



Artificial Intelligence: The Societal Responsibility to Inform, Educate, and Regulate

AI Matters Editors (ACM SIGAI; aimatters@sigai.acm.org)

Alexander D. Hilton (Southern New Hampshire University; alexander.hilton@snhu.edu)

DOI: [10.1145/3362077.3362088](https://doi.org/10.1145/3362077.3362088)

Introduction

Artificial Intelligence (AI) is a rapidly growing field; one that is mysterious to the general public. The mention of the word AI fills the imaginations of many with thoughts of talking robots, jobs being replaced, and possibly even the destruction of mankind. Perhaps imaginations are running wild due to, perhaps driven by the loose definition of AI as systems able to perform tasks that normally require human intelligence that allows Hollywood to take some creative license. The experts in the field tend to work directly with AI and often for large companies, allowing for the imagination and news headlines to be where the public gets their information. Many wonder if this new technology is going to be an overall benefit to society or if it will bring unmitigated disaster. When the imagination runs wild, instead of understanding, news stories can perpetuate concerns and anxieties rather than hope and optimism.

Andrew Ng once said “Just as electricity transformed almost everything 100 years ago, today I actually have a hard time thinking of an industry that I don’t think AI will transform in the next several years.” It may well be the next electricity in the way it will revolutionize and change both industry and even our daily lives. From applications that identify faces in photos on social media to deep learning models meant to help discover new cancer treatments, the potential of AI can impact nearly every person’s life and may already have to some degree. Most of us are already interacting with AI in some form just when we use Google or Facebook, often without our knowing. You may have interacted with one today without even realizing it. When faced with technology that has a large reach and broad scope, it is imperative to consider how the lives of millions could be impacted. We are already seeing individuals stepping up in the

political realm with ideas of how to address safety and privacy concerns associated with AI. Andrew Yang, who announced his bid for the 2020 Democratic Presidential nomination, is running on a platform that stated ‘the robots are coming’ saying that AI will change virtually every part of the global economy. He could be right.

It is certain that AI will begin to play a more prominent role in our daily lives as the field develops and new uses are found for such systems. The change AI can influence our lives in positive ways but could put lives at risk if irresponsibly used. Unfortunately working through governmental bureaucracy and regulation can take time, yet AI continues to advance rapidly. As both individuals and corporations take huge strides forward some governments have begun to realize the impact of easily accessible data that feeds into AI. The European Union implemented the General Data Protection Regulation in 2018 as an effort to inform and gain consent by people to ensure their data cannot be exploited.

The General Data Protection Regulation (GDPR) was an attempt to begin to change how EU citizens interact with their data and offer some protections. AI is thriving in this era of big data, however, unlike data, people will be interacting directly with AI in new and intriguing ways. Yet these interactions will not just be predicated on the quality and behavior of the AI, but also the accessibility and knowledge of the human involved. Interactions are a two-way street and while we need to look at the regulations put on AI, perhaps the most important side of the equation is how society views and interacts with AI. Discussing these requirements going forward in this coming age of AI, which will likely cause a renaissance for society, must consider the humans involved and the technology simultaneously. The quality of the interaction that AI and humans will have will stem from knowledge, access, and

open-mindedness.

Development of these interactions must come from establishing knowledge, establishing trust, and then encouraging responsible use. Allowing for this sort of positive interaction, however, takes more than just governmental oversight, but also investment and a commitment from society. Encouraging these interactions should start with the hardest factor to change. AI can be reprogrammed or retrained, humans are a bit more difficult to change. It is critical to encourage and invest in the human element first before discussing the requirements put on AI.

The Human Factor

Up until recently, I was someone who was about as versed in AI as the average American. Most of my knowledge came from Hollywood movies like “I, Robot”, “The Matrix”, and “Terminator” where AI is depicted as programmed systems coders wrote logic for, which ended up leading to some devastating results. Then I saw the incidents making national headlines of AI crashing cars, chat bots that respond with racist remarks, or smart TVs and speakers listening into conversations and I began to worry.

Most people chalk these problems up to human error, logical issues we could not foresee making it hard to code. Sure, a team of programmers could correct these mistakes, but what about the human error caused by that correction. Humans are fallible so would the hard-coded logic I viewed AI as be the same? Sure, I was optimistic about the prospects of AI, but what if these sci-fi stories and futurists were more foretelling than just stories? Could a rogue AI change everything about our lives for the worse? Human error will not go away, even the best program can have a bug, AI did not seem it would be different. It made me anxious, and I empathize with people who still feel that way.

Studies have shown many Americans to view artificial intelligence in a similar light. While people tend to be cautiously optimistic about how AI can positively impact their lives, they are also very anxious about the prospect of AI. Many people worry about it changing the industries around them, perhaps taking the jobs

of their friends and family. There is concern that AI cannot handle lives as safely as a human, an anxiety which is increased when we hear of car accidents involving AI and the loss of life. These anxieties are real and as AI becomes increasingly prevalent these breaking news stories won't just happen a few times a year, but could occur daily just due to how available AI is becoming. Much of these anxieties revolve particularly around jobs and losing jobs to robots and AI automation ([Gallup, 2018](#)).

Recently, however, I've been able to put my own anxieties regarding artificial intelligence aside. The reason for this was simple, although it was also a bit of a journey. As a data analysis student, I became curious about machine learning and deep learning in their roles in analysis which led me to further my education on the matter. The resources I found came from people like Andrew Ng, Siraj Raval, and Cognitive Class from IBM. Many of these people and classes came at no cost which made it very easy to dive into without having to worry about the financial commitment. It blew me away to learn that most AI we discuss today is not hard coded logic, but rather mathematical operations performed on massive amounts of data. These algorithms, while still susceptible to human error and issues in the data, are learning models that can improve as our data processing abilities become better. Likewise, with many of these neural networks being open source packages like TensorFlow and PyTorch, a massive community of engaged data scientists and programmers can improve how neural networks and AI is designed, in many ways democratizing the process. There isn't just some small group of mad scientists trying to make a humanoid robot, but rather a substantial community that is engaged in trying to make the most of AI.

Neural networks as well as much of modern artificial intelligence can learn and improve far easier than I imagined and the community behind much of AI wants to encourage society to take that next step into the future with an open mind. AI certainly has the capacity for replacing jobs or crashing cars, but these AI can be improved to facilitate our lives offering new opportunities even for the very jobs that they are said to replace. The ease of accessibility to knowledge and the community that discusses

and works on building artificial intelligence and neural networks shows that people want to improve these systems as well as improve the lives of others. After learning about the easily accessible and free learning options pertaining to AI, my anxiety turned more optimistic as worry turned into hope.

Knowledge is power, especially when that knowledge is easily accessible. As fear abates the full capabilities of AI can be revealed. When new technologies for instance are introduced there are always anxieties regarding the new unknowns that have been introduced. When electricity was first taking off, there was often fear related to what it could do. Some claimed it could even destroy the concept of day and night, which some thought could significantly impact our culture, or even our health. While some of these fears could be stoked by incidents like electrical fires, allowing people to become more educated about how electricity works curbs these fears. People can learn how to protect themselves from starting electrical fires with some education growing up with the technology, as well as being told not to shove metal object in the socket. With AI, the principles are much the same, people need to be taught how to adequately protect themselves from harm rather than just give into fear and myth put out by movies or sensationalized news headlines.

However, recent incidents have shown how society reacts to irresponsible use of technology, especially regarding big data which is closely tied to AI. Cambridge Analytica was infamous for using data and machine learning to try to change and influence elections, most notably in Kenya where they used their insights to re-brand a certain political party twice ([Lang'At, 2018](#)). People who were unaware that their data was being used were easy to target. These individuals could be advertised to or presented with information that could have changed their political opinions. When the story broke about how Cambridge Analytica had gathered so much personal information on people globally, the US Senate had hearings and the EU introduced the GDPR in response as well.

The reaction to the news stories regarding the Cambridge Analytica scandal was not unexpected, but it did reveal a lot of societies lack

of knowledge when it came to these issues which the GDPR sought to address. While this was a regulation, the idea of informed consent regarding using data is based on establishing a degree of trust and interactions with data. This concept plays into how AI functions as AI requires data to learn. While these regulations could hinder innovation, it is a key step into establishing trust in these interactions as well as ensuring people be informed about the systems they are interacting with. If humans interacting with devices that contain AI are unaware that these devices are gathering and processing their data, even in a benign way, if (and when) that data gets leaked then there is a huge breach of trust. When humans are caught off guard and made distrustful of these devices, single-event learning can take hold which can turn off people to AI entirely. People do not regain trust quickly so avoiding this sort of situation is paramount.

Nevertheless, simply informing people that they are interacting with AI is only a first step. Being informed is the start to education but going a bit further can put those anxieties to rest while opening doors for new opportunities. Presenting educational opportunities for people to learn how AI works and how they learn and improve is a start. When people are introduced through free resources by some engaged minds and companies in the field, it lets them explore and learn about these systems in an empirical manner. Education has always been a good tool to helping with anxieties that involve natural and artificial phenomena, people are not terrified Zeus will strike them down with a bolt of lightning now that we have a general understanding about weather patterns for instance. Discussing AI should be treated similarly especially as it becomes increasingly more prominent in our day-to-day lives. Now this isn't to say that everyone needs to get a Master's in Computer Science (nice as that would be), but that people have a similar understanding of AI as they do about electricity for instance.

Spreading public general awareness of AI through free, or at least very cheap, educational programs allows for an open approach. While it would be nice to see these groups have a larger reach, they are forging a path in the right direction. Keeping the knowledge out there and accessible to the public helps en-

sure that people will start to learn their own best practices as they watch AI grow. As people start becoming more aware of the AI in their lives and have an easy ability to learn about those systems, then gradually society can adopt AI into their lives responsibly just like we have done with electricity.

Furthering an open approach to help spread public awareness and knowledge should be considered a societal investment. AI has a significant amount of potential, however, if people have anxiety and fear while simultaneously being uninformed can allow for society to create bad policy and bad practice that stifles innovation and growth. While groups are gradually informing the public, a proper investigation into a public awareness campaign and increasing public education on the matter should be considered by industrialized nations. Likewise, investments in education and higher education are a must going forward, not just for people to understand how AI works, but because of its potential to change the job market as well.

Fostering education in AI can come from non-profits, companies, or even governments. Groups like the ACM SIGAI sending out newsletters and offer webinars to members is just an easy way to keep people informed of what is going on in the field and what to be aware of. Other groups do this as well such as Andrew Ng's deeplearning.ai and Siraj Raval's School of AI. Even IBM's Cognitive Class offers free classes for people to learn programming with machine learning and deep learning. These free and inexpensive courses do not preclude anyone from learning about AI. This approach makes learning about AI easier for those who are interested, these groups are all easy to approach and often have a fair bit of free content to help people become educated and even versed in the various aspects of AI, regardless of income. As well, people who post on Github can share their knowledge and expand the reach of understanding. As stated before even the open source packages like Tensorflow open many opportunities by encouraging people to work with deep learning, while Google still profits through their cloud and selling their tensor processing units, the overall accessibility Tensorflow has offered allowed many people to experiment with and learn about deep learning hands-on.

Expanding the accessibility to and education of AI should be a focus; some nations are working to that end. China is attempting to get a one trillion-dollar industry developed. The EU is putting billions of dollars of investment into AI to try to catch up with the US and Asian countries. MIT is in the midst of building a billion-dollar AI college to make AI part of every graduate's education. All this investment will hopefully allow more members of society to be informed about and able to adequate use and interact with artificial intelligence. Likewise, this increases the availability of experts and enthusiasts that can then go out to develop businesses, large and small.

Taking steps to ensure that AI businesses are invested should be a priority. Ensuring that these new technologies are subsidized to be easily accessible by the wider community rather than just being proprietary is important as well as AI then becomes more like a utility rather than a luxury. The accessibility is vital to ensuring that people can start to become educated. Finally, investing in education whether it be through public education initiatives or just through communities forming together to educate people should be encouraged by society in both the form of private investment and governmental assistance.

Though, with all these enthusiasts and business there is a chance mistakes will be made such as the events Cambridge Analytica helped spark. Especially when AI can produce "deep fakes" or other forms of scandals, it should be expected that we will see some horror stories in the coming decades. With an educated populace, the risk of panic and fear can be mitigated substantially and trust in AI will not be broken entirely. Likewise, the appropriate regulations can be discussed early without concern of over regulating and stifling the development of better artificial intelligence.

Regulation and Requirements on AI

For any decisions with AI interacting with humans, the best solution is allowing both society and AI to treat those interactions as learning experiences. Ensuring access to education is a first step here along with ensuring accessibility to AI. This accessibility allows AI to be treated more like a utility which can help minimize damage when accidents occur

and the fear generated by those mistakes. So long as these inevitable errors are small and do not put lives in serious jeopardy, the hope should be that the AI community and the greater global society can inspect and adapt as issues crop up with AI. However, in order to properly carry out this inspection society must be knowledgeable of what they are interacting with.

As we saw with the Cambridge Analytica case and the GDPR that the EU put in place in response, similar mechanisms should be considered for AI. When the GDPR came into effect, two out of three people felt more comfortable sharing their data ([Association, 2018](#)). This statistic alone is important for AI as people who feel more comfortable sharing their data might be more willing to hand that data over, so it can be used for training more complex models. In some ways, despite how the GDPR has been suggested to impact businesses, it may be necessary to the long-term trust in the data collection process, which is incredibly beneficial for AI.

In a similar way, as the GDPR helped build confidence in people's use of data, so too could be similar regulation to keep people informed when they are interacting with AI. There are no surprises that way which increases faith in these systems through informed consent. Perhaps a GDPR for AI is the next logical step for nations to consider. It increases society's faith in these technologies which may mean more data and information can be gathered through those interactions, thus leading to more information to improve the next generations of AI. This regulation could be as simple as just labeling something as containing AI or could go into more detail, such as if a neural network is gathering video data and whether that network is a convolutional neural network or if it is recurrent network which could help identify what the AI is trying to accomplish. Informing people of the process can let them know whether they are dealing with an object recognition program or perhaps one that monitors and identifies activities. Different people might feel better about interacting with one system versus the other, thus should be aware of what they are handling. So long as people are generally aware of what the artificial intelligence is doing, they can volunteer to still work with and use it rather than worry that the AI is doing

something it is incapable of doing.

Trust is key to the long-term usefulness and viability of AI in general. Building and maintaining trust in these systems as mistakes are made will ensure society does not treat AI with concern and anxiety, but as a tool to be utilized to benefit and grow our lives. In the cases of human lives being put at risk as well, trying to ensure that informed consent is given at least creates an understanding of the potential risks as well. These stories become less terrifying as the onus still falls to the human working with the AI rather than people just blaming and worrying about the AI.

Yet society should not simply accept any company or individual putting lives needlessly at risk. It is an advantage to encourage small businesses and individuals to harness AI as the ramifications can be less than if larger companies are the ones to make these mistakes. We know this from examples as well. These mistakes won't be like the Equifax breach in scale, but they can still happen and even a simple car accident with an AI driver can make national headlines. We cannot necessarily expect an AI to be working at Bayes error rate, the lowest possible error rate, every single time. In these situations where lives can be immediately put at risk; the AI should be making errors at similar levels to humans before they are deployed, thus needless risk is not introduced. While this can limit innovation, there are potential work arounds, such as with self-driving cars having a human driver as well. This system decreases the chance of an accident, but mistakes can still happen as we saw in 2018 with a self-driving car crash in Arizona where a human driver was still involved. That incident shows that a higher standard must be set.

Where lives are directly involved AI should only be deployed if the error rate the AI makes is at the same level of humans or performs better than a human. We can expect AI to make errors, but the goal should be to minimize that error rate as much as possible. The struggle is that these systems often learn through gathering massive amounts of data which means at some point they need to be tested in real-life scenarios. Under these circumstances, the AI must have that error rate verified before they enter these situations that

could put lives at risk. Much in the same way we impose driver's licenses on people, so too should AI that could lead to loss of life. Once the human error rate is met, then the AI should be deployed. At this point it can gather data and learn about these real scenarios instead of just learn on simulated data. With any luck, the AI will surpass human error as it learns more in real world scenarios (Jalan, 2017). Perhaps even in the incidents that do led to tragedy, the AI will still improve past the human.

There are even a few potential considerations to have. A self-driving car cannot get drunk and could save lives in the end by preventing incidents like DUIs which can cause thousands of deaths every year (CDC,). Another area could be AI driven drones that, if properly programmed and trained, can be used to target enemy combatants and work to prevent civilian casualties. AI does not have the trouble with fatigue and if it misidentifies, drones powered by AI could take our troops out of harms way while simultaneously making less mistakes such as friendly fire. Along the same lines, long haul trucking or flights could be piloted or assisted by AI without the concern of fatigue or illness potentially imparting cognitive functions. Regulating for what we have never seen or given the opportunity to try though could shut down opportunities to improve the human condition and these AI technologies. The potential opportunity for AI could lead to more lives saved in the end, so long as we can learn the appropriate applications and limitations that is.

Imposing regulation where lives are at risk could be built around the very same concepts used to license human drivers and operators. This regulation would be the most intense as it means creating a system and licensing agency. Previous to the age of big data and the ability to create simulations, this may have been impossible. However, now we can, with a fair bit of accuracy, compare how an AI stacks up to a human in a variety of situations. When the error rates are roughly the same in simulated environments, the company or individual can request for an operating permit to deploy these AI in a real context. While this is massive government oversight, it ensures lives are not put at wanton risk just for the sake of progress or a quick buck.

With these two requirements put on AI, one that requires informed consent for these interactions and another to ensure that where lives are at stake the risks are mitigated, maybe society can start to build trust with these new systems and as more complex AI enter our lives. Perhaps society would even welcome more human-like robots into their lives as these anxieties are curbed through knowledge, consent, and trust.

Conclusion

The key to AI is not unknown to us. Humanity has undergone technological revolutions in the past and there will be more breakthroughs after AI becomes a staple in our daily lives like electricity, computers, smart phones, the internet and many, many other technologies. Like these technologies, the most important element to interaction is knowledge and accessibility. Responsible use cannot just be mandated even though it sometimes feels like the only alternative. The slow process of learning how these new technologies impact our lives requires adaptation by society.

Mistakes will be made, ensuring that these mistakes first happen on a small scale allows society to adapt as these problems come up, rather than regulate for events which have not or might not occur. Educating people also ensures that the minimum level of regulation or societal shift can be made when these mistakes occur so that AI can still be fostered and develop. With any luck this means that AI will be a positive technological revolution that can take humanity forward quickly rather than something people need to be anxious over. As a global society tackling this new industrial revolution with AI, there are a few critical steps we can take now to better society's understandings and ensure responsible use.

These are:

1. Societal (potentially governmental) investment in education of and access to AI
2. Ensure informed consent when people interact with AI
3. Regulate and license AI if lives could be put at risk

Much of the investment and regulation can come from businesses and communities being

respectful of artificial intelligence and the people interacting with it. It can come from allowing transparency through open source packages or informing people a little about the systems they use. People can volunteer to teach and train their communities about AI along with giving information on how to utilize AI to positively impact their lives. Governments may need to get involved if serious incidents occur, but overall the focus should be on encouraging responsible use and providing educational opportunities and access to technologies that utilize AI. Hopefully, this approach will allow people to treat AI as a utility that betters their lives, rather than an enigma to be concerned over.

Putting society's anxieties to rest while fostering knowledge and access should allow for innovation and invention at a rate humanity may not have seen before, akin to that of how electricity changed our lives entirely, even changing how we viewed night and day. Perhaps AI will have a similar impact in our lives and the next generation will see the world in a new light after this coming AI technological revolution. It is hard to say what the changes will be, but if we embrace with understanding and open mindedness, then society could change for the better.

The interactions that take place between humans and AI will set the tone for how these technologies impact our lives and whether they improve the broader society or just a small handful of people. Establishing trust in that technology and spreading knowledge has been humanities solution in the past and protecting lives to maintain that trust is vital. If humanity can succeed at learning and adapting with artificial intelligence, a new era may still dawn.

References

Impaired driving: Get the facts — motor vehicle safety — cdc injury center. https://www.cdc.gov/motorvehiclesafety/impaired_driving/impaired-driv_factsheet.html.

Association, A. . T. D. M. (2018). Gdpr: A consumer perspective.

Gallup (2018). Optimism and anxiety: Views on the impact of artificial intelligence and higher education's response. <https://news.gallup.com/reports/226475/gallup-northeastern-university-artificial-intelligence.aspx>.

Jalan, K. (2017). How to improve my ml algorithm? lessons from andrew ng's experience - ii. <https://towardsdatascience.com/how-to-improve-my-ml-algorithm-lessons-from-andrew-ng-ii/>

Lang'At, P. (2018). Cambridge analytica and kenya elections. <https://www.nation.co.ke/news/Cambridge-Analytica-and-Kenya-elections/1056-4349392-204vmx/index.html>.



The AI Matters Editors

may be contacted at aimatters@sigai.acm.org with questions about the submission process for the AI Matters newsletter. Complete submission instructions may be found on the AI Matters website. If you note any bugs in the AI Matters template, please let the editors know.



Alexander D. Hilton is a Southern New Hampshire University alumnus with his Bachelor's in Data Analytics having graduated the program May of 2019. During his time at SNHU, Alex helped found the ACM Student Chapter at SNHU and worked as a Peer Leader. Alex has shared his passion

for AI and the Data Sciences by creating a study community for students. In his professional life, Alex has worked as an Agile Practitioner, having earned several certifications in Scrum. Most recently, Alex received his Scrum@Scale Practitioner Certification under Jeff Sutherland, a co-signatory of the Agile Manifesto and co-founder of Scrum.