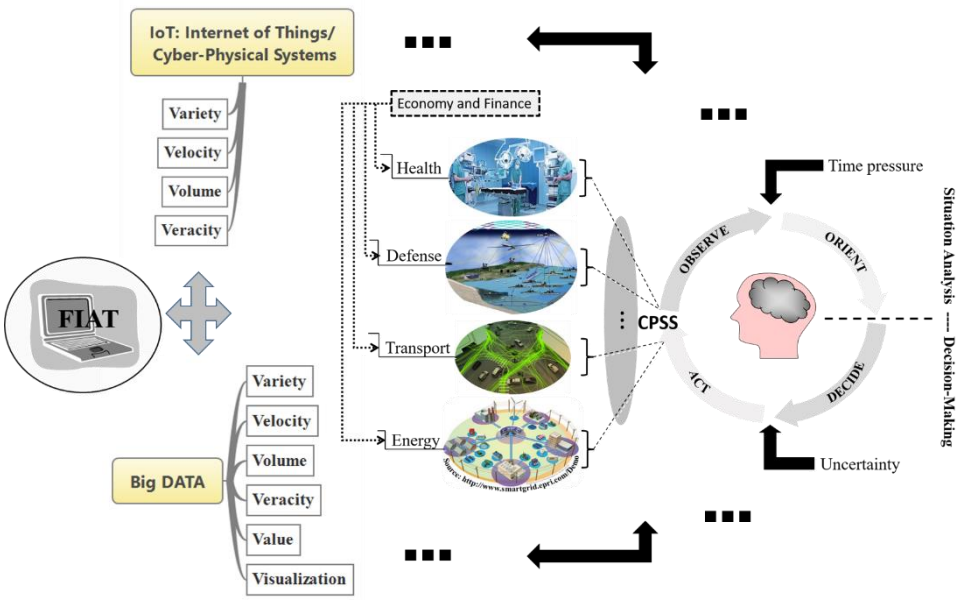


Figure 4 Cyber-Physical and Social Systems (CPSS): (Adapted from Source [12] p.20)



Highly relevant to this discussion on CPSS, Bernard-Weil [45] continues this way: " *…when socially complex organizations and firms have emerged, the pairs of the life model, or their corresponding properties, are found again. … According to this concept, a social system, for instance, would reach an optimal equilibration between ruling powers and developmental trends, if self-organizational processes were introduced in its functioning. This is, however, far from being certain. The*

specialists of management are quite right in claiming that strategies of emergence, though they might well contribute to generate new types of organization, will not necessarily bring forth solutions to the dysfunctions within the firms, if they are not combined with deliberate strategies."

2.2 Decision Support in complex systems: CPSS

Decision making is involved in all aspects of our lives, and it is of particular importance for the critical CPSS in: transport, defense and security, health, and energy. With the advancement of information and communications technologies, CPSS environments, in addition of being composed of networks of CPS or IoT, become more and more complex and amplify the Vs dimensions of Big Data: Volume, Veracity, Value, Velocity, Variety, and Visualization. That presents real challenges in terms of decision support. Understanding the process of decision-making is necessary for the design of decision support solutions (desired zone in figure 2).

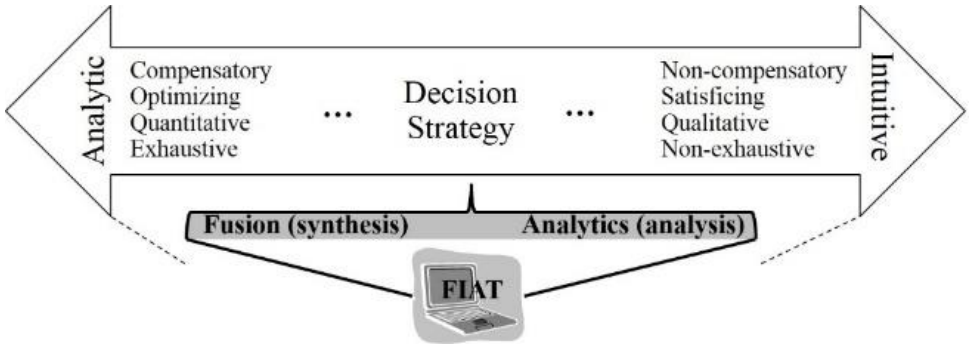
The oversimplified Boyd's Observe-Orient-Decide-Act (OODA) loop [46], illustrated in figure 4, is used to describe the decision process. Although the OODA loop might give the impression that activities are executed in a sequential way, in reality, the activities are concurrent and hierarchically structured. The processes of the loop are typically performed in a very dynamic and complex environment, and are heavily influenced by factors such as uncertainty, information and knowledge imperfections and time stress. The Boyd's OODA loop has been developed for human dynamic decision-making in military environment. It has however being used as well in civilian domains. The OODA applications include information fusion [47],[48], analytics and business intelligence [49],[50], autonomic systems [51], cultural modeling [52], and cyber security [53] to name a few.

The four CPSS environments in figure 4 present cases where interdependent decisions take place in a dynamic environment due to previous actions or decisions and events that are outside of the control of the decision-maker [54]. In addition to conventional one-time decisions, CPSS present dynamic decisions. The latter field is typically more complex than one-time decisions and much relevant to the environments of Big Data and IoT. In addition, dynamic decisions occur in real-time. In complex CPSS, if we wish to provide efficient computer-based decision support, understanding and framing the problem is the most important step. Over the years, multiple efforts have been deployed to better understand and explain the decision-making process in rather complicated environments. The case has now evolved to CPSS.

Among these models, two main influential streams [55],[56] are generally recognized to understand decision-making. The first stream refers to a rational approach that is based on formal analytic processes predicted by normative theories of probability and logic. The second stream, called naturalistic or intuitive theories, is based on informal procedures or heuristics to make decisions within the restrictions of available time, limited information, and limited cognitive processing. Bryant *et al.* [55] insist upon a continuum in decision strategy to adopt

the approach that is best tailored to the situation and may use elements of the two approaches at the same time as illustrated on figure 5. Ideally, it is that continuum of decision strategies that systems based on FIAT must support. Part I of [57] provides detailed descriptions and analysis of the various decision-making models.

Figure 5 A continuum of decision-making strategies: (Adapted from Source [12] p.27)



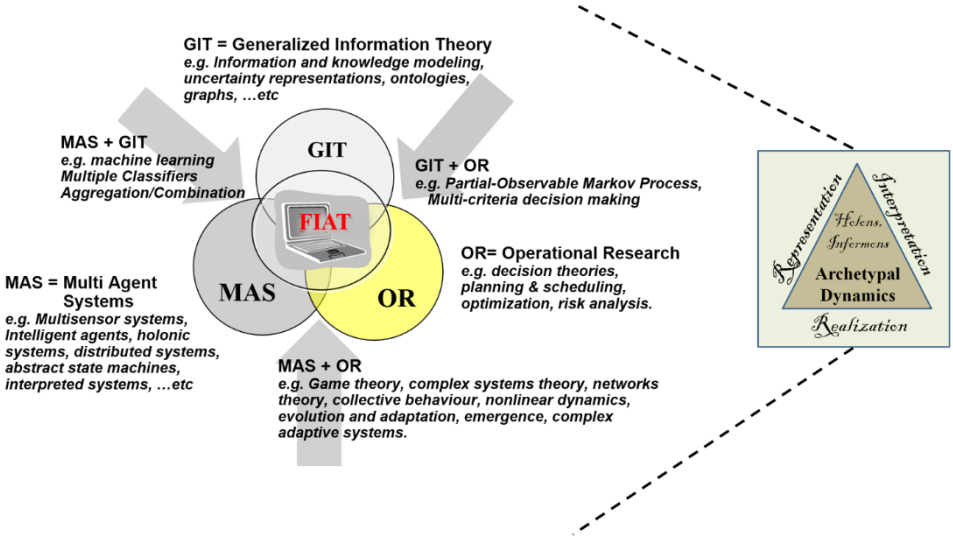
2.2.1 FIAT-based Decision Support

At this point, the question is what is FIAT? The answer is an assemblage of techniques and methods to analyze (analytics) and synthesize (fusion) information from multiple sources to support a decision cycle (e.g. OODA). In figure 6 an assemblage of FIAT is proposed around three main categories to be considered in a holistic fashion (through an integrating framework, e.g. archetypal dynamics) for the design of FIAT-based support systems [58]: 1) Multi-Agent Systems (MAS) theories to formalize the distributed system aspect and the notion of autonomy; 2) Generalized Information Theory (GIT) [59] for knowledge, information and uncertainty representation; and, 3) decision theories (represented by Operational Research (OR) in figure 6) in order to explicitly account for actuation (e.g. decisions-actions and their impact).

We assume here that OR also includes dynamic decision-making methods and techniques. The challenge is to assemble an appropriate set of techniques and methods that will support – measuring-organizing- reasoning-understanding-deciding and acting about/upon situations in complex environments such as CPSS in conditions of data overload and complexity. Note that the three categories of FIAT in figure 6 follow the line of thought of the archetypal dynamics triad (explained later): representation (GIT) – interpretation (OR) – realization (distributed systems - MAS). Taking individually, the techniques in figure 6 can only resolve aspects of the problems associated with complex systems. An assemblage of techniques and methods guided by a sort of transcendence principle, as discussed above, shows a better potential to address the multi-faceted problems

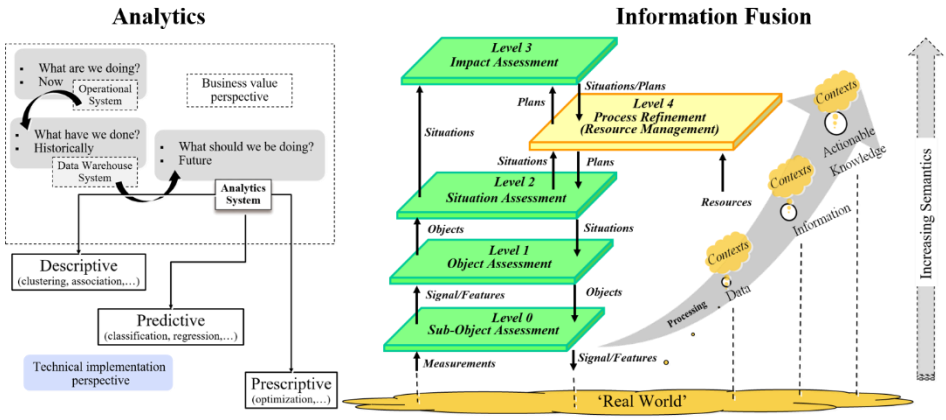
of complex systems. This assemblage is 'Analytics and Information Fusion': a processing chain that "*transform data into actionable knowledge*". Meaning transcends from actuation (contexts of actions). Semantics grows as the transformations (FIAT) progress data towards actionable knowledge. Figure 7 presents the most known data-information fusion model (right) associated with a high-level description of analytics (left).

Figure 6 Main categories of FIAT: (adapted from [12] p.53)



. The data fusion model maintained by the Joint Directors of Laboratories' Data and Information Fusion Group (JDL DIFG) is the most widely-used approach for categorizing data fusion-related functions [60]. The JDL distinction, among fusion 'levels' in figure 7, provides a valuable way of differentiating between data fusion processes that relate to the refinement (semantic levels and growth) of "objects, situations, threats, and processes". The fusion process also progresses through a hierarchical series of inferences at varying levels of abstraction (data-information-knowledge as shown in figure 7). The exploitation of contexts in the inferences processes aims at an increase in semantics that is only obtained by "*what one can do with the information?*" i.e. '*actionable*' knowledge. There are numerous books available that present more definitions, explain concepts in details, develop mathematical techniques and models related to figure 7 [57], [61], [23], [62], [49], [63].

Figure 7 Sheikh's proposal definition of analytics; and the JDL model: (adapted from Source [12] pp.44 et 53)



Eckerson [64] defines analytics as: "*everything involved in turning data into insights into action*." This is a quite broad definition that could include 'data and information fusion' of the preceding section but does not help understanding analytics from an applied point of view. Unlike information fusion community, the analytics community did not benefit from a well-structured organization like JDL to fix the terminology in order to ease communication amongst communities. The joint MIT Sloan and IBM Institute for Business Value report [65] define 'analytics' as "*the use of data and related insights developed through applied analytics disciplines (for example, statistical, contextual, quantitative, predictive, cognitive, and other models) to drive fact-based planning, decisions, execution, management, measurement and learning*."

Operational analytics is "the process of developing optimal or realistic recommendations for operational decisions (real-time and near-real time) based on insights derived through the application of statistical models and analysis against existing and/or simulated future data, and applying these recommendations in operational interactions." The taxonomy associated to Analytics is highly related to BigData. In a recent paper [66], important terms related to "big data" included: "*data integration*" and "*computational science*", "*clouds*", "*cloud computing*" "*Hadoop*", "*MapReduce*", "*big science" , "data sciences" "NoSQL," and "data warehouse", "cyberinfrastructure", "data mining", "data warehouse".*

The definition of analytics associated to Big Data becomes more and more confusing with various vendors, consultants, and trade publications defining and offering new technologies. As Sheikh points out in his book [67], analytics is one of the hot topics on today's technology landscape (also referred as Big Data). Analytics is not new and originates from business intelligence. It has been rejuvenated with Big Data. Figure 7 (left) illustrates Sheikh's proposal [67] to define analytics based upon a business and a technical implementation perspective. The business value perspective looks at data in motion as it is generated through

normal conduct of business. For this data, there are three variations of value: the **present,** the **past**, and the **future**, in the exact order as represented on figure 7. When data is created, referenced, modified, and deleted during the course of normal business activities, it lives in an **operational system**. The operational system at any given time can tell us what we are doing now. The data from day-to-day operations (e.g. selling merchandise, reviewing applications, etc.) is accumulated for record keeping and starts to build history in a **data warehouse.** Reporting can then be done to help understanding how a business did (e.g. total sales) in the last month, quarter, or year. These analysis reports provide managers the tools to understand the performance of their departments. This leads into the question that 'Analytics' should help to answer "*what should we be doing within departments and business units to improve business performance?"*

Any tools, technologies, or systems that help with this question can qualify to be in the analytics space. The technical implementation perspective described in Sheikh [67] is also the tangent adopted by Das [49] in his recent book on computational business analytics. Das' book describes the characteristics of analytics in terms of the techniques used to implement analytics solutions. They are listed in figure 7. In the analytics literature, three general types of analytics make consensus even though terminology may differ: descriptive, predictive and prescriptive. A lot of good reference books are available to describe these techniques particularly under data mining and machine learning fields [68].

In summary, both analytics (from business intelligence and operations) and information fusion (military intelligence and operations) share the same sets of techniques (figure 6). They could be assembled in a different way depending upon the contexts and domains of application. Both analytics and information fusion have the same kind of goals: to create situation awareness for decision makers. In practice, analytics has been more dedicated to collecting, organizing, structuring, storing and visualizing data (variety and volume) under mainly inductive and abductive reasoning processes (e.g. machine learning, data mining) while information fusion has been more on synthesizing information (veracity, velocity and value) under deductive reasoning processes (e.g. estimation and prediction, uncertainty management, impact assessment) both for operational decision support.

2.2.2 Situation Awareness, Situation Analysis, Complex Events and Situations

Situation AWareness (SAW) [69] is a concept around dynamic human decision-making (DM) in both military and civilian complex environments. Situation analysis (SA) is defined as the process that sustains a state of situation awareness for the decision maker(s). The SA process is the provision of decision quality information to the decision maker, thereby enabling timely situation awareness. A recent state of the art on this topic is provided in chapter 2 of [61]. Endsley [69] provides a theoretical model of SAW based on its role in dynamic human decision making. SAW is defined along three main processes: " *1) the perception of the elements in the environment, within a volume of time and space; 2) the*

comprehension of their meaning; and 3) the projection of their status in the future." Part I of [57] presents a detailed analysis of the Endsley's model and its relationship with respect to decision-making models and the cognitive demands of a human to perform a task.

Llinas [70] raises issues involving interdependencies among the situation analysis processes (e.g. analytics, fusion, sense-making) and decision-making processes for complex decision-making environments. He advocates that a critical examination of inter-process (situation analysis and decision making) interdependencies (see figure 4) is needed to design Decision Support Systems (DSS) for optimum performance. He recommends an integrated, multidisciplinary approach (cognitive sciences, human factors, decision sciences and computer sciences), otherwise DSS designs will remain disconnected and suboptimal. The interconnected processes identified by Llinas [70] are: 1) automated FIAT-based situation analysis process ; 2) sense-making and information foraging (for further reading [71],[72]); and, 3) decision-making, a semi-automated process that operates in an analytic, an intuitive or a "hybrid/mixed" decision-making mode to support actuation onto the real-world situation.

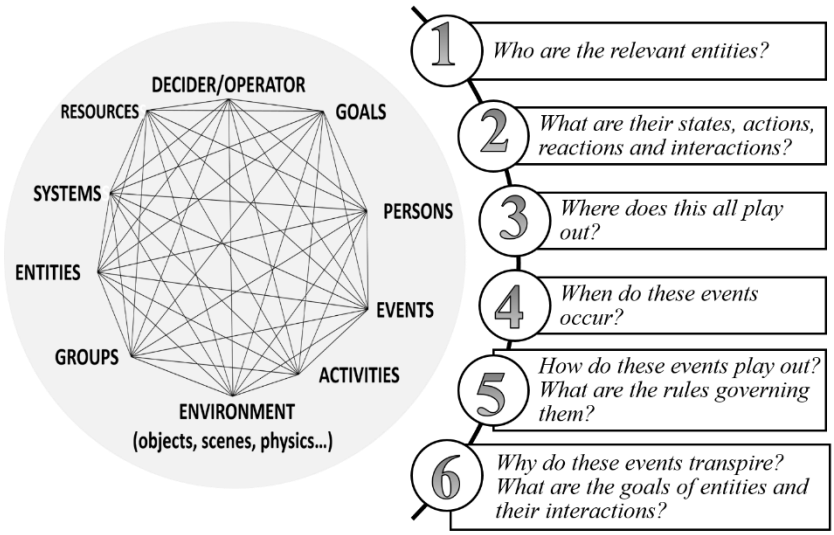
Roy [73] defines Situation Analysis (SA) as: "*a process, the examination of a situation, its elements, and their relations, to provide and maintain a product, i.e., a state of situation awareness, for the decision maker*"; and a situation as: "*A specific combination of circumstances, i.e., conditions, facts, or states of affairs, at a certain moment.*" The SA process is concerned with understanding the world. The situation can be defined in terms of events, entities, scenes, systems, people, etc., and their mutual interactions as pictured in figure 8. *Entities* and *events* are quite important elements of concern for SA.

A complex event is a composition of events that suggest more complicated circumstances. Examples of events may be: "sales leads, orders or customer service calls, news items, text messages, social media posts, stock market feeds, traffic reports, weather reports, or other kinds of data". Complex Event Processing (CEP) methods [74],[75],[76] track and process streams of information (data) from multiple sources about events, establish temporal and causal dependencies between them, infer patterns and derive conclusions (such as opportunities or threats) from all that. This is quite an important set of analytics methods and technologies, part of FIAT, that have applications for system health monitoring, anomalies and fraud detection, Internet of Things, and for situation analysis in general. CEP provides ways to analyze patterns in real-time and can be of prime importance in the composition of services for the enterprise.

Granular Computing [77, 78] is another set of analytics methods and technologies that has emerged in recent years. It concerns the representation, construction and processing of complex information entities called information granules. Information granules are the result of a data abstraction and derivation process of knowledge from data and information. They can be treated as linked collections of objects or entities that, from the numeric level are pinched together

by the criteria of indistinguishability, similarity, proximity or functionality, coherency, or the like.

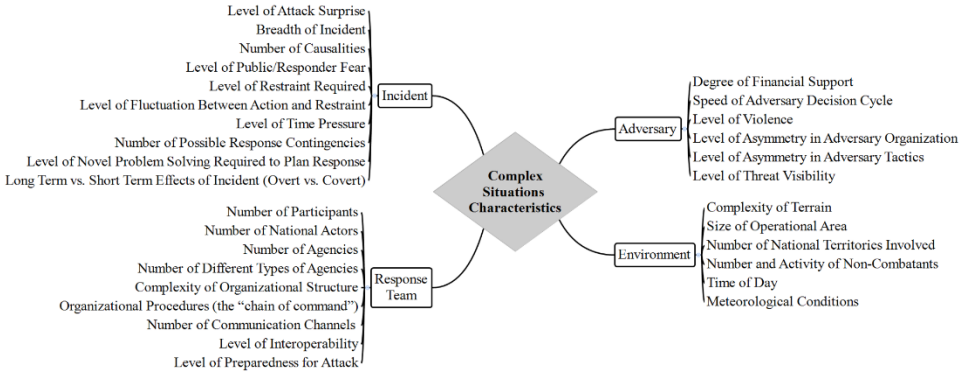
Figure 8 Six epistemic questions for situation assessment: (Source [12] p.32)



The above examples of methods (e.g. sense-making, foraging, granular computing) refer to FIAT-based support systems that can provides answers, in whole or in part, to the six basic epistemic questions excerpted from Sulis [79] that have been listed in figure 8. Phrased in a different way, similar questions appeared in Nicholson [80] with respect to situation awareness: "*what are the objects of interest? where are they? how are they moving? where have they been? and where might they be going?*" Objects could refer to either physical objects, such as vehicles, or symbolic objects such as terrorist events, system faults or business plans.

Finally, '*complex situations'* could be defined by considering the level of complexity of such situations. This is the approach reported in Guitouni [81] as part of an analysis conducted to support the Canadian Crisis and Emergency Management community with respect to complexity: "*compositions and interactions of variables in a situation that affect the cause and the effects, the sources and accuracy of information, the communication and decision making processes, and the activities/actions that are required to achieve a desired end-state.*" Guitouni [81] presented the key variables (figure 9) that have been identified as being influential for situations complexity. For instance, crisis and emergency response plans revealed a number of common variables in emergency situations resulting from terrorist activities and asymmetric threats. They can be grouped into four categories (figure 9) of response team, adversary, environment and incident. Further details on these four categories can be found in [81].

Figure 9 Framework listing the common variables characterizing complex situations



3. FIAT Integrating Frameworks

The literature is quite poor on integrating frameworks or computational models that would ease the design of FIAT-based support systems. FIAT supports a process to 'transform data into actionable knowledge' for sense making, to reduce uncertainty and, thus, to improve dependability in systems. The making-sense process is highly related to actions of any agent being a human or a machine. A goal-driven, or activity-based, approach sounds appropriate for that process. In a recent survey (2010) of information fusion (IF) frameworks, Llinas [82] concluded: "*While a number of efforts have been made to define a robust, extended-domain IF Framework, our review of past works indicates that no such Framework has been well-defined*." Llinas, in coordination with the Board of the International Society for Information Fusion (ISIF) created in 2011, the "Fusion Process Model and Frameworks Working Group (FPMFWG)" recognizing the insufficiency in the current level of advancement of the JDL model to support the design of IF systems for complex dynamic environments. Recently, we see a sort of convergence between the analytics and IF communities about hard-soft fusion [83], [84], [85]. The fusion of hard and soft information comprises analytics technologies.

The development of a framework in which knowledge, information and uncertainty can be represented, organized, and structured is a core requirement. Such integration FIAT framework should: 1) provide means to represent knowledge through well-defined notions of situation and awareness; 2) support the modeling of uncertainty, belief, and belief update; 3) provide the key 'computational model' for FIAT and linkage to actions (users and machines); 4)

provide practical support for system design through modularization, refinement, validation, and verification; 5) offer a good compromise between operational and functional modeling in capturing systems behaviour; 6) enable rapid prototyping and experimental validation of fairly abstract models; and finally, 7) support modeling of multi-agent systems.

A very limited number of information fusion frameworks have been proposed in the literature that offer partial fulfillment of the above mentioned requirements. A number of significant and powerful 'Analytics' tools and techniques have been offered to scientific and engineering communities for organizing, integrating and visualizing large volumes of data. One can think at: MapReduce [86], HADOOP [87], EXALEAD [88] to name a few. In the domain of information fusion alone, noticeable efforts have recently been dedicated to define an integration fusion framework: the State Transition Data Fusion (STDF) model introduced in [89], the Interpreted Systems for Situation Analysis (ISSA) in [90], Abstract State Machine (ASM) methods in [91] for high-level design and analysis of distributed systems, OODA-based agents in [92], the recombinant design approach in [93], fuzzy cognitive maps used for situation assessment [94] and recently, the holonic processing framework of [95],[96] based on Sulis' archetypal dynamics [97].

3.1 The Archetypal Dynamics Framework

Archetypal Dynamics (AD) is almost unknown to the information fusion and analytics communities. Sulis [79] introduced AD as: *"Archetypal dynamics is a formal framework for dealing with the study of meaning laden information flows within complex systems."* Let replicate here what has been written on AD in Solaiman et al. [95]: " *This is a formal framework for dealing with the study of the relationships between systems, frames and their representations and the flow of information among these different entities. The framework consists of a triad of: semantic frame (representation), realizations (system) and interpretation (agent/user). Real systems relate to semantic frames through one of the dimensions of that triad. The viewpoint of archetypal dynamics is that meaning is tight with actions. A semantic frame is an organizing principle that ascribes meaning in a coherent and consistent manner to phenomena that have been parsed into distinct entities, mode of being, modes of behaving, modes of acting and interacting."* The archetypal dynamics triad has been illustrated in the right portion of figure 6. The semantic frame (defined later in the next section) provides partial answers to the six basic epistemic questions presented in figure 8. Full answers would suppose that one know the ground truth. In archetypal dynamics, the way the information is understood is not in the sense of Shannon (i.e. the quantity of information), but in its active sense*: "Information possesses content and elicits meaning.*"

Sulis [98, 99] proposes the mathematical framework of tapestries which is a formal representational system. Tapestries represent information flow by means of multi layered, recursive, interlinked graphical structures that express both geometry (form or sign) and logic (semantics). Sulis [98] presents a detailed mathematical description of a specific tapestry model, the causal tapestry, selected

for use in describing behaving systems. Observables are represented by tapestry '*informons*' while subjective or hidden components (for example intellectual and emotional processes) are incorporated into a reality game that determines the tapestry dynamics. Details on causal tapestries can be found in Sulis' thesis [99], namely in Appendix C and D where the basic ideas of process theory, archetypal dynamics, causal tapestries and emergence are described. An implementation of a causal tapestry for a realistic analytics or information fusion problem is still required in view of comparison and valuation with respect to other frameworks below. A holonic computational model has been proposed in [95, 96] using the ideas of *informon* from archetypal dynamics and *holons* from complex systems theory to progress towards a potential FIAT computational model. Some of the elements are discussed in section 4.

3.2 The State Transition Data Fusion (STDF)

The State Transition Data Fusion, STDF, model introduced by Lambert [89] is a functional model aiming at an unification of the notions of "object, situation, and impact" (Levels 1-3) JDL fusion. In Solaiman et al. [95], STDF is briefly described as: "*It views the world at time k as a composition of states s(k). At time k, the world is understood in terms of the history of its state transitions up to time k. At time k + 1, different states of the world may be sensed by the sensors. This new data is fed into the observation process identifies the detections, normalizes the detections to a frame of reference, and then utilizes the prediction process to match the new observation to one or more previously predicted states of the world. An* ***object*** *instance of the world is represented at time t as a state vector u(t) of the measured values and the understanding of that object at time k is captured as a set of transitioning state vectors. A* ***situation*** *is represented at time t as a set of statements about the world (state of affairs) in some formal language* [100] *and it is understood at time k as a set of transitioning states of affairs. Finally,* ***impact*** *assessment in STDF presents an understanding of the world in terms of scenarios. A scenario instance at time t is expressed as a set of transitioning situations projected into the future. Scenario prediction in STDF involves assessment of intent, awareness, and capability of agents*." Much more details on STDF can be found in chapter 3 of [61] as well as in [89].Similar to causal tapestry, an implementation is due to compare with other frameworks.

3.3 The Interpreted Systems Situation Analysis (ISSA) Approach

Maupin and Jousselme [90] used the notion of Interpreted Systems from Fagin *et al.* [101] and they showed its potential in simple situations comprising few agents. Interpreted Systems is a formal semantic framework for reasoning about knowledge and uncertainty. Maupin et al. [90] proposed to use it as a general framework for situation analysis referred here as ISSA. In Solaiman et al. [95] ISSA is briefly described as: "*The ISSA approach provides a formal framework for reasoning about knowledge and uncertainty and for dealing with belief change*

concepts in a distributed systems context. Unlike STDF, ISSA comes with a computational underlying formal framework for modeling a distributed system. ISSA views distributed computations of concurrent and reactive systems as evolution of states, assuming multiple computational agents interacting with one another as well as with their operational environment representing the external world. The underlying computation model, in ISSA, defines the behavior of a distributed system as the set of all admissible runs originating in a distinguished set of initial system states." Details on ISSA can be found in chapters (4, 14) of [61]. However, the value of ISSA as an integrating framework has yet to be demonstrated since the application of ISSA has been limited so far to a few number of academic examples. It requires further work involving more real-life or realistic scenarios.

3.4 Abstract State Machines (ASM)

Abstract State Machine (ASM) method is a scientifically and industrially recognised systems engineering method for the design and analysis of complex systems. This method guides the development of embedded systems from requirements capture to their implementation. It supports the designer to cope with size, complexity and trustworthiness. This method deserves to be looked at for the design of FIAT-based systems to be embedded in CPSS since, from [102]: "*The method bridges the gap between the human understanding and formulation of real-world problems and the deployment of their algorithmic solutions by code-executing machines on changing platforms. It covers within a single conceptual framework both design and analysis, for procedural single-agent and for asynchronous multiple-agent distributed systems.*"

The ASM method for systems high-level design and analysis builds on the concept of abstract state machines [91, 102, 103] and brings together two tasks of requirement capture and system design. The goal is to improve industrial systems development by integrating high-level abstract modeling into the software development cycle down to executable code. The method presents three essential phases: a) a requirements capture and the development of an abstract operational model: *a ground model* , b) an incremental refinement of the ground model down to the implementation, and c) an experimental validation, by simulation or testing, of the models at each level of abstraction. This approach can be used to design FIAT-based intelligent support systems as demonstrated in [91]. For further details on ASM, see [102],[103].

3.5 Remarks on FIAT integrating frameworks?

None of the above frameworks have been demonstrated so far to be fully appropriate for the design of FIAT-based support systems facing realistic scenarios (i.e. presenting at least some characteristics of complex systems – Table 1). At the basis, the ASM is a universal computation model of distributed systems. The system view of STDF and ISSA, in terms of transitioning states and especially STDF in defining the structure of the states, fits very well with the notion of abstract state machines. The set of states in STDF can be mapped to the global

state of a distributed abstract state machine and the notion of state transition would be captured by the notion of computation steps in ASMs. On the top of this foundation, the information fusion concepts of object, situation, and scenario, present in STDF unified model, can thus be defined. Sensor signals will be modeled by monitored functions that receive input from the environment. The view of ISSA toward knowledge representation, uncertainty, and belief changes enables to model situation awareness of agents as part of their local states. Semantic values in STDF, i.e. the understanding of an *object*, a *situation*, or a *scenario*, are defined based on the whole history of transitioning states of the corresponding entities. In a distributed ASM at any given state, agents have access to the values of functions and terms only in the current state of the machine. However, with the freedom of abstraction in ASMs, one can introduce a notion of '*history of values*' which would expose, in the current state of the machine, the historical values of functions in the previous states of the machine.

A formal computational framework for the design of FIAT-based systems is unavoidable if one is interested in representing and reasoning about dynamic situations and producing solid system designs that can be experimentally validated and systematically verified. Such a framework should not only provide a basic model for FIAT and means to represents knowledge and uncertainty, but it should also offer practical support for systems engineering and experimental validation of models as ASM can provide.. In order to satisfy the set of requirements for such a formal framework, we propose, as of future work, to analyze in details the formal approaches, causal tapestries, STDF and ISSA, with the systems design and analysis approach of ASMs (language of functional programming). This analysis has been started in [91] and needs to be pursued considering the notions of archetypal dynamics [104] and the holonic approach of Solaiman et al. [95]. A strengths and weaknesses analysis of STDF, ISSA and ASM potential integrating frameworks for the design of FIAT-based systems can also be found in chapter 7 of [12].

4. Elements of a Computational FIAT model

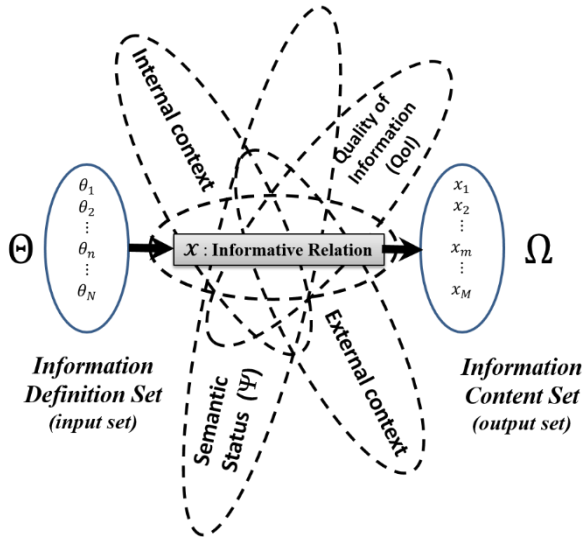
In an attempt to define a FIAT computational model also called information fusion framework, the authors [95, 96, 105], have proposed an approach based on a holonic functional processing. The framework is a goal-driven approach suitable for processing any semantic level of the JDL information hierarchical abstraction model (data-information-knowledge) with notions to take into account quality of information (QoI) [106] for managing the fusion process.

4.1 Definition of an Information Element

The quality of information produced by a fusion process is highly related to the definition of its basic components and the quality of its associated knowledge. As stated in [95]: "*Observing data from a given set is not enough to make it an informative entity. Information hence requires a content set, how its outcome is*

obtained and what it refers to." That leads to the following definition of an information element from [95] illustrated in figure 10:

Figure 10 A basic Information Element ($\mathcal{I}$) structure



Definition, an Information Element is: "*an entity composed of a definition set and a content set linked by a functional relationship called informative relation, associated with internal & external contexts*."

Therefore, as shown in figure 10 and described in [95], the main components of an information element are: " *1) a **Definition Set** giving the potential information input elements (what the information refers to); 2) a **Contents Set** encoding the possible knowledge produced by the information (e.g. measurements or estimations of physical parameters, decisions, hypothesis); 3) an **input-output relationship** representing the "functional link – informative relation" model (e.g. mathematical, physical) that associates the input elements with the produced information contents; 4) an **Internal Context** gathering intrinsic characteristics, constraints or controls about the informative relation itself; 5) an **External Context** containing data, information or knowledge useful to the elaboration of the meaning or the interpretation of the Information Element. The information element ($\mathcal{I}$), is denoted as: $\mathcal{I}$ = (Information definition set, Informative relation, Information content set, Internal Context, External Context).*"

Internal and external contexts, the basic general properties of an ($\mathcal{I}$) such as exhaustivity, exclusivity and incompleteness, along with the characterization of ($\mathcal{I}$) have all been defined in [12, 95]. An information element $\mathcal{I} = (\Theta, X, \Omega)$ is considered as "*fully characterized*" if the following features are known, besides the input and output sets Θ and Ω:

"1) Information *Input Scope* (ISc) indicating how the informative relation considers the elementary objects from the definition set (single-arity, plural-arity,

fuzzy-arity); 2) Information *Output Scope* (OSc) indicating how the informative relation considers the elementary objects from the contents set (single-arity, plural-arity, fuzzy-arity, semantic status Ψ); 3) *Imperfection Status* (*St*) indicating if the information element is considered as Perfect (precise and certain) or Imperfect (Uncertain, Imprecise or Ambiguous); 4) Imperfection available knowledge in terms of *Belief Model* (*BM*); 5) Information *reliability weight* (*R*); 6) Information *consistency degree* (*Cons*);and, a semantic status Ψ ."In this case, a fully characterized information element $\mathcal{I}$ is denoted as (figure 11):

$$\mathcal{I} = [(\Theta, X, \Omega), \mathrm{IS}_C,\ \mathrm{OS}_C, St, BM, R, Cons, \Psi\ ] \quad (1)$$

4.2 Notion of a semantic status Ψ

In the literature, the terms data, information and knowledge are not always clearly distinguished [107]. However, each refers to ideas and concepts that, after examination, will appear differentiable. A distinction, yet clearly apparent in current language when talking about data from a physical problem, information brochure on a commodity, while the term of knowledge remains vested in the mastering of a technic or holding an important function. The terms are distinct since they do not reflect the same reality, and they convey different levels of signified that we call later 'semantic status'. The advances in the development of techniques emphasize that distinction by dedicating to them separate roles and processing. Their conceptual and implementation approaches are different in the architecture of information systems. Finally, the literature has gradually addressed data fusion, information fusion and recently, more and more papers appear on knowledge fusion [108],[109].

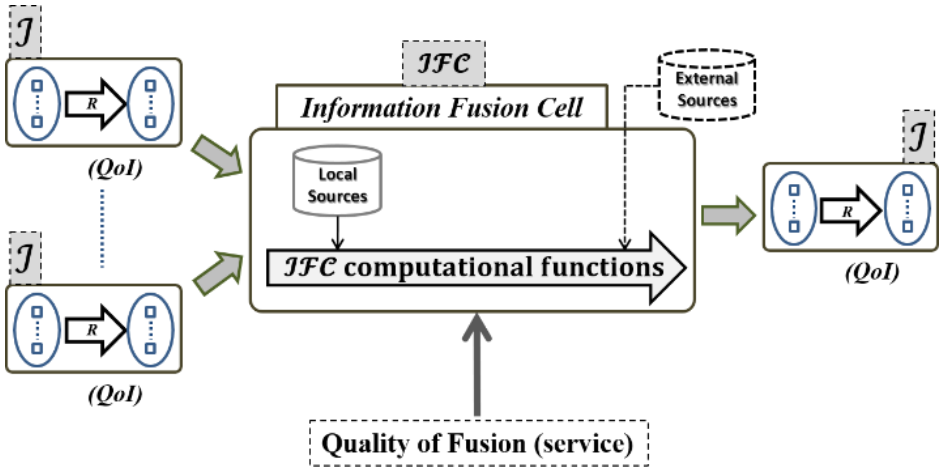
To highlight the importance of the information contents on acting or its impact on the decision-making, it is interesting to introduce here a concept to reflect its position in the context of a semantic field; called semantic status. This will help justify the order on the following: data, information and knowledge. The semantic status is defined as: **Order 0 (data),** when its informative contents remains constant as it can be neither increased nor decreased; **Order 1 (information)**, if its informative contents may change, for instance, under the processing applied to it; **Order 2 (knowledge),** to indicate that its information contents can be significantly enriched through the intervention of cognition and also by an act of human appropriation. The way to measure the semantic status (Ψ), to make the fusion process aware of a semantic status and to represent it, is still an open research question. For now, let us include Ψ in the definition of the input and output scopes: IS_C, OS_C. However, it could very well be part of external context.

4.3 Composition of an Information fusion cell

Information fusion approaches mainly focus on the establishment of adequate *Belief Models* (*BM*) characterizing intrinsic information imperfections, as well as on the development of reasoning tools allowing imperfect information processing

and fusion, through joint "merging" of these belief models. The most widely used belief models dealing with single information elements are: probability distributions $\Pr\{.\}$ [110], evidential mass functions $m(.)$ [111],[112] and possibility distributions $\pi(.)$ [113],[114]. As previously shown, these belief models allow representing, reasoning, processing and merging all forms of intrinsic imperfections that may affect punctual information elements, $\mathcal{I}$. In [95], the concept of *Information Fusion Cells* ($\mathcal{IFC}$) is introduced as the smallest granular component of information fusion systems. $\mathcal{IFC}$ is then considered as an "intermediate conceptual level" between different individual information elements and the global information processing systems. This behavior has been exhibited, in chapters 4-6 of [12], according to several fusion computational strategies.

Figure 11. Information Fusion Cell ($\mathcal{IFC}$) Concept: (Adapted from Source [12] p.140)



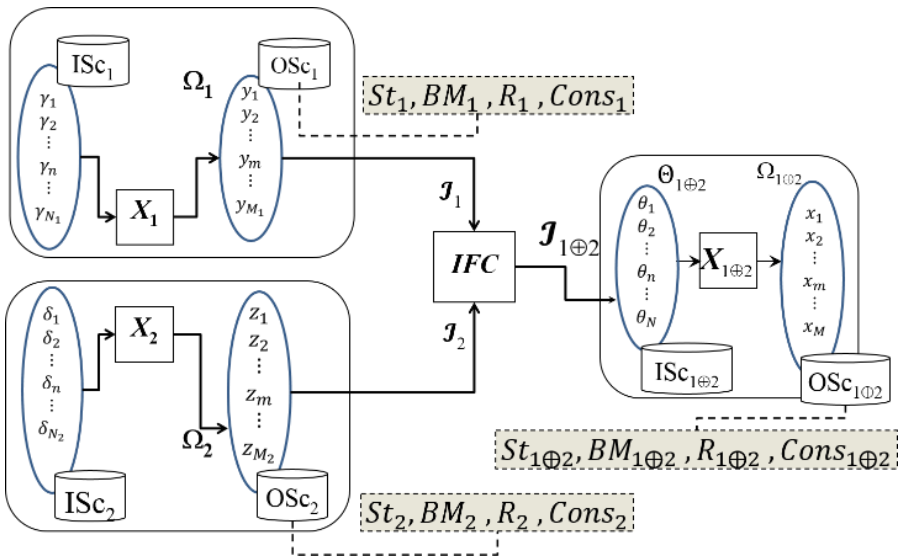
The Quality of Fusion (QoF), as illustrated in figure 11, measures the quality of the fusion process itself and not the quality of the produced information. It may be constraints or indicators relative to the fusion process (not the information which have their own quality description - QoI). Both information fusion and analytics provide services to users (machines or humans). Efficiency and effectiveness are important in delivering those services. The quality of services depends on several criteria derived from notions such as: *" timeliness, accuracy, throughput, confidence, cost, completeness, consistency, correctness, currency, precision, relevance, and various types of uncertainty associated to information.* "

An $\mathcal{IFC}$ is composed of information fusion core functions labelled as '$\mathcal{IFC}$ computational functions' in figure 11. The revised JDL [115],[60], presented a fusion node construct (similar to an $\mathcal{IFC}$) upon three functions: " *Common Referencing, Data Association, and State Estimation.*" Another source, Nicholson [80], used five core functions: "*Detection, Classification, Prediction, Correlation and Assimilation (function of combining)* " for a situation awareness processing

engine. The idea in Solaiman et al. [95] is to propose a more exhaustive set of core functions for an $\mathcal{IFC}$. The 8-computational functions of an $\mathcal{IFC}$ are: "*alignment, detection, partition, combination, veracity, estimation, prediction, association*."

A better characterization of properties of an $\mathcal{IFC}$, formal approaches such as in Kokar *et al.* [116] using category theory, or interpreted systems as with ISSA [90], state-space approach of STDF [89], and 'causal tapestries' of Sulis' archetypal dynamics [97] could contribute to adapt and advance current software engineering methodologies for specifying a computational FIAT model. In [95], [96], the authors state that the core functions are holonic since they can be applied to any semantic status Ψ (data-information-knowledge) provided one can make the process self-aware of Ψ. The core functions have been described as holons. Holons form dynamic hierarchies called holarchies that are networks of $\mathcal{IFCs}$. For instance, meta-agents, at JDL level-4, can manage the 'fusion' holarchy and allow a certain degree of control over lower levels of fusion.

Figure 12. Processing strategies for fusion: (Source [12] p.141)



4.4 Computational Strategies for Information Fusion Cells

Most information processing systems consider information elements as fully reliable (i.e. having unit reliability weights) and fully consistent (having unit consistency degrees). Otherwise, these contextual imperfection indicators are expressed and computed using available *prior* information. An *Information Fusion Cell* ($\mathcal{IFC}$) is defined as the "*basic fusion platform*" allowing the fusion of two information elements (Figure 12):

$$\mathcal{I}_1 = [(\Theta_1, X_1, \Omega_1), \mathrm{IS}_{C1}, \mathrm{OS}_{C1}, St_1, BM_1, R_1, Cons_1] \quad (2)$$

$$\mathcal{I}_2 = [(\Theta_2, X_2, \Omega_2), \mathrm{IS}_{C2},\ \mathrm{OS}_{C2},\ St_2, BM_2, R_2, Cons_2] \quad (3)$$

" The $\mathcal{IFC}$ concept is illustrated in figure 12 with the following notations:
$\Theta_{1\oplus2}$ (resp. $\Omega_{1\oplus2}$) denotes the resulting information definition (resp. content) set;
$IS_{C1\oplus2}$ (resp. $\mathrm{OS}_{C1\oplus2}$) denotes the resulting information input (resp. output) scope;
$St_{1\oplus2}$ is the resulting information imperfection status;
$BM_{1\oplus2}$ is the resulting belief model associated with the information element $\mathcal{I}_{1\oplus2}$;
$R_{1\oplus2}$ (resp. $Cons_{1\oplus2}$) denotes the resulting reliability weight (resp. consistency degree)."

Thus, depending on the $\mathcal{IFC}$ structural and semantic meaning aspects of both the input and the resulting information elements, four $\mathcal{IFC}$ categories, called $\mathcal{IFC}$ types, can be identified. These four types widely corresponding to semantic objectives encountered for fusion strategies are presented in chapters 4-5 of [12]: - Processing Strategy 1: Type 1 (Data Fusion); - Processing Strategy 2: Type 2 (Parallel Belief Fusion); - Processing Strategy 3: Type 3 (Sequential Belief Fusion); - Processing Strategy 4: Type 4 (Competitive Belief Fusion).

A crucial and challenging information fusion tasks concerns the situation where the available belief models $BM_1$and BM_2 come from different mathematical representation models. For instance, in the case where the resulting information element is punctual ($\mathrm{OS}_{C1\oplus2} \equiv Single\ element$), several configurations can be encountered referred as hybrid competitive belief models fusion: Case 1: *Probabilistic* & *evidential* models fusion; Case 2: *Possibilistic* & *evidential* models fusion; Case 3: *Probabilistic* & *possibilistic* models fusion; and the more general, Case 4: *Probabilistic* & *possibilistic* & *evidential* models fusion. To the knowledge of the authors, nothing has been done so far for Case 4. It represents the case where you choose the best representation with respect of the type of the information sources: use *probabilistic* with statistical-type data, use *evidential* with testimony-type, use *possibilistic* with vague and incomplete information. See chapters 4-6 of [12] for more details.

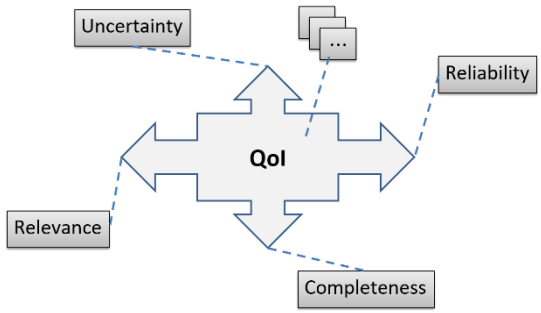
4.5 Quality of Information (QoI): a need for developing measures and metrics

Information and data quality topic has been receiving more attention in the recent years in both civilian and military domains. The problem of representing and incorporating quality characteristics into the design of analytics and information fusion processes is still highly unresolved. There is no unique definition of information quality. In fact, the literature presents several definitions of information quality: "1. *Quality is the totality of characteristics of an entity that bear on its ability to satisfy stated and implied needs*' [117]; 2. '*Quality is the degree to which information has content, form, and time characteristics, which give it value to specific end users*' [118]; 3. '*Quality is the degree to which information is meeting user needs according to external, subjective user perceptions*' [119]; 4. '*Quality is fitness for use*' [120]." Quality of Information (QoI) is a 'user-centric' notion. Users can be either humans or automated agents or models. QoI is a meta-information is 'information about information'.

Measuring the value of this meta-information is through its attributes. In [121], the authors argue that "*without clearly defined attributes and their relationships, we are not just unable to assess* QoI*; we may be unaware of the problem.*". By attributes, we may mean information imperfections. That emphasizes the Veracity dimension in the Big Data problem.

QoI, its representation, interpretation, measuring and processing are probably the most important and difficult tasks of a FIAT-based process. There have been multiple views on QoI, identifying quality attributes and classifying them into broad categories and relations. In [119] (Wang & Strong), data quality was classified into four categories: " *intrinsic, contextual, representational, and accessibility*." In [122], three categories were enumerated: " *pragmatic, semantic, and syntax* " while in [121], four sets were identified: " *integrity, accessibility, interpretability and relevance* ". In [123], they used " *relevance, reliability, completeness and uncertainty* " as illustrated in figure 13. The information quality ontology presented in [124],[125] is one of the first attempt to define information quality and its interrelated dimensions.

Figure 13. Four QoI dimensions that required formalization to be used in FIAT-based designs: (Adapted from Source [12] p.113)



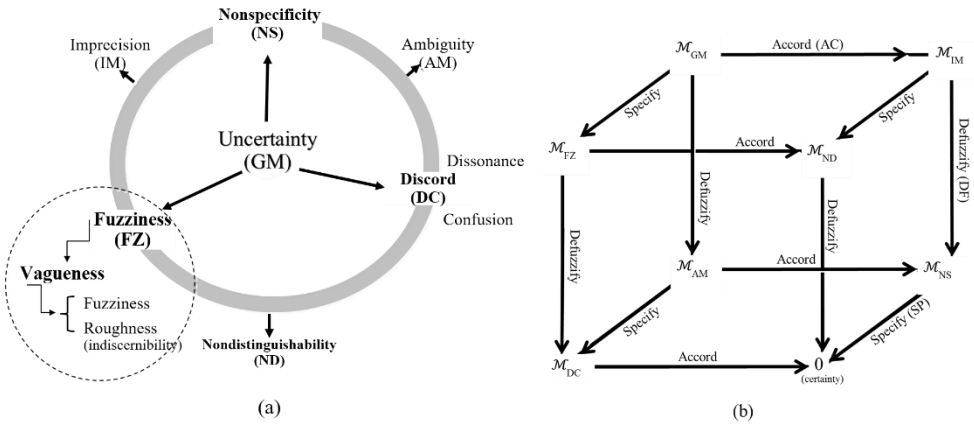
Amongst the four main aspects of figure 13, uncertainty-based information and its measures is certainly the most developed in terms of formalization, measures and its processing.

Measures of uncertainty - More recently, significant efforts [126],[127], [128],[129],[130] have been pursued to address the evaluation of techniques of uncertainty reasoning and to define an ontology for uncertainty reasoning. These efforts contribute towards the understanding of QoI and the formalization of some criteria that could be used for the design of FIAT support systems. However, research is still ongoing and requires additional efforts to get a QoI ontology. In fact, from all the efforts just mentioned, there are mainly four main aspects of information quality that offer formalizations exploitable in computer-based support systems, namely on: uncertainty, reliability, completeness, and relevance [12] shown in figure 13.

In the classical theory of probability, Shannon's entropy is the tool used for quantifying uncertainty. The approach is to verify a set of desirable properties for

probability distributions. In situations where the probabilistic representation is inadequate, the approach is an axiomatic one, by assuming a set of necessary basic properties that a measure must verify. Abellán *et al.* [131, 132] extend the set of required properties originally defined by Klir *et al.* [133] for a total uncertainty measure in Dempster-Shafer Theory. Their extension concerns a monotonicity property that is quite important in the design of a FIAT-based system.

Figure 14. Circular uncertainty typology adapted from [134, 135]



Based on Klir and Wierman's classification, Liu *et al.* [134] developed a circular typology of uncertainty that has been slightly modified as it appears on figure 14. Three general terms are used to designate the combinations of: i) nonspecificity and discord, i.e. ambiguity, ii) nonspecificity and fuzziness, i.e. imprecision, and iii) fuzziness and confusion, i.e. nondisguishability. The term uncertainty refers to the combination of these three basic kinds. The dotted circle on figure 14a, that encircles the notion of fuzziness and vagueness, is plotted to illustrate that the typology of the various kinds of uncertainty may be classified in different ways. For instance, beyond the concept of fuzziness is the broader concept of vagueness which simply means that borderline cases arise when representing set elements. Modelling borderline cases by degrees of truth or membership is only one of many solutions. Vagueness is thus a kind of uncertainty that is detected when borderline cases arise, i.e. objects for which we cannot decide if they belong or not to a given concept. Just like for uncertainty, different kinds of vagueness can be identified.

In the framework of evidence theory (Dempster-Shafer), the belief function can model both nonspecificity and discord. The fuzzy sets theory, representing and managing vague information, deals with fuzziness and nonspecificity as main kinds of uncertainty. The most adequate framework for representing uncertainty when dealing with all three kinds of uncertainty is the combination of the evidence and fuzzy sets theory, i.e. fuzzy evidence theory [135]. Each type of uncertainty

can be quantified in fuzzy evidence theory and a general measure of uncertainty named $\mathcal{M}_{GM}$ is being proposed in [135] along with a new measure of total uncertainty in fuzzy sets theory, named $\mathcal{M}_{IM}$. The general measure of uncertainty is an aggregate measure including all kinds of uncertainty. Figure 14b illustrates the approach used in [135] to show the consistency with respect to existing uncertainty measures in classical set, fuzzy sets, probability, fuzzy probability and evidence theories. The approach is based on the reduction of the uncertainty of to a fuzzy basic probability assignment (*bpa*), from $\mathcal{M}_{GM}$ which quantifies fuzziness, discord and nonspecificity to 0 where no uncertainty can be measured (figure 14b).

Three basic operations aiming at artificially reducing the uncertainty of a fuzzy *bpa* are proposed: (1) defuzzification, (2) specification, and (3) accordance. This scheme leads then to six (6) different ways, as pictured on figure 13b, to make $\mathcal{M}_{GM}$ decreasing to 0, through the different quantities of uncertainty according to the circular typology (figure 13a) with $\mathcal{M}_{AM}$, $\mathcal{M}_{IM}$, $\mathcal{M}_{FZ}$, $\mathcal{M}_{NS}$, $\mathcal{M}_{DC}$, $\mathcal{M}_{ND}$: 1) ***Defuzzification*** transforms a fuzzy *bpa* into a crisp one. When applied to a fuzzy set, *defuzzification* gives a crisp set, while applied to a fuzzy probability distribution, *defuzzification* gives a classical probability distribution. 2) ***Specification*** that transforms a fuzzy *bpa* into a fuzzy probability distribution. When applied to a fuzzy set, specification gives a nonspecific fuzzy set ('*pure*' fuzzy set), while applied to a crisp set, *specification* gives a singleton. 3) ***Accordance*** that transforms a fuzzy *bpa* into a fuzzy set. When applied to a fuzzy probability distribution, *accordance* gives a nonspecific fuzzy set, while applied to a classical probability distribution *accordance* gives a singleton.

Abellán *et al.* [131] as well as Liu *et al.* [135] examine a certain number of measures ($\mathcal{M}_{AM}$, $\mathcal{M}_{IM}$, $\mathcal{M}_{FZ}$, $\mathcal{M}_{NS}$, $\mathcal{M}_{DC}$, $\mathcal{M}_{ND}$, $\mathcal{M}_{GM}$) associated with the circular uncertainty typology (extension of Klir's typology) of figure 13. However, despite of significant recent contributions [136],[131],[134],[137],[138] on measures of uncertainty, great challenges are still unresolved: 1) More investigations are required to attach meaning to those measures. The pragmatics has not been established yet. What those measures are really measuring? 2) Meeting a set of mathematical properties may not mean a meaningful translation to real world systems; 3) Definition of properties that a total measure of uncertainty shall possess in order to be used in the management of FIAT-based support systems.

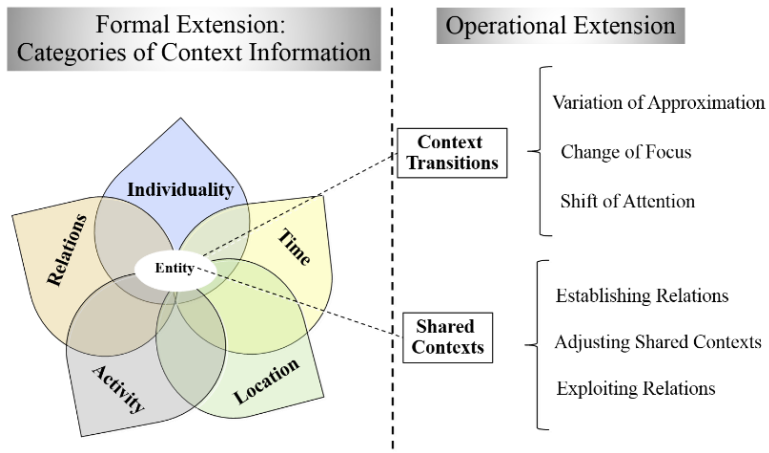
Measures of relevance - Another crucial aspect of QoI is relevance of information. If appropriately formalized, relevance [139],[140],[141] could benefit to any intelligent filtering and context-aware processing system and impacts positively on quality of decisions. Measures of relevance are not presented as such in the literature. The relevance dimension of QoI is about semantics, thus measures of similarity or other semantic measures [142], [143] might be a starting point to define metrics that would help in a FIAT-based design. This dimension of QoI, to be exploitable, would necessitate formalizations. Until now, the literature is presenting too few of contributions on it. This dimension is critical for Big Data and online intelligent applications.

For lack of space, reliability and completeness will not be discussed here. Completeness might be related to aspects of uncertainty. Reliability and trust have received a great deal of attention in the literature since several papers and books have already been published on these topics. Still an open research question is how the four dimensions of QoI are related to the dimensions of system dependability and trustworthiness, the overall objective of a FIAT-based support system (figure 2)?

4.6 Contexts, semantic frames and ontologies

Data acquire meaning through context. Context establishes the basis for discerning meaning of its subjects and may occur at many levels. Exploitation of contextual knowledge is necessary for situation analysis and consequently for its support through FIAT-based systems. The most widely accepted definitions of context is from Dey [145] phrased as follows: "*Context is any information that can be used to characterise the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between the user and the application, including the user and the applications themselves.*"

Figure 15 Definition of context according to Zimmermann et al. [144]



This general definition of context has been extended by Zimmermann *et al.* [144] who introduce a definition comprising three canonical parts: a definition per se in general terms, a formal definition describing the appearance of context and an operational definition characterizing the use of context and its dynamic behavior. Figure 15 shows the five categories of the formal definition as well as the elements of the operational extension. Zimmermann *et al.* [144] extend the definition by the description of the following five categories of elements: "*individuality, activity, location, time, and relations*." These five fundamental context categories determine the design space of context models. The description of the five categories of elements is presented in Zimmermann *et al.* [144] for more comprehensive details. Their underlying motivation is to provide a structure that

bridges the user-developer gap from a general concept (e.g. Dey's definition) easily understandable by the user to its engineering by software developers. This is more than required for the design of a FIAT tools suite to support situation analysis. Dey's definition clearly states that " *context is always bound to an entity and that information that describes the situation of an entity is context.*"

A semantic frame can be thought of as a conceptual structure describing an event, relation, or object and the participants in it. In archetypal dynamics of Sulis, an *informon* acquires meaning through and from contexts. Semantic frames are built of contextual information. The fundamental idea is that a frame represents an object or a concept. Frames are stored as ontologies of sets and subsets of the frame concepts. Frame-based knowledge representation has been exploited in a non-exhaustive list of several applications such as: semantic web [146], object-oriented and markup languages, ontologies, and information fusion.

Ontologies are proving to be effective tools for capturing and specifying categorical information about objects at various levels of granularity, as well as the various sorts of relations that hold between them. Ontologies can be used to model physical objects (including their material composition, attributes and properties), non-physical or psychological objects (e.g. concepts, plans, intentions), temporal events (sequential or scattered processes), and relations between such items (e.g. logical, causal, dependence, internal, external, and intentional kinds of relations). Ontologies have been used to exchange information and knowledge representation in a variety of domains including the four critical CPSS: defense and security, health, transport and energy. Ontologies are part of FIAT to represent domain contextual knowledge.

The distributed nature of most contexts of application requires to relate heterogeneous ontological specifications and to integrate information from multiple diverse sources. The development of Ontology Management Methods (OMM) is then essential to cope with heterogeneity and to enable systems interoperability. OMM are required: for alignment [147],[148],[149]; for merging and translation [150],[151]; for revision and refinement [152], [153],[154].

4.7 Future work toward a FIAT computational model

The elements presented in this section represent a starting point toward a computational model for FIAT. Of course, if the context of application does not have the characteristics elicited in Table 1, the design of FIAT-based support systems might require less sophistication. Nevertheless, CPSS and its complexity are already a reality of our current world. Elements such as those discussed in this section as well as new ones must be developed to get that framework and computational model for a better exploitation of what FIAT can offer.

5. Conclusion

Fusion of Information and Analytics Technologies (FIAT) is a *sine qua non* enabler to cope with Cyber-Physical and Social Systems (CPSS) facing information overload and complexity. FIAT will support rethinking of cyber

interface functions, such as coordination, integration, monitoring, and control and progress decision support concepts to be operational through these interface functions. FIAT-based decision support systems help mainly: 1) in making sense of an ever increasing complexity of our world; 2) to better exploit technological opportunities such as Big Data and IoT; 3) to improve the quality of information (QoI), to reduce and cope with uncertainty; 4) to support distributed decision-making and execution of actions; and, 5) to ensure a proper overall system dependability and trustworthiness. Some of the techniques and methods, taken individually, may be quite mature but can only solve a facet of the problem. The problem is tremendously complex and the current paper has surveyed only a portion of the problem-space. The problem is multidisciplinary by nature and several books and paper would be required to explore the vast solution space. The engineering of an efficient FIAT processing chain to transform data-information-to-actionable knowledge in complex dynamic environments is still quite immature. The authors hope that this paper would motivate the scientific and engineering communities to progress some of the approaches presented here and provide more implementable solutions to cope with that complexity.

Acknowledgement

The authors wish to thank those who have made possible the publication of the book entitled '*Fusion of Information and Analytics for Big Data and IoT*: Artech House, 2016' from which the contents of this paper has been extracted, adapted and updated.

References

[1] P. McFedries, "The coming data deluge [Technically Speaking]," *Spectrum, IEEE*, vol. 48, pp. 19-19, 2011.

[2] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Computer networks*, vol. 54, pp. 2787-2805, 2010.

[3] F. Wortmann and K. Flüchter, "Internet of things," *Business & Information Systems Engineering*, vol. 57, pp. 221-224, 2015.

[4] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future generation computer systems*, vol. 29, pp. 1645-1660, 2013.

[5] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE Communications Surveys & Tutorials*, vol. 17, pp. 2347-2376, 2015.

[6] S. C. Mukhopadhyay and N. Suryadevara, "Internet of things: Challenges and opportunities," in *Internet of Things*, ed: Springer, 2014, pp. 1-17.

[7] A. Whitmore, A. Agarwal, and L. Da Xu, "The Internet of Things—A survey of topics and trends," *Information Systems Frontiers*, vol. 17, pp. 261-274, 2015.

[8] D. Bradley, D. Russell, I. Ferguson, J. Isaacs, A. MacLeod, and R. White, "The Internet of Things–The future or the end of mechatronics," *Mechatronics*, vol. 27, pp. 57-74, 2015.

[9] M. B. Nirmala, "A Survey of Big Data Analytics Systems: Appliances, Platforms, and Frameworks," *Handbook of Research for Cloud Infrastructures to Big Data Analytics*, pp. 393-419, 2014.

[10] M. Chen, S. Mao, and Y. Liu, "Big data: A survey," *Mobile Networks and Applications,* vol. 19, pp. 171-209, 2014.
[11] U. Sivarajah, M. M. Kamal, Z. Irani, and V. Weerakkody, "Critical analysis of Big Data challenges and analytical methods," *Journal of Business Research,* vol. 70, pp. 263-286, 2017.
[12] É. Bossé and B. Solaiman, *Fusion of Information and Analytics for Big Data and IoT*: Artech House, Inc., 2016.
[13] K. Pawar and V. Attar, "A survey on Data Analytic Platforms for Internet of Things," in *Computing, Analytics and Security Trends (CAST), International Conference on*, 2016, pp. 605-610.
[14] T. Dull, "Big data and the Internet of Things: Two sides of the same coin," *SAS Best Practices,* 2015.
[15] N. Jesse, "Internet of Things and Big Data–The Disruption of the Value Chain and the Rise of New Software Ecosystems," *IFAC-PapersOnLine,* vol. 49, pp. 275-282, 2016.
[16] Z. Liu, D.-S. Yang, D. Wen, W.-M. Zhang, and W. Mao, "Cyber-physical-social systems for command and control," *IEEE Intelligent Systems,* vol. 26, pp. 92-96, 2011.
[17] H. Zhuge, "Cyber-Physical Society-The science and engineering for future society," *Future Generation Computer Systems,* vol. 32, pp. 180-186, 2014.
[18] G. Xiong, F. Zhu, X. Liu, X. Dong, W. Huang, S. Chen, *et al.*, "Cyber-physical-social system in intelligent transportation," *IEEE/CAA Journal of Automatica Sinica,* vol. 2, pp. 320-333, 2015.
[19] J. Zeng, L. T. Yang, M. Lin, H. Ning, and J. Ma, "A survey: Cyber-physical-social systems and their system-level design methodology," *Future Generation Computer Systems,* 2016.
[20] P. Jiang, K. Ding, and J. Leng, "Towards a cyber-physical-social-connected and service-oriented manufacturing paradigm: Social Manufacturing," *Manufacturing Letters,* vol. 7, pp. 15-21, 2016.
[21] H. Gill, "From vision to reality: cyber-physical systems," in *Presentation, HCSS National Workshop on New Research Directions for High Confidence Transportation CPS: Automotive, Aviation and Rail*, 2008.
[22] S. Jeschke, "Cyber-Physical Systems - History, Presence and Future," ed, 2013.
[23] M. Liggins, D. Hall, and J. Llinas, *Handbook of Multisensor Data Fusion: Theory and Practice, Second Edition*: Taylor & Francis, 2008.
[24] D. Smith and S. Singh, "Approaches to multisensor data fusion in target tracking: A survey," *Knowledge and Data Engineering, IEEE Transactions on,* vol. 18, pp. 1696-1710, 2006.
[25] G. Navarro-Arribas and V. Torra, "Information fusion in data privacy: A survey," *Information Fusion,* vol. 13, pp. 235-244, 2012.
[26] E. F. Nakamura, A. A. Loureiro, and A. C. Frery, "Information fusion for wireless sensor networks: Methods, models, and classifications," *ACM Computing Surveys (CSUR),* vol. 39, p. 9, 2007.
[27] B. Khaleghi, A. Khamisa, F. O. Karraya, and S. N. Razavi, "Multisensor data fusion: A review of the state-of-the-art," *Information Fusion,* vol. 14, pp. 28-44, 2013.

[28] A. Avizienis, J.-C. Laprie, and B. Randell, *Fundamental concepts of dependability*: University of Newcastle upon Tyne, Computing Science, 2001.
[29] L. Petre, K. Sere, and E. Troubitsyna, *Dependability and Computer Engineering: Concepts for Software Intensive Systems*. Hershey, USA IGI Global, 2012.
[30] M. Johnson and C. N. Dampney, "On category theory as a (meta) ontology for information systems research," in *Proceedings of the international conference on Formal Ontology in Information Systems-Volume 2001*, 2001, pp. 59-69.
[31] (2010, February) Data, data everywhere. *The Economist - A Special Report on Managing Information*. Available: http://www.economist.com/node/15557443
[32] K. Wan and V. Alagar, "Dependable Context-Sensitive Services in Cyber Physical Systems," in *Trust, Security and Privacy in Computing and Communications (TrustCom), 2011 IEEE 10th International Conference on*, 2011, pp. 687-694.
[33] A. Bobbio, L. Portinale, M. Minichino, and E. Ciancamerla, "Improving the analysis of dependable systems by mapping fault trees into Bayesian networks," *Reliability Engineering & System Safety*, vol. 71, pp. 249-260, 2001.
[34] E. A. Lee, "Cyber physical systems: Design challenges," in *Object Oriented Real-Time Distributed Computing (ISORC), 2008 11th IEEE International Symposium on*, 2008, pp. 363-369.
[35] G. Eoyang, "Human systems dynamics: Competencies for a new organizational practice," *Practicing organization development: A guide for leading change*, pp. 446-456, 2010.
[36] C. Gershenson, *Design and control of self-organizing systems*: CopIt Arxives, 2007.
[37] C. Gershenson, "The implications of interactions for science and philosophy," *Foundations of Science*, vol. 18, pp. 781-790, 2013.
[38] M. Niazi and A. Hussain, "Agent-based computing from multi-agent systems to agent-based models: a visual survey," *Scientometrics*, vol. 89, pp. 479-499, 2011.
[39] R. Poovendran, "Cyber–physical systems: Close encounters between two parallel worlds [Point of view]," *Proceedings of the IEEE*, vol. 98, pp. 1363-1366, 2010.
[40] J. Sztipanovits, X. Koutsoukos, G. Karsai, N. Kottenstette, P. Antsaklis, V. Gupta, *et al.*, "Toward a science of cyber–physical system integration," *Proceedings of the IEEE*, vol. 100, pp. 29-44, 2012.
[41] M. A. Solano and G. Jernigan, "Enterprise data architecture principles for High-Level Multi-Int fusion: A pragmatic guide for implementing a heterogeneous data exploitation framework," in *Information Fusion (FUSION), 2012 15th International Conference on*, 2012, pp. 867-874.
[42] N. Agoulmine, *Autonomic Network Management Principles: From Concepts to Applications*: Elsevier Science, 2010.
[43] P. Horn, "Autonomic computing: IBM\'s Perspective on the State of Information Technology," 2001.
[44] M. Parashar and S. Hariri, *Autonomic computing: concepts, infrastructure, and applications*: CRC press, 2006.
[45] E. Bernard-Weil, "Transcendance, an essential concept for system and complexity sciences to spread out," *Complexity*, vol. 6, pp. 23-35, 2000.
[46] C. Richards, "Boyd's OODA loop," *Slideshow. URL: http://www. dni. net/fcs/ppt/boyds_ooda_loop. ppt [Online*, 2001.
[47] E. Blasch, R. Breton, and É. Bossé, "User Information Fusion Decision Making Analysis with the C-OODA Model," in *High-Level Information Fusion Management*

and Systems Design, E. Blasch, É. Bossé, and D. A. Lambert, Eds., ed: Artech House, 2012, pp. 215-232.

[48] E. Shahbazian, D. E. Blodgett, and P. Labbé, "The extended OODA model for data fusion systems," in *Proceedings of the 4th International Conference on Information Fusion (FUSION2001)*, Montreal, 2001.

[49] S. Das, *Computational Business Analytics*: Taylor & Francis, 2013.

[50] J. Taylor, *Decision Management Systems: A Practical Guide to Using Business Rules and Predictive Analytics*: Pearson Education, 2011.

[51] J. Strassner, "Knowledge Representation, Processing, and Governance in the FOCALE Autonomic Architecture," *Autonomic Network Management Principles: From Concepts to Applications*, p. 253, 2010.

[52] E. Blasch, P. Valin, E. Bossé, M. Nilsson, J. van Laere, and E. Shahbazian, "Implication of culture: user roles in information fusion for enhanced situational understanding," in *Information Fusion, 2009. FUSION'09. 12th International Conference on*, 2009, pp. 1272-1279.

[53] D. Dittrich, A. Center, and M. P. Haselkorn, "Visual Analytics in Support of Secure Cyber-Physical Systems."

[54] S. Ghosh, *Algorithm design for networked information technology systems*: Springer, 2004.

[55] D. J. Bryant, R. D. Webb, and C. McCann, "Synthesizing two approaches to decision making in command and control," *Canadian Military Journal*, vol. 4, pp. 29-34, 2003.

[56] A. Guitouni, K. Wheaton, and D. Wood, "An Essay to Characterize Models of the Military Decision-Making Process," in *11th ICCRT Symposium, Cambridge UK*, 2006.

[57] É. Bossé, J. Roy, and S. Wark, *Concepts, models, and tools for information fusion*: Artech House, Inc., 2007.

[58] É. Bossé, A.-L. Jousselme, and P. Maupin, "Situation analysis for decision support: a formal approach," in *Proceedings of the 10th International Conference on Information Fusion (FUSION 2007)*, Quebec City, 2007.

[59] G. J. Klir, *Uncertainty and information: foundations of generalized information theory*: John Wiley & Sons, 2005.

[60] A. N. Steinberg, C. L. Bowman, and F. E. White, "Revisions to the JDL data fusion model," in *The Joint NATO/IRIS Conference*, Quebec City, 1998.

[61] E. Blasch, E. Bossé, and D. A. Lambert, *High-level Information Fusion Management and Systems Design*: Artech House, 2012.

[62] D. L. Hall and J. M. Jordan, *Human-Centered Information Fusion*: Artech House, 2010.

[63] D. L. Hall and S. A. H. McMullen, *Mathematical Techniques in Multisensor Data Fusion*: Artech House, 2004.

[64] W. W. Eckerson, *Secrets of Analytical Leaders*: Technics Publications, 2012.

[65] D. Kiron, R. Shockley, N. Kruschwitz, G. Finch, and M. Haydock, "Analytics: The widening divide," *MIT Sloan Management Review*, vol. 53, pp. 1-22, 2011.

[66] H. W. Park and L. Leydesdorff, "Decomposing social and semantic networks in emerging “big data” research," *Journal of Informetrics*, vol. 7, pp. 756-765, 2013.

[67] N. Sheikh, *Implementing Analytics: A Blueprint for Design, Development, and Adoption*: Newnes, 2013.

[68] I. H. Witten and E. Frank, *Data Mining: Practical machine learning tools and techniques*: Morgan Kaufmann, 2005.

[69] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 37, pp. 32-64, 1995.

[70] J. Llinas, "Reexamining information fusion-decision making inter-dependencies," in *Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), 2014 IEEE International Inter-Disciplinary Conference on*, 2014, pp. 1-6.

[71] P. Pirolli and S. Card, "The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis," in *Proceedings of international conference on intelligence analysis*, 2005, pp. 2-4.

[72] P. L. Pirolli, *Information foraging theory: Adaptive interaction with information*: Oxford University Press, 2007.

[73] J. Roy, "A knowledge-centric view of situation analysis support systems," TR 2005-419) DRDC Valcartier. Technical Report2007.

[74] L. J. Fülöp, G. Tóth, R. Rácz, J. Pánczél, T. Gergely, Á. Beszédes, *et al.*, "Survey on complex event processing and predictive analytics," in *Proceedings of the Fifth Balkan Conference in Informatics*, 2010, pp. 26-31.

[75] G. Cugola and A. Margara, "Processing flows of information: From data stream to complex event processing," *ACM Computing Surveys (CSUR)*, vol. 44, p. 15, 2012.

[76] F. Wang, S. Liu, P. Liu, and Y. Bai, "Bridging physical and virtual worlds: complex event processing for RFID data streams," in *Advances in Database Technology-EDBT 2006*, ed: Springer, 2006, pp. 588-607.

[77] W. Pedrycz, *Granular computing: analysis and design of intelligent systems*: CRC Press, 2013.

[78] W. Pedrycz and S.-M. Chen, "Information Granularity, Big Data, and Computational Intelligence," ed: Springer, 2014.

[79] W. Sulis, "Archetypal Dynamics," in *Formal descriptions of developing systems.* vol. 121, J. Nation, I. Trofimova, J. Rand, and W. Sulis, Eds., ed: Kluwer Academic Publishers, 2003, pp. 180-227.

[80] D. Nicholson, "Defence Applications of Agent-Based Information Fusion," *The Computer Journal*, vol. 54, pp. 263–273, 2011.

[81] A. Guitouni, "A Time Sensitive Decision Support System for Crisis and Emergency Management," NATO RTO-IST-086, C3I for Crisis, Emergency and Consequence Management.

[82] J. Llinas, "A survey and analysis of frameworks and framework issues for information fusion applications," in *Hybrid Artificial Intelligence Systems*, ed: Springer, 2010, pp. 14-23.

[83] J. Llinas, R. Nagi, D. Hall, and J. Lavery, "A multi-disciplinary university research initiative in hard and soft information fusion: Overview, research strategies and initial results," in *Information Fusion (FUSION), 2010 13th Conference on*, 2010, pp. 1-7.

[84] D. L. Hall, M. McNeese, J. Llinas, and T. Mullen, "A framework for dynamic hard/soft fusion," in *Information Fusion, 2008 11th International Conference on*, 2008, pp. 1-8.

[85] R. R. Yager, "Hard and soft information fusion using measures," in *2010 IEEE International Conference on Intelligent Systems and Knowledge Engineering*, 2010.
[86] J. Dean and S. Ghemawat, "MapReduce: simplified data processing on large clusters," *Communications of the ACM*, vol. 51, pp. 107-113, 2008.
[87] M. Olson, "Hadoop: Scalable, flexible data storage and analysis," *IQT Quarterly*, vol. 1, pp. 14-18, 2010.
[88] W. Laura, "A Practical Guide to Big Data: Opportunities, Challenges & Tools," Dassault Systèmes, Paris2012 2012.
[89] D. A. Lambert, "A blueprint for higher-level fusion systems," *Inf. Fusion*, vol. 10, pp. 6-24, 2009.
[90] P. Maupin and A.-L. Jousselme, "Interpreted systems for situation analysis," in *Proceedings of the 10th International Conference on Information Fusion (FUSION2007)* Quebec City, Canada, 2007.
[91] R. Farahbod, U. Glasser, E. Bossé, and A. Guitouni, "Integrating abstract state machines and interpreted systems for situation analysis decision support design," in *Information Fusion, 2008 11th International Conference on*, 2008, pp. 1-8.
[92] P. Valin, E. Bossé, A. Guitouni, H. Wehn, and J. Happe, "Testbed for Distributed High–Level Information Fusion and Dynamic Resource Management," in *Int. Conf. on Info Fusion*, 2010.
[93] M. A. Solano, S. Ekwaro-Osire, and M. M. Tanik, "High-Level fusion for intelligence applications using Recombinant Cognition Synthesis," *Information Fusion*, vol. 13, pp. 79-98, 2012.
[94] S. Chandana, H. Leung, E. Bosse, and P. Valin, "Fuzzy cognitive map based situation assessment for coastal surveillance," in *Information Fusion, 2008 11th International Conference on*, 2008, pp. 1-6.
[95] B. Solaiman, É. Bossé, L. Pigeon, D. Gueriot, and M. C. Florea, "A conceptual definition of a holonic processing framework to support the design of information fusion systems," *Information Fusion*, vol. 21, pp. 85-99, 2015.
[96] H. Paggi, E. Bossé, M. C. Florea, and B. Solaiman, "On the use of holonic agents in the design of information fusion systems," in *Information Fusion (FUSION), 2014 17th International Conference on*, 2014, pp. 1-8.
[97] W. H. Sulis, "Archetypal Dynamics: An Approach to the Study of Emergence," in *Formal Descriptions of Developing Systems, NATO Science Series Volume 121*, 2003, pp. 185-228.
[98] W. Sulis, "Causal Tapestries," *Bulletin of the American Physical Society*, vol. 56, 2011.
[99] W. Sulis, "A Process Model of Quantum Mechanics," Ph.D Ph.D, Physics, University of Waterloo Waterloo, Ontario 2014.
[100] D. Lambert and C. Nowak, "The Mephisto Conceptual Framework," Defence Science and Technology Organisation, DSTO-TR-2162, Australia2008.
[101] R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi, *Reasoning About Knowledge*: The MIT Press, 2003.
[102] E. Börger and R. F. Stärk, *Abstract State Machines: A Method for High-level System Design and Analysis; with 19 Tables*: Springer, 2003.

[103] R. Farahbod, U. Glässer, and M. Vajihollahi, "Specification and validation of the business process execution language for web services," in *Abstract State Machines 2004. Advances in Theory and Practice*, ed: Springer, 2004, pp. 78-94.
[104] W. Sulis, "Archetypal dynamics, emergent situations, and the reality game," *Nonlinear dynamics, psychology, and life sciences,* vol. 14, pp. 209-238, 2010.
[105] É. Bossé, A. Guitouni, and P. Valin, "An essay to characterize information fusion systems," in *Proceedings of the 9th International Conference on Information Fusion (FUSION2006)*, Firenze, Italy, 2006.
[106] J. R. Talburt, *Entity resolution and information quality*: Morgan Kaufmann, 2011.
[107] J. H. Bernstein, "The data-information-knowledge-wisdom hierarchy and its antithesis," *NASKO,* vol. 2, pp. 68-75, 2011.
[108] J. Roy and A. B. Guyard, "A knowledge-based system for multiple hypothesis sensemaking support," in *Information Fusion (FUSION), 2011 Proceedings of the 14th International Conference on*, 2011, pp. 1-8.
[109] C. Laudy, "Introducing semantic knowledge in high level information fusion, Ph. D. Thesis, Université Pierre et Marie Curie, Paris, France," Ph. D. Thesis2010.
[110] J. Pearl, *Probabilistic reasoning in intelligent systems networks of plausible inference*: Morgan Kaufmann Publishers, 1988.
[111] G. Shafer, *A Mathematical Theory of Evidence*: Princeton University Press, 1976.
[112] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," *The annals of mathematical statistics,* pp. 325-339, 1967.
[113] L. Zadeh, "Fuzzy Sets as the Basis for a Theory of Possibility," *Fuzzy Sets and Systems,* vol. 1, pp. 3-28, 1978.
[114] D. Dubois and H. Prade, *Possibility Theory: An Approach to Computerized Processing of Uncertainty*: Plenum Press, 1988.
[115] C. L. Bowman and A. N. Steinberg, "Systems Engineering Approach for Implementing Data Fusion Systems," in *Handbook of Multisensor Data Fusion*, M. E. Liggins, D. L. Hall, and J. Llinas, Eds., ed: CRC Press, 2009, pp. 561-596.
[116] M. M. Kokar, J. A. Tomasik, and J. Weyman, "Formalizing classes of information fusion systems," *Information Fusion,* vol. 5, pp. 189–202, 2004.
[117] I. Standard, "8402," *Terminology. Good mark,* vol. 30, 1994.
[118] J. A. O'brien and G. M. Marakas, *Introduction to information systems* vol. 13: McGraw-Hill/Irwin, 2005.
[119] R. Y. Wang and D. M. Strong, "Beyond accuracy: what data quality means to data consumers," *Journal of Management Information Systems,* vol. 12, pp. 5-34, 1996.
[120] J. M. Juran and J. A. De Feo, *Juran's quality handbook: the complete guide to performance excellence*: McGraw Hill, 2010.
[121] M. Bovee, R. P. Srivastava, and B. Mak, "A conceptual framework and belief-function approach to assessing overall information quality," *International journal of intelligent systems,* vol. 18, pp. 51-74, 2003.
[122] M. Helfert, "Managing and measuring data quality in data warehousing," in *Proceedings of the World Multiconference on Systemics, Cybernetics and Informatics, Florida, Orlando*, 2001, pp. 55-65.
[123] E. Lefebvre, M. Hadzagic, and É. Bossé, "On Quality of Information in Multi-Source Fusion Environments," *Advances and Challenges in Multisensor Data and Information Processing,* p. 69, 2007.
[124] G. Rogova and E. Bossé, "Information quality effects on information fusion," DRDC Valcartier, Tech Rept2008.

[125] G. Rogova and É. Bossé, "Information quality in information fusion," in *Proceedings of the 13th International Conference on Information Fusion (FUSION2010)* Edinburg, UK, 2010.

[126] E. Blasch, P. Valin, and É. Bossé, "Measures of Effectiveness for High-Level Information Fusion," in *High-Level Information Fusion Management and Systems Design*, E. Blasch, É. Bossé, and D. A. Lambert, Eds., ed Boston: Artech House, 2012, pp. 331-348.

[127] E. Blasch, P. C. Costa, K. B. Laskey, H. Ling, and G. Chen, "The URREF ontology for semantic wide area motion imagery exploitation," in *Aerospace and Electronics Conference (NAECON), 2012 IEEE National*, 2012, pp. 228-235.

[128] E. Blasch, A. Josang, J. Dezert, P. C. Costa, and A.-L. Jousselme, "URREF self-confidence in information fusion trust," in *Information Fusion (FUSION), 2014 17th International Conference on*, 2014, pp. 1-8.

[129] E. Blasch, K. B. Laskey, A.-L. Jousselme, V. Dragos, P. C. Costa, and J. Dezert, "URREF reliability versus credibility in information fusion (STANAG 2511)," in *Information Fusion (FUSION), 2013 16th International Conference on*, 2013, pp. 1600-1607.

[130] P. C. Costa, K. B. Laskey, E. Blasch, and A.-L. Jousselme, "Towards unbiased evaluation of uncertainty reasoning: the URREF ontology," in *Information Fusion (FUSION), 2012 15th International Conference on*, 2012, pp. 2301-2308.

[131] J. Abellán and A. Masegosa, "Requirements for total uncertainty measures in Dempster–Shafer theory of evidence," *International journal of general systems*, vol. 37, pp. 733-747, 2008.

[132] J. Abellán and S. Moral, "Measuring total uncertainty in Dempster-Shafer theory of Evidence: properties and behaviors," in *Fuzzy Information Processing Society, 2008. NAFIPS 2008. Annual Meeting of the North American*, 2008, pp. 1-6.

[133] G. J. Klir and M. J. Wierman, *Uncertainty-Based Information: Elements of Generalized Information Theory*: Physica-Verlag HD, 1999.

[134] C. Liu, A.-L. Jousselme, É. Bossé, and D. Grenier, "Measures of uncertainty for fuzzy evidence theory " Technical Report DRDC-Valcartier TR2010-223, 2011 2010.

[135] C. Liu, " A general measure of uncertainty-based information," Ph.D, Electrical and Computer Engineering, Université Laval, Québec, 2004.

[136] J. Abellán, G. Klir, and S. Moral, "Disaggregated total uncertainty measure for credal sets," *International Journal of General Systems*, vol. 35, pp. 29-44, 2006.

[137] A. Burkov, S. Paquet, G. Michaud, and P. Valin, "An empirical study of uncertainty measures in the fuzzy evidence theory," in *Information Fusion (FUSION), 2011 Proceedings of the 14th International Conference on*, 2011, pp. 1-8.

[138] J. Abellán and É. Bossé, "Drawbacks of Uncertainty Measures Based on the Pignistic Transformation," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2016.

[139] R. Breton, E. Bosse, R. Rousseau, and S. Tremblay, "Framework for the Analysis of Information Relevance (FAIR)," in *Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), 2012 IEEE International Multi-Disciplinary Conference on*, 2012, pp. 210-213.

[140] M. Hadzagic, M. St-Hilaire, P. Valin, and E. Shahbazian, "Reliability and relevance in the Thresholded Dempster-Shafer algorithm for ESM data fusion," in *Information Fusion (FUSION), 2012 15th International Conference on*, 2012, pp. 615-620.

[141] F. Pichon, D. Dubois, and T. Denœux, "Relevance and truthfulness in information correction and fusion," *International Journal of Approximate Reasoning,* vol. 53, pp. 159-175, 2012.

[142] M. Gahegan, R. Agrawal, A. Jaiswal, J. Luo, and K. Soon, "Measures of Similarity for Integrating Conceptual Geographical Knowledge: Some Ideas and Some Questions," in *COSIT: Workshop on semantic similarity measurements*, 2007.

[143] L. I. Kuncheva, "Using measures of similarity and inclusion for multiple classifier fusion by decision templates," *Fuzzy sets and systems,* vol. 122, pp. 401-407, 2001.

[144] A. Zimmermann, A. Lorenz, and R. Oppermann, "An operational definition of context," in *Modeling and using context*, ed: Springer, 2007, pp. 558-571.

[145] A. Dey, "Understanding and Using Context," *Personal and Ubiquitous Computing,* vol. 5, pp. 4-7, 2001.

[146] O. Lassila and D. McGuinness, "The role of frame-based representation on the semantic web," *Linköping Electronic Articles in Computer and Information Science,* vol. 6, p. 2001, 2001.

[147] E. P. Blasch, E. Dorion, P. Valin, E. Bossé, and J. Roy, "Ontology alignment in geographical hard-soft information fusion systems," in *Information Fusion (FUSION), 2010 13th Conference on*, 2010, pp. 1-8.

[148] Y. Kalfoglou and M. Schorlemmer, "Ontology mapping: the state of the art," *The knowledge engineering review,* vol. 18, pp. 1-31, 2003.

[149] P. Shvaiko and J. Euzenat, "Ontology matching: state of the art and future challenges," *Knowledge and Data Engineering, IEEE Transactions on,* vol. 25, pp. 158-176, 2013.

[150] A. Zimmermann, M. Krötzsch, J. Euzenat, and P. Hitzler, "Formalizing ontology alignment and its operations with category theory," in *Proc. 4th International conference on Formal ontology in information systems (FOIS)*, 2006, pp. 277-288.

[151] N. F. Noy, "Tools for mapping and merging ontologies," in *Handbook on ontologies*, ed: Springer, 2004, pp. 365-384.

[152] F. J. McNeill, A. Bundy, and M. Schorlemmer, "Dynamic ontology refinement," University of Edinburgh, 2006.

[153] M. Klein, D. Fensel, A. Kiryakov, and D. Ognyanov, "Ontology versioning and change detection on the web," in *Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web*, ed: Springer, 2002, pp. 197-212.

[154] J. Heflin and J. Hendler, "Dynamic ontologies on the web," in *AAAI/IAAI*, 2000, pp. 443-449.