Towards a Pragmatic Approach to Computational Hermeneutics

Charles Bodon

n°6|2024 - LE NUMERIQUE COMME OBJET, LE NUMERIQUE COMME SYSTEME

Towards a Pragmatic Approach to Computational Hermeneutics

BODON Charles

Plan de l'article

Auteur

BODON Charles

Doctorant en Philosophie contemporaine

EXeCO (UMR 8103)

Université Paris 1 Panthéon-Sorbonne
12 Pl. du Panthéon
75 005 Paris
France
bodonbruzel@gmail.com

Citer l'article

Bodon, C. (2024). Towards a Pragmatic Approach to Computational Hermeneutics. *Revue Intelligibilité du numérique*, 6|2024. [En ligne]

Abstract: We propose a perspective article on using computational hermeneutics models, i.e., the analysis of texts assisted by artificial intelligence. It will not be a question here of developing a technical approach to construct these models, but rather of showing how, based on the existing data, some general principles for a pragmatic approach to this field can be proposed using computer ontologies. We will see that these AIs can be used to create hermeneutic models capable of producing formal and semantic representations of sociolinguistic contexts from the analysis of texts. The main suggestion here will be to use these models to reproduce a hermeneutic network, i.e., a collaborative human-machine interaction capable of both interpreting and clarifying the concepts of a field of practice. In conclusion, we will argue that these machines are capable of expressing meaning through symbolic forms that emerge from their interactions with experts.

Keywords: artificial Intelligence, computer ontologies, hermeneutics, human-computer interaction, philosophy of language.

Titre : Vers une approche pragmatique de l'herméneutique computationnelle

Résumé : Nous proposons un article de perspective sur l'usage de modèles d'herméneutique computationnelle, c'est-à-dire d'analyse de textes assistée par intelligence artificielle. Il ne s'agira pas ici de développer une approche technique pour construire ces modèles, mais plutôt de montrer comment, à partir de l'existant, quelques principes généraux pour une approche pragmatique de ce domaine peuvent être proposés en utilisant les ontologies informatiques. Nous verrons que ces IA peuvent être utilisées pour créer des modèles herméneutiques capables de réaliser des représentations formelles et sémantiques de contextes sociolinguistiques à partir de l'analyse des textes. La principale suggestion ici sera d'utiliser ces modèles pour reproduire un réseau herméneutique, c'est-à-dire une interaction collaborative humain-machine capable à la fois d'interpréter et de clarifier les concepts d'un domaine de pratiques. En conclusion, nous défendrons que ces machines sont capables d'exprimer du sens à travers des formes symboliques qui émergent de leurs interactions avec les experts.

Mots-clés : intelligence artificielle, ontologies informatiques, herméneutique, interaction homme-machine, philosophie du langage.

Introduction: Symbols Between Operation and Meaning

There is a gap between being able to use a language and being able to understand it. Particularly in the case of human-machine interaction where the latter can use words that are meaningful for the former but without expressing any thoughts. This dichotomy has been at the foundation of the field of artificial intelligence (AI) since A. Turing (1950) asked in his test if a machine, by playing a language game with humans, would be able to think by itself or to "carry out something which ought to be described as thinking but which is very different from what a man does". This question received a famous negative answer with the thought experiment of the "Chinese room" of J. Searl (1980) where he states that a machine does not "think" by itself or "understand" propositions but only computes logical signs together and that therefore has no intelligence at all. This thought experiment led Searle to distinguish between weak and strong AI: weak AI being the typical kind of system we interact with every day and which is limited to performing specific tasks without any kind of self-awareness, and strong AI being the hypothetical model conscious of itself and able to understand multi-contextual situations as humans do.

Thus, from the beginning of AI's history, the question of meaning has taken a significant place. But moving from a theoretical question in the philosophy of the mind, computational linguistics took the problem of human-computer interaction from a practical aspect. Notably, the question of meaning in AI has been reformulated as the "symbol grounding problem" by S. Harnad (1990), who asks how the semantic meaning of a sign and a proposition could be interpreted by a formal system:

> *"How can the semantic interpretation of a formal symbol system be made intrinsic to the system, rather than just parasitic on the meanings in our heads? How can the meanings of the meaningless symbol tokens, manipulated solely on the basis of their (arbitrary) shapes, be grounded in anything but other meaningless symbols?"*

In other words: how can we fill logical signs with semantic meaning for a machine? Is it possible to inscribe intentional content into data? If machines are mainly syntaxial systems able to analyze with precision logical structures, is there any way that the grammatical structure of a sentence or any proposition read by a machine could capture or at least reflect its meaning? If so, does this distinction between syntax and semantics have to be necessarily maintained between humans and machines? This situation calls into question the problem of interpretation which is also the domain of hermeneutics and for which we can draw some historical lines to understand its link with AI.

A Brief History of Hermeneutics: From Philology to Computers Analysis

Traditionally, hermeneutics is the science of the interpretation of symbols and sacred texts and its purpose is to reveal an esoteric or hidden meaning of the text with a high degree of symbolic interpretation that is supposed to be related to a particular level of reality by using different technic of interpretation (numerology, analogy, etc.). In philosophy, a science of interpretation can be originally located in the *On Interpretation* of Aristotle which shows how ontology and language can be defined by logical and linguistic categories and how words can be symbols of the state of mind.

However, it was particularly between the 18th and 19th centuries that hermeneutics started looking for the foundations of a rigorous methodology to avoid any misinterpretation. To do so, Friedrich Schleiermacher established the principle of the hermeneutic circle: to understand an author (his historical, psychological, environmental context), you need to provide a grammatical analysis of the text (the grammatical features), and to analyze the text you need to understand the author. In other words, this method aims to understand the mind behind the language by explaining how the language is structured in the text. Doing so, this method tries to show that the linguistic context of a text can reflect the historical, sociological, and psychological context (i.e., the contextual features) that influenced its author.

Hermeneutics then took a turn in social sciences during the 19th century with Wilhelm Dilthey who distinguished the methodology of human sciences and natural sciences: the first, more focused on understanding the psychic life by using intuition, the second explaining the causal relation between facts by analysis. Following this distinction, the 20th century saw, with Martin Heidegger and his student Hans-Georg Gadamer, and after them, Paul Ricœur, the emergence of philosophical hermeneutics became more explicitly the interpretation of phenomena, psychic processes, and existential questions, but also textual and symbolic features such as the metaphor.

Also during the 20th with the help of digital technologies, as shown J. W. Mohr, R. Wagner, and R. Breiger (2015), hermeneutics became content analysis procedures consisting of transforming a text into computer-readable digital data sets. These early methods focused on a small, very specific number of units of a text (primary ideas, simple views) to map this textual information into informative units by applying formal methods, regardless of style, nuances of expression, or poetic meaning. On the other hand, humanities pursued a close-reading approach focusing on the complexity of the text (such as its historical context, rhetorical forms, repetitions, syntax, semantics, symbolic mediations, intertextuality, etc.) but coming with a difficult, time-consuming, and always provisional interpretation. With the arrival of Big Data and the presence of new algorithm tools such as AI, the way humans interpret and interact with texts changed drastically. As the authors highlight, the emergence of "computational hermeneutics" changed not only the methods but also the theoretical approach of the text as a whole, by taking into consideration its relation to a social environment. Notably, these models allow applying formal methods to extract the complexity of the text's internal structure and meaning in a short amount of time, but also to represent the external features (contextual and cultural) that come with it.

It is with J. Mallery, R. Hurwitz, G. Duffy (1986) that we can find one of the first suggestions of reinvesting some principles of traditional hermeneutics for machines' text analysis to improve knowledge representation. Inspired by history, the idea of the authors is to reproduce the hermeneutic circle with the help of AI's grammatical analysis. In other words, the idea is that the machine explains the grammatical structure of a text or any kind of proposition to clarify its meaning for humans and makes it possible to infer the contextual components (i.e., psychological, historical, material components) contained in it.

Indeed, for the authors, there is an analogy between the hermeneutic circularity and the computational notion of "bootstrapping". As they explain, "booting" is the process of starting a computer and its software, and involves a loop of steps in which, like a chain reaction, a small program executes a more complex one which then allows another to be launched and which, at the end of the chain reaction, lights up the computer screen. Bootstrapping is in this sense the operation made by the computer to improve itself by interpreting its own programs. As the authors observe, the hermeneutics' circle shares the same idea of a virtuous circle where the understanding of a text is improved by its grammatical analysis and, reciprocally, its grammatical analysis is made possible by its understanding:

> *"Hermeneutics theories and applications also share the idea of the hermeneutic circle [...] Circles or spirals of understanding arise in interpreting one's language [...] in confirming a theory and in distinguishing between background knowledge and facts. [...] The grammatical thrust has a bootstrapping flavor: It places the text (or expression) within a particular literature (or language) and reciprocally uses the text to redefine the character of that literature. The psychological thrust more naïve and linear. In it, the interpreter reconstructs and explicates the subject's motives and implicit assumptions. Thus, claimed that a successful interpreter could understand the author as well as, or even better than, the author understood himself because the interpretation highlights hidden motives and strategies."*

This approach of computational hermeneutics models proposes a hermeneutic circle of understanding-explaining transposed to a network of human-computer interaction. The idea is that, in its collaboration with humans, the machine should be able to realize a process of bootstrapping of expanding its understanding of specific fields of science by learning from the data provided by texts written by humans of the domain. The advantage of this approach is that it finds a tempered position between the one of Searle and Turing about AI's understanding ability. Indeed, by distributing and limiting machines to the task of logical and grammatical analysis of texts, computational hermeneutic respects the facts that AIs are nowadays principally technical tools (as Searle defends it), but also, by implying humans in the process of interpretation, it renews Turing's imitation game and allows the machine to learn to interact with them and to reproduce how they use words.

It is by following this history and idea that our contribution in this perspective article will be to show that the recent state of the art allows us to draw some lines of research for a human-computer interaction based on such a philosophy of language and technology in a pragmatic and hermeneutic approach, i.e., an approach that considers that is from and by practices that AI models can provide efficient content analysis by reproducing a hermeneutic circle with experts.

Problematic and Methodology: How to Make Our Signs Clear?

A hermeneutic project for AI can seem contradictory. How could machines, which are formal systems, be looking for a way to interpret natural language if they only understand formal relations? Indeed, it is only through the inference provided from what they read and the semiotic power of signs that machines can get any information about a world that they don't perceive. In this sense, the facts of the external world are only accessible to them through a specific formalism. Therefore, the first methodological problem for computational hermeneutics models is to justify how mathematical abstractions could adequately represent phenomena that are, by nature, non-formal. Bas C. van Fraassen (2008, p. 240) states the problem in these words:

"How can an abstract entity, such as a mathematical structure, represent something that is not abstract, something in nature? [...] We have only access to structures, meaning relations between things. But what is exactly the relation between the empirical data and the mathematical formalism of the theory?"

Reformulated in our context: how could data correspond to empirical facts such as the propositional content of a text? And how the grammatical structure of a proposition could represent the context from where it emerges?

From this problem of correspondence between empirical data and abstract structures, there is also a problem of application of these structures. As F. Rastier (2004) highlights, there is no perfect natural language translation by formal languages, which makes it impossible to provide a unique and objective model able to analyze every kind of language or text. Every linguistic item (words, terms, expressions, grammatical relations) has a specific nature and the degree of complexity of its relation with others is always dependent on a specific context of use. In other words, languages and texts cannot be analyzed or represented by the same formal relations, otherwise, it would lead to projecting a "structured prejudice" to every kind of language or text without respect to their differences (of genres, uses, meaning).

To answer these two problems, we are going to see that computational hermeneutics models can manage the relation between empirical data and their formal representation by building the models directly in and by the practices, avoiding at the same time any preconceived structure. In other words, these models are specifically built by and through the context of sociolinguistics domains of practices and, therefore, propose an approach to manage the variability of meaning between linguistic practices and respect their differences without trying to unify them in one language or objective structure.

First, we are going to present the ontological foundation for a pragmatic approach. We will see how C. S. Peirce's semiotics theory can express situations and practices in the world, and how this approach can be implemented in the AI context through John Sowa's conceptual graphs. Based on this theoretical background, we will show that meaning between humans and machines is co-constructed directly in their interactions through signs.

Secondly, we will present computer ontologies, AIs that can be used to apply computational hermeneutics models. We will see that these knowledge organization systems share common properties with the philosophy behind hermeneutics, especially about the language-world relation, and can be considered as tools used to provide semantic interoperability between systems and agents.

Then we will present some applications of computational hermeneutics models that combine both principles and tools we will have defined. We will see that by providing formal and semantics graphs of knowledge and statistics on the recurrences of words in a text or any kind of sociolinguistic domain, these models should be able to build and represent precise definitions of how humans use some words according to some context. Also, by reasoning on these data, machines should be able to learn from a database directly built from and in the practices, leading to clear and logical inferences about the concepts that constitute their domain.

This will lead us, in conclusion, to redefine what it means "to mean" for machines and humans according to these analyses.

Semiotics as Ontological Foundation

For an ontological foundation of the computational hermeneutics models that work with a pragmatic program, we can rely on J. Sowa (2000, 2008, 2013, 2015) who developed an approach in AI taking care of the relations between language and the world by applying the logical theory of signs of C. S. Peirce. The field of semiotics itself owes a large part of this codification to C. S. Peirce (1994, p. 362):

"(CP 2.227) Logic has no other general name than semiotics, the formal doctrine of signs. [...] By formal, I mean that we observe the characters of the signs and that we only know this observation, by a process that I will have no objection to calling abstraction [...]."

If the purpose of hermeneutics is to seek the meaning of signs in a text, semiotics is the study of the rules and laws of their manipulation in a logical structure. Also J. Sowa (2000) noticed that, for Peirce, semiotics is the science that studies the use of signs by "any scientific intelligence". J. Sowa considers that Peirce means here "any intelligence capable of learning by experience", including animal intelligence and even mindlike processes in an inanimate matter such as computers. In this sense J. Sowa (2015) shows that computer techniques for AI to manage knowledge bases can meet Peirce's criteria about apprehending signs and his theory could be used as a semiotic foundation for ontology.

And it is indeed true that semiotics and logical thinking are connected for Peirce and not limited to linguistic perspectives. For Peirce, a sign can be an object, a thought, or a judgment, and there are different types of signs that he categorized according to their logical function. Perceiving a sign is obtained by what he calls an "abstractive observation" which Peirce describes as an experience of perception of the thing in the mind. He describes this experiment as akin to mathematical reasoning: the experience of abstractive observation consists of creating a "skeleton diagram" or a "silhouette diagram" of a proposition to represent its structures and logical laws with clarity. By renewing the triadic relation thought-symbol-world of Aristotle, C. S. Peirce (1994, p. 363) proposes a semiotic triangle interpretant-representamen-object to define the structure a sign takes to mean something:

> *"(CP 2.228) A sign, or representamen, is something which stands to somebody for something in some respect or capacity. It addresses somebody, that is, creates in the mind of that person an equivalent sign, or perhaps a more developed sign. That sign which it creates I call the interpretant of the first sign. The sign stands for something, its object. It stands for that object, not in all respects, but in reference to a sort of idea, which I have sometimes called the ground of the representamen."*

The sign is defined by a triadic relationship between i) the representamen, ii) an object, and iii) the interpretant. The representamen is the form that the sign takes, "something that stands for someone in some respect or title": the name "Ulysses" for example is the representamen of a cat called Ulysses; the object (a cat) is what is represented; and the interpretant is a judgment or a representation of an object by a subject applied to the object thanks to its representamen. Here, the semiosis, the process of meaning that constitutes the sign, is involved in this triadic relation: the interpretant becomes a concept or a thought of "cat" thanks to the representamen "Ulysses" which stands for the physical object "cat". In other words, for Peirce, semiosis is a process of perception and logical thinking according to practical effects: a sign does not have the same meaning as another sign depending on the context it is perceived and used.

Based on such principles, he manages to develop "existential graphs" which put in correspondence signs, logical inferences, and knowledge representation of concepts. The interest of these graphs is that they dispose of equivalence with symbolic languages, allowing us to express some concepts analytically and also by a graphical representation.

It is by following this idea of clarifying knowledge through logical and graphical representation that J. Sowa (2008) developed his own "conceptual graphs" for informatics which aim to unify computational linguistics, logic based on semantic networks, and cognitive support of machine reasoning processes. These conceptual graphs reproduce Peirce's existential graphs in that they can also be translated into logical propositions. They have multiple topological forms (hierarchical, cyclic, lists, etc.) or logical relations and allow representing graphically on a computer screen the syntactic structure of the elements composing a proposition in the natural language like: "John takes the bus to go to Boston" which expresses a fact in the world.

For example, in this proposition, knowledge can be represented in a formal structure that is made up of concepts, themselves divided into a semantic network of categories: agent, object, action, places, and else. In the graph, each of the four main [concepts] represents the type of entity to which it refers: [John], [Go], [Boston], or [Bus]. One of them expresses the main action, [Go], that centralized the graph, and the three others have names that identify their referent: "John", "Boston", and "Bus". Also, these concepts are linked to a label in parentheses which represents the (class) to which they belong: agent (Agnt), destination (Dest), person (Person), or instrument (Inst). Then, the graph describes knowledge as follows by connecting concepts and classes together: [John] is an (Agnt) and an instance of the class of (Person), [Boston] is an instance of the class (city) and is a (Dest), and the [Bus] is an (Inst) to [Go] to [Boston]. For (J. Sowa, 2008), this graph can also be translated by the following formula which thus becomes a resource that can also be readable by a machine by paraphrasing it with logical signs: $\exists x \exists y$ (Go(x) $\wedge$ Person(John) $\wedge$ City(Boston) $\wedge$ Bus(y) $\wedge$ Agnt(x, John) $\wedge$ Dest(x, Boston) $\wedge$ Inst(x, y)).

Here, the interest of such conceptual graphs for computational hermeneutics models is that their form expresses the logical relation between signs, making them readable by both humans and machines. As J. Sowa (2013) summarized it:

> *"Graphs have advantages over linear notations in human factors, computational efficiency, and cognitive representation. For readability, graphs show relationships at a glance that are harder to see in linear notations. They also have a highly regular structure that can simplify many algorithms for reasoning, searching, indexing, and pattern matching"*.

Thus, the application of conceptual graphs makes knowledge intelligible and interpretable for both humans and machines:

- Materially: Taken as such, the sign is the material mark that stands for the entities it represents in the proposition. (e.g., the letters x, y, z as variables).
- Relationally: The sign puts one thing in a specific relation with another. (e.g., the relation "Dest(x, y)" where an object x go to a destination y).
- Formally: The sign makes it possible to organize the syntactic entities together to formulate a proposition that can attribute properties to an object. (e.g. "$\exists x Px$" where P is any predicate of an object x).
- Contextually: The sign refers to a situation and tries to put in concepts contextual knowledge of propositions (e.g., the fact in the world that John takes the bus to go to Boston). The definition of the sign then also takes place under the context in which it is applied.

This pragmatic approach of interpreting signs can serve as an ontological foundation for the use of computer ontologies, according to the fact that these AIs models are specifically designed in and by an interaction with linguistics practices to produce formal and semantic knowledge graphs. Indeed, computer ontologies can be built on texts and contexts of scientific fields through collaborative practices among experts in different fields of science. Now, we do not intend to present the technical details of how automated reasoning for higher-order logic can aim at formalizing natural-language argumentative discourse (for this, see : D. Fuenmayor, C. Benzmüller, 2019), but rather we will present how, through the general hermeneutics principles and their application to computer ontologies, we can see a logical analysis and a conceptual explanation of the entities contained in the propositions of a text can be provided.

Knowledge Reasoning and Representation With Computer Ontologies

While the current field of AI seems to be dominated by connectionist approaches, notably large language and generative models (such as ChatGPT), we argue here that "old-fashioned AI", i.e., symbolic AI, can be more efficient when it comes to modeling meaning and sociolinguistic interactions. Since their emergence in the 1960s, symbolic systems have aimed to represent knowledge or situations in the world by modeling their items and relations through logical structures (what M. Minsky referred to as "frames"). Particularly, it was during the 1990s that computer ontologies emerged as tools to structure the entities (e.g., concepts, objects) within a specific domain, specifying their logical and linguistic relationships to provide context to that domain, while also organizing these entities from the most abstract to the most concrete.

These AI models are the most suitable for our purpose, it is because, as F. Neuhaus (2023) explains, "they can provide standardized and controlled vocabulary [...], specify their vocabulary semantic in both machine and human-readable form [...], they may also represent empirical knowledge". Computer ontologies are therefore more explainable and transparent than LLMs, because they manage interoperability between agents (humans and machines) by resolving ambiguities where LLMs (which are statistic systems), most of the time, "navigate through them". Then, it is not a question of putting aside LLMs, but rather putting them at the service of symbolic systems.

However, computer ontologies come with their own problematics and that we can connect to those of hermeneutics. As R. Davis, H. Shrobe, and P. Szolovits1(993) shows it, in knowledge representation (KR²), a question arises about the relation between the external world and the internal system of a computer:

> *"Any intelligent entity that wishes to reason about its world encounters an important, inescapable fact: reasoning is a process that goes on internally, while most things it wishes to reason about exist only externally. A program (or person) engaged in planning the assembly of a bicycle, for instance, may have to reason about entities like wheels, chains, sprockets, handle bars, etc., yet such things exist only in the external world. This unavoidable dichotomy is a fundamental rationale and role for a representation: it functions as a surrogate inside the reasoner, a stand-in for the things that exist in the world. Operations on and with representations substitute for operations on the real thing, i.e., substitute for direct interaction with the world. In this view reasoning itself is in part a surrogate for action in the world, when we cannot or do not (yet) want to take that action."*

Inevitably, there is a dichotomy between the formal internal system of a machine (abstract structures) and the external facts of the world (empirical data). Therefore, the problem here is to find a method that could establish a correspondence between both to allow the machine to reason. But this dialectic also shows us the proximity that KR² and hermeneutics share in their problems and purpose. Here, KR² tries to implement the external facts of the world in an internal formal system called "Knowledge Organization System (KOS)" (e.g., digital libraries, taxonomies, dictionaries, lists, computer ontologies), while hermeneutics try to understand the external and contextual features that influenced the internal intention and inner thought of an author by analyzing the grammatical structures of a text. Therefore, both share this external/internal duality and aim to find a path of linking them through an analysis of language.

G. Hodge (2000) presents in three main points how this relation between external facts and the internal system can be organized into a computer system:

- A KOS imposes a particular view of the world on a collection and the items in it.
- The same entity can be characterized in different ways, depending on the KOS that is used.
- There must be enough in common between the concept expressed in a KOS and the real-world object to which that concept refers that a knowledgeable person could apply the system with reasonable reliability. Likewise, a person seeking relevant material by using a KOS must be able to connect his or her concept with its representation in the system.

KR² is therefore a field that uses formal systems to represent and manipulate general knowledge by formalizing concepts into informatics models. In KR², machines like computer ontologies can be used to semantically represent knowledge.

For computer ontologies, the relations between concepts can be manipulated and interpreted to fix a specific context of science. In the KR²'s context, a computer ontology can be presented as a computer artifact that formally represents in a lattice the common knowledge or the language practice of a field of specialties from a semantic point of view. As S. Staab and R. Studer (2004) explain, "*Ontology is the study of "things that exist", within the domain of computer science, an ontology is a formal model that allows reasoning about concepts and objects that appear in the real world and (crucially) about the complex relationship between them*". Their knowledge base can be made up of a corpus of texts on which it can be applied techniques of hypertext and semantic graphs to structure their data and reflect the "relationship between objects in the real world" by using languages allowing description. For example, one of them, OWL (Web Ontology Language), allows the machine to specify the relation of subsumption ("x is_a y") between entities, classes, or concepts.

The computer ontology is presented in a diagram like a mind map or a semantic network. This form allows contextualizing the relationships that entities of a given domain have with each other. Several types of ontologies (top ontologies, core ontologies, and domain ontologies) can be used separately or in combination with the same or separate domains to specify different kinds of entities and levels of abstraction (for example, metaphysical concepts with top ontologies, interdisciplinary concepts with core ontologies, and concrete and domain-related concepts by domain ontologies). To build them T. Gruber (1994) defines epistemological precepts such as clarity (i.e., avoiding ambiguities), coherence (i.e., avoiding inferences or axioms that imply contradictions), extendibility (i.e., being able to deal with polysemy), minimal encoding bias (i.e., minimizing the dependence of knowledge of a particular formalism) and minimal ontological commitment (i.e., maintaining a correspondence between vocabularies and entities and concepts of a domain) that aims to ensure their correspondence with their external environment and interoperability with other systems and agents.

In other words, ontologies are interesting tools for our purpose because they need but also provide interdisciplinary exchanges between philosophers, AIs, linguists, and experts of a domain to formalize texts and languages in the service of knowledge engineering. Combining their roles in the development of terminologies and the corpus of ontologies, this interdisciplinarity centralized by AI makes possible the learning loop necessary for computational hermeneutics models, thus allowing them to represent with multiple perspectives the different degrees of generalization of the entities that constitute texts and their fields of application.

However, the representative function of these AIs does not provide any realism or reveal the objective conceptual structures of the agents' minds who use them. These systems only help to make intelligible a cultural and social background and help experts understand it and make inferences about the meaning of structured data. Therefore, we defend here the same instrumental position of G. Declerck and J. Charlet (2014), considering that ontologies are principally tools used to enhance human cognitive abilities and play the role of mediator between agents. Now we have defined the principles and the tools we are using we can present some applications of computational hermeneutics models.

Computational Hermeneutics Models as Collaborative Symbolic Systems

As J. Mallery, R. Hurwitz and G. Duffy (1986) put it, these models try to pursue the art of traditional hermeneutics with the help of continual feedback between its internal (grammatical) and external (author's context) features. In other words, computational hermeneutics models aim to apply the traditional hermeneutic circle of understanding-explaining to a network of human-computer interaction where text analysis serves as a basis to help both improve their understanding of a sociolinguistic context. Here, the text is a cultural artifact used to feed a machine with data, and their logical analysis of contextual and intentional features aims to provide experts with a clear overview of their domain. Computational hermeneutics models should be able to accomplish several tasks that can be combined:

- Contextualize meaning: The model should be trained into a hermeneutic network allowing the machine to be trained in a sociolinguistic environment, providing it with several acceptations of the meaning of a word according to the practices of its users (meeting here (L., Wittgenstein, 1958, §43)'principle according to which "the meaning of a word is its use in language").
- Genre classification: The model should be able to classify the entities of a text or a domain to state the type of corpuses and categorize their features, such as the emotion the text expresses (e.g., romantic novel, political discourses, thriller, informatics codes), or the terms and concepts a domain is accepting (e.g., classes, individuals, instances, properties).
- Semantic mapping: The model should be able to learn and represent the grammatical and formal relations and features of the entities by mapping them dynamically into semantic networks or conceptual graphs, just like children are learning the meaning of words from a graphical representation. By this mapping, the machine can formalize the relations of the objects and propositions of a domain or contained in a text and represent them in a comprehensible way for humans.
- Semantic interoperability: The model should be able to describe data in a clear and understandable language for both humans and computers. By operating semantic interoperability, i.e., transmitting signs and maintaining stable and consistent meaning, computational hermeneutics models should transform human-machine

networks into collaborative systems that allow transactions, i.e., exchanges of convertible signs between artificial and natural languages.

To contextualize meaning, this conception of hermeneutics in AI can lead to a "hermeneutic network" proposed by J. Zhu and D. F. Harrell (2009) which aims to analyze the narration of a text through the interaction of authors and machines. The purpose here is to help the machine to acquire an "intentional vocabulary", i.e., a lexicon of words that the humans define according to their understanding of their meaning. For the same word, different users could have a different understanding: then each of them has to inscribe the definition they have in the machine to allow it to enrich its internal vocabulary. The idea is that the more the machine manages this "flexibility of meanings", i.e., the several understandings a word can have, the more it will be able to manage eventual ambiguities and in which context and purposes a word may be used. In the hermeneutic network, users write their own social experiences and cultural backgrounds into the computer system when they interact with the machine to convey to it the meaning of the words they use. In other words, the hermeneutic network reproduces Wittgenstein's concept of "community of language" according to which meaning is collectivity and dynamically built by shared practices.

To classify the genre of a text and its narrative structure H. R. Alker, W. G. Lehneret and D. K. Schneider (1985) propose a model to extract its affective contents by working on contextual units (such as the relationships between the characters, the events they are facing, or the affects they express) and aim to summarize their story by representing their relations. By analyzing the context and the connections between these narrative elements, the affective core of the text can be compared to some others, allowing the machine to state its genre. For example, the authors concluded that the events of Jesus's story conform to a well-known genre which was the romance of self-transcendence.

On the same idea, the intentional content of a text can be exposed by an analysis of its structure, here, based on the recurrences and uses of words. D. Mayaffre (2002) provides the analysis of a discourse of French politics in 1930 from the right-wing based on the occurrence of the verb "having", declined in "to have" and "has". This right-wing discourse shows a passive observation of the social situation in the 1930s where the word "have" is most often conjugated to the past, when by comparison some discourses of the left-wing in the same era are rather conjugated to the future. D. Mayaffre (2002) observes that this speech evokes conservatism and that this affective feature seems to be inscribed in the grammatical structure of the speech given its temporal form, and he infers that this affective trait (conservatism) would be in the mind of the speaker or in what it aimed to mean or to transmit to its auditorium.

For semantic mapping J. C. Mallery and G. Duffy (1986) propose a model of "semantic perception" that allows the machine to realize a semantic analysis of data from syntactic forms. Formally, semantic mapping takes its inspiration from the semantic memory of R. Quillian and A. Collins (1968) which aimed to produce a structural model representing the cognitive activity of the association of ideas in the mind. Here each node represents concepts or ideas and every arc is the connection between them. By formalizing lexical items into different classes of words, this model takes into consideration the polysemy and intentional structure of communicative situations. It aims to represent the grammatical relations that words could manage together and, by a process of mapping, the model formalized words into relations that a machine could handle. The graph can also be used to represent the association of ideas and their relations that a domain of knowledge maintains, just as the conceptual graphs.

For semantic interoperability, a language like Information Economy MetaLanguage (IEML), developed by P. Lévy (2023), can enrich data and graphs through a formal metalanguage. From a pragmatic perspective, such a language is crucial for analyzing and structuring metadata across multiple levels of granularity to resolve ambiguities, and operating the translation between humans and machines. For instance, IEML can encode concepts while specifying the syntagmatic roles of elements (e.g., subject, verb, predicate) using a standardized grammar, making it possible to model the diverse meanings of a sentence depending on its context. Thus, when combined with the analytical capabilities of conceptual graphs, which logically model propositions, IEML provides a dynamic framework for representing semantic relationships that are flexible and generative. In this way, machines potentialize meanings in a text, which humans actualize through interpretation.

With the help of a common metalanguage, computational hermeneutics can account for the evolving dynamics of a collective intelligence, emerging "from a heterogeneous network involving people, technical devices and messages (composed of symbols)" (Lévy, 1997).

All of these methods aim to provide a model of a domain of discourse through a formalization of its uses by the means of an exchange between humans and machines. As D. Fuenmayor and C. Benzmüller (2019) put it, experts are involved in a "dialectical exchange with the computer" in which it extends its database from the axioms inscribed by the experts, and by the inferences and graph representation that it provides in return, it allows experts to clarify logical relations between entities.

Also, we can see that these models respond to Moravec's paradox that states that what is easy for the human is difficult for the machine (e.g., understanding contextual meaning), and what is easy for the machine is difficult for the human (e.g., making complex computations). Here, the explanatory power of machines is used in tandem with the natural understanding of humans to answer any question about meaning. Therefore, we can attempt to define the philosophy of these models in these terms:

Def.: A computational hermeneutic model is a collaborative symbolic system between humans and machines to analyze textual data. To understand the contextual meaning of words, machines need humans, and to explain logical and grammatical structure, humans need machines. In this circle, both improve their ability to understand and explain the meaning of the text and the contextual characteristics of the propositions of a language thanks to a balanced distribution between formalization (machines) and interpretation (humans). The objectivity of the model is then ensured by the interaction it involves between agents.

Conclusion: What Does It Mean "to Mean" for a Machine?

We have presented in this article some possibilities for computational hermeneutics models, their purpose, and principles by reproducing Schleiermacher's hermeneutic circle in a digital context. We have investigated the feasibility of hermeneutics in AI through the principles of Peirce's semiotics and pragmatics applied to conceptual graphs with J. Sowa's models (2000, 2008, 2013, 2015), and we also saw, following Wittgenstein, how the uses of words can be defining their meaning in a hermeneutic network J. Zhu and D. F. Harrell (2009). By using ontologies Fuenmayor D. and Benzmüller C. (2019) and logometric analyses D. Mayaffre (2002) to provide semantic representations J. C. Mallery and G. Duffy (1986), gender classifications H. R. Alker, W. G. Lehneret and D. K. Schneider (1985), and semantic interoperability P. Lévy (1997, 2023), we have seen how AI can support the cognitive activity of interpretation of humans by analyzing the uses of words, by building dictionaries applicable to specific or different contexts, or by avoiding projecting a genre pre-structured on a text by making it emerge directly from its composition. Through these distributed-symbolic systems, meaning is preserved for humans. But in which sense is that the case for machines?

To be able to mean something seems, at first sight, to be able to use a language to intentionally signify and express something (an idea, a proposition, an emotion, or else) that can be perceived by the senses or represented in the mind (a word, a sign, a mental picture). But machines are essentially mathematical tools and not intentional beings. They are only computing and the way they express something is by automatically and mechanically responding to requests or accomplishing tasks: they don't have any psychic life that could produce meaningful ideas. Also, machines do not have biological organs nor can perceive things directly: they only recognize and represent things through formal structures. In other words, the only thing that "exists" for AI systems is, as T. Gruber (1993) noticed, "what can be represented", i.e., what is formally inscribed in its internal system. This position implies, as B. Bachimont (2022) notes, that there are no "words" or "propositions" or even such a thing called "text" for a machine because a text is an object whose content always refers to an external, existential, and meaningful ecosystem where humans are living. We can't even properly say that machines "read" anything, or only if by "reading" we assume that what computers only do is "computing numerical signs".

Thus, it seems imprecise to use the verb “to mean” for AI, because it implicitly implies an intentional feature. However, if we agree on the fact that machines do not signify things, we can state that they express results that can have meaning for humans. Therefore, we suggest that the question “What does it mean “to mean”?” when it comes to machines in the context of computational hermeneutics models should be replaced by the question “What does it mean to express something?”, which is easier to answer because it is more specified and put aside the intentional implicit of the verb “to mean”. We can also clarify the problem more accurately by following the formulation of C. Taylor (1985, p. 219) about meaning and how it is expressed:

> *“What is meant by 'expression' here? I think it means roughly this: something is expressed when it is embodied in such a way as to be made manifest. And 'manifest' must be taken here in a strong sense. Something is manifest when it is directly available for all to see. It is not manifest when there are just signs of its presence, from which we can infer that it is there, such as when I 'see' that you are in your office because of your car being parked outside. [...] Expression makes something manifest in embodying it."*

Here, Taylor emphasizes the fact that language isn’t just a mechanical function but also produces a range of symbolic forms, i.e., structures of meaning that need to be embedded in a medium to be interpreted. The relation between meaning and its expression is organic: it can only be expressed when it is embodied and structured to be manifest.

We can transpose this definition into our models by considering that meaning circulates in the hermeneutic network when humans inscribe data into the machines’ system that they structure into a symbolic form. That is to say: “to mean” for an AI, in the context of computational hermeneutics models, is providing a symbolic form or formal bodies (e.g., semantic networks, conceptual graphs, computer ontologies) of the data inscribed in its system to make it clear for experts.

Defined that way, these symbolic forms are no longer pure abstract structures but inscriptions that carry with them the environment from which they come. It would be an “inscription error”, as B. C. Smith (1998) calls them, to think that the data provided by computers are not influenced by the dynamics and actions by which they are built. In other words, no formalism can claim to be absolutely separate from sensibility or culture, or, as J. Cavaillès (1997, p. 53) highlights, a “theory of science can be clarified and specified by formalization, it is not constituted by it.” (our translation).

Now we can respond to the problem of correspondence of van Fraassen. From a pragmatist point of view, the dichotomy between the external world and the internal system of the machine vanishes, since there is a continuum between mathematical structures and the ecosystem of practices from which they inherit their meaning.

Also, the symbol grounding problem of Harnad finds a solution that way. If it is not possible to inscribe meaning directly into the data, it is because it is already in it. However, it is only with the conjoin forces of humans and machines that we can make it manifest through symbolic structures.

In other words, if AIs are not intelligent, conscious, or organic entities capable of perceiving the world, their computational activity is not meaningless nonetheless. By entering data from a domain into their systems, AI necessarily improves their analytics capacities by learning from it. And, by structuring these data, they provide symbolic forms that experts interpret and can restructure to feed again the machine to bring out refined data. This learning loop between humans and machines causes new knowledge to emerge precisely from the human-machine interaction and this process can be reproduced until reaching a fixed point. From a pragmatic point of view, the hermeneutic network can then be considered as an activity of documentation and re-documentation about itself, trying to understand its own linguistic and epistemological ambiguities, cultural and methodological biases. In other words, computational hermeneutics models imply a meta-hermeneutic process.

Acknowledgment

We warmly thank Adrien Barton and Jean Charlet for their proofreading, comments, and helpful feedback.

Bibliography

Alker Jr. H.R., Lehnert W.G., Schneider D. K. (1985). Two Reinterpretations of Tonbee's Jesus: Explorations in Computational Hermeneutics. In *G. Tonfoni, ed. Artificial Intelligence and Text Understanding. Quaderni di Ricerca Linguistica*, 6, 49-94

Bachimont B. (2022). La complexité herméneutique à l'épreuve du calcul. *Interfaces numériques*, 10(3), ⟨10.25965/interfaces-numeriques.4686⟩. ⟨hal-03906427⟩

Cavaillès J. (1997). *Sur la logique et la théorie de la science*, Vrin

Declerck G., Charlet J. (2014). Pourquoi notre sémantique naïve n'est pas formalisable et pourquoi c'est (presque) sans conséquence sur l'ingénierie ontologique. In: *Intellectica. Revue de l'Association pour la Recherche Cognitive*, n°61, Philosophie du Web et Ingénierie des Connaissances. pp. 143-188

Fraassen (van) Bas C. (2008). *Scientific Representation: Paradoxes of Perspective*. Oxford University Press, Clarendon Press

Fuenmayor D., Benzmüller C. (2019). A Computational-Hermeneutic Approach for Conceptual Explicitation. arXiv:1906.06582v2 [cs.AI]

Gruber T., (1994). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal of Human-Computer Studies*, 43(5-6)

Harnad S., (1990). The Symbol Grounding Problem. *Physica D: Nonlinear Phenomena, Elsevier*, Volume 42, Issues 1-3, June 1990, Pages 335-346

Hodge G., (2000). *Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files*. The Digital Library Federation Council on Library and Information Resources 1755 Massachusetts Avenue, NW, Suite 500 Washington, DC 20036

Lévy P. (1997). *Collective Intelligence: mankind's emerging world in cyberspace*. New York: Plenum Press

Lévy P. (2023). Semantic computing with IEML. *Collective Intelligence*, Volume 2, Issue 4, October 2023, https://doi.org/10.1177/26339137231207634

Mallery J., Hurwitz R., Duffy G. (1986). Hermeneutics: From Textual Explication to Computer Understanding. Massachusetts Institute of Technology, *Artificial Intelligence Laboratory, A.I. Memo No. 871*

Mallery J. C., Duffy G. (1986). A Computational Model of Semantic Perception. *AI Memo No. 799, Artificial Intelligence Laboratory*, MIT, Cambridge

Mayaffre D. (2002). L'Herméneutique numérique. *L'Astrolabe, Recherche littéraire et Informatique*, 2002, pp.1-11. hal-00586512

Mohr J. W., Wagner R., Breiger R. (2015). Toward a computational hermeneutics. *Big Data & Society*, 2(2)

Neuhaus F. (2023). Ontologies in the era of large language models – a perspective. *Applied Ontology*, 18(4):399-407

Peirce C. S. (1994). *Collected Papers of C. S. Peirce* (CP). electronic edition reproducing Vols. I-VI ed. Charles Hartshorne and Paul Weiss (Cambridge, MA: Harvard University Press, 1931-1935), Vols. VII-VIII, ed. Arthur W. Burks (same publisher, 1958)

Quillian M. R., Collins A. M. (1968). Retrieval Time from Semantic Memory. Bolt Beranek and Newman, Inc., Cambridge, Massachusetts, *Journal of Verbal Learning and Verbal Behavior*, 8, 240-247

Rastier F. (2004). Ontologie(s). C.N.R.S., *Revue des sciences et technologies de l'information, série : Revue d'Intelligence artificielle*, vol. 18, n°1, p. 15-4

Searle J. R. (1980). Minds, Brains and Programs. *The Behavorial and Brain Sciences,* Vol. 3, Cambridge University Press

Smith C. B. (1998). *On the Origins of Objects*, The MIT Press

Sowa J. (2000). *Ontology, Metadata, and Semiotics*. ICCS'2000 in Darmstadt, Germany, Published in B. Ganter & G. W. Mineau, eds., Conceptual Structures: Logical, Linguistics, and Computational Issues, Lecture Notes in AI #1867, Springer-Verlag, Berlin, pp. 55-81

Sowa J. (2008). *Conceptual Graphs*. preprint of Chapter 5 of the Handbook of Knowledge Representation, ed. by F. van Harmelen, V. Lifschitz, and B. Porter, Elsevier, pp. 213-237

Sowa J. (2013). From Existential Graphs to Conceptual Graphs. *International Journal of Conceptual Structures and Smart Applications*, 1(1):39-72

Sowa J. (2015). Signs and reality. *Applied ontology*, 10 (3-4):273-284

Staab S., Studer R. (2004). *Handbook on Ontologies. International Handbooks on Information Systems*, Springer

Taylor C. (1985). *Human Agency and Language*. McGill University, Montréal, Cambridge University Press

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, Vol. 59, No. 236 (Oct., 1950), pp. 433-460

Uschold M., Grunninger M. (1996). Ontologies: Principles, Methods and Applications. *Knowledge Engineering Review*, vol.11, n°2

Wittgenstein L. (1958). *Philosophical Investigations*. Translated by G. E. M. Anscombe, Blackwell

Zhu J., Harrell D. F. (2009). *System Intentionality and the Artificial Intelligence Hermeneutic Network: The Role of Intentional Vocabulary*. UC Irvine, California Digital Library, University of California