Philosophy in Informatics VI

Frontiers of philosophy of computing and information

Organizers:

Chair of History & Philosophy of Science, Pontifical University of John Paul II

Comission on Philosophy of Science, Polish Academy of Arts and Sciences

Warsaw University of Technology

Krakow – Warsaw (online), Dec 16-17th, 2021

Department of History and Philosophy of Science, Faculty of Philosophy, Pontifical University of John Paul II in Krakow, Commission for Philosophy of Science, Polish Academy of Arts and Sciences and Faculty of Administration and Social Sciences, Warsaw University of Technology invite proposals for the presentations on the 6th Conference on Philosophy in Informatics. This year the theme of the conference is **Frontiers of philosophy of computing and information**. The possible topics include but are not limited to:

- Non-Turing computational paradigms
- Natural Computing
- General Theory of Information - interpretations and consequences
- Ontology of information
- Epistemology and ontology of computer simulation
- Ontology of virtual reality
- Morphological computation
- Biosemiotic inspirations in artificial Autonomous Systems
- Information and Computation in autopoietic systems
- History of AI
- Explainable AI
- Ontology of AI systems
- Ontological gap in autonomous AI systems
- Philosophical foundations of AI paradigms
- Judgment and phronesis as AI paradigm?
- Artificial General Intelligence (AGI) - philosophical prerequisites
- Super intelligent systems - conceptualizations
- Technology beyond the horizon - threats and possibilities

The proposals for the presentations should be submitted for the anonymous review to <phil.in.info.2021@gmail.com>. On the separate page the author(s) should submit the full name, affiliation and a short CV (up to 150 words). The proposals should not exceed 300-400 words (including references). We accept presentations in English or in Polish.

Important dates:

Nov 30th - Deadline for abstracts
Dec 5th - The authors of accepted abstract are notified.
Dec 12th - Program finalized; Zoom contact sent out.
Dec 16-17th - Conference online (ZOOM)

The conference will be organized using ZOOM platform. There are no fees for attending the conference. The list of accepted abstracts will be published before the conference. The authors of selected abstracts will be invited to submit the paper (up to 6000 words) for publication in the special issue of the philosophical journal "Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)" 2022.

The conference Web site is https://calculemus.org/fi6/
Inquiries should be sent to conference secretary Dr Roman Krzanowski <phil.in.info.2021@gmail.com>

Abstracts

(in alphabetical order of first Autor family name)

Abstracts

INVITED LECTURE

Gordana Dodig Crnkovic,

Natural Information, Computation, Cognition and Intelligence: In the Search of Common Framework for Natural and Artificial Cognition and Intelligence

This talk presents a conceptual analysis of cognition and intelligence, natural and artificial, based on the study of information structures and processes, informed by the up-to date knowledge of related scientific fields. It suggests two-way, biomimetic learning: from models of nature to constructive study of new artifacts and back from increasingly sophisticated artifacts to models and theories of natural systems (such as brains, swarms, or unicellular organisms). Building such an info-computational framework requires generalization of concepts of information, computation, cognition, intelligence in relation to the (computational) process of evolution as extended evolutionary synthesis which explains processes of emergence of new more complex cognitive/intelligent organisms. Possible applications of new cognitive biomimetic architectures are in medicine, new cognitive and intelligent computing technologies, regenerative medicine as well as micro- and nanorobotics.

About the Author

Gordana Dodig-Crnkovic is Professor in Interaction Design at Chalmers University of Technology, and Professor in Computer Science at Mälardalen University. She holds PhD degrees in Physics and Computer Science. Her current research is in Morphological computing and the connection between computation, information and cognition via interacting agents on different levels of organization - from physics to biology and cognition, to morphogenetic and biomimetic computational design. Dodig-Crnkovic is member of Faculty Board, Chalmers AI Ethics committee , European Network For Gender Balance in Informatics , and Ethics4EU project.

ABSTRACTS

in alphabetical order of first authors

Oliver Buchholz,

A Means-End Account of Explainable Artificial Intelligence

(Eberhard Karls Universität Tübingen, Ethics and Philosophy Lab)

Explainable artificial intelligence (XAI) seeks to make machine learning (ML) methods explainable for humans. However, there is considerable disagreement about what this means and how to achieve it. In this talk, I draw on insights of means-end epistemology to structure the field: different authors pursue different ends and different means are appropriate to achieve them.

The first part of the talk gives a short introduction to XAI by providing several examples from the field. I show that contributions to the XAI literature differ widely, both on a conceptual and on a methodological level. This raises the question whether the field is scattered into various disconnected subprojects or whether there is a common structure that is shared by the variety of approaches pursued in the literature.

The second part outlines relevant aspects of means-end epistemology. *Means-end epistemology* takes epistemology to be a normative discipline; it is based on a principle of instrumental rationality. Rational agents are assumed to have certain epistemic ends and they ought to adopt certain means if and only if they further these epistemic ends (Huber, forthcoming; Schulte, 1999).

The third part combines the two introductory sections. It describes the means-end account of XAI and discusses both its normative and descriptive component. I argue that problems of XAI can be framed as problems of means-end epistemology: given an epistemic end (making something explainable, transparent, etc.), XAI tries to provide means that further the given end. This account has several important consequences. First, it explains why disagreement arises in the field: the divergence in instruments of XAI follows from the disagreement on epistemic ends. Second, it structures the field: there is a common methodology of developing appropriate means for given ends. Third, this structure has a normative component: according to means-end epistemology, different means ought to be rationally adopted to achieve different epistemic ends. Therefore, the ends of an explanation normatively constrain the set of admissible means to achieve it. The means-end account thus reveals how the suitability of particular instruments of XAI is prescribed by the ends for which an explanation is sought. Fourth, this structure also has a descriptive component: different authors specify different ends and different means are appropriate to achieve them. This allows for a taxonomy which classifies existing contributions to the field along the specific means-end relations that are considered.

About the Author

Oliver Buchholz is a PhD student in philosophy at University of Tübingen's Cluster of Excellence "Machine Learning: New Perspectives for Science". He workson philosophical aspects of explainable artificial intelligenceand more generally on methodological issues arising atthe intersection of philosophy of science and the philosophy of machine learning. Prior to becoming a PhD student, Oliver obtained master's degrees in both Philosophy and Economicsfrom University of Tübingen.

Hajo Greif,

Models, Mechanisms and Organisms in Turing and Ashby

(Warsaw University of Technology)

This paper will outline the differences in approaches to and resources of "producing models of the action of the brain" (Turing 1946) in Alan M. Turing and W. Ross Ashby, who were in conversation on these topics as members of the "Ratio Club". Ashby (1960) explicitly committed himself to building analogue machine models of the adaptive behaviours of brains and other systems, their functions and their relationships to their environments, all understood in explicitly Darwinian terms. However, he restricted his focus to the origins of adaptive behaviour by learning, leaving aside "genic" adaptation, and therefore the organic basis of that behaviour. Conversely, Turing developed a notion of idealized theoretical machines, known as "logical computing machines", which originally served metamathematical purposes but informed the concrete design of the digital computer. He used his theoretical machines for inquiries into a varied set of phenomena, from proto-connectionist models of the brain via simulation of conversational behaviour to pattern development in organisms. Notably, in the latter (1952) he relied on the non-Darwinian account of morphogenesis in Sir D'Arcy Thompson's On Growth and Form (1942). We will broadly outline the state of biological theorizing on which Turing and Ashby relied at the time of their writing, and ask how their specific biological commitments may have influenced their choice of modelling approach.

About the Author

Hajo Greif is a philosopher working in the fields of history, philosophy and social studies of science and technology. He received his doctorate from TU Darmstadt and was habilitated at the University of Klagenfurt (AAU), Austria. Hajo's main research areas are the history and philosophy of cognitive science (AI vs cybernetics), situated cognition and cognitive artefacts ('4E' cognition, smart environments), and naturalism in the philosophy of mind (teleosemantics, evolution of cognition). After positions at AAU and the Munich Center for Technology in Society, Technical University of Munich, he is now Research Assistant Professor at the Department of Philosophy and Ethics in Administration, Warsaw University of Technology.

Phillip Hintikka Kieval,

Unsupervised Discoveries, Understanding, and Semantic Opacity

(University of Cambridge)

State-of-the-art machine learning (ML) models based on deep neural networks (DNNs) exhibit predictive accuracy that far outstrips that of previous hand-coded models. This success has generated widespread optimism at the prospect of using DNNs to enhance the march of scientific progress in a diverse range of fields. By using these models to explore massive collections of data, scientists might be able to produce novel discoveries with DNNs. However, DNNs are opaque in ways that preclude scientific understanding. Thus, their predictive accuracy comes at the cost of one of the central aims of scientific inquiry. This tension has brought about increased interest in explainable AI (XAI), a growing discipline that aims to understand and explain how these models work. However, I argue that XAI as it is now conceived cannot deliver on the promise of understanding in the scientific contexts in which DNNs purport to show the greatest promise, namely in exploratory contexts. Traditional mathematical and computational models typically depend on parameters whose values represent the properties of their target system. By contrast, DNNs lack any such interpretable mapping between parameters and properties of real systems. Instead, these models depend on hyperparameters that determine how a model will learn to generate input/output mappings in a way that optimizes some measure of performance. Currently, XAI aims to explain *how* these models optimize performance and make decisions. Yet, in exploratory contexts where scientists hope to produce discoveries, it will be the very content of a model's output that requires understanding. In such contexts, scientists will deploy unsupervised models, which generate clustering patterns from unlabeled data. Given that the hope is to explore domains of inquiry that are not already well understood, we will lack the interpretative machinery to decide if the clusters of data produced by such a model correspond to real, scientific kinds rather than jerry-rigged ones tied to spurious correlations in big data. In exploratory contexts we lack the conceptual frameworks necessary to say what the output clusters of unsupervised learning models *mean*. I call this phenomenon *semantic opacity*. When confronted with semantic opacity, the knowledge required for interpreting the decisions of a model depends on theoretical assumptions about the very domain of inquiry about which we had hoped the model could teach us. I argue that this should be the understood as the more pressing problem facing XAI in the context of science.

Roman Krzanowski & Paweł Polak,

Meta-Ontology of AI Paradigms

(Pontifical University of John Paul II in Krakow, Faculty of Philosophy)

Each AI paradigm postulates some kind of ontology (let us call a synthetic ontology) of the real world. The type of ontology determines how well AI paradigm may internally represent the reality and how well it can interact with its external environment. The failures of AI in creating AGI, ethical robots, or socially acceptable AI systems may be traced to the gap between the AI ontology and the ontology of the real world. This paper discussed ontological assumptions of past and current AI paradigms, their limitations stemming from these ontological models, and suggests the directions of the future development of AI systems that would narrow the gap between the reality and the synthetic ontologies of AI constructs. We need to clarify basic concepts to avoid possible misinterpretations of the discussed topic. The concepts such as Ontology, meta-ontology, paradigm, AI paradigms, may have, and have, many different interpretations so we better state what we mean by them in this study. We define knowledge as a representation of facts, facts are treated as primitive concept, about the specific domain; the domain may be the whole world or its selected aspects. Ontology is defined as the study of the nature of existence and what is real. Ontology of AI is defined as the computer representation of knowledge about the world with a set of concepts and relationships that exit among those concepts. Meta-ontology explores concepts and assumptions in different ontologies (of real or synthetic domains). Meta-ontology is interested in, among other things, in what the ontological commitments of a given theory are. We are investigating the theories of ontology of the real world not of virtual worlds that is why we use the term synthetic ontlogies. Paradigm is a set of assumptions, concepts, values, and practices that constitutes a way of viewing reality for the community that shares them, especially in an intellectual discipline. AI Paradigm is a set of assumptions, concepts, values, and practices that determine how AI model the real world. Meta-Ontology of AI Paradigms is the inquiry into the ontological commitments of AI paradigms i.e., what kind of ontology AI paradigms imply. Further. We distinguish three AI paradigms

(1) The content of the world can be determined (exhaustively) by the a priory rules and procedures (Descartian view)

(1a) First AI wave – GOFAI. Knowledge systems, XCON, DENDRAL, ACRONYM, MYCIN, etc. (1950-1980**) [logical intelligence]**

(1b) Probabilistic/ Bayesian reasoning **[probabilistic intelligence]** Reasoning with uncertainty

(2) The content of the world can be determined by generic procedures with a given external objective targets (Evolutionary view)(1980-2010)

Emergent intelligence, swarm intelligence, EA, GA, **[emergent Intelligence]**

(3) The content of the world can be determined by learning from existing information (Data) (2000 –on) (Cognitivist views) ML, DNN, Big Data, pathways **[neural models of intelligence]**

Realizing the ontology of the real world or some form of ontology close to it, is not the guarantee of the AI system to be human-like, to have AGI. But it seems to be a necessary condition. Autonomous system be "like-us" must "see" what we "see"; for the same reasons dogs and other creatures will never be "like-us". An ontology, i.e., the representation of reality, may be conceptualized as "problem space". What is needed then is the proper search or decision algorithm that is able to find out the best solution over this problem space. It seems again that such an algorithm to be successful should approximate the concepts of phronesis or practical judgment – Aristotelian term for human decision making. Phronesis is unfortunately, not algorithmic in the sense of GOFAI, emergent intelligence or NN. How Ai system may realize phronetic principles is an ongoing research project.

About the Authors:

Roman M. Krzanowski – MA, Ph.D, D. Phil. He is an expert in Ethernet networking technology spatial information systems, information processing and philosophy of computing and information. He published books in information science and network technology. His interests in philosophy include the philosophy of information and informatics, ontology and metaphysics of nature, ethics and ethical problems created in information society. He published papers on robotic ethics, foundations of ethical AI, phronesis in robotics and Ethical testing of autonomous robots.

Paweł Polak – professor of philosophy at Pontifical University of John Paul II in Krakow, MA in telecommunication at AGH University of Science and Technology. Editor-in-chief of periodical "Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)"; member of Commission on the History of Science of the Polish Academy of Arts and Sciences, secretary of the Commission on the Philosophy of Science (the same Academy). Published in "Studies in Logic, Grammar and Rhetoric," "Studia Historiae Scientiarum," "The Philosophy of Science", "European Journal of Science and Theology." He published books about philosophical aspects of scientometrics and about philosophical reception of Special and General Relativity in Poland. His interests in philosophy include the history and philosophy of computing/informatics, history and philosophy of science, history of Polish philosophy, ethics of autonomous robots, philosophy *in silico*, science-religion studies, as well wine philosophy.

Adam Paweł Kubiak,

What's Wrong with the Theory of Evolution from the Turingian Perspective?

(Warsaw University of Technology)

We scrutinize Turing's scarce (one published article, and three other works posthumously extracted and reconstructed based on his manuscripts) but influential contribution to biology having the form of a development of the theory of morphogenesis. We ask the question why he preffered to work on the morphogenetic explanations of structures of organisms rather than on evolutionary explanations theirof.

One can approach this topic by stating that Turing had some specific biological commitments and that they influenced the choice of modelling approach. Some aspects of the topics and content of morphogenetic theory can be regarded as primary factor of his biological preferences. Nevertheless, this paper concentrates on seeking for modelling commitments as primary factors of his inclinations to morphogenesis rather than to the theory of evolution.

We explain how ideals of universality, recursiveness and automation present in Turing's vision of computing machine relate to characteristics of his approach to modelling present in his works on morphogenesis. We provide arguments why it is hardly/im- possible to model biological evolution with Turing's modelling approach. We find this fact to be plausible explanation of Turing's lack of interest in dealing with the theory of evolution. We relate this outcome to the to-date topic of the so-called genetic algorithms. We indicate that they cannot be conceived as something analogical to biological evolution

About the Author

I am interested in philosophical problems of scientific methods and scientific claims. I respect, accept and appreciate them despite their existential and philosophical vulnerability and hopelessness; these weaknesses are probably the most exciting aspect of science. My background is in natural science (University of Łódź) and philosophy (Catholic University of Lublin). Since the beginning of October 2021, I am a postdoctoral researcher at the Faculty of Administration and Social Sciences, Warsaw University of Technology, where I explore the exciting world of "Turing, Ashby and 'the Action of the Brain'" together with Hajo Greif (project's PI), Paweł Stacewicz (project's Co-I) and other companions.

Josh Lalonde,

Information and ontogenesis in Simondon

While today there is a growing literature on the ontology of information, one of the first philosophers to consider this question, Gilbert Simondon, is little-known in the Anglophone world. In his *Individuation in Light of Notions of Form and Information*, first published (in part) in 1964 and recently translated into English, Simondon seeks to develop an account of what he calls "tension of information". This notion extends the quantitative measure of information introduced by the mathematical theory of communication into the domain of the qualitative. Without contesting the validity of the quantitative measure in its technical domain, Simondon notes a paradoxical property of this notion of information: a highly-structured message such as a text in a natural language contains *less* information than a random sequence of characters, even though the former would be intuitively regarded as "informative" and the latter as "uninformative". It is this more intuitive sense of information that Simondon seeks to capture with the notion of tension of information.

Furthermore, the mathematical theory of communication treats only the *transmission* of information, leaving unexamined the *generation* of information in the first place. Simondon builds on the notion of tension of information to develop an informational ontology; information is, on this account, an essential principle in explaining the ontogenesis of physical entities. In particular, Simondon ties his account of physical individuation to Louis de Broglie's "double solution" theory of quantum physics.

I will first outline Simondon's criticism of the mathematical theory of communication and alternate notion of tension of information, then sketch the ontology he develops on this basis. Finally, I will compare Simondon's account of the individuation of quantum entities with the notion of "active information" introduced by David Bohm in a similar context

Wojciech Mamak,

Using Correspondence Theory of Semantic Information to Understand Neural Representational Similarity

(Polish Academy of Science)

The aim of this talk is to demonstrate the usefulness of a recent novel account of information in terms of structural similarity offered by Miłkowski (2021) in philosophical analyses of representational similarity analysis (RSA) in neuroscience (Kriegeskorte & Kievit 2013).

The relationship between information and representation is hugely important yet notoriously elusive in philosophy of neuroscience. Nevertheless, following Dretske's (1981) dictum that 'information is a rung on a ladder that gets one to representation', the conviction among philosophers and neuroscientists that the recipe for understanding cognitive representation should be based on a preceding naturalistic theory of neural information-processing remains strong (Sprevak 2020). This main idea is also a core tenet in scientific practice of experimental studies of neural decoding/encoding as theories of neural representations' content (Ritchie et al, 2019).

But traditional (e.g. Shannonian, teleosemantic) theories of information famously run into trouble when asked to account for the content of neural representations (Cao 2012). To remedy that, I propose in this talk to use a recent alternative Correspondence Theory of Semantic Information [CTSI] put forward by Miłkowski (2021), in which information is cashed out in terms of structural similarity. This theory formalizes the intuitive, yet undervalued in the literature on information idea that it is similarity relation between two objects (A and B) that affords information-bearing between A and B. It enables us to draw inferences about A when we only have access to B. In other words, similarity warrants us to perform valid inferential operations in the form of surrogative inference (Swoyer 1991). This idea is formalized using Barwise and Seligman's (1997) logic of information flow using the notion of infomorphism, conceptualized as a bidirectional relation between two typetoken classifications. To show its promise for philosophy of neuroscience, I turn to Representational Similarity Analysis (RSA).

RSA is a research program in neuroscience that uses a host of methods for mapping representational geometries and measuring their structural similarity to establish information-bearing relations between them (analogously to the theoretic import of CTSI). For example, RSA was successful in cross-region and cross-species representational comparisons, being able to show how some semantic relations between the schemas (e.g. taxonomic clustering; Kriegeskorte et al. 2008) are retained. RSA was also robust in predicting neural content, which lends credence to its promise of uncovering the structural makeup of neural representation. However, I argue that RSA's conceptual relation to traditional accounts of neural information (cashed out in terms covariance, correlation, or nomological necessity) is problematic and should be remedied. The main reason for that is what I suggest to informally dub 'format-friendliness'.

The core idea of RSA is that it is intentionally flexible, i.e. can demonstrate the existence of isomorphic relations between given representational geometries (mapped by the matrices), regardless of their format ('format-friendliness'). Isomorphic relations may obtain even when there is no statistical covariance or causal link between them. Traditional theories of neural information struggle here, whereas this crucial property lays at heart of CTSI.

I conclude then that CTSI theory can do a better job for laying solid philosophical foundations for RSA, as both share common understanding of similarity as format-agnostic, structure-preserving relation that affords valid *surrogative* inferences about the contents of representational schema.

Bibliography

Barwise, J., & Seligman, J. (1997). *Information flow: The logic of distributed systems*. Cambridge University Press.

Cao, R. (2012). A teleosemantic approach to information in the brain. *Biology & Philosophy*, *27*(1), 49–71

Dretske, F. I. (1981). *Knowledge and the Flow of Information*. MIT Press

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, *2*, 4.

Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: Integrating cognition, computation, and the brain. *Trends in cognitive sciences*, *17*(8), 401–412

Ritchie, J. B., Kaplan, D. M., & Klein, C. (2019). Decoding the brain: Neural representation and the limits of multivariate pattern analysis in cognitive neuroscience. *The British journal for the philosophy of science*, *70*(2), 581–607..

Miłkowski, M. (2021). Correspondence Theory of Semantic Information. *The British Journal for the Philosophy of Science*. https://doi.org/10.1086/714804

Roskies, A. L. (2021). Representational similarity analysis in neuroimaging: Proxy vehicles and provisional representations. *Synthese*, 1–19.

Sprevak, M. (2020). Two kinds of information processing in cognition. *Review of Philosophy and Psychology*, *11*(3), 591–611.

Swoyer, C. (1991). Structural representation and surrogative reasoning. *Synthese*, *87*(3), 449–508.

Alice Martin & Mathieu Magnaudet,

Modelling Interactive Computing Systems: Do We Have a Good Theory of What Computers Are?

(Interactive Computing Team (ENAC), France)

Computers are more and more interactive. They are no more transformational systems producing a final output after a finite execution. Instead, they continuously react in time to external events that modify their course of execution. While philosophers have been interested in conceptualizing what a computer is for a long time [5, 6], they seem to have paid little attention to the specificities of interactive computing, thus running the risk of not offering an adequate conceptualisation. In this paper, we propose to tackle this issue by a survey of the literature in theoretical computer science where one can find some explicit proposals toward a model for interactive computing. In that field, the formal modelling of interactive computing systems has been brought down to whether the new interaction models are reducible to Turing Machines (TMs). To our knowledge, there are three areas where interaction models are framed.

In all of them the comparison between TMs, oracle machines [7] and interactive system models is systematically at stake. These areas are namely work (i) on concurrency by Milner and his followers [3], (ii) on Reactive Turing Machines [1] and (iii) on interaction as a new computing paradigm [8].

For each of the three identified models, we:

• present the motivation behind it,

• sum up its account for interaction (its expressiveness and possible equivalence with a TM),

• identify how it has been used and criticised.

The survey shows two difficulties. First, Turing reducibility of interaction models is left unclear. Second, such analyses only focus on formal equivalence between models. They prevent the philosopher from getting a clear-cut comprehension of what makes a computing machine interactive. We suggest another path for philosophers: the view of computing from a cyber-physics perspective [2], where key concepts of time and causality help fleshing out the concept of interactive computing. This amounts to switching explanatory focus [4].

References

[1] Baeten J. et al. 2013. Reactive Turing Machines

[2] Lee EA. 2016 Fundamental Limits of Cyber-Physical Systems Modeling

[3] Milner R. 2006 Turing, Computing and Communication

[4] Milkowski. 2016. A Mechanistic Account of Computational Explanation

[5] Piccinini. 2008. Computers

[6] Smith B.C. 2002. The Foundations of Computing

[7] Soare RI. 2009 Turing Oracle Machines, Online Computing, and Three Displacements in Computability Theory

[8] Wegner P, Goldin D. 2003 Computation beyond Turing Machines

About the Authors

Alice Martin holds a master's degree in philosophy from Institut Jean Nicod (Paris) and is currently a PhD student at ENAC in the Interactive Informatics Team. She is working on a bidisciplinary subject in philosophy of computing and computer science She is interested in understanding interactive computing systems, and how they fit into the classical framework of computability. She also works on a visualization tool to support interaction programming.

Mathieu Magnaudet is an associate professor at ENAC, in the Interactive Informatics Team. His main area of research is in programming languages for interactive systems. He is strongly involved in the development of the Smala programming language. He also works on the philosophical topics related to the field of computer science especially regarding the limits of the Turing paradigm to model interactive computing systems.

Łukasz Mścisławski,

Is information something ontological, or physical or perhaps something else? Some remarks on R. Krzanowski approach to concept of information

(Wrocław University of Science and Technology)

As one can easily see, the title of our presentation is a little bit provocative. We have found the proposal of R. Krzanowski approach to the problem of information very interesting [1], [2], [3]. Our aim is to point out that it has some advantages when it comes to answer some fundamental questions in philosophy of physics, metaphysics and, as we suppose, philosophy of information and computer science as well. As the issue of great importance we see a proposal of introduction some subtle distinctions in relation between ontological and epistemological information, that can be seen as analogous to G.F.R Ellis analyses of passage of time in his concept of Crystallizing Block Universe[4]. This step could be found useful in further study of relations between different types of information.

We also would like to formulate some subjects of further study, for which proposal of R. Krzanowski can be very solid foundation and good start point to examine traditional metaphysical issues using classical philosophical doctrines combined with new approach.

References

[1] R. Krzanowski, „What Is Physical Information?", Philosophies, t. 5, nr 2, Art. nr 2, cze. 2020, doi: 10.3390/philosophies5020010.

[2] R. Krzanowski, „Does purely physical information have meaning? A comment on Carlo Rovelli's paper: Meaning = information + evolution [arXiv:1611.02420]", arXiv:2004.06716 [physics], kwi. 2020, Dostęp: lis. 30, 2021. [Online]. Dostępne na: http://arxiv.org/abs/2004.06716

[3] R. Krzanowski, „Why can information not be defined as being purely epistemic?", ZFN, nr 68, s. 37–62, wrz. 2020.

[4] G. F. R. Ellis i R. Goswami, „Space time and the passage of time", arXiv:1208.2611 [gr-qc], sie. 2012, Dostęp: cze. 10, 2014. [Online]. Dostępne na: http://arxiv.org/abs/1208.2611

Kiran Pala,

Information Emergentism and Transformations

(Boston University, Boston, USA and University of Eastern Finland, Finland)

Recently a range of practices that have occupied the periphery of scholarly thinking about information emergence, has yielded many insights on artificial systems. By exploring plausible reasons, why these practices are so often overlooked while at the same time potentiality reveals how, attending to 'quality of things' enables new ways of thinking about knowledge transformation and supervenience in artificial systems. Yet despite the potential, this article holds for enriching philosophical discussions of artificial systems, and their transformativeness functionality continue to remain neglected themes in philosophical practices of information and computing. Beyond, The Continuity of Mind by Michael Spivey (2008), Cognitive Systems and the Extended Mind By Robert D. Rupert (2009), a few works in intelligence philosophy, and The Dynamics of Control by Fritz Colonius, Wolfgang Kliemann (2012) on emergence feature a sustained engagement with the topic of knowledge transformation, and their based applications.

All the more surprising, considering this transformation has emerged from the studies of cognition and perception. An increased interest in the theories on nature of the knowledge sources and quality of things are relevant to the current and long standing discussions in philosophy of information and computing. In addition, the recent information emergentism in artificial systems also served to highlight a range of issues which, despite having received little attention by philosophers, nonetheless hold important philosophical implications, such as the role of 'quality of things' in enabling the transformativeness functionality. Given the great potential studies of emergence hold for enriching philosophical discussions of information and computing. The present aim of this study is to initiate a meta-analysis on historical assessment of epistemic symmetry between introspection and emergence within framework of modern control theories: what inferential causes of information emergence as a consequence, ie. transitions of "things" in view of the adaptation, at least in principle. While we do not deny the obvious insufficiencies of current methods, I hypothesise that they can be overcome by future scientific, methodological and technological developments: from the devices used in houses to large technical systems which surround us.

About the Author

Kiran Pala, PhD – his research background is in philosophy of mind and science, consciousness, artificial systems and control theory, and philosophy of cognitive science.

Paula Quinon & Jens Ulrik Hansen,

The Role of Expert Knowledge in Big Data and Machine Learning

(International Center for Formal Ontology, Warsaw University of Technology)

According to popular belief, Big Data and Machine Learning provide a *brand-new* approach to science that has the potential to revolutionize scientific progress (Hey et al., 2009; Kitchin, 2014). The extreme version of this belief is illustrated by Anderson's claim that Big Data and Machine Learning in science will lead to the "end of theory" (Anderson, 2008). The idea behind this extreme version of the belief is that advanced Big Data and Machine Learning algorithms enable us to mine vast amounts of data related to a given problem without prior knowledge and we do not need to worry about causality, as correlation is all that is needed.

The extreme version of this belief is not seriously held by many philosophers of science, but there are several serious attempts to determine the extent to which Big Data and Machine Learning imply a resurgence of inductive methods (Pietsch, 2021) or agnostic science (Napoletani et al., 2021). Without claiming that "the theory came to its end", these approaches advocate new scientific methods that can be applied to various fields in a similar way, without the need for domain-specific knowledge.

Two questions arise in connection with these views: Where did these ideas come from? and to what extent are they justified? The first question could be addressed by following the hype around Big Data and Machine Learning in industry and how easily conversations about new innovations and disruptions translate into conversations about paradigm shifts in science. The leakage of the style of argumentation and the attraction to hype, from industry to science, should be carefully watched and frequently questioned. In these times of pressure to bridge the gap between business and science, it can be difficult to distinguish valuable insights and ideas from superficial buzz-talk. The case of Big Data and Machine Learning is one of the areas, which shows how important it is to clearly map the flow of knowledge between business and science. We do not, however, make any scientific progress towards this question, we only sketch our observations here.

Regarding our second question, we argue that using methods from Big Data and Machine Learning is not a passive mode where you feed raw data into a Big Data or Machine Learning algorithm and wait for the algorithm to detect correlations between features of a massive dataset. We argue that there is always work in manipulating data, cleaning data, etc. that requires significant domain knowledge of scientific applications.

We agree that expert domain knowledge used in Big Data and Machine Learning is of a different kind from that required by traditional methods. The question we ask is about the required amount of expert knowledge needed for Big Data and Machine Learning methods to function in an efficient manner. By carefully assessing examples of Big Data and Machine Learning applications in science, such as skin cancer detection (Esteva et al., 2017), protein folding (Senior et al., 2020) and language generation (Brown et al., 2020), we assess which knowledge plays a role in each of these cases. We observe that significant domain knowledge is needed, not so much in theory and model building, but in data preparation and validation of Machine Learning models. Data needs to be labelled appropriately, decisions need to be made about which data to include, and new features might need to be created. Furthermore, we need to know what constitutes a well-functioning Machine Learning algorithm for a given problem, what we should measure and what is a good enough value to constitute a solution to the problem.

In addition to expert knowledge about data samples, specific knowledge of Machine Learning is also often needed, as algorithms cannot be applied blindly in practice. A promising model architecture needs to be selected, appropriate data augmentation techniques need to be applied to increase the performance of the algorithm, and the algorithm needs to be tuned and adjusted to get good enough Machine Learning performance.

We do not intent to dismiss new scientific methods or new lines of research, but to disclose what work and knowledge the new methods require, and thus show in more detail what is new and what is business as usual. Big Data and Machine Learning methods may be used in a more agnostic way, but they do not lead to completely agnostic science. It is not a question of changing or revolutionizing science, but of expanding the methodological toolkit. This will certainly not necessarily revolutionize all science, but it could lead to changes in subfields and provoke the emergence of new fields or endeavors, such as digital humanities or computational social science.

In our talk, we will discuss concrete case studies (skin cancer detection, protein folding, language generation), where we present methods that are used and we highlight those moments where expert knowledge is involved. We will attempt to classify various aspects of expert knowledge involved in the application of Big Data and Machine Learning methods, for instance, the expert knowledge necessary at the training data sample preparation or the expert knowledge necessary for choosing algorithms. We will also suggest that, depending on the field, the range of traditional methods vs agnostic methods varies, leading us to believe that the process of "agnosticization" is different from field to field, and that the possibility to reach the "no-theory" stage varies depending on domain. Consequently, we observe that the way in which Big Data and Machine Learning methods enter scientific methodology involves continuous small conceptual shifts rather than a rigid paradigm shift in Kuhn's sense.

References

Anderson, C. (2008). The End of Theory: The Data Deluge Makes the ScientificMethod Obsolete. *Wired Magazine*, *16*(7). https://www.wired.com/2008/06/pb-theory/

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., … Amodei, D. (2020). Language Models are Few-Shot Learners. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, & H. Lin (Eds.), *Advances in Neural Information Processing Systems* (Vol. 33, pp. 1877–1901). Curran Associates, Inc. https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115–118. https://doi.org/10.1038/nature21056

Hey, T., Tansley, S., & Tolle, K. (Eds.). (2009). *The fourth paradigm: data-intensive scientific discovery*. Microsoft Research.

Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1), 2053951714528481. https://doi.org/10.1177/2053951714528481

Napoletani, D., Panza, M., & Struppa, D. (2021). *The Agnostic Structure of Data Science Methods*. https://arxiv.org/pdf/2101.12150.pdf

Pietsch, W. (2021). *Big Data*. Cambridge University Press.

Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., Žídek, A., Nelson, A. W. R., Bridgland, A., Penedones, H., Petersen, S., Simonyan, K., Crossan, S., Kohli, P., Jones, D. T., Silver, D., Kavukcuoglu, K., & Hassabis, D. (2020). Improved protein structure prediction using potentials from deep learning. *Nature*, *577*(7792), 706–710. https://doi.org/10.1038/s41586-019-1923-7

Marcin Rabiza,

The Dual-Process Attribution of Artificial Intelligence Agency

(Polish Academy of Sciences, Institute of Philosophy and Sociology)

In recent years, research on machine agency has been particularly fruitful, as it was driven by the rapid and widespread development of artificial intelligence (AI) technology. Because of the perceived similarity to humans in terms of the capability to act in the environment, we may experience a shift of AI from technical artifacts to artificial agents. Along with the increasing presence of intelligent machines in our daily practices, the interest of researchers in the problem of AI agencies is growing.

In this light, the aim of the paper is to propose a dual-process model of agency attribution phenomenon. Two general modes of thinking about machine agency are differentiated, and possible relation to AI interpretability is indicated. The author argues that even if a strong ontological rationale behind granting AI agential status cannot be reconstructed, agency attribution may be epistemically relevant, as an instrumentally rational mental strategy for interpreting, predicting, and explaining the behavior of AI actors. The preliminary dual-process theory of attribution may provide an input for further research in not only philosophy of mind or commonsense psychology, but also in human-computer interaction and user experience studies.

References

1. Ciardo, F.; Beyer, F.; De Tommaso, D.; Wykowska, A. "Attribution of intentional agency towards robots reduces one's own sense of agency". Cognition 2020, 194.
2. Cowley, S.; Gahrn-Andersen, R. "Drones, robots and perceived autonomy: implications for living human beings". AI Soc. 2021.
3. Dennett, D.C. The Intentional Stance, 1st ed.; MIT Press: Cambridge, Mass., 1987.
4. Kahneman, D. Thinking, Fast and Slow, 1st ed.; Farrar, Straus and Giroux, 2011.
5. Nomura, O.; Ogata, T.; Miyake, Y. "Illusory agency attribution to others performing actions similar to one's own". Sci. Rep. 2019, 9.

6. Rammert, W. “Where the Action is: Distributed Agency between Humans, Machines, and Programs”. In Paradoxes of Interactivity: Perspectives for Media Theory, Human-Computer Interaction, and Artistic Investigations; Seifert, U.; Kim, J; Moore, A., Eds.; 2015; pp. 62–91.
7. Rose, J.; Jones, M. “The Double Dance of Agency: A Socio-Theoretic Account of How Machines and Humans Interact”. Systems, Signs and Actions 2005, 1, 19–37.
8. Swanepoel, D. “Does Artificial Intelligence Have Agency?” In The Mind-Technology Problem: Investigating Minds, Selves and 21st Century Artefacts; Clowes, R.; Gartner, K.; Hipólito, I., Eds.; Studies in Mind and Brain, Springer, 2021; pp. 83–104.

About the Author

Doctoral candidate in Philosophy at the Institute of Philosophy and Sociology of the Polish Academy of Sciences. Graduate of the Adam Mickiewicz University in Poznan in Philosophy (BA in 2019) and Ethics (MA in 2021), as well as from the Poznan University of Technology in Robotics (BEng in 2019) and Computer Science (MSc in 2020). Information technology professional currently employed at IBM, and former academic teacher at the Collegium Da Vinci. Interested in philosophy of mind and artificial intelligence, researching mind-technology problems.

Marcin Rządeczka,

Computational psychiatry from an epistemological perspective. How computational paradigm can facilitate the analysis of complex mental disorders?

(Maria Curie-Sklodowska University in Lublin, Institute of Philosophy)

In 2010, the US National Institute of Mental Health formulated a new conceptual model to facilitate psychiatric research known as the Research Domain Criteria (RDoC), which takes a fundamentally different perspective than the dominant Diagnostic and Statistical Manual of Mental Disorders and tries to answer a different set of proximate and ultimate questions.

The main aim of RDoC is to deepen the understanding of psychopathology through pathophysiology by building upon the advances in the computational neurobiological sciences. The RDoC hypothesizes that behaviors cannot be understood without taking into account the variety of individual developmental trajectories and environmental influences upon behavior. Within this paradigm, each mental disorder is a dimensional construct from illness to health, without a specific well-defined demarcation line.

The RDoC can easily serve as a basis for the nascent field of computational psychiatry. In theory, neural processes can be modelled by algorithmic representations that describe information processing in the complex multi-level neural system. Computational psychiatry describes the structures and mechanisms of the nervous system in terms of information processing. For example, impairments in the processes involved in predictive coding could, in theory, explain a variety of psychopathological phenomena, ranging from the impoverished theory of mind in autism spectrum disorder to peculiar abnormalities of smooth-pursuit eye movements in schizophrenia spectrum disorder.

Integrating computational modelling into psychiatry can facilitate research in several fundamental and novel ways. What are the fundamental biopsychological components involved in mental disorders and what are the mathematical relationships between these components? How do local dysfunctions of the endocrine or immune system create complex interactions with the nervous system and finally lead to some mental illness? Why natural selection has not eliminated many gene variants responsible for some of the most debilitating mental disorders, such as schizophrenia, autism, bipolar disorder or depression. These are only preliminary questions that require the computational paradigm due to their sheer complexity and the interdisciplinary nature of the research involved.

Last but not least, computational psychiatry creates an interesting opportunity for an epistemologist to revaluate computational theories of mind, which have been discarded due to the neurobiological turn. From such a research perspective mental disorders can be analysed as suboptimal algorithms running by the computational mind and resulting in dysfunctional behavior.

About the Author

Marcin Rządeczka holds a Ph.D. in philosophy and currently works at the Institute of Philosophy of Maria Curie-Sklodowska University in Lublin. He has a keen interest in philosophy of science, with particular focus on the philosophy of biology and the philosophy of computer sciences. He is the author of one book and several papers and the member of the Polish Philosophical Society and the Polish Cognitive Science Society.

Ábel Ságodi,

Aligning computational models of temporal phenomena with data through category theory

(Champalimaud Center for the Unknown)

In computational science, data used to be compared to simulations by mere visual comparison. This usually relies on some intuitive understanding of empirical adequacy of computational models through the similarity version of representation (Giere, 2004).

The first scientific theories were formulated about planetary motion in the form of lines drawn on paper to represent this motion through space. The Newtonian models represent planetary motion with solutions of ordinary differential equations, i.e. integral curves. These curves are not the same as the drawn lines to record the motion itself, because they are one dimensional while the drawn lines have dimension two (the also have a width). A similar issue arises when we consider the infinite precision of analog-to-digital converters in computers such as the successive-approximation ADC or when we consider computational models. From an epistemological point of view, we can better represent measurements as (non-zero measure) subsets of our state space rather than single points, which shifts our conception of what similarity should entail. We introduce a framework that can account for different temporal phenomena in this representation in terms of an associated model that explains them. We will consider the representation of temporal processes as objects in a category and a possible way to conceptualize what could count as an explanation.

This will be achieved by considering families of dynamical systems and looking which is empirically adequate for the observations we have at hand in this representation. The Conley-Morse graph summarizes the global dynamics of a system by showing the Conley index of all the separate invariant sets. For the current purposes it is sufficient to consider the category of Conley-Morse graphs with as objects Conley-Morse graphs and with as morphisms the identities. As the logical space we consider the set of all local flows defined on open subsets X _ Rn, constrained by the theory of the target system (assumptions about its composite parts, their ranges and interactions, etc). We define the category LocFlow to have as objects local flows (with possible constraints) and the morphisms to be topological equivalence. The existence of an explanation for a temporal process as represented by a Conley-Morse graph (as inferred from measurements), can be then defined in terms of the functor, which gives us a better notion of explanation in this similarity picture.

References

Giere, R. N. (2004). How models are used to represent reality. Philosophy of science, 71(5):742{752.

Krzysztof Sołoducha,

The project of coherent extrapolated volition as a tool for building confidence for decisions made by autonomous machines

(Military University of Technology, Warsaw, Poland)

The aim of my presentation is to consider the problem of whether we can build a database of decision-making patterns for autonomous machines that would allow them to make socially acceptable decisions in situations of moral dilemmas.

The reasoning I have done has two basic assumptions:

1. Modern autonomous machines can have full situational awareness, which allows them to fully analyse their environment and make appropriate decisions

2. The decisions made by autonomous machines are accepted by users from an ethical and legal point of view. The level of this acceptance and trust in their decisions may determine their use or not by users from different cultures and societies. Therefore, my question is whether we are able to build universal patterns for making such decisions ?

The basic statement underlying these considerations is that we are currently in a situation where the statistical paradigm is in use in the field of artificial intelligence. This means that it operates on the basis of large data sets, which are analysed through the use of statistical tools, and based on these, patterns are created in the form of predictive algorithms.

The use of this statistical method to construct moral patterns, however, only allows machines to reach the level of conventional morality in Kohlberg's model of development. It is not possible to achieve a level of universal morality - one that can be accepted cross-culturally.

Because machines operating in the statistical paradigm are unable to create such supra-conventional values on their own, we will use Elizier Yudkowski's and Nick Bostrom's suggestion that they should be instilled by introducing a bottom up ethics project.

I am going to present some consequences of such an approach in the field of ethics, functionality and architecture of autonomous machines and confidence in their decisions.

References

Print

1. Nick Bostrom, Superinteligence, Oxford 2015
2. Philippa Foot, The Problem of Abortion and the Doctrine of the Double Effect, https://www2.econ.iastate.edu/classes/econ362/hallam/Readings/FootDoubleEffect.pdf, access 17.06.2021
3. Jarek Gryz, Sztuczna Inteligencja: powstanie, rozwój, rokowania, Copernicus Center,
https://www.youtube.com/watch?v=3ZDfVgC897k, access 17.06.2021
4. Piotr Makowski, Gilotyna Hume'a, w. Przegląd Filozoficzny – Nowa Seria Nr 4 (76) 2011, s. 1-15
5. Carl Schmitt, Teoria partyzanta, Warszawa 2016
6. Mirosław Szreder (2021), Różne oblicza istotności statystycznej, Copernicus Center https://www.youtube.com/watch?v=QsJLCxHY794&t=2436s access 15.06.2021
7. Judith Jarvis Thomson, The trolley problem, HeinOnline (http://heinonline.org), http://www.psy.vanderbilt.edu/courses/hon182/thomsontrolley.pdf, access 17.06.2021
8. Elizier Yudkowsky (2004), Coherent extrapolated volition. The Singularity Institute, San Francisco

Digital

1. Sztuczna inteligencja zachęca do samobójstwa, https://wgospodarce.pl/informacje/73035-sztuczna-inteligencja-zachecala-do-samobojstwa - access 17.06.2021
2. https://pl.wikipedia.org/wiki/Lawrence_Kohlberg, access 15.06.2021
3. https://www.britannica.com/biography/Francisco-de-Vitoria

Kristina Šekrst and Sandro Skansi,

Machine learning and essentialism

(University of Zagreb)

The goal of this paper is to establish the metaphysical essentialist position both for machine learning and deep learning, a stance not often researched in philosophy and computing (cf. rare examples like Pelillo and Scantamburlo 2013, Scantamburlo 2014, Duin 2015). In *machine learning*, properties used for supervised learning and dataset tagging will be compared to *minimal* and *maximal essentialism* in metaphysics, but also with componential and semantic analysis in linguistics and psychological and linguistic theory of prototypes. To better understand different kinds of essentialism in computing, we will differentiate between *supervised learning* (using labeled data with pre-defined features of interest) and *unsupervised learning* or finding relevant groups or clusters without predefined properties (cf. Henning 2015).

In various applications of machine learning, different *features* can be seen as essential properties. For example, in handwriting recognition, we can talk about various pixels and recorded movements, while in computer vision and pattern recognition we can talk about higher-level objects such as blobs, where regions in a digital image that differ in properties are compared to the surrounding regions. *Feature engineering* in computer science refers to the process of extracting features (i.e. essential properties) from raw data. In this case, we can see that either a computer can serve as a certain oracle regarding what properties are considered essential for machine learning to take place, or a human being can choose a predefined set of such features or Humean bundles.

We will connect different feature engineering strategies to different metaphysical essentialist theories, in order to see how metaphysics can be reflected in machine-learning applications. Such strategies will serve the purpose of establishing that a philosophical approach in machine learning is not a matter of rejecting or supporting essentialism but choosing a specific essentialist stance for a specific application. Namely, in the background, we are dealing with ontologically different phenomena that require different metaphysical and computing analyses.

Selected bibliography

Duin, Robert P. W. (2015). The dissimilarity representation for finding universals from particulars by an anti-essentialist approach. *Pattern Recognit. Lett*. 64: 37-43.

Henning, Christian (2015). What are the true clusters? *Pattern Recognit. Lett*. 64: 53-62.

Pelillo, Marcello and Teresa Scantamburlo (2013): How Mature is the Field of Machine Learning. *AI*IA*: 121-132.

Scantamburlo, Teresa (2014). *Philosophical Aspects in Pattern Recognition Research*. PhD thesis. Ca' Foscari University, Venice.

About the Authors

Kristina Šekrst is a late-stage PhD student in Logic at the University of Zagreb. She holds a bachelor's degree in General Linguistics, and in Croatian Language and Literature, and master degrees in Comparative Linguistics, Cognitive Linguistics, Croatian Language and Philosophy. She has been teaching philosophical and linguistics courses since 2013, and her areas of interest include logic, artificial intelligence philosophy of computing, computational complexity and philosophy of language and linguistics. She has published one book (*The Ancient Egyptian Language*, 2014, in Croatian) and more than 30 papers, book chapters and conference presentations in fields of philosophy, computer science and linguistics.

Sandro Skansi is an Assistant Professor in Logic at the University of Zagreb. His research interests include logic (propositional satisfiability, inductive logic programming, complexity) and artificial intelligence (deep learning and propositional satisfiability, applied deep learning, cybernetics). He is a member of the Association for Symbolic Logic, the SAT Association, The Association for the Advancement of Artificial Intelligence, and the Croatian Philosophical Society. He has authored two books: *Introduction to Deep Learning* (Springer, 2018) and *Logic and Proofs* (in Croatian, 2016), and edited and co-authored a volume *Guide to Deep Learning Basics* (Springer, 2020).

Mariusz Szynkiewicz,

Wojna w epoce cyfrowej: Procesy automatyzacji i robotyzacji pola walki a wzorce kultury obronnej Zachodu

(Adam Mickiewicz University in Poznań)

Relacje łączące społeczny fenomen wojny z obszarami nauki i techniki stanowią przykład swoiście rozumianej symbiozy. Z jednej strony to co określić możemy mianem technicznego zaplecza wojny korzysta bezpośrednio z wyników badań podstawowych oraz arsenału techniki cywilnej. Z drugiej zaś osiągniecia techniki wojennej niejednokrotnie przyczyniały się do radykalnego postępu w zakresie badań podstawowych oraz szerokorozumianego postępu technicznego. Wojna – w takim ujęciu - może więc być traktowana, jako jeden z głównych zewnętrznych determinantów rozwoju nauki i techniki.

W swoim wystąpieniu przeanalizuję wybrane aspekty relacji występującej pomiędzy wymienionymi uprzednio aspektami kultury ludzkiej, przyjmując za punkt odniesienia nowoczesne technologie informatyczne. W szczególności interesować mnie będą zastosowania narzędzi informatycznych w obszarach automatyki i robotyki. W referacie skupię się więc na dwóch zasadniczych kwestiach. Po pierwsze: analizie najnowszych tendencji obecnych w rozwoju techniki wojskowej, które jak będę argumentował, odzwierciedlają system wartości typowych dla tzw. kultury obronnej zachodu.

Po drugie, postaram się udzielić odpowiedzi na pytanie, na ile program robotyzacji i automatyzacji (autonomizacji) współczesnego pola walki uznać należy za zjawisko wyjątkowe w całych dziejach rozwoju techniki wojskowej.

Prezentowane rozważania prowadzone będą w perspektywie współczesnej filozofii techniki, filozofii informatyki z uwzględnieniem perspektywy polemologicznej.

Timothy Tambassi,

Between Open and Closed World Assumptions. The Epistemology of Information System Ontologies

University of Salerno

According to Goy and Magro (2015), one of the main roles of information system ontologies [ISOs] for the (Semantic) Web is

> [1] supporting communication and mutual understanding between human beings, human beings and software systems, and software systems.

[1] does not, however, imply that human beings and software systems understand ISOs' entities in the same way. But if so, what would the difference involve?

This talk aims to account for such a difference. Firstly, we maintain that while human being can have access to entities represented in ISOs, software applications cannot. Secondly, we argue that the difference also involves the (Semantic web) languages by which ISOs are developed. More precisely, some of those languages are based on the open-world assumption, according to which everything that cannot be inferred as false from an ISO must be considered unknown. Conversely, in many programming languages and formal systems of logic used for knowledge representation, some ISOs' approaches adopt the closed-world assumption, according to which, everything that is not known to be true, in the system, must be considered as false.

Main References

Baader F, Horrocks I, Lutz C, Sattler U (2017) An Introduction to Description Logic. Cambridge University Press, Cambridge

Breitman KK, Casanova MA, Truszkowski W (2007) Semantic Web: Concepts, Technologies and Applications. Springer, London

Goy A, Magro D (2015) What Are Ontologies Useful for? In: Khosrow-Pour M (ed) Encyclopedia of information science and technology. IGI Global, pp. 7456–7464

Gruber TR (2009) Ontology. In Liu L, Özsu MT (eds) Encyclopedia of Database Systems. Springer, Boston (MA). https://doi.org/10.1007/978-0-387-39940-9_1318

Guarino N, Musen M (2015) Applied Ontology: The Next Decade Begins. Applied ontology 10:1–4

Lutz C, Seylan I, Wolter F (2012) Mixing open and closed world assumption in ontology-based data access: Non-uniform data complexity. In: Kazarov Y, Lembo D, Wolter (eds) Proceedings of the International Workshop on Description Logics, DL 2021, CEUR Workshop Proceedings Volume 846

Porello D, Endriss U (2014) Ontology Merging as Social Choice: Judgment Aggregation Under the Open World Assumption. *Journal of Logic and Computation 24(6):11229–1249*

Razniewski S, Savkovic O, Nutt W (2016). Turning the Partial-Closed World Assumption Upside Down. In: Pichler R, Soares da Silva A (eds) Proceedings of the 10th Alberto Mendelzon International Workshop on Foundations of Data Management, Panama City, Panama, May 8-10, 2016. CEUR Workshop Proceedings 1644, CEUR-WS.org 2016

Turner R (2019) Computational Artifacts: Towards a Philosophy of Computer Science. Springer, Berlin, Heidelberg

Javier Toscano,

Deconstructing an algorithm. Intentionalities and socio-computing infrastructures

(APRA Foundation, Berlin, Germany)

Computing does not only imply a logical interaction with and through machines, but also –maybe more poignantly– a way of thinking. As historians of technology acknowledge, computing meant in the past so much as counting, or even reasoning (e.g. Leibniz 1890). But in this sense, the history of computing has a much earlier beginning than what has been popularly thought. The first machines that we can recognize as abstract computers were imagined by Charles Babbage in the 19th Century, but the first algorithms were assembled centuries before, to be performed by social machineries. Of course, in order for this account to unfold, we need to precise what an algorithm and programming are (or can be), and how they articulate together a social structure to produce cultural meaning, a certain dynamic and a given output. Drawing on early definitions of programming by mathematicians von Neumann and Goldstine (1947), as well as pioneer logicians Newell, Simon and Shaw (1958), but especially on a recent human-evolutionary hypothesis by cognitive scientist Michael Tomasello (2014), which maintains that humankind relies on a cooperative mechanism that ultimately resulted in a particular modality of social thinking, this article explores the deep historical foundations of computing and coding from a reinterpretation of specific cultural practices.

Articulating a notion of code, Newell et. al. wrote that "the appropriate way to describe a piece of problem solving behavior is in terms of a program [...]. Computers come into the picture only because they can, by appropriate programming, be induced to execute the same sequences of information processes that humans execute when they are solving problems." (1958). In this sense, the machine is contingent, and only the processes of defining, framing and solving a problem become relevant. This allows to shift our attention from the object to the interrelated and normative cultural practices that sustain a socio-computing infrastructure. On that track, this research implies a reassessment of historico-

cultural materials following Schutz's (1979) and Garfunkel's (2008) empirical phenomenology, in order to analyze underlying social structures along what Tomasello terms collective intentionalities. On that path, this project inspects ancient coding technologies in religious, magical, legal and proto-scientific domains, realigning them through shifting categories in order to explore socio-historical functions of command and control, calculation and ordering, knowledge and transfer, protection and safeguarding, contemplation, synthesis and regeneration.

About the Author

Javier Toscano (PhD in Philosophy) is currently a research fellow at the APRA Foundation in Berlin, Germany. He is an interdisciplinary researcher in the fields of philosophy of technology, visual sociology and political philosophy. His work has involved a continuous search to understand and collaborate with minorities, communities and groups with disabilities towards the production of practice-based knowledge, while considering narratives of self-affirmation and vital exploration. He has published a number of peer-reviewed articles and was awarded the *Mouton d'Or* Prize in 2019 from the *Semiotica* Journal, for the best article on the field that year.

Kazimierz Trzęsicki, Turing Paradigm

(University of Bialystok, Institute of Informatics)

Scientific knowledge is acquired according to some paradigm. Galileo said that the Book of nature is written in mathematical language and it cannot be understood unless one first understands the language and recognizes the characters with which it is written. We argue that the seeds of algorithmic paradigm was planted by Turing. Turing's paradigm says that the Book of nature is written in algorithmics language and the aim of science is to acquire the knowledge how the algorithms change the physical, social and human universe. Some sources of Turing paradigm are pointed. A few examples of application of Turing paradigm are shortly discussed.