Developing an Ontology-oriented Framework to Enable Peace-building Solutions Based on Fairplay Algorithms

Tania C. D. Bueno, Hugo C. Hoeschl
Institute of E-government, Intelligences and Systems -i3G
Florianópolis, Brazil
{tania.bueno, hugo.hoeschl}@i3g.org.br

Milena B. Hoeschl
Federal University of Santa Catarina -UFSC
Florianópolis, Brazil
milena_bueno@hotmail.com

***Abstract*—The objective of this work is to develop an ontology-framework that will support intelligent systems in decision-making in small and large-scale projects with the application of Peace Engineering. For this, it is necessary to define and delineate moral values, so they can be interpreted by A.I. However, that is not an easy task, and turning them into machine-readable rules is even more challenging. For this reason, we have organized an ontology structure based on the underlying principles of Fairplay algorithms to offer a plausible output for engineering areas. Four main domains are designed to represent the context of Peace Engineering: Artificial Intelligence Safety research, Ethical Engineering, Diversity and Sustainability.**

Keywords—Peace Engineering; Ontology- oriented Framework; Artificial Intelligence; Diversity; Sustainability

I. INTRODUCTION

Peace engineering has become a topic of great relevance today due to the need for a more effective social approach to engineers' education combined with the rapid development of automation and artificial intelligence and its social consequences. Yampolskiy [1] emphasizes that the last decade has seen a boom of new subfields in computer science concerned with the development of ethics in machines (Machine ethics, computer ethics, robot ethics, machine morals, computational ethics, roboethics, robot rights, and artificial morals, etc.).

Thus, defining algorithms capable of promoting moral values like social equality becomes a core issue to our current society. The technology exists, but sadly there are no relevant researches in Peace Engineering and the perceived abundance of research in intelligent machine safety is misleading [1][2]. One major issue in the area is that the great majority of published papers concern themselves with discussing and defining why machine ethics are necessary and especially - what set of ethical beliefs should be used to model them.

It is important to note that ethical norms are not universal. Yampolskiy [1] points out that a single “correct” deontological code based on any predefined abstract principles could never be selected over others to the satisfaction of humanity as a whole; nor could the moral values of a single person or culture be elected to represent all of humanity.

Thus, the deontological ethics do not exist for the machine, but the ontological ethics does. Based on this premise, this study focused on the development of an ontological framework to support the decision-making processes of algorithms in intelligent systems involving peace engineering. To define the ontological content of ethical algorithms, it is necessary to determine what kind of complete actions can simulate a moral effect. The moral algorithm will work towards a lessening of the impacts over the environment, the protection of fragilized groups and always prioritizing human life.

Besides, other methodologies and norms were used as guide for the development of this work. Many academics and companies are committed to improving the current engineering teaching model hoping to make the professional resulting from this process more open to innovation, with political and social skills included. The following recommendations “Preparing for the future of Artificial Intelligence”, "Engineering Criteria 2000" and “Ethically Aligned Design (EAD)” [3][4][5] provide ethical criteria to guide the Ontology-oriented Framework.

In previous works [6][7] an ontology-oriented framework for electrical sector projects was created. The method for analyzing the content of a package of research and development projects identifies the context and the most important indicative expression of the domain to identify key consequences of a proposed R&D Projects and provides quantitative information about them. The method for analyzing the content of a package of R&D projects, and for comparing it to different criteria to extract the information is called Mind Engineering [7]. This method identifies the context and the most important indicative expression of the domain in order to identify key consequences of a proposed R&D Project and provides quantitative information about them. Therefore, all submitted projects can be measured in a semi-automatic way in the Knowledge Based System allows arbitrary notions, such as innovation, to be more easily identified. Case-based Reasoning algorithms were also used to aid in the evaluation and selection of subjective criteria. For this reason, in this present work we have chosen to use the same methodology to define an ontology framework that allows the inclusion of language related to Peace Engineering. We also opted for the introduction of Fairplay algorithms.

Here, a new approach will be presented on how ethical ontological framework can influence algorithms in knowledge-based systems. The state-of-the-art of the Peace Engineering, Artificial Intelligence and Fairplay algorithms are presented in the Chapter II, in Chapter III, methodologies are presented and in Chapter IV the results are developed. Future and Conclusions will be presented in chapter V.

II. BACKGROUND

A. Peace Engineering

Defining Peace Engineering is a great challenge as very few published works are dedicated towards this purpose. For this reason, and due to Peace Engineering encompassing differentiated and rapidly changing contexts, an ontological framework was developed to support decision-making processes in engineering projects, where the need for ethics is ever-growing. Peace Engineering is the first program dedicated to preventing and reducing violent conflict through education and research that integrates innovative technologies, approaches and policies into the studies and practices of peacebuilders. Peace Engineering also examines ethical engineering concepts, including the engineer's duty to society; engineering relationship with war and peace; and the engineer's commitment to the environment. The Peace Engineering program also aims to develop engineers to use their technical skills to promote social justice [8].

B. Artificial Intelligence Safety Research

Even if we are successful at designing machines capable of passing a Moral Turing Test [9], human-like performance means immoral actions might be undertaken, which should not be acceptable from the machines we design. In other words, we don't need machines debating about what is right and wrong, we need our machines to be safe and law abiding [1].

In each domain, even as AI continues to deliver important benefits, it also raises important ethical and social issues, including privacy concerns. Robots and other AI technologies have already begun to displace jobs in some sectors. As a society, we are now at a crucial juncture in determining how to deploy AI-based technologies in ways that promote, not hinder, democratic values such as freedom, equality, and transparency [10].

The complex relationships between Artificial Intelligence and socioeconomic performance, just as it happens with information technologies - ICTs [11] are not fully understood and their causality not fully established. However, our hope in this work is to highlight the opportunities offered by Artificial Intelligence and Autonomous Systems and provide an indication of how they are transforming economies and societies around the world. One issue that raises controversy is related to the high autonomy of robots, and their inclusion in one of the existing legal categories, with its own entity and consequences in terms of the attribution of rights and obligations, including liability for damages that may be inflicted.

We will soon reach a time when computers can solve their own problems—what scientists call the "technological singularity [12]." At that point, computers would become smarter than humans and could design new computers that are even smarter, which would then design computers that are smarter still.

C. Ontology-based Knowledge Management Systems

The Knowledge Management is widely acknowledged as one of the most prominent applications of ontologies in the ontology community. Nevertheless, ontologies still must find their way into the practice of Knowledge Management (KM) in the social justice field. Both these research areas seem to evolve in parallel without exploiting the synergy they could have in their application areas and supporting technology. The natural synergy between the KM and ontology communities could be better exploited in order to benefit both fields. The perspective for the study of KM could achieve result from combining the perspectives on knowledge (epistemology) and social reality (ontology) [7].

Establishing a common vocabulary plays a key role in enabling organization's members to communicate and collaborate with each other. Finding the right balance and supporting a common understanding is critical for the success of Ontology-based projects.

There is a great potential for ontology-based methods supported by KM technology in the context of Peace Engineering.

D. Fairplay Algorithms

We live in the age of the algorithm. Increasingly, the decisions that affect our lives—where we go to school, whether we get a car loan, how much we pay for health insurance—are being made not by humans, but by mathematical models. In theory, this should lead to greater fairness: everyone is judged according to the same rules, and bias is eliminated. However, for O'Neil [2] an algorithm is a model coded with math, and models are built using data with outcomes that are defined by conditions for success. She concludes - referring the teacher-ranking algorithms used in firing and tenure decisions in USA:

"…the algorithm itself is just badly designed; it's just not designed well enough to accurately predict anything at all."

Automated systems are making decisions that are too complex for humans to be involved, so we need to ensure they distribute resources fairly. Meanwhile it is important to note that AI technologies will encode the biases, flaws and prejudices of their creators. This is a problem for all computers: Their output is only as good as their input. An AI system that is fed information inflected by race is at risk of putting out racist results.

Several initiatives are already analyzing this issue, including amongst others the IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems, the Partnership on AI, and the AI Now Institute [5]. At the level of national governments, many countries are also looking at means to regulate AI [3] [14].

Ethics by Design concerns the methods, algorithms and tools needed to endow autonomous agents with the capability to reason about the ethical aspects of their decisions, and the methods, tools and formalisms to guarantee that an agent's behavior remains within given moral bounds.

However, not only is it hard to program "Fair Play" into AI, another complication is that ethics, or "moral rules" change with time and from group and individuals.

A few studies about the application of Fairplay algorithms towards defining what is right or wrong, ignore the possibility of building a sense of Fairplay in computation. Those works highlight just how complicated it is to translate these ethical principles from natural language to formal rules. In such cases, applying the construction of ontologies into this translation can improve the comprehension of rules in natural language for text interpretation and decision-making [15].

III. Methodologies

Not enough models exist for bringing engineers and designers in contact with ethicists and social scientists, both in academia and industry, so that meaningful interdisciplinary collaboration can shape the future of technological innovation. The process of utilizing multiple ethical approaches to provably aligned end user values will provide a key competitive differentiator in the algorithmic economy by prioritizing respect for individuals above exponential growth. AI engineers and design teams often fail to discern the ethical decisions that are implicit in technical work and design choices, or alternatively, treat ethical decision-making as just another form of technical problem solving [5]. A methodology for bridging the need of a truly interdisciplinary and intercultural education of the intricacies of technology and its effects on human society for the engineers who develop them is required especially regarding the immediacy of the ethical considerations of Artificial Intelligence and Autonomous Systems.

A. Metaethics

Metaethics [13] tries to determine the ontological status of the basic concepts in ethics, such as right and wrong. For example, are matters of morals and ethics more like matters of fact or of opinion? Who determines whether something is good or bad? Is there a divine being who stipulates what is right or wrong, or a Platonic realm that provides truth-values to ethical claims, independently of what anyone thinks? Is ethics merely in our heads, and if so, how can anyone moral outlook be better than any other? As engineers bestowing ethical qualities to robots (in a manner soon to be explained), we are automatically confronted with these metaethical issues, especially since we're given the power to determine a robot's sense of right and wrong.

B. Ethically Aligned Design (EAD)

In order to develop successful Autonomous Intelligent Systems (AIS) that will benefit our society, it is crucial for the technical community to understand and be able to embed relevant human norms or values into their systems. The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems [5] has taken on the broader objective of embedding values into AIS as a three-pronged approach, that is to help designers: 1) Identify the norms and values of a specific community affected by an AIS; 2) Implement the norms and values of that community within the AIS; and, 3) Evaluate the alignment and compatibility of those norms and values between the humans and the AIS within that community.

The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems [5] is an incubation space for new standards and solutions, certifications and codes of conduct, and consensus building for ethical implementation of intelligent technologies. Ethically Aligned Design includes eight sections, each tackling a specific topic related to AI/AS that has been discussed at length by a specific committee of The IEEE Global Initiative. The initiative provides the opportunity to bring together multiple voices in the Artificial Intelligence and Autonomous Systems communities to identify and find consensus on timely issues. The result was the construction of the following document: Ethically Aligned Design: A Vision for Prioritizing Human Well Being with Artificial Intelligence and Autonomous Systems (EAD, Version 1) represents the collective input of over one hundred global thought leaders from academia, science, government and corporate sectors in the fields of Artificial Intelligence, Ethics, Philosophy, and Policy.

C. Mind Engineering

In the era of digitized information, researchers of different areas of the knowledge face a new subject: the semantic organization of the data. This has become a major issue as the majority of existing digital information is made available in the form of non-structured data. It is known that the structuring of data is a complex problem that can be solved through the construction of formal models and of languages of the Computer Science. It is necessary, however, to observe that those specific areas possess their own culture and a singular way of communicating. Observing the involved factors in this research, a methodology was developed for the understanding and the computational representation of such communication barriers. This allowed each community or work atmosphere to viably express their knowledge domain. The methodology was called Mind Engineering, and its concept of 'Ontological' became an important 'key' for the structuring of data and for the construction of Knowledge Based Systems.

The Mind Engineering [7] is defined by the Knowledge Engineering and Ontological Engineering as a process of synchronization of the knowledge developed with the purpose of fastening conceptual models in relation to the objective and to the application of the administration model based on the Artificial Intelligence (AI), and to identify and systematize the intellectual abilities of the system's development team. Aiding them in the perception of the problem with quality and creativity.

Mind Engineering identifies emotion as an aspect of the indispensable human intelligence in the stage of the representation of the knowledge in the Artificial Intelligence, more specifically in the development of Knowledge Based Systems (KBS) in cooperative work atmospheres in the Web. The larger the synchronicities between the Knowledge Specialists and the Knowledge Engineers are, the more effective the system will be.

The foundations of this methodology go by the perspective of the neuroscience - one of the most respected and competed field of today's science - understanding the structure of the language and of the human unconscious, together with the mechanisms used by the mind, to define the relevant knowledge. This process includes the people's study, processes and technologies, through the sharing of the knowledge, of the visualization and of the definition of relevance. The synchronization of those factors aims to allow the understanding of the knowledge, or expertise, of a certain domain in the totality, through a computational system - more specifically a tool of Ontological Engineering - that acts as a mechanism of knowledge acquisition.

It is important to put the user of technology at the center of this issue and discuss the ethical questions. Only when we truly understand the individuals using the technology - their behaviors, needs, and problems - can we create better solutions, solve bigger problems, and achieve significant Artificial Intelligence change.

IV. Results

For the construction of an ontology framework on the context of peace engineering, this work used the methodology of Mind Engineering, previously defined in item III.

This study focused on the use of the ontology engineering methodology, developed by three premises: 1. Knowledge Sharing; 2. Visualization; 3. Definition of Relevance.

In general, four types of inventory are made in the Mind Engineering methodology: processes, technology, content and people. Human capital, the mission, methodologies, technologies and knowledge available in the institution are identified seeking motivation, flexibility, versatility, expert adaptabilities and users related to context.

From these inventories comes the concept mapping of processes, technologies, content and people. The content maps are the basis of ontology engineers, but the maps of processes, technologies and people are the ones that define the relevance and coherence of ontological relations. On this phase the controlled vocabulary of the Peace Engineering domain is developed. The controlled vocabulary is separated into domains (themes and subthemes). The ontologies are organized in context belonging or not to a domain or subdomain. The domain definitions are essentially related to the expected results of the system. The domains form a hierarchical tree, grouping the subjects in order to facilitate analysis of information from the knowledge base.

In this case, the goal is to index ontologies through the evaluation criteria defined in the management model as relevant (See table 2 and 3).

The team identified the main concepts and sources used in the domain evaluation and selection for Peace Engineering projects. Then, the team reached an agreement regarding the strategic criteria established by diversity and sustainability questions.

Inventories of content, processes and people were conducted following the premises of knowledge sharing, visualization and definition of relevance, and their developments will be described in the following items.

Peace Engineering can be defined as a set of actions that have as an aim the improvement of the social well-being with permanent and effective actions. Other than that, it must prioritize groups in situations of risk, promote egalitarian actions, and rank anti-egalitarian conduits.

Peace Engineering envelopes very wide contexts in the promotion of social justice, which here served to help identify and organize concepts and themes to build an ontological framework. These are such themes: the advances of AI, ethics in engineering, diversity and sustainability.

Peace Engineering's four pillars were organized and associated to the main theme based on a list of indicative expressions of the domain (see table 1). The associative frequencies of these terms in researches, texts and norms allow you to fit them into one of these four domains. The list of indicative expressions of each domain is called the subdomain, which is shown in table 2. The ontological relations meronymy, holonomy and synonymy were not developed in the present work.

TABLE 1. PRIORITY THEMES FOR PEACE ENGINEERING

Priority Themes
EE - Ethical Engineering
AISR - Artificial Intelligence Safety Research
SU - Sustainability
DI - Diversity

In the methodology, after conducting the domain inventories, the next step and expand the concepts themes, through the creation of a list of expressions indicative of the domain - subtheme (see table 2).

This ontological model will allow the learning algorithms to better interpret the context. In previous works, as the project evaluation of electric sector companies [6], there was an effective improvement in the classification of the projects through the evaluation of the achieved results.

V. CONCLUSIONS

This work is only at its beginning and, for this reason, an evaluation of its effectiveness will not be possible. In a following work, the framework will present a broader representation of the Peace Engineering context and of all the complexity surrounding the theme. Only then can the ethical questions referring Peace Engineering be answered, be it on the area of technology, diversity, or that of sustainability.

- Why is every pillar individually and contextually important?
- What are diversity's keywords?
- What are the biggest challenges for social justice and its promotion?
- What should an ethical algorithm encompass?

TABLE 2. PRIORITY SUB-THEMES FOR PEACE ENGINEERING.

Theme	Subthemes – Controled Vocabulary
EE	Prevent any harm to humankind Social dimensions elective Social justice Conflict Management Peacebuilding skills Toolsfor humanistic thinking
AISR	Technological singularity Roboethics knowledge Engineering with Bayesian Networks Robot Rights Superintelligent agent Virtual World Simulation Machine ethics Fairplay Algorithms Artificial Intelligence Regulations Unbiased developers
DI	Most innovative designs Absent a diverse engineering team More creativity Diverse work force Underrepresented minorities Toolsfor humanistic thinking Barriers determined by socially constructed identity Ideas and potential innovations Lack of diversity
SU	Sustainable Development Goals (SDGs) Social justice in the city planning Community-based design Reshape a City Build more roads to increase capacity; Automatically raise rates as traffic increases Establish HOV (high-occupancy vehicle) Encourage carpooling Infrastructure shapes a city City's sustainable future

Peace Engineering actions promote effective e permanent results observing equality, diversity and sustainability using intelligent ethical systems. Nearest neighbor algorithms can generate better results than learning algorithms, since the latter can easily reproduce biases. The ontology can help the algorithms to must contain procedural behaviors that allow for a ranking and categorization of actions. There is great need for a theoretical base.

If Engineering puts human well-being as a focus through automating Fairplay rules in the design of projects, these projects will be more social and balanced. Maybe the cost will increase, maybe the profits will decrease, but that is a choice that we should make as humanity.

References

[1] Yampolskiy, Roman V., and Joshua Fox. 2012. "Safety Engineering for Artificial General Intelligence." Topoi. doi:10.1007/s11245-012-9128-9 https://intelligence.org/files/SafetyEngineering.pdf

[2] O'Neil, Cathy.Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown. 2016. ISBN 9780553418811.

[3] USA NSTC's Subcommittee on Machine Learning and Artificial Intelligence,. "Preparing for the future of Artificial Intelligence." Washington DC, October 2016.

[4] L. Lattuca, P. Terenzini, J. Fredricks Volkwein. Engineering Change: A Study of the Impact of EC2000. ABET, 2006. http://www.abet.org/wp-content/uploads/2015/04/EngineeringChange-executive-summary.pdf

[5] The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems. Ethically Aligned Design: A Vision For Prioritizing Wellbeing With Artificial Intelligence And Autonomous Systems, Version 1. IEEE, 2016. available at http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html.

[6] Bueno T.C.D., Donatti F.T., Oliveira T.P.S., Zotti A.I., Hoeschl H.C. (2013) Analytic Hierarchy Process to Improve the Evaluation Process in R&D Management System. In: Abramowicz W. (eds) Business Information Systems Workshops. BIS 2013. Lecture Notes in Business Information Processing, vol 160. Springer, Berlin, Heidelberg.

[7] Bueno, Tania C. D.. (2005) "Engineering of Mind: a methodology of knowledge representation for the construction of ontologies in knowledge-based systems," Ph.D. dissertation, Dept. Production Engineer, Federal University of Santa Catarina, Florianópolis, SC. https://repositorio.ufsc.br/handle/123456789/102910?show=full. Acessed: July, 19, 2018.

[8] Vesilind, P.A.. (2010). Engineering peace and justice: The responsibility of engineers to society. 10.1007/978-1-84882-674-8.

[9] Allen, C., Varner, G., Zinser, J.: Prolegomena to any future artificial moral gent. Journal of Experimental and Theoretical Artificial Intelligence 12, 251–261 (2000)

[10] P. Stone, R. Brooks, E. Brynjolfsson, R. Calo, O. Etzioni, G. Hager, J. Hirschberg, S. Kalyanakrishnan, E. Kamar, S. Kraus, K. Leyton-Brown, D.Parkes, W. Press, A. Saxenian, J. Shah, M. Tambe, and A. Teller. "Artificial Intelligence and Life in 2030." One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, Stanford University, Stanford, CA, September 2016. Doc: http://ai100.stanford.edu/2016-report. Accessed: April 8, 2017.

[11] S. Dutta, T. Geiger, B. Lanvin, "The Global Information Technology Report 2015, " World Economic Forum. http://reports.weforum.org/global-information-technology-report-2016/Acessed: April 7, 2017.

[12] R. Kurzweil, The Singularity is Near: When Humans Transcend Biology. 2005. http://www.kurzweilai.net/the-singularity-is-near-selected-chapters.Acessed: April 7, 2017.

[13] P. Lin, K. Abney, and G.Bekey. Robot Ethics The Ethical and Social Implications of Robotics. The MIT Press Cambridge, Massachusetts. http://kryten.mm.rpi.edu/Divine-Command_Roboethics_Bringsjord_Taylor.pdf

[14] General Data Protection Regulation (GDPR). https://gdpr-info.eu/

[15] Allen, C., Varner, G., Zinser, J.: Prolegomena to any future artificial moral gent. Journal of Experimental and Theoretical Artificial Intelligence 12, 251-261 (2000)