Garbage In, Cyber Truck Out:

A proposal for Ethical Artificial Intelligence Sustainability Impact Statements (E.A.I.S.I.S.) in Autonomous Vehicle Manufacturing

By Ronny Bogani

I. Introduction

Aiming to address the increasing demand for data presented by Artificial Intelligence (AI). This essay proposes adopting an Ethical Artificial Intelligence Sustainability Impact Statement (E.A.I.S.I.S) through the various stages of the Artificial Intelligence Development Cycle (AIDC). The following analysis treats data as a limited resource requiring sustainable exploitation and management. Following a brief discussion about sustainability and AI, the different stages of the AIDC and accompanying factors for an ethical impact evaluation are addressed through proposed questions. These inquiries are intended to identify, disclose and address the data's toxicity level, or potential ethical harm for use.

The proposed approach to data as a public good benefits both the digital environment (DE) and conserves natural resource demand in the real world (IRL). Data is currently a commodity commanding prices higher than oil within a booming data brokerage industry.[1] As with other commodities, derivative markets will form with the price for data lowering proportionally to its level of impurity, or debasement. Through using an E.A.I.S.I.S. approach, certain ethical issues attached to data can be flagged, evaluated, and provided to all users and redeployed in a transparent manner. The following discussion addresses factors to be considered in creating an E.A.I.S.I.S. in the development of an autonomous vehicle (AV). As a novel approach to AIDC, the issues raised are intended to serve as a framework for later discussion, research, development, and application.

[1] Devin Coldeway, TechCrunch, 'AI is more data-hungry than ever, and DefinedCrowd raises $50M B round to fund it.' (26/5/2020) < https://techcrunch.com/2020/05/26/ai-is-more-data-hungry-than-ever-and-definedcrowd-raises-50m-b-round-to-feed-it/?guccounter=1> [accessed 27/5/2020]

II. Sustainable Artificial Intelligence

Sustainability

Sustainability is grounded in respecting the environment, properly balancing between performance and cost, while maintaining social attention and care.[2] This model allowing economic growth, innovation and conservation of resources is applicable to data and AIDC. Prior to engaging AIDC as the digital alternative to traditional AV life cycle of raw materials, processing, pre-assembly, assembly, distribution and end of life disposal and introduction to AI is helpful.

Artificial Intelligence

To begin, an avenue for understanding and demystifying AI is helped by breaking down the term 'Artificial Intelligence'. An artefact is something non-naturally occurring made by humans and intelligence denotes the ability to make the right decision at the right time within a certain context.[3] Artificial versus natural intelligence is simply a difference of origin. Like its organic counterpart, it makes decisions based on observable information or data. Through data and machine learning, AI properly evaluates the given information within that context and executes the appropriate function, or 'right' decision. For example, assume an intelligent agent was instructed to feed a stranger's pet. Although seemingly simple, neither human or machine can perform this task without additional heuristics or data. Suppose the agent is informed the pet is domesticated, has four legs, requires regular feeding and is located in North America. Given this data, certain assumptions can be made and certain possible animals can be eliminated like a blue whale. While a turtle and a dog both fit the description, confusing their care instructions would be disastrous. These are the types of issues that arise when relying on bad or incomplete data.

Human's mistakes and erroneous data impact AI functionality. Even when functioning completely autonomously, human input and data are necessary in the training process. The ability for a machine, either carbon or silicone to make a good

[2] Setchi, Rossi., Robert J. Howlett, Ying. Liu, and Peter. Theobald. Sustainable Design and Manufacturing 2016. Smart Innovation, Systems, and Technologies 52. Cham: Springer International Publishing, 2016

[3] Joanna J Bryson, 'The Artificial Intelligence of the Ethics of Artificial Intelligence: An Introductory Overview for Law'. (2020) Oxford Handbook for the Ethics of Artificial Intelligence

decision is directly proportional to the amount and quality of this data provided within the given context. However, with impugnable data and a sufficiently narrow context, AI can outperform even the most exceptional humans at certain tasks.[4]

This ability is astounding, but AI is still just an algorithm. Demystification requires recognizing 'algorithms' date back to 2500 B.C. and something used by people every day with each other.[5] An algorithm is simply a set of instructions for a computer follow in a certain order. For example, cooks regularly create and execute algorithms when passing or making family recipes. Given sufficient data, AI at its core is an instructional guide for the program's actions within a certain context. However, like the recipe, inserting the wrong ingredients (data) or processing it incorrectly can result in a cake ranging from inedible to toxic for human consumption.

The Digital Environment

It's time to recognize the divide between the digital environment and IRL is largely illusory. Globally, the internet penetrates 58% of the population with a growth rate of over 1000% and accessibility in 100% of countries.[6] Currently, with the majority of countries employing shelter in place orders, reliance on the digital environment has shifted from a convenience to a survival necessity. When a global threat arose IRL, people retreated to the safety of their ever-expanding colonies in the digital new world. Even prior to this, the digital environment supplied educational resources, coordinated travel and assisted within finding take-out options, friends and spouses.[7]

AI and digital growth depends on IRL resources to meet its voracious appetite for data and power consumption.[8] As pioneers in the new digital frontier, the trail blazers owe a duty of care to the next generation. Humans have spoiled every frontier they have encountered, littering or contaminating land, sea, air and space.[9]

---

[4] Stuart Russell and Peter Norvig, *Artificial Intelligence A Modern Approach Third Edition* (2010).
[5] Fauvel, John. "A History of Algorithms: From the Pebble to the Microchip (Book Review)." The Mathematical Gazette 84, no. 499 (2000): 168-69.
[6] United Nations OHCHR (2020) https://www.un.org/en/development/desa/population/index.asp [accessed 22/5/2020]
[7] Turkle, Sherry. Alone Together : Why We Expect More from Technology and Less from Each Other. New York: Basic Books, 2011
[8] Power drain articles
[9] Lewis, Alan., Polly. Turton, Thomas. Sweetman, and Ben. Caldecot. Litterbugs : How to Deal with the Problem of Littering. London: Policy Exchange, 2009.

Through ignorance or avarice this cycle continually repeats. Although a plan for ethically sustainable AI may not solve all problems, ignorance should not be a valid excuse to repeat the past.

Prior to discussing digital resources preservation for sustainable AI, an important distinction between AI and its physical embodiment should be made. The AI is a program, while the car is a conduit to the physical world for the program. Simply put, a car without AI is a fancy toaster and AI without access to the physical world is a great sudoku solver. Without a physical embodiment, AI's physical influence in IRL is limited. The remainder of this paper focuses on AI as located in the digital environment.

III. **Ethical Artificial Intelligence Sustainability Impact Statement (E.A.I.S.I.S.)**

This section proposes factors and stages for developing a sustainable digital environment for AIDC. While focusing entirely on the data needs of AI, at all times humans are the intended beneficiary of this analysis. The aim of this section is to identify data toxicity its potentially negative impact on human end users and IRL resources.

A. Artificial Intelligence Developmental[10] Cycle (AIDC)

The Artificial Intelligence Developmental Cycle (AIDC) model approaches data as a limited resource that is critical for AI development. As AI is intended to help all of humanity, a sustainable data approach should be considered a public good. Toward this, the following AIDC requires accepting the following principles:

- Humans are the moral patient. AI is a tool meant to serve humans. Artificial Intelligence functioning should only occur in the absence of or to protect against threats to the safety and welfare of humans[11]

- Artificial Intelligence is a product.[12] Although an intelligent tool, AI is still only a tool. AI bears no more responsibility or moral accountability for negative

[10] Recognizing the ever-present threat of anthropomorphism, the term 'life cycle' should be prohibited with AI
[11] UKRI Principles of Robotics, < https://epsrc.ukri.org/research/ourportfolio/themes/engineering/activities/principlesofrobotics/> [accessed 28/5/2020]
[12] Joanna J Bryson, Mihailis E Diamantis and Thomas D Grant, 'Of, for, and by the People: The Legal Lacuna of Synthetic Persons' [2017] Artificial Intelligence and Law.

consequences of fulfilling its programmed functions than a toaster can be found guilty for burning bread.

- AI should never be trusted.[13] As an artifact resulting from humans, due diligence should include considerations of human malfeasance, misfeasance, motivations and moral practices. This includes questioning the materials employed and processing employed during the AICD.

- Artificial Intelligence exists in the digital world but impacts IRL.

- The digital world is a shared resource requiring certain protections for the common good.

- Sustainable Artificial Intelligence is threatened by scarcity of resources (data) in the digital environment

Grounded in these assumptions, the following stages are proposed for sustainable AIDC discussion:

B. AIDC Stages

Data is a foundational and critical component for producing AI. The creation, functioning and future adoption of AI depends on quality data. Defined as 'factual information (such as measurements or statistics) used as a basis for reasoning, discussion, or calculation'[14], it informs AI and people alike. However, invoking the first commandment of programming: 'garbage in, garbage out',[15] it is apparent the value of data needs to be accurate, consistent and contextually appropriate for proper machine functioning.

Stage 1: Dumb Data Acquisition (Raw Material and Sourcing)

[13] Joanna Bryson, United Nations University, Science, Technology, Innovation and Governance 'No One Should Trust Artificial Intelligence' (2018) < https://ourworld.unu.edu/en/no-one-should-trust-artificial-intelligence> [accessed 29/5/2020]
[14]Merriam Webster (2020) < https://www.merriam-webster.com/dictionary/data> [accessed 28/5/2020]
[15] Seland, Darryl. "GARBAGE IN GARBAGE OUT." Quality 57, no. 5 (2018): 6.

Although a relatively new field, AI data acquisition shares many characteristics in with the more traditional product life cycle. However, the unique nature and specifically tailored needs of data for machine learning require a novel and hybrid approach for this new commodity. Currently the data brokerage industry is fast growing with data brokers desperately attempting supply the insatiable appetite for machine learning.[16] As the industry exponentially demand on IRL for resources increases. The cost in human labor exploitation and natural resource consumption to support the head-spinning pace of this disruptive technology is sobering.[17] While recognizing the potential harmful impacts and IRL costs to the environment and disadvantaged populations, this sustainability analysis focuses on the data itself as a sustainable product.

As the demand grows so will the commodity price of data and likelihood of inadvertently acquiring toxic data. This will likely follow the same pattern as other industries. Secondary derivative markets will form to trade a debased and inferior version of the original commodity at lower prices, in this case contaminated data. Whether processing new data or reusing established data the following ethical concerns the following ethical concerns should be considered during Stage 1 data capture:

Ethical Identification:

- Is identification and labeling process conducted in a manner free from undue influence or manipulation?
- Has identification process considered applicability to the digitally vulnerable?[18]
- To what degree is identification robust and universally applicable?[19]
- Is identification translatable/interpretable cross-culturally?

Ethical Capture:

[16] ibid (no. 1)
[17] Andrew Ross (2018) Nexidia *Information Age* 'A Perfect Storm: The environmental impact of data centers'< https://www.information-age.com/a-perfect-storm-the-environmental-impact-of-data-centres-123474834/> [accessed 30/5/2020]
[18] Those individuals or groups of people lacking sufficient or equal access to the internet regardless of cause, at times called the 'digitally invisible'.
[19] Universality allows easier transfer of resources and ability to advance the digitally vulnerable and narrow the digital divide

- Was data freely observable 'in the wild', requiring no personal information, consent or individual interaction for capture?
- If consent was required, was it obtained in a manner consistent with local, national and international laws and regulation?
- Was consent given through digital duress under a 'take it or leave it' adhesion contract possibly subject to later challenge?
- Were the intellectual property rights of the original data creator/compiler/subject respected through appropriate compensation or other sufficient consideration?
- Has transparency regarding data origin through chain of commerce been properly documented and maintained?

Toxicity Level:

- The above serve as framing questions for an in depth E.A.I.S.I.S.. The varied end user sophistication, local regulations, intended use and need to protect naïve users must also be factored in any analysis.

Stage 2: Domestication Training for Smart Data[20]

The previous stage focused on data as a raw material acquired through capture in 'the wild' or possible reuse. This section addresses the preparation and use of the data in a sustainable manner for future reuse. The process of refining captured or brokered 'Dumb Data'[21] to ethically filtered 'Smart Data' for AI training and future use raises many significant issues. As the data forms into functional programs, it receives general user training before learning personalized lessons. In the 'book smarts' portion of AI learning, the programmer represents the greatest ethical concerns for data manipulation or misuse. Consider the following:

- To what degree do we consider the bias, motivations or simple skill level of the programmer?
- Should a user's decisions be limited by a programmers judgement?
- To what degree will ethical decisions be automated by factory preset programming?

[20] Processing/Subassembly/Assembly/Pre & Post Distribution
[21] Data freely observable in the public domain and holding no legal or moral deficiencies of title (raw data)

- Are programmers ethical decisions incentivized by corporations?

These questions will require further exploration in a E.A.I.S.I.S. analysis of the AIDC. These types of questions are best addressed and weighed for impact on sustainability during this stage.

The different stages of machine learning coincide with the developer's amount of dominion and control over the applications use and exposure to outside data, or ability to maintain a controlled environment. Prior to acquiring user specific preferences through human machine interaction, programmers are preparing it for a specific application for targeted users. At this point as the data is manipulated and algorithms are created for reaching an intended goal at peak efficiency. The programmer teaches the AI how to 'book smarts' through executing functions in a simulated environment without exposure to the real time challenges of daily use.

During this phase of the AIDC, the AI is under complete control of the manufacturer or developer. The data employed for training and machine learning should be in its best form for its intended purpose, hold a toxicity rating and maintain a clear record detailing the chain of custody prior to using for development. Although seemingly a relatively secure refuge from the risk of toxification, this stage is ethically challenging for the programmers. An Accountable, Responsible and Transparent Artificial Intelligence (ART-AI) ethical analysis framework brings some of these issues into focus:

Accountability

- Was a record of AIDC leading to this point accurately maintained to identify liable party(ies) for product related injuries?
- What guarantees or safeguards are in place to assist consumer in seeking help and identifying a human to hold accountable?
- What steps can be taken to minimize embedding conscious and unconscious bias in programming and data refinement?
- How will this application and use of data impact individual or group human rights?

Responsibility

- Are safeguards in place including insurance or other monetary funds to compensate for product defects or to mitigate damages?
- Have the programming and design team maintained a model for continuing education on ethical Artificial Intelligence implementation?
- Has an independent outside body or designated Chief Technology and Ethics Officer conducted a thorough E.A.I.S.I.S.?
- Are sufficient resources dedicated to educating and assisting public regarding risks of AI and data sharing?

Transparency

- Can the AI algorithmic decision making be monitored explained to the consumer?
- Where was the origin of the data and how has it been changed?
- Will explicit consent be obtained from user ongoing data capture during use of product?
- What safeguards are in place to avoid AI deceiving naïve users from confusion or anthropomorphized emotional manipulation?
- Is the chain of command and responsibility for any product defect clear and accessible to the public?
- Has toxicity level been disclosed conspicuously and clearly?
- Have factory AI motivational settings and risks to consumer been clearly disclosed?

As the AI moves through AIDC, the levels of future toxicity will increase as the trained applications are released back into the wild. These stages holding different threats for the ethical sustainability of AI. During Stage 3, the AIDC enters the period of greatest learning potential and risk of toxicity.

Stage 3: Return to the Wild & The Ecosystem of Trust

AI is meant to benefit and serve humans through the ability to learn and process decision autonomously. Placing AI in consumer's hands accelerates this process. This is when AI learns 'street smart' lessons, both beneficial and toxic. The utility of AI depends on its ability to learn from these lessons and offer consistent,

accurate and predictive solutions for the user. Mass adoption and the ability to maintain AI as an embedded and reliable resource relies on maintaining good data for this purpose. Sustainable AI depends on maintaining quality data through an ecosystem of accuracy and trust.

As this portion of conducting an E.A.I.S.I.S. is largely speculative this early in the process AI and AV adoption process, the missteps of social media can teach us some lessons on maintaining public trust. At the forefront of establishing trust in any interactive technology, the consumer needs reassurance against invasion of privacy, inaccurate information (fake news) and potential manipulation. The success and accuracy of ability of predictive systems to make the right decision for the right users depends on candor. Largely depending on self-reported or observed behaviour, AI requires true and accurate data to function properly. This requires trust and security in revealing intimate personal information.

Maximizing the capability of AI personal data points. However, consumer trust is continually eroded by data breaches, ransomware attacks, improper digital information capture or simply company policies considered bad faith actions undermining already shaky corporate trust. It results in a Catch-22 scenario wherein trust is needed to establish trust, but the big data companies largely squandered what little they previously held .

During Stage 3, the AI has the greatest ability to learn the appropriate response for a given context based on observations. This is not a new practice with most internet users encountering personalized dynamic advertisements and targeted news feeds daily, if not hourly or even more frequently. These are created through combinations of data mining, web scraping, emotional recognition and locational monitoring amongst other factors.[22] Although seemingly intrusive and despite protests of data gathering over-reach, it is a simple barter system. Data of commercial value is traded for a service or convenience of personal value.[23] Increased use of domestic robots has also increased opportunities for data capture.

[22] Kuempel, Ashley. "The Invisible Middlemen: A Critique and Call for Reform of the Data Broker Industry." Northwestern Journal of International Law & Business 36, no. 2 (2016): 207-35.

[23] Recognizing many issues exist regarding duress or adhesion contracts, further discussion is outside the scope of this paper. See ibid

As the novelty of owning a robot, like Alexa, wanes and dependency on it as a digital resource increases, familiarity of use can provide domestic robots candid and intimate information about their users.[24] This will most likely be compounded with AV. Drivers can imagine themselves as invisible when traveling alone, often inadvertently singing, dancing or practicing even more private habits in public view.[25] An ongoing E.A.I.S.I.S. model allows users security in knowing the fate of their sensitive personal data reinforcing the ecosystem of trust required for sustainable AI.

Stage 4: Toxic Data Mining: Remove, Retire, Rescind, Redisribute or Redress

The final stage of AIDC relies on the previous toxicity rating and evaluation for future human use. The existing toxicity is determined by levels of inaccuracies, socially unacceptability or danger to human interests. These factors, weighed against the ability to and difficulty in removing toxicity can be used to determine future reuse. The acceptable level of data toxicity will depend on the intended future use and foreseeable danger for unsophisticated end users. The following mitigation approaches are suggested:

Remove: Removing data should remain a decision for the supplier of the raw material, data. Currently, the Right to Be Forgotten from the Google Spain Case[26] and the recent findings against Volkswagen for DeiselGate[27]are paving the way for legal avenues to invoke this right.

Retire: Rather than outright removal or deletion, personalized data can be placed in a data trust for transferability to other smart devices allowing portability of personalized systems.

Rescind: As discussed, data capture is largely a barter system wherein data is traded for a digital service or product. The internet is forever but the conditions of a

[24] Lau, Josephine, Benjamin Zimmerman, and Florian Schaub. "Alexa, Are You Listening?: Privacy Perceptions, Concerns and Privacy-seeking Behaviors with Smart Speakers." Proceedings of the ACM on Human-Computer Interaction 2, no. CSCW (2018): 1-31.
[25] https://www.healthline.com/health/nose-picking
[26] Post, Robert C. "DATA PRIVACY AND DIGNITARY PRIVACY: GOOGLE SPAIN, THE RIGHT TO BE FORGOTTEN, AND THE CONSTRUCTION OF THE PUBLIC SPHERE." Duke Law Journal 67, no. 5 (2018): 981-1073.
[27] Krity Pladsen (25/5/2020) 'German Court rules automaker must pay dieselgate compensation' <https://www.dw.com/en/volkswagen-scandal-top-german-court-rules-automaker-must-pay-dieselgate-compensation/a-53556743 > [accessed 26/5/2020]

barter contract change. Rescission allows users to limit the unauthorized exploitation their data.

Redistribute: As AI is adopted and learns individual or culturally specific lessons, these heuristics can be helpful to others wanting to train their machines in like manners. For example: One AV driver may place greater weight on the risk of being stranded alone at night over the fuel and time savings of traveling on sparsely populated roads. These types of heuristics and personalizations should be capable of redistribution. Intellectual property protection should also be extended for any unique approaches to algorithmic reasoning that may become widely adopted and potentially profitable.

Redress: Data is a commodity and data creation of future use should be compensable. If personal data is employed for commercial enterprise, the originator should receive payment to reduce future liability claims or other challenges to its legitimate capture. Alternatively, if personalized data is mishandled resulting in harm to an individual, a standard system for redressing grievances including a well-defined compensation structure based on liquidated damages clauses should be considered.

Through employing one or a combination of the above methods to detoxify data would allow capture of valuable information while protecting personal interests. Remembering the goal of maintaining data in a manner free from human harm, if successfully applied an E.A.I.S.I.S. approach can preserve useful data while reducing toxicity.

IV. Conclusion

The mass adoption and increased demand for data requires a sustainable approach for AIDC. Reducing demands on natural resources, human labour and maintaining a digital environment capable of supporting a healthy, reliable and safe AI ecosystem is a public good. The risk of integrating bad data in AI application may result in small errors like recommending the wrong take out option, or more existential risks of AI misinterpreting data to disastrous effects.[28]

[28] "SUPERINTELLIGENCE." In The Technological Singularity, 85. Cambridge, Massachusetts; London, England: MIT Press, 2015.

The AV industry will serve as a leader in the field of human and intelligent machine interaction. Currently, an industry leader like TESLA stands in the best position to benefit from a commitment to employing an E.A.I.S.I.S approach. Implementation of ethics certifications, transparent and public E.A.I.S.I.S. disclosures and public engagement would allow a first-mover advantage and opportunity to acquire an association with Intelligent Data Responsibility that Volvo has similarly enjoyed with safety for years.

V. Bibliography

- Devin Coldeway, TechCrunch, 'AI is more data-hungry than ever, and DefinedCrowd raises $50M B round to fund it.' (26/5/2020)
- Joanna Bryson, United Nations University, Science, Technology, Innovation and Governance 'No One Should Trust Artificial Intelligence' (2018)
- Joanna J Bryson, Mihailis E Diamantis and Thomas D Grant, 'Of, for, and by the People: The Legal Lacuna of Synthetic Persons' [2017] Artificial Intelligence and Law
- John Fauvel "A History of Algorithms: From the Pebble to the Microchip (Book Review)." The Mathematical Gazette 84, no. 499 (2000): 168-69
- Ashley Kuempel "The Invisible Middlemen: A Critique and Call for Reform of the Data Broker Industry." Northwestern Journal of International Law & Business 36, no. 2 (2016): 207-35
- Josephine Lau, Benjamin Zimmerman, and Florian Schaub. "Alexa, Are You Listening?: Privacy Perceptions, Concerns and Privacy-seeking Behaviors with Smart Speakers." Proceedings of the ACM on Human-Computer Interaction 2, no. CSCW (2018): 1-31
- Alan Lewis, Polly. Turton, Thomas. Sweetman, and Ben. Caldecot. Litterbugs : How to Deal with the Problem of Littering. London: Policy Exchange, 2009
- Krity Pladsen (25/5/2020) 'German Court rules automaker must pay dieselgate compensation' <https://www.dw.com/en/volkswagen-scandal-top-german-court-rules-automaker-must-pay-dieselgate-compensation/a-53556743 >

- Robert C. Post, "DATA PRIVACY AND DIGNITARY PRIVACY: GOOGLE SPAIN, THE RIGHT TO BE FORGOTTEN, AND THE CONSTRUCTION OF THE PUBLIC SPHERE." Duke Law Journal 67, no. 5 (2018): 981-1073

- Andrew Ross (2018) Nexidia *Information Age* 'A Perfect Storm: The environmental impact of data centers'

- Darryl Seland "GARBAGE IN GARBAGE OUT." Quality 57, no. 5 (2018): 6.

- Rossi Setchi, Robert J. Howlett, Ying. Liu, and Peter. Theobald. Sustainable Design and Manufacturing 2016. Smart Innovation, Systems, and Technologies 52. Cham: Springer International Publishing, 2016

- Sherry Turkle, *Alone Together : Why We Expect More from Technology and Less from Each Other*. New York: Basic Books, 2011