

Fundamentals of Human-Centric Artificial Intelligence (A.I.):

Comparative Analysis of Europe and the U.S. Landscape

****DRAFT****

Torré A. Williams, Twill049@odu.edu

Center for Cybersecurity Education and Research
Old Dominion University

Cesar A. Pinto, cpinto@odu.edu

Engineering Management & Systems Engineering
Old Dominion University

Submitted in partial fulfillment of the requirements with the
COVA CCI Undergraduate Research Project

December 2020

I. Abstract

This research is a comparative analysis of human-centric Artificial Intelligence (A.I.) in Europe and the U.S. This research establishes fundamentals that are critical to what makes A.I. human-centric. This research contains eight phases: 1) Lawful A.I.; 2) Robust A.I.; 3) Ethical A.I.; 4) Human-centric A.I.; 5) Current State of A.I.; 6) A.I. in Europe; 7) A.I. in the U.S.; 9) Importance of Human-centric A.I. This research shows that there are still ongoing changes with having a human-centric A.I. and why it is very important to society. This research is the beginning of the making of a successful and reliable human-centric A.I.

II. Introduction

Throughout our history, the concept of what is Artificial Intelligence (A.I.) has been changing since the creation of A.I. However, the core of what A.I. remains the same. The idea of A.I. is thought up as a machine being created that can think and perform like humans (Marr, 2020). The early stages of A.I. started between the 1950s- 1960s. Many researched symbolic and problem-solving methods, as well as beginning to train computers to mimic the basic responsibility of humans and learn human reasonings (Goodnight, 2020). Throughout this research, three important fundamentals are benefits to having a successful human-centric A.I. Those fundamentals are lawful, robust, and ethical, as illustrated in Figure 1. Without these three important factors, there will never be a successful human-centric A.I.

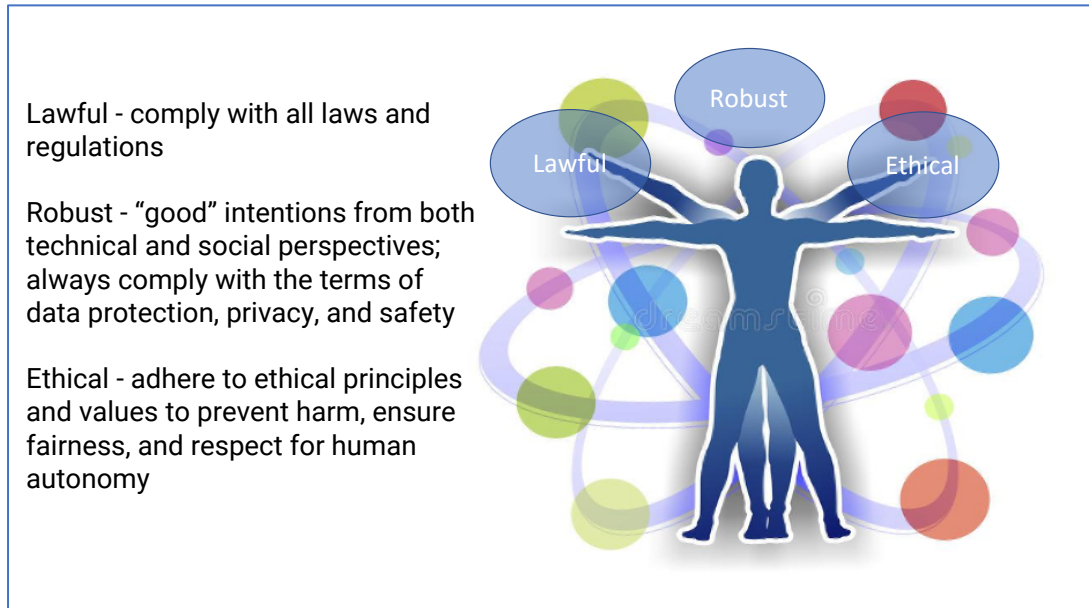


Figure 1. The three fundamental principles of human-centric AI: Lawful, Robust, and Ethical.

a. Lawful A.I.

For A.I. to be legal around the world, some laws must be complied with and regulations that must always be followed without the chance of being illegal (Smuha, 2019). A.I. is often used in physical forms, i.e. robots, smart devices, or machines (Mathys, 2019). Due to the physical form, it creates liability so therefore, it is subject to adhere to all laws (Mathys, 2019). On the other hand, if companies misuse A.I., they will be responsible for any consequences that will bring a negative impact on the company (Imperva, 2019). When this happens, it will then bring out the question of whether the company had the proper training to implement the use of A.I. systems. If lawful consists of making sure everything is legal, then unlawful would be A.I. would be doing something illegally. Things illegally would be consisted of drugs and everything that would subject you to disciplinary actions under the law. As of right now, there are drafts

of regulations/ rules that A.I. would have to follow when it comes to government agencies. Here are some of the following considerations of regulations that are being drafted: public participation, risks assessment and management, and benefits and costs (Nantais, 2020). Public participations allow oversight and participation from the public for their opinions when discussing new or modified rules regarding A.I. (Nantais, 2020). Risk assessment and management means that the approach should be an A.I. risk-based approach (Nantais, 2020). Benefits and costs would mean that the benefit of having certain rules should outweigh the cost of purchasing those rules.

b. Robust A.I.

When A.I. is robust, it must have good intentions from both a technical and social perspective (Smuha, 2019). Regarding A.I., the robustness must be complying in terms of data protection, privacy, and safety (Hamon, 2020). Also, security is very important when we talk about how robust something is. Evaluations of robustness are to show the level of the security of a model to make sure everything complies with all terms and regulations. If the robustness of A.I. is not where it needs to be, some methods that have been proposed to help with making the robustness of A.I. where it needs to be (Chang, 2020). To establish something robust, some rules are designed to coincide with the whole idea of what is considered robust. It is believed and researched that the rules for robust would well perform in economic environments (Taylor, 2010). If a variety of models being evaluated by rules and policies, this can give the advantage to be able to point out robust policy characteristics (Taylor, 2010). Examples of evaluation for robustness is having a good number of researchers to come together to compare policies

effects of rules in different models. Some rules of robustness are determined by the concept of fault tolerance (Taylor, 2010). What does it mean to be fault tolerance? Fault tolerance is given to a system by the ability to which a system can operate without the interruption of one or more failed components (Imperva, 2019). The goal of creating this fault tolerance system is to ensure business continuity and high availability while preventing disruption from a single failed system's component (Imperva, 2019). For the concern of whether a system is robust, society and individuals must have some confidence and trust that these systems will not cause any harm unintentionally (Smuha, 2019). Any environment that A.I. is used must always create security, safety, and make sure that the systems are fully robust to avoid malicious use (Smuha, 2019).

c. Ethical A.I.

The last component of A.I. is being ethical. Being ethical means that it ensures adherence to ethical principles and values (Smuha, 2019). On the other hand, unethical issues may arise from inequality, wrongful security, or even bias (Bossmann, 2016). Ethical A.I. is rooted in the following principles to have prevention of harm, to have fairness and explicability, and to have respect for human autonomy (Smuha, 2019). Another type of ethics called robot ethics or roboethics is a common principle that is used to design systems of A.I. using special codes to ensure that these systems are working and responding ethically (Walker, 2020). Codes are used to ensure that these machines are aligning with the goal of the ethical framework (Walker, 2020). A.I. systems can often show and serve many purposes of doing an "ethical reflection" to ensure that these systems are aligning with the A.I. framework (Smuha, 2019). These

reflections ensure that groups and individuals are protected at any level (Smuha, 2019). Secondly, reflections seek innovations that portray these ethical values (Smuha, 2019). As a result, having these ethical reflections can promote equality within social, economic, and political opportunities (Smuha, 2019). Having an ethical A.I. can sometimes be difficult establishing due to the high amount of bias problems (Walker, 2020). Opacity, privacy, and manipulation are the three other issues that arise when dealing with ethical A.I.(Müller, 2020). Other challenges that A.I. systems often face have to do with how technology, people, and society would react with each and its impact, which can alter their decision-making about safety (Smuha, 2019). Though the goal is to always ensure that A.I. remains ethical and adhere to all rules, there is a side where there are instances where A.I. can be unethical (Macciola, 2019). Unethical A.I. comes in the form of mortgage lending, education, searches, and human resources (Macciola, 2019). Establishing a good foundation for ethical A.I. promotes several benefits for both communities and businesses. Ethical A.I. guarantees that deployment and developments are safe, accountable, and ethical (Fourtané, 2020).

III. Human-centric A.I

Along with the three components of A.I., it needs to be human-centric (Smuha, 2019). What it means to be human-centric is that it needs to be resting on the commitment to their use in the service of humanity and the common good, with goals of improving human welfare and freedom (Shuma, 2019). Also, to mean human-centric is that systems is improving constantly due to human activity while comparing and conducting experiments between robots and humans (Analytics Insight, 2020). From a business

standpoint of having a human-centric A.I., it has solutions that require “thick” data that will help them deeply understand different behaviors in the marketplace (Analytics Insight, 2020). Also, a human-centric A.I. for long-term stability in society is very needed. On the other hand, the business’ standpoint feels there are benefits with having a human-centric A.I. Benefits that include, informed decision making, successful product-building and software, and scalability and reliability (Analytics Insight, 2020). All three benefits come intending to hopefully be able to fully understand the behavior of human nature (Byson, 2019). With all requirements and expectations of A.I., the goal is to maximize the benefits of A.I. systems while preventing and minimizing their risks (Smuha, 2019). Without A.I. systems unwanted, consequences may happen, and their updates might be hindered (Smuha, 2019).

IV. The current state of A.I.

Just like with anything, regulations and rules are meant to assure everything is working as intended, and minimize unintended and negative consequences. A.I. has seven regulations that it must follow to be successful. Here are the following regulations: Human agency and oversight, Technical robustness and safety, Privacy and governance, Transparency, Diversity, non-discrimination, and fairness, Environmental and societal well-being, and Accountability (Smuha, 2019). A.I. systems need to not only comply with all laws and regulations, but they must ensure equality and due process (Smuha, 2019). If there is more public insight about their opinions, it could potentially “reshape” A.I. regulations to gain more trust in A.I. with the public (Nantais, 2020). As regulations and procedures of A.I. are being drafted there are a couple of things that need to be kept

in mind when moving forward to a better A.I. system (Nantais, 2020). This new update what A.I. will be applied to all private or non-private government sectors or agencies, but these new updates will not control them but be a guide to further promote safety and trust (Nantais, 2020). With government organizations and sectors as the forefront of the innovation of A.I., the government will be actively looking for any roadblocks that may not be aligned with the goal of progression of A.I. and will not proceed with the progression unless every roadblock is cleared or it is needed (Nantais, 2020). Along with the current state of A.I., there are at least six current trends of the use of A.I. that should be on the lookout. Predictive analytics, real-time marketing activities, A.I. customer support and assistance, enhanced customization, A.I. powered chatbots, and data access enabling ubiquity- all of these trends are programs that companies seek to use to try and build a strong foundation on loyalty and a good reputation (Analytics Insight, 2020). The following sections highlights the current state of human-centric A.I. in Europe and the U.S., and summarized in Figure 2.

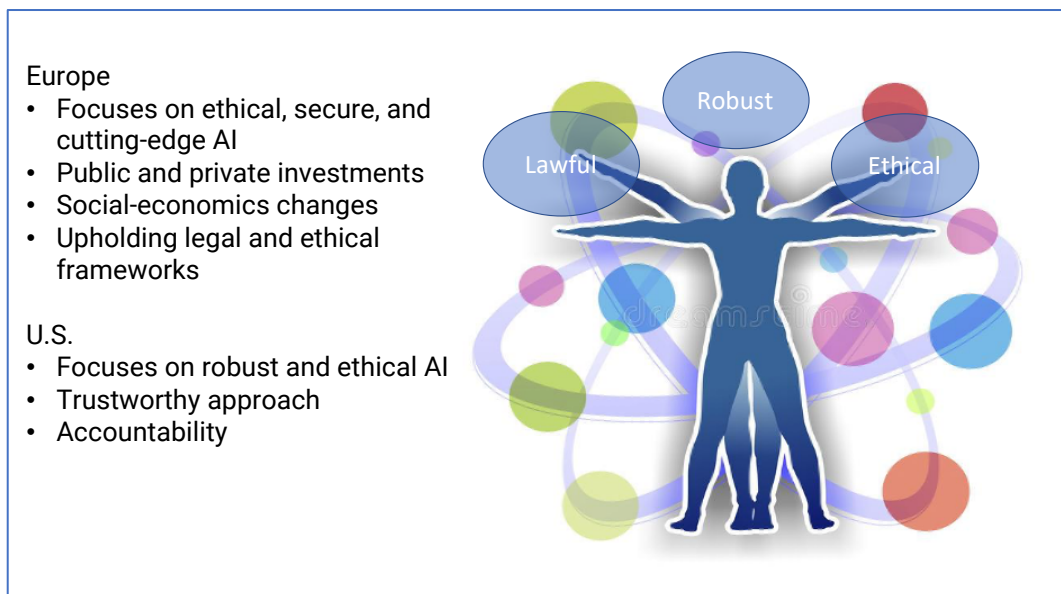


Figure 2. Highlights the current state of human-centric A.I. in Europe and the U.S.

a. A.I. in Europe

In Europe, a vision is shared from the commission of the plans on how A.I. will function and how A.I. will better benefit the community. Although the original support of A.I. will remain its focus on an “ethical, secure, and cutting-edge” platform, there are three more provisions and goals for the use of A.I. in Europe (Smuha, 2019). An increase in investments whether it is private or public in A.I., changes in socio-economics, and adhering to the legal and ethical framework to ensure that the values of Europe are held to high standards- these are the three goals that the commission is working on to ensure the trustworthiness of A.I. in Europe (Smuha, 2019). It is said to believe that the potential use of A.I. is intended to drastically change society to help nurture humans for the common good and bring innovation (Smuha, 2019). Europe has the intention that the use of A.I. should understand the values and have respect for human rights, the rule of law, and democracy (Smuha, 2019). Having a trustworthy A.I. established in Europe is critical to how it can have a positive or negative influence on the public. Being trustworthy means that it relies on principles that are ethically based on human rights (Smuha, 2019).

b. A.I. in the US

The framework of A.I. is used to offer guidance on the second and third components which are robust and ethical (Smuha, 2019). Along with guidance from A.I., there is encouragement for all stakeholders to be able to work towards a framework that is

trustworthy for A.I. (Smuha,2019). With trust, frameworks are to ensure the accountability for all dimensions and development of any use of A.I. systems (Smuha, 2019). There are a few artificial frameworks when A.I. is being used. PyBrain, which stands for Python-Based Reinforcement Learning, Artificial Intelligence, and Neural Network Library, a tool that is used for reinforced learning and learning that is unsupervised (Analytics Insight, 2020). Secondly, the first-ever framework that was created in the 1900s, is Shogun (Analytics Insight, 2020). A popular framework that is known for standard algorithms that are often used by A.I.(Analytics Insight, 2020). Another framework of A.I. is Caffe. Caffe is a learning framework that works with extensible hard coding while fostering the development of different projects (Analytics Insight, 2020). When thinking about these frameworks of A.I., they must be able to comply with all laws and be able to abide by the values and principles that are ethical (Smuha, 2019).

V. Importance of Human-centric A.I.

The use of human-centric Artificial Intelligence being adopted by many small or large corporations is very wanted, however, adopting a human-centric Artificial intelligence brings many concerns (Guszcza, 2018). The need of having a human centric A.I. is vital. A.I. must create human-centric applications that can understand human psychology (Guszcza, 2018). On the other hand, it brings a psychological impact that affects humans' daily activities. The use of A.I. in media and entertainment applications can be detrimental to the well-being of humans (Guszcza, 2018). For example, the compulsive checking on social media can lead to a greater level of unhappiness and feeling as if they

are missing out on certain parts of life (Guszcza, 2018). The addiction to A.I. technologies is quite a concern among scholars because they worry about the minds of people being “hijacked” (Guszcza, 2018). Other reasons why A.I. is important are the following: through neural networks, A.I. achieves great accuracy; A.I. adds more intelligence to products that already exist; and A.I. can adapt to “progressive” algorithms, etc... (Goodnight, 2020).

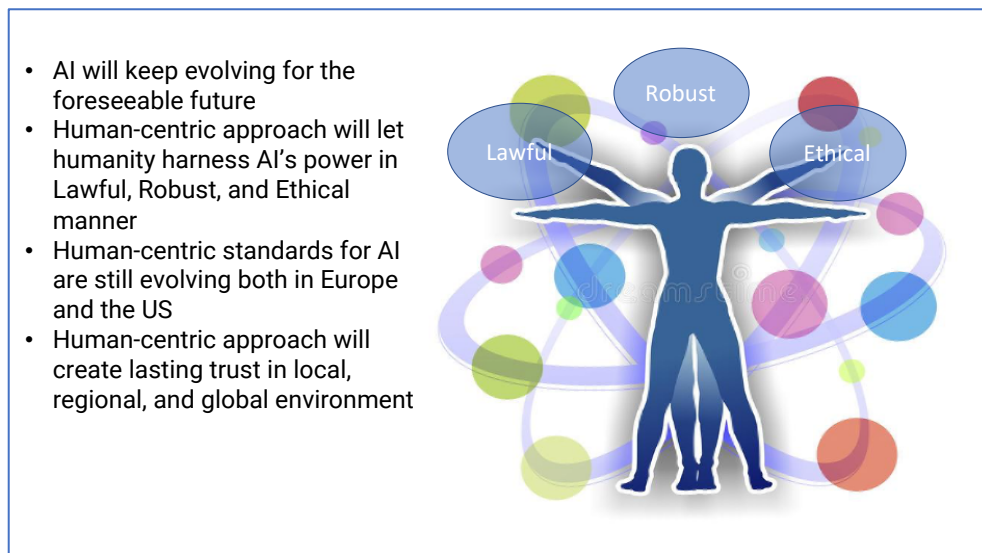


Figure 3. Summary of the future of human-centric AI in the US and Europe.

VI. Conclusion/Summary

As the technology of A.I. continue to develop, new uses of this technology should always be in the interest of the people, and having their safety as the priority. A.I. should always remain legal, robust, and ethical to ensure that the trust is kept. The progression of regulations on the use of A.I. in Europe is currently on a faster pace than in the U.S. The U.S. is still formulating new regulations ensuring the safety of the people around A.I. technologies. As this technology is developing, its uses are also evolving and new

ones evolve – and principles such as robustness, lawfulness, and ethical needs to be upheld to keep the technology trustworthy, as summarized in Figure 3.

VII. References

Analytics Insight. Artificial Intelligence Trends in 2020 (2020, October 03).

Retrieved November 18, 2020, from <https://www.analyticsinsight.net/artificial-intelligence-trends-in-2020/>

Bossmann, Julia A (2016). Top 9 ethical issues in artificial intelligence. Retrieved October 07, 2020, from <https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence>

Bryson, J. J., & Theodorou, A (2019). How society can maintain human-centric artificial intelligence. In *Human-centered digitalization and services* (pp. 305-323). Springer, Singapore. <https://www.cs.bath.ac.uk/~jjb/ftp/BrysonTheodorou19.pdf>

Chang, C. L., Hung, J. L., Tien, C. W., Tien, C. W., & Kuo, S. Y (2020, October). Evaluating Robustness of AI Models against Adversarial Attacks. In *Proceedings of the 1st ACM Workshop on Security and Privacy on Artificial Intelligence* (pp. 47-54). https://dl.acm.org/doi/pdf/10.1145/3385003.3410920?casa_token=Wh0QHJOFRxAAAAA:A:CzteqVR4MyvyO-9ai9G5Tpd8Q0pxdqoia62kINw68FbIODTdq1bSqrDIVSX5eevU1OYvW1t6kgyQgw

Fourtané, S (2020, August 27). Ethics of AI: Benefits and Risks of Artificial Intelligence Systems. Retrieved November 11, 2020, from <https://interestingengineering.com/ethics-of-ai-benefits-and-risks-of-artificial-intelligence-systems>

Goodnight, J (2020). Artificial Intelligence – What it is and why it matters. Retrieved December 03, 2020, from https://www.sas.com/en_us/insights/analytics/what-is-artificial-intelligence.html

Guszcza, J (2018). Smarter together: Why artificial intelligence needs human-centered design. Retrieved November 25, 2020, from <https://www2.deloitte.com/us/en/insights/deloitte-review/issue-22/artificial-intelligence-human-centric-design.html>

Hamon, R., Junklewitz, H., & Sanchez, I (2020). Robustness and Explainability of Artificial Intelligence. Retrieved October 07, 2020, from https://publications.jrc.ec.europa.eu/repository/bitstream/JRC119336/dpad_report.pdf

Imperva (2019, December 30). What is Fault Tolerance?: Creating a Fault Tolerant System: Imperva. Retrieved November 04, 2020, from <https://www.imperva.com/learn/availability/fault-tolerance/>

Macciola, A (2019, August 29). Bad, biased, and unethical uses of AI. Retrieved November 11, 2020, from <https://enterpriseproject.com/article/2019/8/4-unethical-uses-ai>

Marr, B (2020). What is AI? Retrieved December 03, 2020, from <https://bernardmarr.com/default.asp?contentID=963>

Mathys, R., Klaus, S., Burrus, L., & Hari, O (2019, October 15). Legal Aspects of "Artificial Intelligence" (AI). Retrieved November 18, 2020, from <https://www.lexology.com/library/detail.aspx?g=786f8d42-c0dc-4df4-89e0-106bd52bc0c2>

Müller, V (2020, April 30). Ethics of Artificial Intelligence and Robotics. Retrieved November 11, 2020, from <https://plato.stanford.edu/entries/ethics-ai/>

Nantais, J (2020, January 22). Federal Government Regulation of AI. Retrieved October 07, 2020, from <https://towardsdatascience.com/federal-government-regulation-of-ai-4fa08b7bd99a>

Smuha, N (2019). INDEPENDENT HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE. Retrieved October 7, 2020, from https://ai.bsa.org/wp-content/uploads/2019/09/AIHLEG_EthicsGuidelinesforTrustworthyAI-ENpdf.pdf

Smuha, N (2019). Press corner. Retrieved October 07, 2020, from https://ec.europa.eu/commission/presscorner/detail/en/IP_19_1893

Taylor, J. B., & Williams, J. C (2010, April). SIMPLE AND ROBUST RULES FOR MONETARY POLICY. Retrieved November 03, 2020, from https://www.nber.org/system/files/working_papers/w15908/w15908.pdf

Top 10 Artificial Intelligence Frameworks & Tools (n.d.). Retrieved October 14, 2020, from <https://www.cumulations.com/blogs/177/top-10-ai-frameworks-tools>

Walker, D (2020, October 20). What is ethical AI? Retrieved November 11, 2020, from <https://www.itpro.co.uk/technology/30736/what-is-ethical-ai>

What Is Human-Centered Artificial Intelligence?: Cognizant (n.d.). Retrieved October 14, 2020, from <https://www.cognizant.com/glossary/human-centered-ai>