

From Social Computing to Reflexive Collective intelligence: The IEML Research Program

Prof. Pierre Lévy, CRC, FRSC, University of Ottawa

22/08/2009

Abstract

The IEML research program promotes a radical innovation in the notation and processing of semantics. IEML (Information Economy MetaLanguage) is a regular language that provides new methods for semantic interoperability, semantic navigation, collective categorization and self-referential collective intelligence. This research program is compatible with the major standards of the Web of data and is in tune with the current trends in social computing.

The paper explains the philosophical relevance of this new language, expounds its syntactic and semantic structures and ponders its possible implications for the growth of collective intelligence in cyberspace.

Keywords: collective intelligence, IEML, semantic space, semantic tagging, semantic interoperability, translation, metalanguage, Web of data, philosophy.

Introduction

Collective intelligence (CI) is the capacity of human collectives to engage in intellectual cooperation in order to create, innovate and invent [42, 68, 71]. It can be applied at any scale, from work teams to huge networks or even to our whole species. Collective intelligence is a determining factor in competitiveness, creativity and human development in a knowledge based economy, or in an *information economy*. The more our society depends on the creative management of knowledge the more this capacity becomes of fundamental importance [3]. There is a growing feeling that there exists a strong correlation between communities collective intelligence and the degree of their human development. CI can be seen as a driving force of human development and within this conceptual framework, conversely, human development provides CI with an environment for growth [47]. As digital technologies give us more and more powerful tools to augment personal and collective cognitive processes, it becomes essential to understand how the collective intelligence processes can be multiplied by digital networks [45]. The topic of the augmentation of CI through digital networks is an emerging research area [33], as the extensive body of literature of knowledge management shows [57, 76, 56, 13], and as the Web 2.0 or social computing attests [74, 70, 60, 65, 49].

Douglas Engelbart can be considered as the main founder of this field [17]. I have been involved in its development since the end of the 1980s with an invention [40], several publications [38, 39, 41, 42, 43, 44, 45, 46] and with the creation in 2002 of the first academic research center exclusively devoted to this subject (the Canada Research Chair in Collective intelligence, or CRC-IC, at the University of Ottawa). This article summarizes the research I have pursued at the CRC-IC for the last seven years. I have to warn the reader that these researches are philosophical in nature, that their style is mainly theoretical, not to say highly speculative, and finally that they are still in progress. The technical developments that have been achieved so far by my small team (see acknowledgements at the end of this article) were mainly for "proof of concept" (by demonstrating the feasibility of the computations implied by the theoretical ideas) and not yet for adoption by any large community of users. Nevertheless, the research program developed at the CRC-IC is now mature enough to be presented to the scientific community. This research program is based on the development and technical exploitation of an artificial language called IEML (Information Economy MetaLanguage) that is supposed to be developed and used collaboratively for the purpose of augmenting human collective intelligence by technologies. More precisely, IEML semantic space will help

solving two important issues: the semantic interoperability problem and the problem of self-reference of digital-based collective intelligence. Before introducing the syntax and semantics of IEML, I will put it in context. I will first evoke this context in the framework of information engineering and then in the framework of humanities and social sciences.

Part I

IEML in the Context of Information Technologies

1 The Problem of Collective Categorization in Cyberspace

In cyberspace, for the first time in human history, our species is growing a universally interconnected common memory where ubiquitous data can be accessed and transformed by automatic symbol manipulators. Since the Web only became public around 1995, it is less than one generation old and we are just beginning its techno-cultural exploration. Prior to the Web, there already were intellectual technologies tapping into digital computation power, like spreadsheets, multimedia interactive simulations or hypermedia. But my hypothesis is that the main developments into the full symbolic and cognitive exploitation of the global digital memory are still to come. No generation before ours has been confronted with the challenge of organizing and exploiting an inexhaustible amount of shared data including the various cultural productions of diverse past and present communities. As shown by recent R&D activity addressing the improvement of social tagging [16], or the construction of a "Web of data" organized by ontologies[20, 27], the augmentation of our collective categorization power is a key issue in the development of new symbolic systems. It can be argued that categorization is indeed at the core of cognitive processes, and particularly in the case of those human cognitive processes that are driven by (cultural) symbolic systems [37]. Database design and management are becoming the main scientific activity [6] as well as the essence of digital art [52]. The problem of useful data categorization is also at the core of collective intelligence management in companies and businesses. The participation in several social networks with hundred of contacts and the access to data and metadata through global sharing systems like Twitter (hashtags), Delicious, Twine, YouTube or Flickr, make the issue of categorization in personal and collective cloud management urgent [67].

My main hypothesis is that natural languages, as well as notation systems invented before the 21st century, are not appropriate for the current and future scale of the social categorization process. They are not fit to exploit the new interconnected global digital memory and its unprecedented power. Natural languages are in accord with human brain processing, they were not designed to be automatically manipulated. Old notation and writing systems correspond to heavy and slow physical storage and retrieval processes, and not to ubiquitous high speed automatic computing. Ontologies get around the difficulty of categorizing the Web by building rigid logical relationships between opaque chains of symbols (URIs) and strings of characters borrowed from natural languages. But there are several disconnected ontologies. Search engines and advanced folksonomic design bypass the problem of accurate semantic representation of data by operating a statistical analysis on strings of characters originally meant to represent sounds (not meaning). Trying to synchronize and optimize such diverse and massive categorization processes in cyberspace by relying on natural languages and old notation systems is like trying to find powerful algorithms for the manipulation of roman numerals or numbers written in alphabetic notation instead of looking for a better symbolic system. The IEML research program suggests that we should adopt a symbolic system for the notation and manipulation of concepts designed from the very beginning for massively distributed social computation in a technically interconnected global memory. As a support for automatic computing, this symbolic system must be a regular language [11].

2 The Nature of IEML

2.1 Differences between IEML and Natural Languages

IEML is a symbolic system for the notation of meaning that is “semantic content oriented” rather than “instruction oriented” like programming languages or “format oriented” like data standards. In that sense, it is closer to natural languages than to computer languages.

But there are also several differences between IEML and natural languages.

First, it is an *artificial* language that is mainly designed to be written and read instead of being spoken. It can be thought of as a scientific notation for meaning. As a writing system, IEML is an ideography and not primarily a phonetic system. This means that its symbols denote ideas and not sounds, as is the case for contemporary numerical and mathematical notations. The fact that the Roman alphabet can be used to read and write IEML expressions is mainly a user interface device intended to exploit the availability of keyboards and the existing habits of a majority of potential users.

Secondly, natural languages have not evolved to be processed by machines but by human brains in cultural, social and emotional contexts. Unlike them, the syntax of IEML is essentially regular and IEML semantics harnesses as much as possible this regularity: this is what makes IEML semantics “computable”. But, at the same time, IEML can be understood and processed by human brains (of trained users). Being both semantic content oriented (not just a pure data format for logical inference) and computable (it has a regular syntax and a formally definable semantics), IEML is a putative language for the interface between the human mind and the computing power of cyberspace. Direct human manipulation of IEML expressions will be the province of semantic engineers, information architects, specialists and power users. The general public can reap the benefits of IEML coding while interacting with multimedia and/or natural languages representations.

By the way, Esperanto and other artificial languages with universal intent have been designed to be spoken and, in general, to be used like natural languages. The case of IEML is completely different. It is neither supposed to be spoken, nor to be learned, written and read directly by its end-users. Only specialists (information architects, knowledge engineers...) need to learn IEML. Besides, Esperanto and other languages of the same kind are neither regular languages, nor groups of transformations (in the mathematical sense) and are not designed to be manipulated automatically by machines. As it is designed from the beginning to be manipulated automatically, IEML can be translated automatically in different natural languages. It is also fit for the measurement of semantic distances, the calculus of intersections, unions and differences between complex concepts coded in USLs, and for the automatic generation of meta-texts (USLs translatable in natural languages) and networks of meta-texts.

In the rest of this paper, I will refer to any descriptor of an IEML expression in natural language by the expression: **tag* (STAR-tag).

2.2 Differences between IEML and OWL

In my understanding, the main purpose of the Web of data is to decompartmentalize databases on the Web. As a standard of the Web of data, OWL allows *automatic reasoning on the basis of data* across the Web, provided that these data conform to the same ontology (or to a mapping between the ontologies). Once expressed in OWL, an ontology can become the basis for automatic inferences on the Web. OWL is the continuation of previous work in knowledge representation by researchers in artificial intelligence. As a knowledge representation format, OWL has no content in itself. One must specify and use a particular ontology before anything can be meaningfully represented. In a way, OWL can be seen as a *file format* for description logics. The description logic is itself based on first order logic, one of the favorite formal tools of artificial intelligence. Therefore, by their very nature, OWL and IEML are two very different kinds of languages. Even if IEML has a formal syntactic structure and even if an IEML dictionary can be described as a formal network of concepts, IEML is neither a data format nor a programming language but a full language that has not only a grammar with inflections but also a whole dictionary of nouns and verbs that allows the construction of phrases. It is indeed possible to use IEML inside the angle brackets of OWL ontologies (and so to decompartmentalize distinct ontologies), or to describe in OWL the complex network of concepts of the IEML syntax and dictionary (and so to automatize reasoning based on information coded in IEML).¹ Moreover, the IEML tags on the Web will be coded in RDFa.

¹The international research network on IEML semantic space is already working on such an ontology.

At this point, some readers may insist and say that IEML is “nothing more than another super-ontology”. To understand why this is not the case, they can consider the fact that the dictionary of any natural language can be organized like an ontology, or at least as a formalized network of concepts. This is exemplified by Wordnet[21] that is something like “the explicitation of the ontology inherent to the English language”. But because the dictionary of a language can be considered or constructed as an ontology, it does not follow that a language is “nothing more than an ontology”. A language (of the semantic content oriented kind, including IEML and natural languages) is much more than an ontology because its rules allow for the construction of phrases and texts that can express endlessly any kind of meaning. An ontology is a graph, a language is a generative machine.

In conclusion, OWL and IEML have different natures (as shown here) and different goals (as will be shown in the following section), but they can - and will be - be interfaced.

3 Toward Semantic Space

3.1 Solving the Semantic Interoperability Problem

The IEML research program shares an important goal with the Web of data: it aims at decompartmentalizing online information. But its approach is not to standardize data formats since this is already done by the W3C and other standardization institutions. There are indeed *other obstacles* to information exchange and human collective intelligence in cyberspace *than diverse data formats*: diverse ontologies reflecting different contexts and area of practice, diverse classification systems, diverse folksonomies emerging from social tagging in various social media [25] and, last but not least, multiple natural languages. The ambition of the IEML research program is not to impose some unique super-ontology (since IEML can be used in the context of ontologies with very different hierarchies of concepts) but rather to provide a pivotal language into which at least the higher classes of any ontology and classification system as well as the most popular tags of user-produced folksonomies can be translated. As the IEML dictionary can provide *tags for diverse natural languages, IEML could also be used to establish automatic correspondences between *tags expressed in different natural languages. Once expressed in IEML, a complex concept - the meaning of a *tag - can be automatically translated to any natural language supported by the IEML dictionary.

In addition to transparent information exchange and semantic interoperability, the IEML research program pursues other interdependent goals that can be divided into three main categories: 1) the establishment of a consistent and transparent semantic addressing system, 2) the development of a powerful notation of meaning for exploiting the computing power and the immense storage capacity of cyberspace, 3) the provision of a symbolic tool for self-observation and self-reference of collective intelligence.

3.2 Opaque or Transparent Semantic Addressing System ?

IEML expressions are not only meaningful and translatable automatically to any natural language but, because of their very nature, they also form a *group of transformations* involving unions, intersections, differences and diverse complex algebraic operations [48]. The good news is that, in IEML, these automatable algebraic transformations on symbols correspond to automatable algebraic transformations on significations (on “semantics”). This property opens the way toward a potentially universal “semantic space” structured by a group of transformations in which each distinct IEML expression play the role of a distinct semantic address.

One of the limitations of the current Web of data is the opacity by design² of its ultimate universal addressing system: the URIs. As an abstract addressing system, the IEML semantic space is in principle independent of the URI address space just as it is independent of any physical or telecommunication addressing system. However, from a synergistic perspective, IEML can bring to the system of URIs a general semantic interconnection and a full group of transformations on semantics. Any IEML expression can indeed be represented by a URI in a straightforward manner, simply by taking any prefix and then adding a URI-encoded syntactic representation of the IEML expression to it. Thus, such IEML-URIs can be directly used as concepts in RDF or OWL and it can offer an alternative grounding to the entities of the Web of data, mapping URIs to such IEML-URIs. Therefore the power of IEML can be leveraged by the existing standards of the Web of data. Symmetrically, the expressive and algebraic properties of

²See:<http://www.w3.org/TR/webarch/#uri-opacity>

Cyberspace Evolution

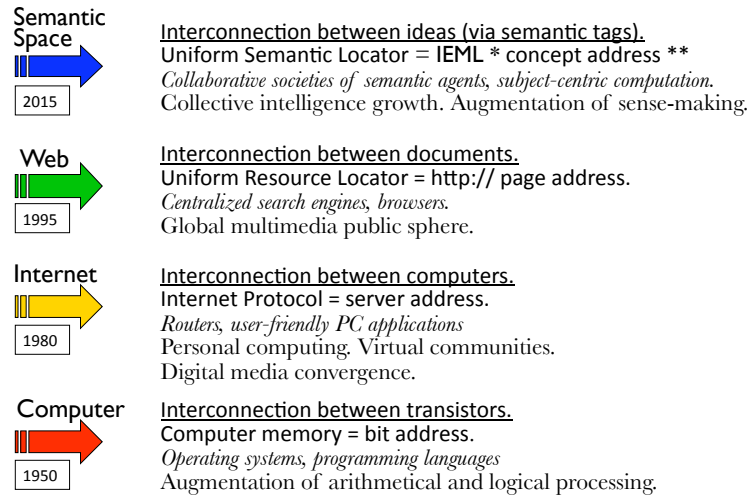


Figure 1: Semantic Space Addressing Layer in the Evolution of Cyberspace

IEML can leverage the current Web of data by providing it with a novel grounding that can make it more "semantic".

3.3 Successive Addressing Layers in the Evolution of Cyberspace

In order to understand the germane nature of the IEML research program, it helps to view it in the context of a progressive build-up and usage of the digital memory at the service of human collective intelligence. As illustrated in Figure 1, I see cyberspace being constructed from addressing layers, each depending on the one directly below it. In Figure 1, for each layer, the underlined text describes the type of interconnection, the second line states the addressing system itself, the text in italic suggests the main applications or technical tools developed on this layer and the last lines evoke the principal socio-cultural effects.

The first layer concerns the information bits contained and addressed in a computer's memory. The key point here is that the data and the programs are addressed - in principle - by the same storage system. This design has already been anticipated in the early works of Turing (in the concept of the *universal* Turing machine, the programs and the data are recorded on the same abstract tape)[72], as well as in the first plans of the EDVAC of von Neumann [73, 14]. One of the main purposes of operating systems is to manage the memory of the computers at the bit level.

The second layer, called the Internet layer, identifies the servers in digital networks. The Internet Protocol is an addressing system - now universal - that allows the communication between computers or between digital processing devices. Technically, the Internet was introduced at the end of the sixties, but its growth and wide public adoption dates back to the 1980s, in parallel with the progressive adoption of personal computers by the public. The Internet and the PC allowed networks of users and institutions to participate in cyberspace by contributing information and by being able to navigate in it.

The third layer, the Web layer, is made of the addresses of the "pages" of hyperdocuments and, by the same token, allows the identification of hyperlinks between those pages [4]. Even if the principle of HTTP was invented in 1991, the Web became a public phenomenon around 1995 thanks to the availability of browsers and numerous websites. The direct consequence of this last layer is the emergence of a new global multimedia public sphere. The recent growth in social computing and social tagging, experimentations in large-scale automatic natural language processing[21], as well as continuous research and development in the "Web of data" (or the so-called semantic web)[27, 20], is indeed part of the Web layer but is quick to point towards the next addressing layer, related to semantic computing.

The objective of the IEML research program is to establish this new addressing layer, where significations, or subjects, have addresses ("subject-centric computing" [?]). As suggested in Figure 1, in semantic space, USLs (Uniform Semantic Locators) are used to uniquely address distinct significations. IEML (Information Economy MetaLanguage) is the finite regular language that I have invented for the notation of USLs.

Every USL in IEML can be considered as a "point" in an abstract semantic space. When it can be translated into natural languages, any point of the IEML semantic space is also a distinct meta-text. The IEML semantic space is a "group of transformations" where each point is a functional transformation of any other point. Unlike in usual Euclidian space, each point of the semantic space (each USL) is a qualitative singularity, the mathematical notation of a configuration of ideas. Figures, graphs and complex dynamics can be drawn in semantic space by computable functions. The IEML semantic space can be used as a *coordinate system* to represent economic, social and cultural phenomena. More generally, semantic space can be considered as the abstract "place" where human collective intelligence is unfolding.

Today, there are several graphic representations of the flows of information in the internet, but the positions of the nodes refer to geographic positions and the size of channels represent quantities. When we represent semantic networks it is always from an ad hoc point of view, in the limited context of one particular natural language, classification or ontology. Because we lack a common semantic coordinate system, we are neither able to track the information flows between subjects, ideas or disciplines, nor the complex evolutionary patterns of meaningful information exchange at a global interdisciplinary scale. In the future, thank to IEML, it will become possible to track the *topics* of the flows of information in cyberspace in a common semantic framework, whatever the natural language of the users. The global computing network will function as a "mirror" of human collective intelligence at any scale. I think that the real "singularity"³ in human history will not come from some super-intelligent self-programming software but from the ability of human collective intelligence to become reflexive in cyberspace. At this time, we will witness a scientific revolution in humanities and social sciences based in the cultural maturity of cyberspace. This future revolution could be compared to the scientific revolution in natural sciences that unfolded in the 17th century after the invention and growing social use of the printing press.

Part II

IEML in the Context of Humanities and Social Sciences

4 The Role of Media and Symbolic Systems in Cognition

There is no doubt that human cognition is grounded in a brain structure and neural activity that is biologically determined[10, 18].⁴ Nevertheless, in the recent decades, an impressive body of research has been devoted to the subject of intellectual technologies and symbolic tools [30, 22, 58, 23, 5, 26, 77, 8, 51]. The main idea behind this interdisciplinary research is that culture-driven collective memory apparatus, communication media and symbolic systems all play a prominent role in shaping personal and social cognitive abilities [28].

The invention of writing has allowed the development of systematically organized knowledge (lists, tables, archives, accountancy, complex hermeneutical procedures) beyond the lore of oral cultures, organized around myth, narratives and rituals [22, 58, 23, 26]. The invention of phonetic, or alphabetic, writing systems limited to more or less thirty signs (as opposed to systems requiring thousands of ideographic signs or mixed systems) led to the social extension of writing and reading abilities and fostered the development of abstract conceptual thinking [29, 55, 50]. The invention of the indo-arabic numerals including notation by position and the zero, made arithmetic simpler and easier mainly by allowing the use of uniform algorithms [34]. Making a multiplication using roman numerals instead of using indo-arabic numerals is to understand the importance of notations and symbolic systems in the performance of cognitive tasks. In addition to supporting a wide dissemination of information and knowledge, the invention of the printing press led indirectly to the development of several standard scientific notation systems including accurate

³See the Wikipedia entry for "technological singularity" http://en.wikipedia.org/wiki/Technological_singularity

⁴Note that I don't say that human cognition is *determined* by neural activity, but that it is *grounded* in neural activity.

maps with geometric projections of parallels and meridians, biological classification systems, chemical and mathematical notations [19]. Print technology also fostered the development and progressive formalization of linguistic studies [1] and the creation of metadata systems for the organization of libraries and archives [69]. Note that the development of new symbolic systems did not show up immediately after the invention of the printing press: it took several generations to assimilate and exploit the cognitive possibilities opened by this new medium. In general, cultural evolution follows technological evolution. By analogy, we can easily imagine that the full symbolic exploitation of the new communication and processing environment provided by computer networks has not been achieved.

Even if these historical remarks may seem remote from the central point of this paper, they are indeed very close. They suggest that major advances in human cognition are related to inventions of media and symbolic systems. IEML is precisely a symbolic system designed to exploit the cognitive possibilities - and particularly the avenues for collective intelligence - opened up by the new digital medium.

5 Empowering Readers and Writers in the Digital Realm

From calligraphy to hermeneutics, the humanities have always been related to the arts of reading and writing. It is an understatement to say that digital technologies are transforming these ancient arts. In one or two generations, “digital humanities”⁵ will become a pleonasm. There will be no other humanities than digital because the digital medium will ensure an unprecedented power to the operations of analysis, synthesis, search, filtering and information extraction from the global memory. These operations will be augmented by a mix of direct automatic computing and collaborative / social computing. On the creative side, document generation and immersive environment design will be further automated and made even more collaborative than today. The distinction between author, editor, publisher, critique (assessment) and librarian (categorization) will continue to blur. The role of IEML has to be thought in the context of this *generalized digital humanities*.

By contrast with natural languages, the syntax and semantics of IEML obey to strict syntactic rules (without exceptions) and offer a group of algebraic transformations that enable composition of complex functions. IEML makes it easy to program *writing functions*, i.e. automatic generation of IEML expressions, instead of writing one expression after another. This feature could be used, for example, in the context of hypermedia digital story-telling or game scenario meta-programming. For the same reasons, it becomes possible to assist IEML users to program *reading functions* through IEML-tagged documents or databases. These reading functions can augment navigation, analysis, synthesis, interpretation, filtering and connection to relevant information. Once they are programmed, these reading and writing functions can be exchanged, reused and modified by IEML users. They will pave the way for a new kind of search, both collaborative and decentralized, boosting a trend already perceptible in the contemporary social Web. Moreover, applications based on IEML could provide some help for the definition of problems and the design of complex systems⁶ by fostering the exploration of semantic blind spots and suggesting ideas.

6 Setting the Stage for Collective intelligence Self-Awareness

Human collective intelligence in cyberspace is still in its infancy. In my judgement, the main obstacle that prevents human collective intelligence from crossing the next cognitive threshold is the current absence of systematic self-awareness. The ethical drive behind the IEML research program is the hypothesis according to which crossing the threshold of *reflexive* collective intelligence will lead to a blooming of human development (economic prosperity, health, education, transmission of cultural heritages, human rights, control of ecosystems, peace, security, research, innovation...). According to this hypothesis, the highest goal of the IEML research program is to provide a symbolic framework for the making of digital tools that can help human collective intelligence observe its own activity in cyberspace, and therefore improve human development.

Ideally, collective intelligence self-awareness can be reached from various semantic perspectives, at various time-scales and for various communities both large and small. This is why IEML has been conceived

⁵http://en.wikipedia.org/wiki/Digital_humanities

⁶Problem definition goes beyond problem solving and design implies more than getting answers to questions.

primarily for information synthesis rather than for automatic reasoning.⁷ More precisely any piece of IEML-tagged information will always be understood or visualized as a form on a background. The *form* is a set of semantic addresses, possibly endowed with weights or quantities (computed from truth values, statistics, votes, prices...), and the *background* is a zone of the semantic space that contains or displays the form.⁸ As forms change over time, it becomes possible to represent, analyze, synthesize, compare and simulate the dynamics of collective intelligence in semantic space. The fact that collective intelligence reflexivity is a priority of the IEML research program has an important consequence for the design of the IEML dictionary. IEML and its current lexicon is much more oriented toward the representation of human affairs, broadly the objects of humanities and social sciences, than toward the representation of the objects of the natural sciences. Furthermore, because of its fragmentation of disciplines, theories and methodologies, it is probably the humanities and social sciences that are more in need of a common and open semantic framework to exploit collaboratively the interconnected digital memory of human race.⁹

7 The Future Scientific Revolution in Humanities and Social Sciences

As a growing proportion of human interaction, communication and memory use the medium of cyberspace, it becomes in principle feasible to ground interdisciplinary research in social sciences and humanities in a common body of digital data. But as different disciplines and even different schools in the same disciplines have incompatible theoretical frameworks, the new opportunities offered by the extension of cyberspace for the study of human or cultural phenomena are not yet fully exploited.

Between the 16th and 20th centuries, the natural sciences acquired a unique and infinite physical space [?], equipped with a system of universal coordinates and units of measurement. The observational instruments in the natural sciences today are very elaborate in their engineering, and undergo constant progress. The symbolic and conceptual instruments of natural sciences is highly formalized, logically coherent, and largely shared within the scientific community. Mathematicians have their sets, relations, numbers, and functions. Physicists have their mass, energy, and particles. Chemists have their elements, molecules and reactions. Biologists have their biomolecules, DNA, and their intracellular and intercellular pathways of exchange. The theories may abound and diverge, but the language, just like the system of coordinates and measures, remains common to them all, enabling dialogue, controlled testing and an articulated accumulation of discoveries. In terms of knowledge management, we can say that natural sciences have been successful in making a significant portion of their knowledge explicit, so that it could be shared and thus offering mutual enrichment.

By contrast, the humanities and social sciences do not share a cultural space that is unique, infinite, coordinated and measurable. The disciplines are fragmented. Within those disciplines, conflicts between paradigms often limit fruitful dialogue. It is sometimes even difficult to agree on the nature of the disagreements. The observational instruments are not well developed in terms of engineering. Except in certain highly formalized sub-disciplines, the calculability, predictive capacity, and testability of the theories are weak. The main consequences of this situation is that the greater part of the considerable knowledge accumulated by the community of researchers in the humanities and social sciences remains “implicit”. That is to say, in terms of knowledge management, the knowledge and expertise accumulated by the humanities are difficult to share in contexts that differ from the initial environment in which they emerged. And yet, the resolution for the difficult problems confronting contemporary humanity demand the effective collaboration of all cultural sciences.

The notion of collective intelligence is a good candidate for the federation of humanistic disciplines and social sciences in a joined effort to make the most of this new universal digital memory. There are at least three reasons to support this claim. First, collective intelligence is the basic engine of human development. Secondly, a comprehensive model of CI can involve any dimension of human society and culture in a scalable, dynamic and holistic way. Thirdly, the objects of social and humanistic sciences are also personal and collective subjects, or expressions of subjects. So, a collective intelligence theoretical framework is a good fit to tackle the subjective, intersubjective and hermeneutical nature of the objects

⁷This does not prevent automatic reasoning in IEML, based on hierarchies of classes and sub-classes, symmetry groups, grammatical roles and other computable or explicit semantic relationships.

⁸This background is named “category” in the technical vocabulary of IEML syntax (see Section 12).

⁹See, for example, the international research project “Digging into Data” at <http://www.diggingintodata.org/>

of human sciences, especially if it is constructed from the beginning to describe distributed, fractal and self-referential social cognition.

CI processes should be observed scientifically, and the results of this observation should help various communities to improve dynamically and collaboratively their own collective intelligence. My claim is that this goal cannot be met in the absence of a common semantic coordinate system. This system must be created in a such way as to permit various communities, networks and institutions to express their dynamic collective intelligence activity without forcing them to conform to a unique mould, but allowing them comparisons and exchanges. In this perspective, IEML semantic space can be considered as a conventional geometric framework for the observation of the (practically) infinite abstract universe where collective intelligence processes occur and intertwine. As the printing press, along with new instruments of measurement and observation of the physical world, were the basis of the revolution in natural sciences, I think that the digital medium, along with new instruments of measurement and observation of the information economy in cyberspace, will be the basis of a revolution in cultural sciences. This is why the IEML research program should bring together computer scientists and scholars from humanistic disciplines.

The IEML research program will enhance human collective intelligence in cyberspace and contribute to general human development if it is able to achieve all of its interdependent goals in a synergistic way: empowering social categorization of the global digital memory, solving the semantic interoperability problem, providing a transparent semantic addressing system (both meaningful and computable), enhancing the impact of collaborative writing and reading in cyberspace and setting the (symbolic) stage for self-aware collective intelligence.

Part III

IEML

8 Introduction to IEML Notation and Syntax

8.1 STAR-IEML Notation

There are several ways to represent IEML expressions. *Binary IEML*¹⁰ is for automatic computation. *XML IEML* is for standard exchange between Web services. *STAR-IEML* is for manipulation by human beings (STAR stands for: Symbolic Tool for Augmented Reasoning). In this paper, IEML expressions will be generally written in STAR.

Expressions are represented in STAR-IEML with one star at the beginning and two stars at the end as delimiters of IEML expressions. Everything that is between these two delimiters can be parsed and translated into other IEML representations at <http://starparsr.ieml.org/>.

In STAR-IEML, a capital letter followed by a colon represents an expression at layer zero: these symbols can be assembled into higher, more complex layers of expression.

8.2 *Tags

The star before an English expression marks a *tag, a natural language descriptor of an IEML expression. A *tag holds the place of an IEML expression by *suggesting* its meaning rather than uttering the IEML expression. The meaning of a *tag has to be understood from the singular place that its corresponding IEML expression occupies into the network of IEML semantic relations.

8.3 Syntax Overview

All IEML expressions are built from a syntactically regular combination of six symbols, called its *primitives*.

In IEML a *sequence* is a succession of 3^λ single primitives, where $\lambda = (0, 1, 2, 3, 4, 5, 6)$. λ is called the *layer* of a sequence. For each layer, the sequences have respectively a length of 1, 3, 9, 27, 81, 243 and 729 primitives.

¹⁰See an early version of binary IEML at: <http://www.ieml.org/spip.php?article138>

From a syntactic point of view, any IEML expression is nothing else than a *set of sequences*. As there is a distinct semantic for each distinct sequence, there is also a *distinct semantic for each distinct set of sequences*. The main result is that any algebraic operation that can be made on sets in general can also be made on semantics (significations) once they are expressed in IEML. This is why IEML semantic space (i.e. the set of all sets of IEML sequences) is a *group of transformations*. The dictionary that assigns semantics to IEML expressions by providing the correspondence between IEML sequences and *tags is arranged in order to take full advantage of this feature.

9 Primitive Distinctions

IEML is made of a small number of building blocks, six primitive symbols, that are combined and recombined in a regular way in successive layers of composition. I came to these six primitives after twenty years of reading and meditation on meaning, language and symbolic manipulation. I can only give here an abstract of the result instead of a detailed account of all my explorations of the problem. The following exposition is organized as a succession of basic distinctions.

9.1 Empty and Full

First, in any expression of signification we must be able to distinguish between something (whatever it may be) and nothing, because some syntactic roles must sometimes be left unplayed. Therefore, we need a symbol for blank or empty meaning, just as there are symbols for silence in musical notations and a zero symbol in arithmetic. So the first semantic dialectic of IEML will be between *emptiness, noted *E:** and *fullness, noted *F:** in STAR.

The fact that there is nothing, or that a syntactic role is unplayed, is of important information. Emptiness *E:** does not “mean nothing”, on the contrary, it means “nothing”, which is a subtle, but major distinction. The possibility that any primitive may play a given role is noted *I:**, *information.

$$*I:** = *(E: + F:)**$$

9.2 Verbs and Nouns

Now, we must make a distinction within *fullness. The terms used in natural language expressions are traditionally divided between verbs and nouns, and a similar distinction can be made (in cognitive terms) between entities and processes. This distinction is very basic in human cognition; it exists in all natural languages. Moreover, the verb/noun distinction is not only useful at the lexical level but also at the propositional level: there are verb phrases and noun phrases. So this distinction is embedded in the deep structure of IEML. That is why *fullness is organized by a dialectic of *verbishness, noted *O:** and *nounishness, noted *M:**. In STAR, every verbal sequence begins by a vowel (*O:** being the prototypic vowel) and every nominal sequence begins by a consonant (*M:** being the prototypic consonant).

$$*F:** = *(O: + M:)**$$

9.3 Virtual and Actual

If we try to synthesize the dialectic of process (*O:**), we can say that it is very often organized around a movement or a circulation of energy between two poles. Various cultural traditions have named these two poles differently, but the bipolarity itself is fairly universal: yin and yang, negative and positive, intelligible and perceptible, transcendence and immanence, mind and body, type and token, etc. These two poles are represented in IEML by *U:** (*virtual) and *A:** (*actual).¹¹ As we are speaking about verbs, we can also describe *U:** and *A:** like two symmetric *processes* of virtualization and actualization. The IEML dictionary will take advantage of the *virtual / *actual dialectic to represent every semantic structure that can be organized by a bipolar symmetry (male and female, singular and plural, and so on). Note that *U:** and *A:** are vowels.

$$*O:** = *(U: + A:)**$$

¹¹See [43] for a philosophical account of the virtual / actual dialectic

9.4 Sign, Being and Thing

What are the inner distinctions of the nounish (*M**) realm? As verbishness (*O**) shelters bi-polarities, nounishness (*M:**) covers ternarities. The main ternarity lies in the structure of semiotic representation which needs three interdependent poles.

The first pole may be called the *signifier*, for example the sound or the visual representation of the noun “car”.

The second pole is the *signified for an interpreter*: the notion of car. We need a (generally human) mind for the interpretation of a linguistic signifier. There is no “notion” outside the context of a living cognitive system able to manipulate symbols. For example, English signifiers do not mean anything for ants or for the sun. Despite the fact that we cannot have access to any signified without any representation of it, signifieds exist. A signified is of course something abstract, a concept, that is different from the perceptible signifier. As an illustration, we can say that “car”, in English, has the *same* meaning (or signified) as “voiture” in French. If there can be two (or more) signifiers for one single signified, signified and signifier must be different entities.

The third pole is the *referent*, the actual thing that is referred to by the signifier: real cars in general, or this particular car, that have weight, a price on the market and can be driven.

Some versions of the semiotic process use only two terms. By contrast, I adopt here a ternary semiotic that includes the signified for an interpreter (the *sense* according to Frege [31]), and the referent (some objective context). This basic decision has deep philosophical implications. Some reference-oriented semantic theories consider only signifiers and referents, bypassing the interpretation of symbols to explain their ability to represent something. My inclusion of the interpreter asserts the necessity of a mind for any full account of the signification process. This choice opens the door for freedom and dissent in interpretation.

Saussure’s symbolic duality [63], that includes only signifier and signified, works well for the words of a dictionary, but not for full enunciations in context. The latter have not only a conceptual meaning but also an effective reference. My inclusion of a referent pole takes into account performative effects in context and some actual existence outside of the linguistic or logical realm.

The tradition of ternary semiotics begins with Aristotle¹² and is developed by medieval philosophers (*vox, conceptus, res*) from various linguistic and religious backgrounds [62, 59, 15]. It is pursued by classical philosophers or grammarians and finally refined by C. S. Peirce [61] and contemporary linguists.

This ternary dialectic¹³ is expressed in IEML by three primitive symbols: *sign (or signifier) *S:**, being (or signified for an interpreter) *B:** and *thing (or referent) *T:**. In general, *sign will be used to connote symbols, documents, language, and representations; *being to connote people, personality, interiority and community and finally *thing to connote efficiency, power, technology, reality and the material world. We can also use *M:** in IEML to mark triadic symmetries that cannot be derived from the initial semiotic ternarity like: future (*S:**), present (*B:**) and past (*T:**) for tenses; or first (*S:**), second (*B:**) and third (*T:**) persons for verbs and pronouns inflections.

Every semantic structure that can be organized by a ternary symmetry will take advantage of the *sign / *being / *thing dialectic to be represented in IEML. For example: proposition (*S:**), judgement (*B:**) and state of things (*T:**) in logic; legislative (*S:**), judiciary (*B:**) and executive (*T:**) in politics; price (*S:**), property (*B:**) and utility (*T:**) in economy; teachings (*S:**), community (*B:**) and ultimate reality (*T:**) in religion, etc.

$$*M:** = * (S: + B: + T:) **$$

9.5 Six Primitives and Four Primitive Dialectics

All the basic distinctions and structural symmetries that I have just explained can be expressed with only six primitives: *T:**, *B:**, *S:**, *A:**, *U:** and *E:**. The four other symbols introduced so far (*I:**, *F:**, *O:** and *M:**) are only abbreviations denoting sets of primitives, as shown in Figure 2

The combining, alternating and permutation of emptiness/fullness dialectic (*E:** / *F:**), virtual/actual dialectic (*O:**), sign/being/thing dialectic (*M:**), or even pentadic dialectic *(T: + B: + S: + A: + U:**) lead to an open-ended construction of regular structures and complex symmetries.

¹²See the very beginning of “On Interpretation” [2].

¹³Etymologically, the word “dialectic” comes from the Latin *dialectica* (art of reasoning with method) deriving from the Greek *dia-legein* (through-saying) that means “reasoning” and that is related to the notion of dialogue. So “dialectic” does not derive from any Greek or Latin root meaning “two” or “duality”. Therefore a ternary dialectic is indeed possible.

IEML PRIMITIVES

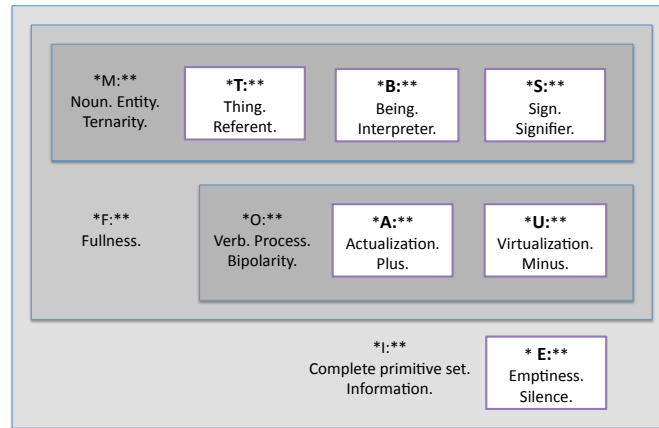


Figure 2: IEML primitives

10 Sequences

10.1 Recursive Construction of Sequences

Now that we have our six building blocks, the primitives, how are we going to assemble them in a meaningful way? The answer lies in two basic IEML notions: the recursive construction of six layers and the three syntactic roles. The primitive building blocs (sequences of layer 0) are used to construct building blocks of layer one: sequences of three primitives. Then, building blocks of layer one are used to construct building blocks of layer two: sequences of three sequences of layer one, that is to say successions of nine primitives. The same process is reiterated until layer six, where sequences are made from three successive sequences of layer five, that is to say successions of 729 primitives.

Why is there only seven layers? The terms of the dictionary belong to layers 0 to 3. There are rules to create inflected words from these terms, to create sentences from inflected words and to create relations between sentences by using some terms as conjunctions. Given these rules, it is possible to express any *network of relations between sentences* by using sequences up to layer 6. As USLs are not supposed to represent concepts so complex that they cannot be expressed by networks of relations between sentences, there is no need for an eighth layer. Note that USLs can contain as many sequences as needed to express *several* networks of relations between sentences and that more complex meanings (like relations between networks of sentences) can be expressed by networks of USLs.

10.1.1 Examples of Sequences

- ***B:B:E:**** and ***A:S:E:**** are sequences of layer 1
- ***B:B:E:A:S:E:E:E:E:-**** and ***A:S:E:B:B:E:E:E:E:-**** are sequences of layer 2
- ***S:S:E:U:B:E:E:E:E:-B:B:E:U:B:E:E:E:E:-S:S:E:A:S:E:E:E:E:-**** is a sequence of layer 3

These sequences are not given for comment and interpretation of their meaning. Before any explanation of the semantics of IEML sequences can be understood, the pre-requisite notions of seme, layer and etymologic relation have to be assimilated. For the moment, all that is required from the reader is to observe the structure of these sequences.

10.1.2 Layer Marks

In STAR-IEML notation, terminating punctuation marks, called *layer marks*, indicate the layer of a sequence. So, in STAR-IEML, the letters manifest the semantic content of a sequence while punctuation marks show the interlocking of layers and the syntactic structure of a sequence. Here is the meaning of all layer marks in STAR-IEML.

- : indicates the end of an expression of a sequence of layer 0
- . indicates the end of an expression of a sequence of layer 1
- indicates the end of an expression of a sequence of layer 2
- ' indicates the end of an expression of a sequence of layer 3
- , indicates the end of an expression of a sequence of layer 4
- _ indicates the end of an expression of a sequence of layer 5
- ; indicates the end of an expression of a sequence of layer 6

10.1.3 Semantics of the Layers and Morphological Rules

In IEML, the lower the layer of a sequence, the more general or undefined its semantics. Conversely, the higher the layer of a sequence, the more detailed and precise its semantics. The general principle is that the user of the metalanguage can add specificity and complexity to an expression by augmenting its length. So, the most basic verbish and nounish notions belong to layer 1. Roots of verbs and nouns, or short verbs and nouns, as well as pronouns, adverbs, prepositions and inflections (cases, conjugations) belong to layer 2. Complex conjunctions and long word-roots belong to layer 3. Layer 4 is for complex inflected words and simple sentences. Layer 5 is for complex sentences or relations between simple sentences and layer 6 for relations between complex sentences. The complete set of rules for the formation of inflected words, sentences and relations between sentences will not be explained here because of a lack of space. It will only sketch below the basic principles.

10.2 Syntactic Role Players: Semes

We can see in the previous examples that each sequence of layer n ($1 \leq n \leq 6$), is composed of three sequences of layer $n-1$. If each distinct sequence has a distinct meaning. *A:S:E:**, for example, will have a different semantic than *S:A:E:**. Therefore, the first, second and third sequences of layer $n-1$ that compose a sequence of layer n each play a different syntactic role. Since IEML is regular not only at the syntactic level but also at the semantic level, the three roles of the sequences have the same general semantic interpretation at all layers and regardless of the grammatical nature of the expression (verb phrases or noun phrases). In IEML, the syntactic role players of a sequence are called *semes*. I name “*substance*” the seme of the first role, “*attribute*” the seme of the second role and “*mode*”, the seme of the third role. Substance, attribute and mode are not to be taken here as ontological¹⁴ realities but as playings of syntactic roles since the same sequence of layer $n-1$ can play any combination of the first, second or third role in a sequence of layer n . Moreover, “substance”, in its technical IEML sense, is not to be interpreted as excluding the notion of process, since the substance of a sequence can indeed be a verb or a process.

10.2.1 First Seme: Substance

The substance determines the main subject and the grammatical orientation of the expression. If the first primitive of a substance belongs to *O:** = *(U: + A):**, the substance is a verb or a verbal root and the whole sequence itself is a verb or a verbal phrase. If the first primitive of a substance seme belongs to *M:** = *(S: + B: + T):**, the substance is a noun, a noun root or an adjective and the whole sequence itself is a noun or a noun phrase. Adverbs, prepositions, conjunctions, pronouns and inflections (cases and conjugations) begin by *E:**, their substance is “empty”.

¹⁴I use “ontology” here in the philosophical sense.

10.2.2 Second Seme: Attribute

The attribute is a quality, a feature of a complement of the substance.

In the terms of the dictionary, that go until layer 3, the substance indicates the main notion of the semantics of the sequence and the attribute indicates some variation on this notion.

In verb phrases, the attribute can be the subject, the object or any kind of complement of the verb that is in substance. The precise grammatical function of the attribute (in other words: the relation between the substance and the attribute) is indicated by the mode.

In noun phrases, the attribute can be the “attribute” (in the ordinary grammatical meaning of the term), a genitive of any complement of the noun in substance. Again, the exact grammatical function of the attribute is indicated by the mode.

10.2.3 Third Seme: Mode

In the *terms of the dictionary*, the mode is used to indicate variations on the meaning of a root in substance-attribute.

In the *construction of the words*, the mode is used to mark the inflections like gender, number, persons (first, second or third), moods, tenses, cases and so on.

In the *construction of sentences* the mode is used to define the grammatical relations between the substance and the attribute.

There are already more than 150 cases, conjugations, pronouns, adverbs and prepositions¹⁵ of layer 1, 2 and 3 in the current IEML dictionary. They have been designed to be used mainly as modes.

10.2.4 Basic Principle of the Construction of Phrases

The simple creation of a grammatical relation between a verb and its subject (for instance) is made in one single sequence, called a *clause*. For example:

wu.d.-d.a.-E:E:S:-’* is a sequence of layer 3, where the three semes of layer 2 are terms of the dictionary and constitute therefore undecomposable units of meaning. ***wu.d.-**** means “observe”, ***d.a.-**** means researcher and ***E:E:S:-**** marks the nominative case, indicating that ***d.a.-**** is the subject of ***wu.d.-****. So the whole clause means: “A researcher observes”. Note that, to keep the example simple, I did not add any inflection to the words. Also, as will be explained below, lowercase letters are used in STAR notation as abbreviations for sequences of uppercase primitives of layer 1. For instance ***a.** = *A:B:E.****

If I want to say now that a researcher observes *a regularity*, I have to use a second clause and to create a relation of *anaphor* between the two clauses. The anaphoric relation is created just by re-using the same verb in the second clause.

((wu.d.-d.a.-E:E:S:-’)+(wu.d.-t.we.-E:E:T:-’)* In the second clause, ***t.we.-**** means “regularity” and ***E:E:T:-**** marks the accusative, creating a verb-object relation between the substance and the attribute. The whole expression is a *phrase* made of two clauses in anaphoric relation that means: “a researcher observes a regularity”.

10.3 Etymological Relations

The IEML dictionary strives to derive the semantics of a sequence of layer n from the semantics of its semes of layer $n-1$: the substance, attribute and mode of the sequence. If a sequence’s meaning is built on the meaning of its role players, it is called a *derived* sequence. There is an *etymological relation* between a derived sequence of layer n and the semes of layer $n-1$ of which it is composed. In general, sequences of layer n ($n \geq 4$) should have an etymological relation with their semes. This means that the semantics of IEML “phrases” should be easily understood in terms of the meaning of their component “words” of layer 2 and 3, as modified by “inflections” of layer 1 or 2. At lower layers, many sequences of layer n ($1 \leq n \leq 3$) of the current dictionary have also an etymological relation with their non-empty semes. But it is not possible to derive etymologically all conceivable semantics from only six primitives, specially at layers 2 and 3: the “words”. The meaning of some sequences of layer n ($1 \leq n \leq 3$) is *not* derived from

¹⁵Post-position would be more accurate, since the mode comes after what it modifies

the meaning of their semes of layer $n-1$. These sequences are called *prime* sequences, by analogy with prime numbers, and serve as the origins of the semantics of the sequences at layer $n+1$ in which they are semes.

10.4 From Sequences to Sets of Sequences

IEML expressions in general consist of *sets* of sequences. So far I have been discussing the relatively simple cases of sets with a single member (or two members in the above example of a phrase, see Subsection 10.2.4). I will now explore *categories*, sets of sequences whose members are all at the same layer, along with an important kind of category, called *generated categories*, that can be expressed in a very compact way using IEML notations. Then I will introduce *ordered sets of categories of same layer*, called catsets. Finally, I will define USLs as sets of n ($1 \leq n \leq 7$) catsets of distinct layers, n corresponding to the number of IEML layers at which their are any sequences. A USL can consist of one single sequence or one single category, or one single catset. The USLs are the addresses of the “points” in the IEML semantic space. They can be processed as the operands and be the results of all kinds of selection operations including combinations of unions, intersections and differences. USLs can as well be processed as operands and results of an unlimited variety of automatable functions for USL generation and USL semantic networking.

11 Primitive Categories

11.1 Definition of Primitive Categories

First, let’s consider sets of sequences at layer 0. We know that sequences at layer 0 are composed of one single primitive. Since there are six primitives, there are $2^6 = 64$ possible subsets, including the null set \emptyset . This null set of primitives should not be confused with the empty primitive *E:** that means “emptiness”. The null set has no meaning at all, but it is needed for the group of transformations on IEML expressions. Each of the 63 meaningful subsets of primitives is called in IEML a *primitive category*. Every distinct primitive category has a distinct meaning. I cannot explain here the semantics of all of them, so I will only introduce six new primitive categories (see Figure 4) in addition to the ten primitive categories that we have already seen (see Figure 3).

11.2 Binary Semantic Characters

For purposes of computation, a primitive category can be represented in binary IEML by a *semantic character* that is a 6 bit unsigned integer (one bit per primitive, in the order: TBSAUE). For example, *A:** is expressed in binary notation by 000100, *O:** by 000110 and *(S: + U):** by 001010. Through binary IEML, there is a bijection between the set of primitive categories and the set of natural numbers from 1 to 63 ($1 \leq n \leq 63$). The order of the primitives (bits) in the binary semantic character is one of increasing “materiality” from right to left: *E:** is less material than *F:**, *O:** (processes) is less material than *M:** (entities), and inside each of these two last subsets, *U:** is less material than *A:**, *S:** is less material than *B:** that is itself less material than *T:**.

*T:** (100000), corresponding to binary $32 = 2^5$ is numerically more significant than *B:** (010000), that corresponds to binary $16 = 2^4$, and so on until *E:** (000001) that corresponds to binary $1 = 2^0$. So, on account of their numerical coding, primitives represented by larger numbers have more “material” semantics and vice versa. The dictionary is designed to arrange, as much as possible, a correspondence between the numerical value of semantic characters and the semantics of the categories in which they appear. This feature is used for an automatic ordering of IEML categories that has more semantic relevance than the traditional alphabetic order.

A semantic character can be represented as a hexagram following the rule represented in Figure 3. Starting from the bottom, each line of a hexagram symbolizes a primitive: 1) *thing, 2) *being, 3) *sign, 4) *actual, 5) *virtual, 6) *emptiness. Unbroken lines mean that the corresponding primitives are "ON" (corresponding to 1) and broken lines that the corresponding primitives are "OFF" (corresponding to 0).

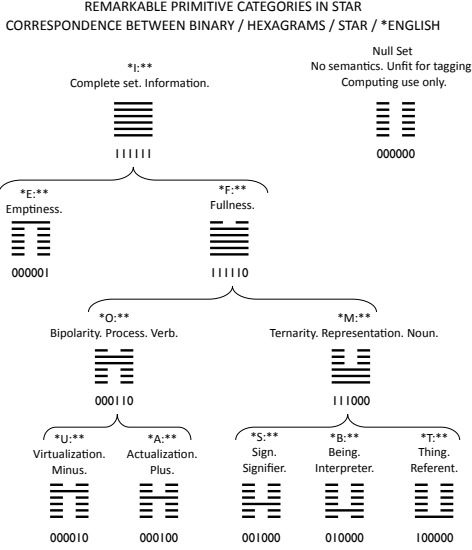


Figure 3: Ten Primitive Categories

11.3 The Six Networks of Collective intelligence

The six primitive categories: $\{*(U:+ S:)**, *(U:+ B:)**, *(U:+ T:)**, *(A:+ S:)**, *(A:+ B:)**, *(A:+ T:)**\}$ are called the six networks of collective intelligence. These primitive categories illustrate two important properties of IEML: first, that new semantic symmetries emerge from the combination of primitive dialectics, as stated in section 9.5, and second its ability to represent some models of collective intelligence, even at the primitive layer (layer 0).

Figure 4 highlight the symmetry of the relationship between two dialectics. Vertically, the virtual/actual bipolar dialectic juxtaposes and joins the two complementary 3-tuples: networks of documents/persons/bodies and networks of knowledge/will/power. Horizontally, the ternary dialectic sign/being/thing juxtaposes and joins the three complementary pairs: networks of documents/knowledge networks of persons/will and networks of bodies/power.

I have adopted here the network or actor-network theory [7, 36] that is broadly used in human and social sciences [24, 9, 75]. This theory leads to the integration of the mathematical tools of graph theory in the humanities. The diagram shows that a sustainable collective intelligence implies a continuous exchange of resources between the six networks of collective intelligence, that can also be called the six human development capitals.

12 Generated Categories

12.1 Definition of a Generative Operation

As already discussed, all sequences of layer n (except for layer 0 where a sequence is a single primitive) have three role players called semes: the first is the substance, the second is the attribute and the third is the mode. These semes are sequences of layer $n-1$. In this context, a generative operation is an operation that takes three sets of semes as its operands:

1. first operand: a category of substances,
2. second operand: a category of attributes,
3. third operand: a category of modes.

The result of a generative operation is the category of all distinct sequences of layer n that have a substance seme of layer $n-1$ from the first operand, an attribute seme of layer $n-1$ from the second operand and a

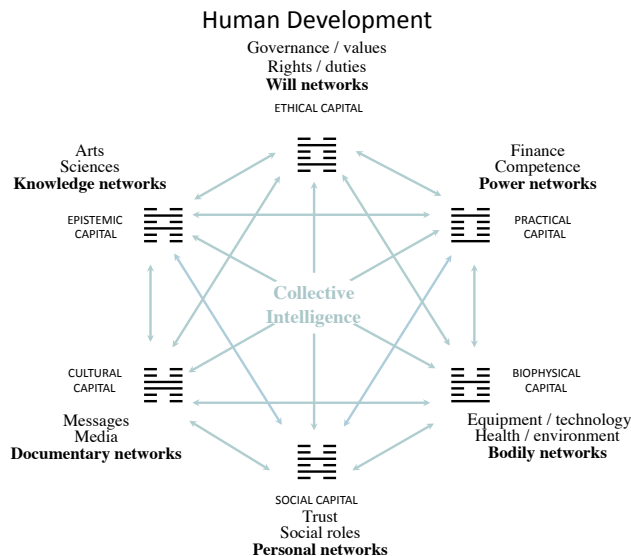


Figure 4: A Model of Collective Intelligence in the Service of Human Development

mode seme of layer $n - 1$ from the third operand. The generative operation can be thought of as a cubic Cartesian product of three categories of layer $n - 1$ that produces a category of layer n . I have also shown in [48] that a generative operation can be expressed formally as a finite automaton.

In the technical vocabulary of IEML, a *generated category* is a category that can be expressed as the result of a generative operation. *Sequence* can now be redefined as a particular case of a category; it is a generated category that has only one member.

12.2 The Generated Category as a Notation Tool

This notion of generated category plays an important role in IEML. First, instead of writing explicitly the set of sequences that compose a category, it is possible to write it compactly as a generative operation. In STAR notation, the operator of the generative operation is implicit. For example, $*(U:S:E:. + U:B:E:. + U:T:E:. + A:S:E:. + A:B:E:. + A:T:E:.)^{**}$ can be noted compactly as: $*O:M:E:.*^{**}$ (See Figure 5).

Not all categories can be expressed as the result of a generative operation. For example: $*U:S:E:. + S:A:T:.*^{**}$ is a set of sequences that cannot be the result of a generative operation, consequently it is a *composed category*. Typically, a composed category is expressed as a union (noted “ + ”) of generated categories.

12.3 Extension of the Notion of Seme

As already discussed, when a sequence acts as syntactic role player, it is called a *seme*. When a category acts as a role player, it is also called a seme. In a generated category, a seme is a set of sequence-semes that play identical syntactic roles. For example, in $*O:M:E:.*^{**}$, $*O:.*^{**}$ is the substance seme, $*M:.*^{**}$ is the attribute seme and $*E:.*^{**}$ is the mode seme. The $*E:.*^{**}$ seme contains only one sequence.

Figure 5 shows the workings of a generative operation in the STAR, hexagrammatic and binary representations. In the hexagrammatic diagram, the set of generated sequences is represented as the set of threads that exhaust all possible paths between the three semes by crossing one “ON” primitive.

12.4 The Generated Category as a Tool for Semantic Coherence

Apart from compactness of notation, another reason why the notion of generated category plays an important role in IEML is that every sequence translated by a **tag* in the IEML dictionary is a member of several interlocked generated categories. The meaning of these categories, that can also be expressed in

Differents Representations of a Generated Category

*O:M:E:** = *(U:S:E:. + A:S:E:. + U:B:E:. + A:B:E:. + U:T:E:. + A:T:E:.)**

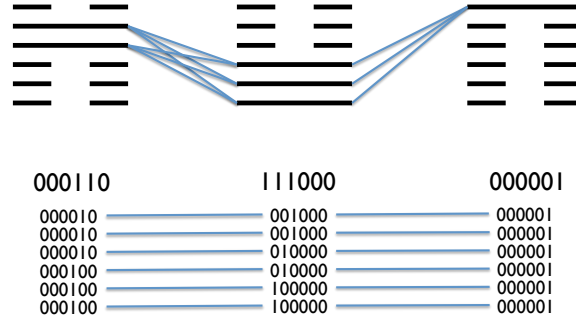


Figure 5: Generated Category and Generative Operation

natural language by *tags, is equivalent to the union of the meanings of the *tags of its sequence members. This fundamental feature of the IEML dictionary will be abundantly illustrated below. The whole IEML dictionary is *semantically* organized by inclusion and intersection between generated categories. This is why, even in case of prime categories (see sub-section 10.3), relationships between categories and their *tags are not arbitrary. Semantic networks of *tags are structured by networks of syntactic relations between categories. Thanks to its organization by generated categories, the IEML dictionary is not a long chaotic list of point-to-point relations between sequences and their *tags. On the contrary, it is designed from the start as a complete *relation between structures*. Moreover, a generated category (that can always be displayed as a 2D or 3D matrix) can be used as a background or as a kind of coordinate system to locate subcategories or sets of sequences. This feature is very important for the synthetic function of IEML as well as for its function as a design and analysis tool, since every particular sequence or subcategory always appears as a possibility in a visible set of other semantically related possibilities.

13 Examples of Categories of the IEML Dictionary

13.1 Introduction: Collaborative Building and Use of the IEML Dictionary and Databases of USLs

The IEML dictionary contains generated categories (and only generated categories) at layers 0, 1, 2 and 3. These categories have three main purposes.

First, they can be used as semes to build other categories of the dictionary at higher layers, up to layer 3. The choice of a *tag (see sub-section 8.2) to impose a meaning on a category should reflect strong constraints of semantic coherence, re-usability and generality. So, collaborators in the construction of the dictionary should be people who have been trained in the syntactic and mathematical structure of IEML as well as in its semantic and philosophical principles. The basic training should be broadly equivalent to a graduate course of one semester. I have already given such a class and I know that it is possible.

Second, the categories of the dictionary can be used directly and freely at their own layer by anybody to tag data, documents and Web pages. In this case, the categories can of course be used in union with other categories in the context of a USL (that is to say: a set of categories from different layers, see sub-section 10.4).

Third, the categories can be used freely by anybody to build categories of higher layers in USLs: inflected words, phrases and relations between phrases. Such category building can take place in the

	O:M:. act		
	O:S:. act according to sign	O:B:. act according to being	O:T:. act according to thing
U:M:. act virtually	U:S:E: y. know	U:B:E: o. want	U:T:E: e. can
A:M:. act actually	A:S:E: u. express	A:B:E: a. commit	A:T:E: i. do

Figure 6: *O:M:**

context of manual USL production, or in the context of automatic generation, or the programming of search functions (see the notion of *power notation* at sub-section 5). In this third case, as in the preceding, users will compose USLs or power notation programs *by selecting categories from the dictionary*.

By contrast with the IEML specialists who develop the dictionary, the collaborative building of databases of USLs combining the categories of the dictionary at all layers can be completely open, as the tagging of documents, data and Web pages with these USLs.

Through the following examples of categories of the current dictionary, I will illustrate some principles of IEML that have been exposed above. I will also explain some new ideas that could not be explained in the absence of an understanding of primitives, layers, syntactic roles and generative operations.

13.2 *O:M:E**

13.2.1 Deciphering *O:M:**

In the generated category *O:M:** described in Figure 6 the lowercase letters are abbreviations for the full expressions above them. There are 25 such abbreviations in STAR notation for the most-used sequences of layer one; but for the sake of brevity I will present only 15 of them in this paper.

*O:M:** is a verbal category since its substance seme is *O:** (see sub-section 9.2). As verbishness is invoked in STAR notation by vowels, all the sequences of *O:M:E:** are abbreviated by vowels. These letters denote sequences of layer 1, as indicated by the trailing layer mark “.” (see sub-section 10.1.2). In Figure 6, the expressions in English below the lower case IEML expressions are *tags (see sub-section 8.2) intended to evoke the meaning of the sequences.

Finally, when all the primitive categories until the end of a sequence, a generated category or a seme, are identical to *E:**, the string of *E:** can simply be omitted. The layer of the sequence, generated category or seme is automatically recognized by the STAR-IEML parser thank to the trailing layer mark.¹⁶ In the case of *O:M:E:**, the string of *E:** counts only one member.

13.2.2 Semantic Analysis of *O:M:**

*O:M:** is a category that has a verbal seme as substance, a nominal seme as attribute and an empty mode. *O:M:** can be used as a verb in phrase. The meaning of *O:M:** is the union of the meanings of all the cells of the table. In Figure 6 the columns display sequences with an identical substance while the rows display sequences with an identical attribute.

The three columns represent three distinct categories that are subsets of *O:M:**.

*O:S:** has a verbal substance and a *sign attribute. It is the title of the first column and it can be used as a verb the meaning of which is the union of the meaning of *U:S:** and *A:S:**

¹⁶See <http://starparser.ieml.org/>

M:M: basis			
	M:S: basis for sign	M:B: basis for being	M:T: basis for thing
S:M: sign in basis	S:S:E: s. thought	S:B:E: b. language	S:T:E: t. memory
B:M: being in basis	B:S:E: k. society	B:B:E: m. affect	B:T:E: n. world
T:M: thing in basis	T:S:E: d. truth	T:B:E: f. life	T:T:E: l. space

Figure 7: *M:M:**

O:B:* has a verbal substance and a *being attribute. It is the title of the second column and it can be used as a verb the meaning of which is the union of the meaning of *U:B:** and *A:B:**

O:T:* has a verbal substance and a *thing attribute. It is the title of the second column and it can be used as a verb the meaning of which is the union of the meaning of *U:T:** and *A:T:**

The two rows represent two distinct categories that are also subsets of *O:M:**:

U:M:* has a *virtual substance with a nominal attribute. It is the title of the second row and it can be used as a verb the meaning of which is the union of the meaning of *U:S:**, *U:B:** and *U:T:**

A:M:* is an *actual substance with a nominal attribute. It is the title of the second row and it can be used as a verb the meaning of which is the union of the meaning of *A:S:**, *A:B:** and *A:T:**

*O:M:** is a *derived category* (see 10.3) since the meaning of its sequences is related to the meaning of its semes. The *O:M:** category is organized by several set-subset relationships, where the *syntactic relations* between sets and subsets of sequences *parallel semantic relations* between the meanings of these sequences and sets of sequences. Also, the category is clearly organized along two symmetries: a *virtual/*actual symmetry and a *sign/*being/*thing symmetry.¹⁷ Since these semantic symmetries have a counterpart in explicit syntactic symmetries, they can be exploited computationally. Finally, there is a semantic order of increasing materiality, from left to right and from top to bottom, that has its syntactic counterpart in the numerical order of the semantic characters (the substance character for the vertical order and the attribute character for the horizontal order).

All the tables of the IEML dictionary are more or less organized the same way and display comparable semantic relations (etymology, set-subset, symmetry, order) that are supported by computable explicit syntactic relations.

13.3 *M:M:**

13.3.1 Deciphering *M:M:**

As in the previous example (see Figure 6) the lowercase letters of Figure 7 are abbreviations for the full expressions above them. Since its substance seme is *M:**, *M:M:** is a nominal category (see 9.2). As nounishness is invoked in STAR notation by consonants, all the sequences of *M:M:** are abbreviated by consonants. And again, as these letters denote sequences of layer 1, they are followed by the layer mark “.”

¹⁷Note how Figure 6 resonates with Figure 4

13.3.2 Semantic Analysis of *M:M:**

As for Figure 6, rows and columns of Figure 7 manifest subcategories of the main category *M:M:**. Rows correspond to subcategories having the same substance and columns to subcategories having the same attribute. Note that *(s. + m. + l.)**, corresponding to a *diagonal* of the table, are three sequences with an *identical seme* as substance *and* attribute. STAR-IEMML syntax includes a *diagonal operator* allowing the generation of categories the sequences of which share the same semes. For example, the diagonal of *M:M:** can be noted *{M:}{M:}**. Remark also that the three pairs *(k. + b.)**, *(d. + t.)** and *(f. + n.)** *symmetrically exchange* their substance and their attribute.

The following is an explanation of how the meaning of the nine sequences of *M:M:** is derived from the meanings of their semes.

s.* In *thought, the substance as well as the attribute is *sign. Deduction, induction, interpretation, imagination, and so on combine signs.

b.* In *language, as a mean of communication, a *sign substance is oriented toward a *being attribute.

t.* In *memory, *sign is reified, taking the attribute of *thing. This sequence manifests the elementary gesture of inscription or recording: *sign is transformed into *thing. Without inscription of some sort there is no memory, and thus no notion of time.

k.* In *society, a collective of *beings organizes itself through *signs. We commit ourselves to promises and contracts. We obey the law. The members of a clan have the same totemic animal. We fight for the same flag. We exchange economic goods by agreeing upon their value. Our obligations and conventions tie us to shared symbols. In all these way we form *society.

m.* In *affect, a *being is oriented toward other beings. Desire, love, hatred, indifference, compassion, equanimity are emotional qualities that circulate among beings.

n.*, In *world, human beings (*being substance) express themselves through their physical environment (*thing attribute). They inhabit it, they work there with tools, they name objects and endow them with values, creating an ordered world, or cosmos.

d.* In *truth a substantial reference (*thing) expresses itself through the attribute of *sign. A discursive message is true if it contains a correct description of a state of things.

f.* In *life, a substantial *thing takes the attribute of a *being, thus *f.** evokes the physical incarnation, or the *body*, of a living creature. Eating and drinking are among the most obvious ways by which a thing is transformed into being.

l.* In physical *space, an objective *thing is related to other *things. Physical space is constructed of relations and proximities.

13.3.3 Philosophical Foundation of *M:M:**

The *M:M:** category corresponds to nine major philosophical orientations.

1. The Western idealist tradition (Plato, Hegel, etc.) - like a large part of the major (meditational) philosophies of India - start from the category of *thought.
2. The philosophy of *language and communication specializes in an approach founded on the word, the script, and the media. Wittgenstein made his entire philosophy revolve around the problem of the limits of language.
3. Bergson (the author of *Matter and Memory* [12]) and with him the evolutionists, or the philosophies of impermanence and karma (such as Buddhism) underscore the passage of time and its inscription in *memory. Likewise, the grand universal religious traditions are based on sacred writings that fall within the archetype of inscription.
4. Rousseau, Comte, Marx, Durkheim, Mauss, and Weber, like sociology in general, highlight the essential symbolic function that consists of forming *society.

5. After the poets, the pious, and the actors, Freud, psychoanalysis and a large part of clinical psychology have insisted on the importance of the *affect and emotional functions.
6. Nietzsche (who accorded a central role to the creation of values), and all anthropological thought, primarily base their approach on the notion of *world, or the organized cosmos.
7. The empirical, phenomenological, and biological traditions base their ontological point of view on the living body and its functions (*life).
8. The logical tradition and analytical philosophy are primarily interested in the notion of *truth and in problems tied to reference. Likewise, epistemology and the cognitive sciences place the construction of true knowledge at the foundation of their approach.
9. Finally, geometers, topologists, atomists, materialists and physicists base their concepts on *space.

In a way, IEML is like all natural languages: it allows the expression of any philosophy. But as it is an artificial metalanguage, or a “philosophical language”, I had to design it from its core to avoid any philosophical blind spot. In IEML, each philosophy appears as the partial exploration of an integral semantic space that accommodates them all. A given philosophy represents an “ontological point of view” that is equally as valid as any other. *M:M:** can be seen as an illustration of the philosophical design principles that shape IEML.

13.4 *M:M:.O:M:.-**

13.4.1 Deciphering *M:M:.O:M:.-**

*M:M:.O:M:.-** is an abbreviation for *M:M:E:.O:M:E:.E:E:E:.-**. The “-” at the end of the expressions specifies that they are at layer 2 (see 10.1.2).

The *M:M:** category plays the substance role, so the whole *M:M:.O:M:.-** category is nominal. *O:M:** is used as attribute, providing a kind of semantic declension of the *M:M:** substance.

In Figure 8, the rows display subcategories that share the same attribute and the columns display subcategories that share the same substance. *M:M:.O:M:.-** is called "Basis of human development" because it provides for categories of layer 2 that are involved (as semes) in the generation of categories of layer 3 that describe at length the six human capitals of Figure 4.

13.4.2 Semantic Analysis of *M:M:.O:M:.-**

Figure 8 shows three levels of set-subsets relationships.

1. The first one is between the greater category (*M:M:.O:M:.-**) and the three meta-columns as well as between the greater category and the two meta-rows. The three meta-columns are organized along a *sign/*being/*thing symmetry at the substance of the substance and the two meta-rows are organized by a *virtual/*actual symmetry at the substance of the attribute.
2. The second level of set-subset relationship is between the meta-columns and the three columns that each one includes, on the one hand; and between the two meta-rows and the three rows that each one includes, on the other hand. The groups of three columns are organized by a *sign/*being/*thing symmetry at the attribute of the substance and the groups of three rows are structured by a *sign/*being/*thing symmetry at the attribute of the attribute.
3. The third level of set-subset relationship is between the columns and their cells, as well as between the rows and their cells. The columns are organized by a six terms symmetry corresponding to the six sequences of *O:M:** and the rows are organized by a nine terms symmetry corresponding to the sequences of *M:M:**.

As in the *O:M:** and *M:M:** categories previously analyzed, the set-subset relationships and the symmetries existing *at the syntactic level of the arrangement of symbols* have a parallel *at the semantic level of the meaning of the categories*.

In the category *O:M:.M:M:.-* (not displayed here) substance and attribute of *M:M:.O:M:.-** are symmetrically exchanged. This symmetrical verbal category is used to describe ideas that are the opposites of those of *M:M:.O:M:.-**. For example, the sequence *y.l.-** means *ignore_limits while the symmetric

M:M:O:M:- basis of human development										
S:M:O:M:- sign in basis for HD			B:M:O:M:- being in basis for HD			T:M:O:M:- thing in basis for HD				
M:M:y- intelligence	s:O:M:- thought actualized	b:O:M:- language actualized	t:O:M:- memory actualized	k:O:M:- society actualized	n:O:M:- world actualized	d:O:M:- truth actualized	f:O:M:- life actualized	I:O:M:- space actualized		
M:M:U:M:- basis for HD into the <i>virtual</i>	s:y- organized knowledge	b:y- formal intelligence	t:y- memory management	k:y- relational intelligence	m:y- emotional intelligence	d:y- knowledge of facts	f:y- body intelligence	l:y- awareness of limits		
M:M:o- value	s:o- concern for thought	b:o- taste for language	t:o- right time	k:o- desire for social bond	m:o- emotional involvement	d:o- truth value	f:o- living body / sentient being	l:o- space organization		
M:M:e- ability	s:e- mastery of codes / grammar	b:e- reasoning ability / dialectic	t:e- ability to convince / rhetoric	k:e- self-control / grammar	m:e- negotiation ability / dialectic	d:e- technical know- how / grammar	f:e- sensorimotor ability / dialectic	l:e- engineering competence / rhetoric		
M:M:u- semiotic function	s:u- image	b:u- speech	t:u- mark	k:u- symbol	m:u- symptom	d:u- pointer	f:u- sensation	l:u- trace		
M:M:a- social function	s:a- interpreter	b:a- storyteller	t:a- scribe	k:a- chief	m:a- parent	d:a- researcher	f:a- healer	l:a- guardian		
M:M:i- technical function	s:i- interface	b:i- medium	t:i- container	k:i- gift	m:i- toy	d:i- measurement instrument	f:i- organ	l:i- knot / connection technique		

Figure 8: *M:M:O:M:-*

	Ability in *Sign (*S:** in substance of substance)	Ability in *Being (*B:** in substance of substance)	Ability in *Thing (*T:** in substance of substance)
Grammar (*S:** in attribute of substance)	*s.e.-** Mastery of codes	*k.e.-** Self-control	*d.e.-** Practical Know-how
Dialectic (*B:** in attribute of substance)	*b.e.-** Reasoning ability	*m.e.-** Dialogue and negotiation	*f.e.-** Sensori-motor refinement
Rhetoric (*T:** in attribute of substance)	*t.e.-** Ability to convince	*n.e.-** Leadership	*l.e.-** Engineering

Table 1: Basic Abilities in IEML (*M:M:.e.-**)

sequence *l.y.-** means **awareness_of_limits*. Similarly, *i.m.-** means **use_weapons* while *m.i.-** means **toy*, etc.

Note that role symmetries in IEML are not always semantically relevant and that, when they are, they do not necessarily denote contrary meanings (they may, for example, indicate complementarity). The semantic network of the dictionary display only the *semantically relevant* relationships between categories.

Table 1 shows the inner structure of the row *M:M:.e.-**. The organization of study through the “Trivium” (grammar, dialectic, rhetoric) is a long tradition dating back from the Greeks, the Romans and the European medieval University. But the classical Trivium was generally limited to linguistic, logical and literary studies, that are the realm of *sign. I have generalized the classical Trivium to include the abilities concerning the inner and social realm of *being and the material, bodily and technical realm of *thing. The nine basic abilities of *M:M:.e.-** are used for the construction of more complex and detailed tables of abilities in categories of layer 3.

The meanings of the subcategories of *M:M:.O:M:-**, as well as the meanings of their sequences, are derived loosely from the meanings of their semes. Therefore, *M:M:.O:M:-* as a whole is a derived category.

I will explain the relationships between the meanings of sequences and the meanings of their semes only on the *M:M:.i.-** subcategory (technological functions). The relationship between this technological subcategory and *i.** (*do) in *attribute* position is connected to the concreteness of the verbal seme *i.**. The *M:M:.i.-** row, having *i.** as attribute, makes a system with the other subcategories having *y.** (kind of intelligence, types of knowledge), *o.** (values, will orientations), *e.** (types of abilities), *u.** (semiotic functions) and *a.** (social functions) as attribute. I will now focus on the relationships of the sequences of *M:M:.i.-** with their *substance*.

*s.i.-** **interface*: because of its role in translation and conversion, the interface function is close to **thought*.

*b.i.-** **medium* covers the technologies of **language* and **signs*.

*t.i.-** **container*: by preventing liquids from flowing and by protecting goods, the container is a material ally of **memory*.

*k.i.-** **gift*: by weaving social relationships between people, **gift* is a material support of a **society’s* fabric.

*m.i.-** **toy*: a material object can convey emotion or **affect*.

*n.i.-** **transformation_technology*: mastery of fire is the first condition of a human **world*, which is fundamentally connected with cooking, pottery, metal transformation, chemistry and energy production.

*d.i.-** **measurement_instrument* is the material tool of **truth*.

*f.i.-** **organ* is the tool of bodily **life*.

		E:F:.O:M:.- conjugation				
		E:O:.O:M:.- participation mood		E:M:.O:M:.- tense / time / when		
		E:U:.O:M:.- passive	E:A:.O:M:.- active	E:S:.O:M:.- future	E:B:.O:M:.- present	E:T:.O:M:.- past
E:F:U:M:.- declarative mood	E:F:.y.- indicative mood	E:U:.y.- known	E:A:.y.- knowing	E:S:.y.- will know	E:B:.y.- know now	E:T:.y.- has known
	E:F:.o.- optative mood	E:U:.o.- wanted	E:A:.o.- willing	E:S:.o.- will want	E:B:.o.- want now	E:T:.o.- would
	E:F:.e.- abilitative mood	E:U:.e.- enabled	E:A:.e.- abling	E:S:.e.- will be able	E:B:.e.- can now	E:T:.e.- could
E:F:A:M:.- performative mood	E:F:.u.- expressive mood	E:U:.u.- expressed	E:A:.u.- expressing	E:S:.u.- will express	E:B:.u.- express now	E:T:.u.- has expressed
	E:F:.a.- promissive mood	E:U:.a.- committed	E:A:.a.- committing	E:S:.a.- will commit	E:B:.a.- commit now	E:T:.a.- has committed
	E:F:.i.- imperative mood	E:U:.i.- be done	E:A:.i.- be doing	E:S:.i.- will execute	E:B:.i.- execute now	E:T:.i.- should

Figure 9: *E:F:.O:M:.-**

*l.i.-** *connection_technique covers knots as well as cutting and fastening instruments that connect and disconnect things in physical *space.

13.5 *E:F:.O:M:.-**

As already explained above and illustrated again in Figure 9, when the attribute and mode of a seme are completely empty, they can be omitted in STAR notation. Categories that have an empty initial primitive are used to code auxiliary terms, that are neither verbs nor nouns: conjugations, cases, adverbs, pronouns, prefixes, etc. The category *E:F:.O:M:.-** is an example of an auxiliary category devoted to conjugation.

*E:F:.O:M:.-** shows how *M:** = *(S:+B:+T:)** can be used to represent ternary structures that are *not* related to the semantic area of signifier/interpreter/referent. Here, *(S:+B:+T:)** is used to represent future/present/past. In the same vein, the dialectic between *U:** and *A:** is not used to represent a virtual /actual opposition but to represent a passive/active opposition.

Since the meaning of the attributes of the sequences of *E:F:.O:M:.-** are indeed related to the meaning of the corresponding sequences of *O:M:** (see 13.2.2), the *E:F:.O:M:.-** category is only partially derived: it has an etymologic relationship with its attribute but no etymologic relationship with its substance.

The meanings of the subcategories of Figure 9 are self-explanatory. The reader will observe the set-subset relations between the categories and the various symmetries that organize the table.

13.6 *M:M:.a.-M:M:.a.-f.o.-**

The category *M:M:.a.-M:M:.a.-f.o.-**, as well as all of its subcategories and sequences, ends with the layer mark “ ’ ”, indicating layer 3. This category has *M:M:.a.-** as substance *and* attribute (see the fifth row of Figure 8) and each of its subcategories corresponds to identical seme substances (rows) or identical seme attributes (columns).

The meanings of the sequences and subcategories of *M:M:.a.-M:M:.a.-f.o.-** are loosely related to the meanings of their semes; it is a derived category.

The mode *f.o.-** (*living_body, see 8), identical for all the subcategories and sequences, indicates that this category is about real persons that embody institutions or play social roles. There is a category *M:M:.a.-M:M:.a.-** that contains the corresponding abstract functions or institutions that are embodied by *M:M:.a.-M:M:.a.-f.o.-**. For example, when *s.a.-m.a.-f.o.-** means *spouse, *s.a.-m.a.-** means

*marriage; when *k.a.-k.a.-f.o.-’** means *citizen, *k.a.-k.a.-’** means *city_political_unit, when *l.a.-l.a.-f.o.-’** means *warrior, *l.a.-l.a.-’** means *war, etc.

13.7 Analysis of a Phrase of Layer 4

I will now give an example of a complex sentence in IEML. In English, the sentence sounds like: “My neighbor’s child is often authorized by a physician not to go to school.”

```

*/(
(u.A.:E:U:A:.-E:.- (E:U:.e.- + E:B:T:.-)’ d.a.-m.a.-f.o.-’ E:E:S:.-’, )
+
(u.A.:E:U:A:.-E:.- (E:U:.e.- + E:B:T:.-)’ h.y.-’ E:T:.d.-’, )
+
(d.a.-m.a.-f.o.-’ m.a.-k.a.-f.o.-’E:E:A:.-’, )
+
(m.a.-k.a.-f.o.-’ E:.-’ E:U:.we.-’, )
+
(m.a.-k.a.-f.o.-’ (E:S:.wo.- + E:S:.wu.-)’ E:E:A:.-’, )
+
(u.A.:E:U:A:.-E:.- (E:A:.e.- + E:B:T:.-)’ n.a.-f.a.-f.o.-’ E:E:S:.-’, )
)**

```

Analysis of the phrase, clause by clause

/u.A.:E:U:A:.-E:.- (E:U:.e.- + E:B:T:.-)’ d.a.-m.a.-f.o.-’ E:E:S:.-’, *

This clause means “The child is often authorized not to go”: It is a verbal clause.

The substance category *u.A.:E:U:A:.-E:.- (E:U:.e.- + E:B:T:.-)’** is a verb word that can be analyzed into its substance *u.A.:E:U:A:.-E:.-** and its mode *(E:U:.e.- + E:B:T:.-)**. The attribute of the substance is empty. The substance of the substance is generated from the verbal morpheme *u.A.:** that means “go” and the auxiliary morpheme marking the negation *E:UA:.**. So *u.A.:E:U:A:.-** means: “not to go”. In the mode of the substance, *E:U:.e.-** is a conjugation auxiliary (see Figure 9) (abilitative and passive mood) and *E:B:T:.-** is an adverb meaning “often”.

*d.a.-m.a.-f.o.-’** in attribute means “child” and *E:E:S:.-’** in mode means that the child is the subject of the verb.

/u.A.:E:U:A:.-E:.- (E:U:.e.- + E:B:T:.-)’ h.y.-’ E:T:.d.-’, *

This verb clause means “is often authorized not to go to school”: *h.y.-’** in attribute means “school” and the auxiliary *E:T:.d.-’** “to” indicates that *h.y.-’** is a complement of the verb.

/d.a.-m.a.-f.o.-’ m.a.-k.a.-f.o.-’ E:E:A:.-’, *

This noun clause means “neighbor’s child”. *m.a.-k.a.-f.o.-’** means “neighbor” and the auxiliary *E:E:A:.-’** marks the genitive (“of”).

/m.a.-k.a.-f.o.-’ E:.-’ E:U:.we.-’, *

This category is not a clause but a word meaning “*this feminine neighbor*”. *E:U:.we.-’** , the inflection of the word, indicates a “feminine demonstrative”.

/m.a.-k.a.-f.o.-’ (E:S:.wo.- + E:S:.wu.-)’ E:E:A:.-’, *

The noun clause means “my neighbor (me-masculine)”. *m.a.-k.a.-f.o.-’** (neighbor) is the substance, *E:S:.wo.-** marks the first person in singular number and *E:S:.wu.-** marks a masculine first person. The auxiliary *E:E:A:.-** indicates a genitive: “the neighbor *of me*”, so: “my neighbor”.

u.A.:E:U:A:.-E:.- (E:A:.e.- + E:B:T:.-)’ n.a.-f.a.-f.o.-’ E:E:S:.-’, *

The verb clause means “The physician gives authorization not to go.” *n.a.-f.a.-f.o.-’** means “physician” and the auxiliary in mode *E:E:S:-’** indicates that *n.a.-f.a.-f.o.-’** is the subject of the verb. Note that the two words *u.A.:E:U:A:.-E:-(E:U:.e.-+E:B:T:-)’** (verb in substance of the two first clauses) and *u.A.:E:U:A:.-E:-(E:A:.e.-+ E:B:T:-)’** (the verb in substance of the last clause) are in anaphoric relation, even if they are not identical. In the first verb, the authorization is in passive mood (*E:U:.e.-**) - when the subject is the child - but in the second verb the authorization is in active mood (*E:A:.e.-**) - when the subject is the physician. It is sufficient for the creation of an anaphor that the roots of the words are identical, and the root *u.A.:E:U:A:.-** “not to go” is here identical for the two words.

If all the clauses are considered as vertices of a graph and if all the anaphoric relations between clauses are considered as edges between these vertices, the composed category that has just been analyzed can be represented as a *connected graph* of clauses and form a phrase.

Besides, the uniform rules for the creation of phrases allow for another connected graph representation of phrases in which the vertices are inflected words (mainly verbs, nouns, adjectives and pronouns) in substance and attribute roles and the edges are labelled by the grammatical relations between these words as marked by auxiliary words in mode role. This last feature could provide an interesting support for the automatic translation of IEML into natural languages.

14 USLs

As said earlier, a USL is composed of up to seven catsets of different layers. In the STAR notation of USLs, the categories are separated by “/”. Below is a USL that could be used as a *tag for “wikipedia” and that uses only terms that already exist in the current dictionary.

14.1 *WIKIPEDIA

*

```

/ (U:+S:)
/ d. / t.
/ wo.y.- / wa.k.- / e.y.- / s.y.- / k.h.-
/ a.u.-we.h.-' / n.o.-y.y.-s.y.-'
/ ((n.o.-y.y.-s.y.-' s.o.-k.o.-' E:E:A:-') + (n.o.-y.y.-s.y.-' E:-' b.i.-b.i.-E:T:l.-'))
/ u.e.-we.h.-' m.a.-n.a.-f.o.-' E:E:S:-',
/
(e.-'s.e.-k.u.-' E:E:T:-',
+
e.-'((b.a.-b.a.-f.o.-')+(t.a.-b.a.-f.o.-')) E:E:S:-',
+
((b.a.-b.a.-f.o.-')+(t.a.-b.a.-f.o.-')) E:-' E:M:wa.-',
+
((b.a.-b.a.-f.o.-')+(t.a.-b.a.-f.o.-'))E:-'l.o.-m.o.-s.u.-',
+
(s.e.-k.u.-E:B:s.-'b.o.-u.u.-E:F:wa.-'E:T:x.-'),
+
k.i.-b.i.-t.u.-'b.x.-'E:E:U:-'[wiki],
+
s.e.-k.u.-E:B:s.-'k.i.-b.i.-t.u.-'E:T:x.-')[wikipedia]
**

```

L0: knowledge networks

L1: truth / memory

L2: gets one’s bearing in knowledge / act for the sake of society / synthetize / organized knowledge / collective creation

L3: opening public space / encyclopedia

L4: collective intelligence encyclopedia in cyberspace / volunteers producing didactic material / enables plural and open authorship and editing in several languages, by the medium of a software collaborative environment

The expressions between square brackets in the USL are called *instantiators* in STAR-IEMML syntax and are used for specialized terms that are not supported by the IEMML dictionary, like proper names (of people, places, brands, etc.), numbers, prices, measures, etc., for instance */t.u.-n.u.-'[Suzan Smith]** where */t.u.-n.u.-'** means "proper name", or */t.o.-n.o.-s.u.-'[18/03/1956]** where */t.o.-n.o.-s.u.-'** means "date".

14.2 Set Operations on USLs

It is possible to perform set operations on USLs. Let's imagine two USLs composed by different users: *XML and *Web_of_data.

14.2.1 *XML

```
*
/ (A:+S:)
/ b.
/ we.g.- / we.b.-
/ e.o.-we.h.-' / b.i.-b.i.-' / t.e.-d.u.-'
/
((i.i.-we.h.-' s.a.-t.a.-' E:E:T:.-'),
+
(i.i.-we.h.-' b.o.-j.j.-E:A:g.-' E:T:.x.-'),)
**
```

L0: document networks

L1: language

L2: unify the documentation / cultivate information system

L3: establishing norms and standards / cyberspace / meeting information needs

L4: guarantee compatibility of data through same formal structure.

14.2.2 *Web_of_data

```
*
/ (U:|S:) / (A:|S:)
/ d.
/ s.x.- / x.j.-
/ e.o.-we.h.-' / b.i.-b.i.-'
/ e.o.-we.h.-' b.i.-b.i.-' E:E:B:.-',
/
((t.e.-t.u.-wa.e.-' k.i.-b.i.-t.u.-' E:E:A:.-'),
+
(k.i.-b.i.-t.u.-' E:.-'b.e.-E:.-E:A:.x.-'),)
/ s.a.-t.a.-' b.i.-l.i.-t.u.-' E:E:U:.-',
**
```

L0: knowledge networks / document networks

L1: truth

L2: technical project / evolution of computation

L3: establishing norms and standards / cyberspace

L4: design of software bots with reasoning ability / linked data

14.2.3 Results of Two Operations on the Previous USLs

The set operations on USLs are always performed on sets of categories (and therefore on sets of sequences) *layer by layer*.

The *intersection* of the two USLs of the previous examples gives a new USL (*Internet_standards?) with two categories of layer 3

```
*
 / e.o.-we.h.-' / b.i.-b.i.-'
**
```

L3 establishing norms and standards / cyberspace

The *union* of the two USLs gives the following resulting USL (*Great_Web_of_data?)

```
*
 / (A:+S:) / (U:+S:)
 / b. / d.
 / we.g.- / we.b.- / s.x.- / x.j.-
 / e.o.-we.h.-' / b.i.-b.i.-' / t.e.-d.u.-'
 /
 ((i.i.-we.h.-' s.a.-t.a.-' E:E:T:.-'),
 +
 (i.i.-we.h.-' b.o.-j.-E:A:g.-' E:T:x.-'),)
 / e.o.-we.h.-' b.i.-b.i.-' E:E:B:.-',
 /
 ((t.e.-t.u.-wa.e.-' k.i.-b.i.-t.u.-' E:E:A:.-'),
 +
 (k.i.-b.i.-t.u.-' E:.-'b.e.-E:.-E:A:x.-'),)
 / s.a.-t.a.-' b.i.-l.i.-t.u.-' E:E:U:.-',
**
```

L0: document networks / knowledge networks

L1: language / truth

L2: unify the documentation / cultivate information system / technical project / evolution of computation

L3: establishing norms and standards / cyberspace / meeting information needs

L4: guarantee compatibility of data through same formal structure / design of software bots with reasoning ability / linked data

15 Operations and Relations

15.1 Some Types of Automatable Operations on IEML Categories

Any set operation and composition of set operations can be performed on USLs. The mathematical consequences of this possibility are developed in the document “Algebraic Structure of IEML Semantic Space”, available online¹⁸. As already emphasized, the main result is that IEML semantic space (the set of all USLs) is a group of transformations.

Many operations can be defined for automatic generation, automating networking or selection of categories and USLs, including:

- generation / recognition of the members of the powerset of a category;
- generation / recognition of categories exchanging symmetrically the generative role of their semes like for example *a.y.-** (“teach”) and *y.a.-** (“learn”);
- partition of a category;

¹⁸Algebraic Structure of IEML Semantic Space: <http://www.ieml.org/spip.php?article152>

- “rotation” of a category of layer n (successive replacement of a seme of layer $n - m$ by the members of a catset of layer $n - m$);
- generation / recognition of categories sharing an identical seme, at the same syntactic role or at different syntactic roles;
- generation of trees of categories and subcategories;
- linear ordering of categories using the binary characters (see Sub-section 11.2)

The international research team on IEML led by the CRC-IC at the University of Ottawa is currently developing a specification language that will formally define such operations for implementation purposes.

15.2 Types of Semantic Relations between the Categories of the Dictionary

The aforementioned operations can be used to automatically generate sets of categories and impose some semantic relations on them. This feature is especially useful for the construction of the dictionary. From the semantic relations imposed on the categories of the dictionary, it is possible to compute complex semantic relations between USLs. There are four types of semantic relations between categories in IEML:

1. *Set-subset relations* connect categories at a given layer that are related by inclusion (category / subcategory relation)
2. *Symmetric relations* connect subcategories of same layer that are included in the same category or categories that exchange the roles of their semes (like *y.a.-** “learn” and *a.y.-** “teach”). A set of categories in symmetric relationship form a symmetry group because their substitution conserves a constant semantic feature.
3. *Etymological relations* connect categories of layer n with their semes of layer $n - m$ ($1 \leq m \leq n$) when the meaning of the category of layer n derives from the meanings of its semes.
4. *Linear order relations* organize sequences of categories on the basis of semantic criteria. Computing of order relations is ultimately based on the numerical value of the binary semantic characters that generate categories.

The construction of the dictionary is one of the main challenge of the IEML research program. An important constraint for this construction - as well as a major evaluation criterium of its quality - will be the maximization of *semantically relevant* relations between IEML categories.

Conclusion

16 One or Several IEML Dialects?

The mathematical structure of IEML (6 primitives, 3 generative roles, 7 layers, sequences, categories as sets of sequences of same layer, USLs, basic set operations and complex operations for the automatic generation and recognition of categories) can be interpreted as *virtual circuitry* for a new kind of semantic computing. In this context, the STAR language, its dictionary and its grammar (the rules for the construction of phrases) can be understood as one particular *semantic operating system* exploiting the computational features of this virtual circuitry. It is possible to separate the syntactic and mathematical structure of IEML from the current STAR dialect and therefore from the general philosophy that inspired the elaboration of the metalanguage. This is why I cannot prevent the possibility that, in the future, concurrent IEML dialects could propose different ways to exploit IEML’s virtual circuitry, with possibly a notation different from the STAR notation, a dictionary obeying different principles than the STAR dictionary and a different organization of phrases. In this case, there would be *several* semantic spaces (each one corresponding to a different IEML dialect) instead of *one common* semantic space. But there is considerable as yet unallocated space in the current dictionary. Since its ongoing construction will be transparent, transcultural and collaborative, I hope that future efforts, at least in the coming decades, will be devoted to the augmentation of the current STAR dialect instead of the elaboration of other ones.

Because of its *finite* nature, IEML semantic space is not concerned with the limitation theorems of Turing and Gödel. Nevertheless, it is an immense set: even if the whole universe were a quantum computer, the complete collection of USLs would be impossible to record. This leaves abundant room for semantic differentiation and innovation!

17 Compatibility and Innovation

The IEML research program is a novel approach to the interoperability challenges that face semantic technology and collective intelligence research.

The approach differs from traditional Semantic Web ontologies, but is compatible with it: it only requires a translation to IEML *tags of the concepts expressed in natural languages and defined in ontologies. Furthermore, it is our intention to code the *tags in RDF and to model in OWL the syntactic and grammatical rules as well as the current dictionary of IEML. Equipped with IEML, the semantic web technologies could become the main tool for the transformation of the current Web into a worldwide knowledge management system for human collective intelligence at any scale.

The IEML research program is also complementary to automatic natural language processing. On the one hand, it can benefit from automatic text analysis for the IEML tagging of documents and Web pages. The conceptual networks extracted from a document by a natural language processing program (e.g. Open Calais) can of course be translated into networks of IEML *tags. On the other hand, IEML can provide a pivotal artificial language in order to improve the efficiency and relevance of automatic translation between natural languages, at least for metadata, key words and abstracts.

The IEML research program is perfectly aligned with the current trend toward human-centric social computing. This trend is evidenced by the large number of blogs, wikis, social media, collaborative applications and sites using social tagging, as well as by the popularity of microformats (semantic tagging in HTML) [35], that allow users to stamp documents with semantic metadata and to share those tags in collaborative networks [66]. But current techniques for collaborative filtering based on social tagging face some problems[32, 53]: polysemy, synonymy, heterogeneity of levels of categorization, sparse data and, above all, multiplicity of natural languages. The use of IEML can help solving these problems in a radical manner and, in general, could improve the “programming of collective intelligence”[64] by strengthening its semantic side.

Beyond its compatibility with current research programs and cultural trends, the IEML research program promotes a *radical symbolic innovation* that could have a deep positive impact on humanities, social sciences and, beyond academic disciplines, on a future global civilization of collective intelligence. It is my deepest conviction that this innovation is needed in order to allow the increasing pervasive ability of computation power to be exploited in the service of human development.

Acknowledgements

The work presented here was mainly funded since 2002 by the Canadian Federal Government through its Canada Research Chairs programs. I also received two grants from the Social Sciences and Humanities Research Council of Canada. I would like to thank for their collaboration Darcia Labrosse, Michel Biezunski, Steve Newcomb, Andrew Rocznik, Christian Desjardins and Samuel Szoniecky.

References

- [1] Auroux, Sylvain, *La Révolution technologique de la grammatisation*. Mardaga, Liège, 1994
- [2] Aristotle, *On Interpretation*. Translated by E. M. Edghill, <http://classics.mit.edu/Aristotle/interpretation.1.1.html>
- [3] Benkler Yochai, *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. Yale UP, 2006
- [4] Berners-Lee Tim, *Weaving the Web*. Harper, San Francisco, 1999
- [5] Bottero Jean, *L’Ecriture la raison et les Dieux*. Gallimard, Paris, 1987
- [6] Bowker Geoffrey, *Memory Practices in the Sciences*. MIT Press, 2005

- [7] Callon Michel (dir.) *La Science et ses réseaux. Genèse et circulation des faits scientifiques*. La Découverte, Paris, 1989
- [8] Carruthers Mary, *The Craft of Thought, Meditation, Rhetoric and the Making of Images, 400, 1200*. Cambridge UP, 2000
- [9] Castells Manuel, *The information Age*. Blackwell, Cambridge MA, 1996, 97, 98
- [10] Changeux Jean-Pierre, *Neuronal Man: The Biology of Mind*. Pantheon Books, NY, 1985
- [11] Chomsky Noam; Schützenberger, Marcel P. "The algebraic theory of context free languages", in Braffort, P.; Hirschberg, D.: *Computer Programming and Formal Languages*. Amsterdam: North Holland, 118-161. (1963)
- [12] Bergson Henri, *Matière et mémoire*, Alcan, Paris, 1908 (first edition: 1896)
- [13] Dalkir Kimiz, *Knowledge Management in Theory and Practice*. Elsevier, NY, 2005
- [14] Davis Martin, *Engines of Logic: Mathematicians and the Origin of the Computer* (1st ed.). W. W. Norton & Company, NY, 2000
- [15] De Libera Alain, *La Querelle des universaux, De Platon à la fin du Moyen-Age*. Seuil, Paris, 1996
- [16] Dichev Christo, Jinsheng Xu, Darina Dicheva, Jinghua Zhang, "A Study on Community Formation in Collaborative Tagging Systems", *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence*, Sydney, Vol.3 p. 13-16, 2008
- [17] Engelbart Douglas, *Augmenting Human Intellect*, Technical Report. Stanford Research Institute, 1962
- [18] Edelman Gerard, *Neural Darwinism*. Basic Books, NY, 1987
- [19] Eisenstein Elisabeth, *The Printing Revolution in Early Modern Europe*. Cambridge University Press, 1983
- [20] Feigenbaum L., et al., "The semantic Web in Action". *Scientific American*, dec 2007, p. 90-97
- [21] Fellbaum Christiane (ed.), preface by George Miller, *Wordnet, An Electronic Lexical Database*, MIT Press, 1998
- [22] Goody Jack, *The Domestication of the Savage Mind*. Cambridge UP, 1977
- [23] Goody Jack, *The Logic of Writing and the Organization of Society*. Cambridge UP, 1987
- [24] Granovetter Mark, "The Strength of Weak Ties." *American Journal of Sociology*, 78 (May): 1360-1380, 1973
- [25] Halpin Harry, Robu Valentin & Hannah Shepherd, "The Complex Dynamics of Collaborative Tagging". *Proceedings of the 16th International WWW Conference*, 2007, p. 211-220
- [26] Havelock E. A., *The Muse Learns to Write, Reflections on Orality and Literacy from Antiquity to the Present*. Yale UP, 1988
- [27] Hendler Jim & Allemang Dean, *Semantic Web for the Working Ontologist*. Morgan Kaufmann, Burlington MA, 2008
- [28] Hutchins E., *Cognition in the Wild*. MIT Press, Cambridge, Mass., 1995
- [29] Innis Harold, *Empire and Communication*. Toronto UP, 1950
- [30] Jaynes Julian, *The Origin of Consciousness in the Breakdown of the Bicameral Mind*. Houghton Mifflin, Boston, 1976
- [31] Frege Gottlob, ed. Peter Geach and Max Black, *Philosophical Writings of Gottlob Frege*. Blackwell, Oxford, 1952

- [32] Herlocker J. L. , Konstan J. A., Terveen L. G. and Riedl, J. T. "Evaluating Collaborative Filtering Recommender systems". *ACM Transactions on Information Systems*, 22, 1, 2004, p. 5-53
- [33] Kapetanios Epaminondas, "On the Notion of Collective intelligence: Opportunity or Challenge?" in *Inter. Journal on Organisational and Collective intelligence (IJOICI)* Vol. 1, Number 1, Idea Group Publishing, 2009
- [34] Kaplan Robert, *The Nothing That Is, a Natural History of Zero*. Oxford UP, 1999
- [35] Khare R., "Microformats: the next (small) thing on the semantic Web?". *IEEE Internet Computing*, Volume: 10, Issue: 1, Jan.-Feb. 2006 p. 68- 75
- [36] Latour Bruno, *Science in Action*. Harvard UP, 1987
- [37] Lévi-Strauss Claude, *La pensée sauvage*. Plon, Paris, 1962 / *The Savage Mind*. Chicago UP, 1966
- [38] Lévy Pierre, *La Machine Univers. Création, cognition et culture informatique*. La Découverte, Paris, 1987
- [39] Lévy Pierre., *Les Technologies de l'intelligence. L'avenir de la pensée à l'ère informatique*. La Découverte, Paris, 1990
- [40] Lévy Pierre & Authier Michel (preface of Michel Serres), *Les Arbres de connaissances*, La Découverte, Paris, 1992
- [41] Lévy Pierre, *De la programmation considérée comme un des beaux-arts*. La Découverte, Paris, 1992
- [42] Lévy Pierre, *L'Intelligence collective. Pour une anthropologie du cyberspace*. La Découverte, Paris, 1994 / *Collective intelligence: Mankind's Emerging World in Cyberspace*. Perseus Books, Cambridge, Mass. 1997
- [43] Lévy Pierre, *Qu'est-ce que le virtuel?* La Découverte, Paris, 1995 / *Becoming Virtual. Reality in the Digital Age*, Plenum Trade, New York, 1998
- [44] Lévy Pierre, *Cyberculture*. Odile Jacob, Paris, 1997 / *Cyberculture*. University of Minnesota Press, 2001
- [45] Lévy Pierre, *World Philosophie: le marché, le cyberspace, la conscience*. Odile Jacob, Paris, 2000
- [46] Lévy Pierre, *Cyberdémocratie: Essai de philosophie politique*. Odile Jacob, Paris, 2002
- [47] Lévy Pierre, "Société du savoir et développement humain". In *Le Canada et la société des savoirs*, sous la direction de Patrick Imbert, CR Université d'Ottawa: Canada, enjeux sociaux et culturels dans une société du savoir, 2007, p. 115-175
- [48] Lévy Pierre, *Algebraic Structure of IEML Semantic Space*. CI Lab Technical Report, 2009. <http://www.ieml.org/spip.php?article152>
- [49] Li Charlene & Bernoff Josh, *Groundswell. Winning in a World Transformed by Social Technologies*. Harvard Business Press, 2008.
- [50] Logan Robert K., *The Alphabet Effect*. Cresskill NJ: Hampton Press, 2004 (1st edition 1986. New York: Wm. Morrow)
- [51] Logan Robert K., *The Extended Mind: The Emergence of Language, the Human Mind and Culture*. U. of Toronto Press, 2007
- [52] Manovitch Lev, *The Language of New Media*. MIT Press, 2001
- [53] Marlow C., Naaman M., Boyd D. and Davis M. "HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, To Read". *Proceedings of the 17th Conferemce on Hypertext and Hypermedia*, 2006, 31-40
- [54] McLuhan Marshall, *The Gutemberg Galaxy*. University of Toronto Press, 1962
- [55] McLuhan Marshall, *Understanding Media: The Extensions of Man*. New American Library, NY, 1964

- [56] Morey D., Maybury, M. & Thuraisingham, B. (Ed.): *Knowledge Management, Classic and Contemporary Works*. MIT Press. Cambridge Mass., 2000
- [57] Nonaka I. & Takeuchi, H., *The Knowledge-creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford UP. NY, 1995
- [58] Ong Walter, *Orality and Litteracy, The Technologizing of the Word*. Methuen, London & NY, 1982
- [59] Panaccio Claude, *Les Mots, les concepts et les choses, la sémantique de Guillaume d'Occam et le nominalisme d'aujourd'hui*. Bellarmin-Vrin, Paris-St Laurent (Qc), 1992
- [60] Pascu Corina, *An Empirical Analysis of the Creation, Use and Adoption of Social Computing Applications*, European Commission, Joint Research Centre and Institute for Prospective Technological Studies. 2008
- [61] Peirce Charles, *The Essential Peirce: Selected Philosophical Writings*, Indiana UP, 1998
- [62] Rastier François, « La triade sémiotique, le trivium et la sémantique linguistique » *Nouveaux actes sémiotiques*, n# 9, 54 p., Presses Universitaires de Limoges, 1990
- [63] Saussure Ferdinand, *Cours de linguistique générale*. Payot, Lausanne-Paris, 1916
- [64] Segaran Toby, *Programming Collective intelligence, Building Smart Web 2.0 Applications*, O'Reilly Media, Sebastopol, CA, 2007
- [65] Shirky Clay, *Here Comes Everybody: The Power of Organizing Without Organizations*. Penguin, 2008
- [66] Smith Gene, *Tagging: People-powered Metadata for the Social Web*. Peachpit New Riders, Berkeley, 2007
- [67] Spivack Nova, <http://novaspivack.typepad.com>
- [68] Surowiecki James, *The Wisdom of the Crowds*, Random House, London, 2004
- [69] Svenonius Elaine, *The Intellectual Foundation of Information Organization*. MIT Press, 2000
- [70] Tapscott Don, Williams A. D., *Wikinomics, How Mass Collaboration Changes Everything*. Portfolio, 2007
- [71] Tovey Mark (Ed.), *Collective intelligence: Creating a Prosperous World at Peace*. Oakton, VA: EIN Press, 2008
- [72] Turing Allan, "On Computable Numbers, with an Application to the Entscheidungsproblem" *Proc. London Math. Soc* (2) 42 (1936 – 7), p. 230 – 265.
- [73] Von Neumann John, *First Draft of a Report on the EDVAC*. Contract No.W-670-ORD-4926, between the United States Army Ordnance Department and the University of Pennsylvania. Moore School of Electrical Engineering. Partially reproduced in *The Origins of Digital Computers: Selected Papers*, edited by Brian Randell, Springer Verlag, Berlin-Heidelberg, 1973 (p. 375-382)
- [74] Weinberger David, *Everything Is Miscellaneous: The Power of the New Digital Disorder*. Henri Holt and Cie, USA, 2007
- [75] Wellman Barry, *Computer Networks as Social Networks*. Science, 293 (14 September), 2031-2034, 2001
- [76] Wenger Etienne, *Communities of Practice: Learning, Meaning, and Identity*. Cambridge UP, 1998
- [77] Yates Frances, *The Art of Memory*. University of Chicago Press, 1974
- [78] Zara Olivier, *Le Management de l'intelligence collective*. M2 Editions, Paris, 2004