ETHICALLY ALIGNED AI

REPORT ON AI ETHICS

CANADIAN MARKET SCAN

ETHICALLY ALIGNED AI | June 2021

"THE WORLD HASN'T HAD THAT MANY TECHNOLOGIES THAT ARE BOTH PROMISING AND DANGEROUS" – BILL GATES

Artificial Intelligence (AI) is a game changing technology that is already impacting our lives in ways we don't always see. AI technology is already almost ubiquitous in its reach into our lives, from autonomous vehicles to the Internet of Things (IoT), to a wide range of AI enabled algorithms and robots. But as with all new technologies, advancement can outpace regulation, unleashing a range of ethical concerns about possible harm done by the careless deployment of AI.

Ethically Aligned AI is a company that provides education, tools, and consulting services to organizations that aspire to build and/or deploy ethically responsible artificial intelligence and machine learning (AI/ML) systems.

This report will present a market scan and analysis conducted during March and April 2021. It includes secondary research into the field of AI ethics, and primary research interview surveys conducted with individuals at Canadian organizations representing a variety of industry sectors.

This research was conducted to explore the market surrounding the impact of ethics on AI where people and technology converge. Analysis of the research assessed the state of market readiness; interests of the surveyed companies with regard to potential regulatory compliance; and services and product offerings to address these issues.

The names of both the companies and interview participants in this study are being kept confidential.

Part 1: Introduction to AI Ethics

The rapid and ubiquitous embrace of AI raises a host of questions and issues around how AI is developed and deployed. There are concerns about bias, fairness and transparency, as well as safety. Issues around privacy and control of data also need to be examined. Data is a key input in training artificial intelligence systems. Reams of data are being gathered and used to train systems in a "wild-west" approach. There is a lack of transparency about what is being gathered, by whom and for what purposes. The people generating this information are not told how their information may be used and what datasets are being combined.

AI systems are being deployed that are making decisions which, in some cases, can impact lives. The people being impacted by these technologies often don't know, or understand the implications. There are several documented cases where decisions made by AI technologies have adversely impacted people's lives without a clear explanation of why and how these decisions are being made. It's not clear who is setting the agenda and what standards or rules, if any, are being applied. If AI is going to impact all of us, shouldn't we have the opportunity to weigh in on the discussion?

One example that illustrates these issues in practice is the case of Amazon's hiring algorithm that inadvertently created a bias against women.

One example that illustrates these issues in practice is the case of Amazon's hiring algorithm that inadvertently created a bias against women. "Everyone wanted this holy grail...the company's experimental hiring tool used artificial intelligence to give job candidates scores ranging from one to five stars - much like shoppers rate products on Amazon. They literally wanted it to be an engine where I'm going to give you 100 resumes, it will spit out the top five, and we'll hire those".[1] However, the company recognized that the new system was not recommending female candidates for certain technical jobs, an error caused by a decade worth of AI training data that reflected a male gender bias for these roles. Even after attempting to correct the situation, the system still found ways to discriminate. Its decision-making process wasn't even transparent to those who designed it. The project was eventually scrapped in 2017, three years after the project launched.

Core issues

There are many issues that could be examined in terms of AI ethics. Our focus is on bias, discrimination, fairness, power imbalances, increasing inequality, transparency and privacy. These are issues that are impacting people right now and need to be addressed to correct current deployments of AI technologies, as well as prevent future technologies from exhibiting dangerous design flaws or being used in ways that perpetuate inequality. Broadly speaking, these issues can be addressed through

[1] Dastin, J. (October 9, 2018*) Amazon scraps secret AI recruiting tool that showed bias against women*. Reuters.

creating conditions for diversity, inclusion, accountability and trust in both the development and deployment of AI technologies.

1. Bias and Discrimination

A briefing document prepared by the technology research organization, Forrester, titled *The Ethics of AI: How to Avoid Harmful Bias and Discrimination* states: "By their very nature, machine learning algorithms can learn to discriminate based on gender, age, sexual orientation or any other perceived differences between groups of people".[2] The document outlines how bias can be baked into machine learning models through various types of bad data:

- **Incomplete datasets**, such as a lack of sufficient data to accurately reflect the population. For example, facial recognition programs that don't have enough data for darker skinned individuals can create racist outcomes.
- **Data sets containing errors**. Most data sets contain errors and have to be "cleaned." Cleaning can be a time-consuming and expensive process, thus some companies would prefer to "look the other way."
- **Historical bias** that gets encoded when a model uses a proxy for race, age or another discriminator. For example, the majority of single parents in the US are female, thus the resulting model can use single parent as a proxy for females, creating gender bias in the model.

The report goes on to describe a FAIR model, which stands for creating learning models that are Fundamentally sound, Assessable, Inclusive and Reversible. It's one example of how companies developing AI might approach correcting their datasets, if necessary, and ensuring they can explain their models. There is self-interest in this approach for companies whose brand, reputation and revenue are ultimately at risk if they don't attempt to address these issues.

The tech industry in general suffers from an underrepresentation of female, Black and Hispanic employees.

2. Inclusion and Diversity

Inclusion is part of a sustainable solution to address bias, and it was a key issue at the G7 Multi-stakeholder Conference on Artificial Intelligence in Montreal in December 2018. Some of the questions outlined in a discussion paper on this topic were:

- How do we make sure gender, social and cultural diversity fuel AI design and development?
- What are the best practice examples of inclusion and diversity in AI and how can we replicate them?
- What can the government, academia, and industry do to promote an inclusive use of AI technology that will benefit a diverse society?[3]

One of the best ways to ensure bias doesn't enter into the design of AI technologies is to ensure the people designing the technologies represent a diverse perspective. "The tech industry in general suffers from an underrepresentation of female, Black and Hispanic employees."[4]

[2] Purcell, B. (February 27, 2018) *The Ethics of AI: How to avoid harmful bias and discrimination*, p.6.
[3] AI for Society. (December 2018) *Inclusion in AI Development and Deployment*. [PDF].
[4] Purcell, 2018, p. 6.

In a recent report, the AI Now Institute calls for fairness, accountability and transparency of the "full stack" supply chain. This is an attempt to account for all aspects of the components that make up the AI system. "For meaningful accountability, we need to better understand and track the component parts of an AI system and the full supply chain on which it relies; that means accounting for the origins and use of training data, test data, models, application program interfaces (APIs), and other infrastructural components over a product life cycle."[5]

3. Data Privacy

The topic of data privacy is vast and could easily be the focus of an entire paper. It needs to be mentioned as part of the ethical considerations pertaining to AI ethics because data is a key input for training AI systems. As it relates to AI ethics, the issue is one of using data for training AI without the proper consent or visibility from those whose data is being used. In general, most countries do not have strong data protection laws and as such, companies are gathering data and using it with little to no oversight.

4. Transparency, Auditability and Accountability

Two core ethical questions that need to be addressed are how a decision is made and who is responsible or accountable. These questions relate to issues of transparency and auditability – being able to see and understand how the process works, so that there can be accountability. This is an important concept not only from an ethical standpoint, but from a legal standpoint as well.

Accountability is about a clear acknowledgement and assumption of responsibility and 'answerability' for actions, decisions, products and policies, building explainability into the AI systems, determining which individuals or groups are accountable for the impact of AI algorithms, and as a feature of the broader socio-technical system that develops, procures, deploys and uses AI.
Currently there is an accountability gap between those who develop and profit from AI and those most likely to suffer the consequences, and that gap is growing larger.

Finally, it's not enough to apply any of these issues in a "one and done" fashion because AI systems learn and evolve. Thus, what might be onside at the time it's deployed, may, with new datasets and new learning, move out of alignment. Therefore, it's important to continually monitor these systems to ensure that they remain in alignment with the intended goals and objectives. This monitoring aspect is part of the ethical framework.

This is by no means an exhaustive list of ethical issues but it does serve to highlight some of the major problem areas that need to be addressed.

Mapping the Regulatory Environment

In general, there seems to be agreement that Canada needs to improve laws on the use of data and automated systems.[6] There is also an understanding that Canada needs to catch up with some international laws, such as the European Union's General Data Protection Regulations (GDPR) which came into force in May 2018. The GDPR applies to any organisation operating within the EU, as well as

[5] Whittaker, M. et al (2018) *AI Now Report 2018*. AI Now Institute, NYU, p. 5.
[6] Hodgett, S., Liu, T., & Perey, A. (2021). AI, machine learning & big data 2021|Canada. Global Legal Insights.

any organisations outside of the EU which offer goods or services to customers or businesses in the EU. In other words, almost every major corporation in the world needs a GDPR compliance strategy.[7]

1. Canadian Government

There are a few government initiatives in Canada taking place at the federal level. The Treasury Board created a directive on the use of Automated Decision Systems used by the federal government, which came into effect on April 1, 2019 and applies to Automated Decision Systems procured after April 1, 2020.

The objective of this Directive is to ensure that Automated Decision Systems are deployed in a manner that reduces risks to Canadians and federal institutions, and leads to more efficient, accurate, consistent, and interpretable decisions made pursuant to Canadian law.
The expected results of this Directive are as follows:

- Decisions made by federal government departments are data-driven, responsible, and comply with procedural fairness and due process requirements.
- Impacts of algorithms on administrative decisions are assessed, and negative outcomes are reduced.
- Data and information on the use of Automated Decision Systems in federal institutions are made available to the public, where appropriate.[8]

Additionally, the Digital Charter Implementation Act (DCIA) was proposed in 2020 to update Canada's privacy laws to increase protections to Canadians' personal information by giving more control and greater transparency when companies handle personal information. It lays out 10 general principles around the use of digital technologies, including the right to an explanation of AI decision-making, or even the right to opt out of having personal data collected. The Charter would implement steep fines for violations, and strengthen enforcement and oversight.[9]

2. Provincial Governments

Various provincial jurisdictions are also starting to release "guidance" and other directional but non-binding statements, signalling that more localized regulation for AI may be on the horizon. For example:

- Alberta's privacy commissioner made a statement suggesting a need for regulation for the use of facial recognition and other AI technologies.[10] The Alberta Government is also in the process of drafting an AI strategy and roadmap.
- BC is in the process of starting discussions based on a number of principles including trust, transparency, inclusivity, and embedded ethics.[11]
- Ontario, responding to business pressure and increasing AI use/awareness, recently held consultations to build an AI framework that will be accountable, safe and rights based. The report will be released in July 2021.[12]

---

[7] Gaytandjieva, I., et al. (2019). Comparing privacy laws: GDPR v. PIPEDA [PDF]. OneTrust DataGuidance.
[8] Directive on automated decision-making. (2019). Government of Canada.
[9] Innovation, Science and Economic Development Canada. (2021). *Canada's Digital Charter: Trust in a digital world*. Government of Canada.
[10] Clayton, J. (2021, February 3). Commissioner's Full Statement on Clearview AI Investigation [Press release].
[11] Digital principles for the Government of British Columbia (version 1.0). (n.d.). BC Government.
[12] Consultation: Ontario's trustworthy artificial intelligence (AI) framework. (2021, May). Ontario Government.

3. Canadian Legal Community

The legal community in Canada is also beginning to analyse the implications of AI from a variety of litigation perspectives.[13] Currently, there appear to be few relevant cases but many identified gaps, and a sense that AI changes much about many of the assumptions on which the law is founded, that whole areas of law may need to be reworked. Examples include:

- Contract law generally assumes that consenting humans "make a deal". When an AI drafts or adjusts a contract, that working rule of thumb no longer applies in the same way[14]
- Intellectual property law generally assumes that humans create ideas and inventions. When an AI makes something a human did not anticipate, concepts of ownership, property and related purposes of laws are upended. (E.g. patents as an incentive to create)[15]
- If an AI makes a decision, or causes an accident,[16] who is responsible in the eyes of the law?

In general, the legal profession mostly appears to be considering AI in its business of law, as opposed to public policy implications.[17]

Investment and Consumer Market Trends

There are some socio-cultural trends that are impacting the market place and the ways in which investors are evaluating companies with which they wish to do business. These trends are of particular interest to markets that focus on emerging technologies.

1. Socially Responsible Investing (ESG)

A notable shift that is rapidly taking place in the global financial industry is the consideration of environmental, social, and governance factors of a company when making investments (MSCI). Commonly referred to as ESG investing or socially responsible investing (MSCI), there has been a steady increase in investments of this nature since 2004, which suggests that the trend is here to stay.[18] Larry Fink, CEO of Blackrock, notes in his most recent letter to the CEOs that throughout 2020, we saw a global 96% increase in dollars invested in sustainable assets. He then goes on to express his belief that this is only the "beginning of a long but rapidly accelerating transition" in investing practices.[19]

---

[13] Martin-Bariteau, F., & Scassa, T. (Eds.). (2021). *Artificial Intelligence and the Law in Canada*. LexisNexis Canada.
[14] Martin-Bariteau & Scassa (2021), chapter 3.
[15] Martin-Bariteau & Scassa (2021), chapter 1.
[16] Martin-Bariteau & Scassa (2021), chapter 4
[17] Christian, G. (2019). Predictive Coding: Adopting and Adapting Artificial Intelligence in Civil Litigation. *Canadian Bar Review* 97(3).
[18] Henisz,w., Koller, T., Nuttall, R. (November 14, 2019). *Five ways that ESG creates value.* McKinsey & Company.
[19] Fink, L. (2021). *Larry Fink's 2021 letter to CEOs.* Blackrock.

People want to know that their money is going towards something good, and investors are recognizing that ESG elements are essential to the long-term sustainable success of a company. While there are few official regulations to enforce ESG investing, Norway has recently announced ESG guidelines for their 1.3 trillion-dollar sovereign-wealth fund, signifying that increased regulation is looming. The Norwegian finance minister points out that "there is a growing expectation, not only in Norway, but internationally that there will be a high level of ethical awareness" when investing.[20]

It will be essential for companies developing and deploying AI to consider the ethics underlying their technology, in order to uphold positive social and governance practices.

As the AI industry grows in tandem with trending ESG considerations, and regulation emerges in each area, there will undoubtedly be overlap. It will be essential for companies developing and deploying AI to consider the ethics underlying their technology, in order to uphold positive social and governance practices, and ultimately maintain a sustainable competitive advantage. Companies using AI systems either directly or as part of their supply chain have an opportunity to get ahead of these issues by addressing algorithmic risk and ethical concerns now.

2. Millennials, Generation Z & Corporate Social Responsibility

It's important to note that pairing closely with the trend in ESG investing is the shift in approach to consumption and workplace selection that is occurring in the Millennials (born 1980 to 1994) and Gen Z (born 1995 to 2012) generations. From a consumer standpoint, Gen Z no longer judges products solely based on quality, and instead focuses on the ethics and social impact of the company that makes them.[21]

In a study done by McKinsey, 70% of the Gen Z respondents said that they try to purchase from companies they would consider to be ethical.[22] Rather than selecting products only for function, the consumer behaviour of these generations is based on company alignment with their own personal values, increasing the importance of factors like sustainability, ethics, diversity, and inclusion in the corporate world.[23]

[20] Taraldsen, L. (April 9, 2021). Masters of $1.3 Trillion Fund See ESG Dominating for Decades. Bloomberg.
[21] Gomez, K., Mawhinney, T., Betts, K. (n.d.). *Welcome to Generation Z*. [PDF]. Deloitte.
[22] Francis, T., & Hoefel, F. (2018, November 12). *'True Gen': Generation Z and its implications for companies*. McKinsey & Company.
[23] Gomez et al, n.d.

This approach to consumption is reflected similarly in how Gen Z and Millennials select their place of work. A recent Cone Communications Millennial survey found that 64% of millennials will not work for a company without a strong corporate social responsibility (CSR) mandate.[24] There is an increasing expectation within the upcoming workforce that corporations will put effort towards initiatives beyond profit, many willing to take a pay cut to work for those who do.[25] It's also important to recognize that Gen Z ,in particular, requires proof. This generation does not stop at reading the mission statement, instead going further to understand where and how products are made, and demanding action to follow CSR policies.[26]

70% of Gen Z say they try to purchase from ethical companies.
64% of millennials will not work for a company without a strong corporate social responsibility (CSR) mandate.

Focusing on the field of AI, whether they are investing or consuming products, these generations expect the companies they associate with to be inclusive, sustainable and ethical. As the notion of biased and blackbox AI surfaces, it will undoubtedly be questioned. Particularly with the onset of regulation, it will be essential for AI companies to be prepared to demonstrate transparency, and a clear prioritization of social good.

Summary

AI is a technology that is already having profound impacts on society. Yet, AI does not exist separate or outside of society, but within it. AI is constructed by people using data and models, all of which exist within, and are informed, by a socio-cultural context. There are numerous examples of how AI systems have served to encode harmful power structures by further amplifying inequity, discrimination, and bias. These systemic issues can be linked to the data, models and ultimately, the people involved, as well as the business models and social systems which currently govern how AI is constructed and regulated.

Given the push from Europe's leadership on the GDPR, our own governments' moves towards creating guidance and legislation for AI technologies, and the increasing interest among both consumers and investors in social responsibility, EAAI sees opportunities in the marketplace. Therefore, it conducted primary research to ascertain the needs and attitudes of Canadian organizations currently deploying or considering the use of AI.

[24] Aziz, A. (March 7, 2020). The Power of Purpose: The Business Case for Purpose (All The Data You Were Looking For Pt. 2). *Forbes.*
[25] Cone Communications. (2016). *Millennial Employee Engagement Study.* [PDF] Concomm.
[26] Francis & Hoefel, 2018

Part 2: Market Survey Interviews

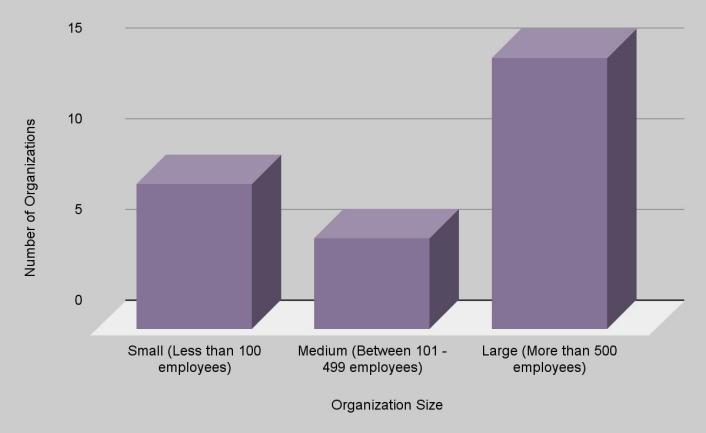
Methodology
We interviewed individuals from a number of companies in a wide variety of market sectors to provide a high-level environmental scan of the use of AI/ML in a variety of applications. We used a purposeful sampling strategy to connect with people who either worked with, or had high level knowledge of the AI/ML systems at their organizations.

Philosophically, it was decided to "go wide not deep." It was anticipated that this process would give a "snapshot in time" of the use of automated technologies in industry and government to provide insight on the market understanding of how, and if, artificial intelligence and machine learning are being used, and the eventual impact it has in creating a culture using technology to make decisions. Further, the issue of ethical alignment within the building of datasets and algorithms was considered to be of critical importance as decision-making would occur that could significantly affect people's lives.

The interviewing process included 28 individuals of diverse backgrounds in terms of gender, race and levels of responsibility around the use, or potential use, of AI within the company. Individuals ranged from board members or executives to mid-level management, and included 9 women and 6 people of colour.

INTERVIEW QUESTIONS

1. *Does your company use AI/ML? If yes, does it focus on people, including Personal Identifiable Information (PII) or industrial data?*
2. *What problem are you trying to solve with AI/ML?*
3. *What do you consider the most important risks to your company? How do you define ethical use of AI/ML?*
4. *How do you manage those risks? Who is responsible in your company for risk management? What is the impact if risks are not mitigated?*
5. *What are the current products and/or services you are using today related to AI/ML?*
6. *How do you see the use of AI/ML in the future?*
7. *What product and service needs do you anticipate in the future?*



Representation was provided from both the public and private sector, as well as non-profits. Specific market sectors included agriculture, education, finance, healthcare, infrastructure, telecommunications, and technology consulting. We spoke with 8 representatives from small businesses that have under 100 paid employees, 4 from mid-sized companies with staff of 100 to 499 and 13 individuals from large organizations with 500 or more

employees.[27] We also spoke with 3 individual technology consultants who spoke about broader sector trends. Geographic representation came mostly from Alberta, but also Ontario and British Columbia.

Analysis of Emerging Themes

The data was gathered via video conferencing interviews, and analysis of the discussions revealed a number of themes with respect to the use of data, automated systems and attitudes towards ethics and potential regulation of AI technologies.

1. Data Collection & Processing

Data collection and interpretation is commonly used throughout all the sectors interviewed. A number of companies use a variety of sources to collect data, process and provide findings based on the data. However, the source of the data is proving problematic. In other words, data is collected in very unsystematic ways across various jurisdictions and is lacking a common set of standards.

One individual provided an example of someone who was responsible for inputting data and "the person just checked off male as a gender if the name sounded male and the same with women, rather than confirming the actual gender" of the person. Her point being that bad data in any capacity can lead to bad outcomes in any kind of data processing or analysis. "Trust is the most important element for us so anything that could damage that is a major risk to both the individual organizations and the entire sector," she added.

2. Understanding of AI/ML

It appears a considerable amount of education needs to be done to assist the marketplace in understanding "artificial intelligence" and "machine learning" technologies. There also appears to be blurred understanding between general data processing, deep mining data analysis and AI/ML. Several of our interviewees made comments similar to one individual who suggested there was a definite "market for education about the tech itself, as well as its ethical use and the potential risks associated with bad data. I can also see a need for consulting services to help create guidelines and ethics codes."

> Across the board, there is understanding and concern about reputational risk and financial harm should AI/ML processes go wrong.

3. Use of AI/ML in the Market

Our findings indicate there is interest in AI/ML across all sectors included in the survey. As one technology consultant put it, AI/ML can help with the "need by every industry to gain insights, understand the market better, predict trends and analysis. That's needed in all sectors." However, it appears the AI/ML use is largely in its infancy, with financial, healthcare and telecommunications being the most advanced. Other sectors, such as agriculture and education have attempted to integrate the technology into their business models with varying degrees of success due, in part, to high costs and insufficient, or bad data. Governments at both the municipal and provincial levels are also embracing AI/ML technology in a variety of ways. Representatives from the federal government were not included in the survey, but a literature review suggests that interest is also strong at the national level as well.

[27] Innovation, Science and Economic Development Canada. (2020). *Key Small Business Statistics — 2020*. Government of Canada.

It's also worth noting that several of the companies surveyed used AI/ML systems from third party vendors, suggesting that responsible use of AI needs to be considered throughout the chain of supply as risk ultimately lies with the end user.

4. AI/ML Ethics, Governance & Risk

There appears to be considerable blurring between AI ethics and data governance, especially with regard to existing privacy and information protection legislation. However, data governance is frequently considered the first step towards responsible data use, analytics and machine learning. Across the board, there is understanding and concern about reputational risk and financial harm should AI/ML processes go wrong.

As an interviewee in a government space put it. "When creating these kinds of ML systems, broken information and the bias of developers can be a problem. If there is no good data governance system and common philosophy in approaching it, then we are just going to end up with bad data being fed by poor ideology or poor thought processes, then we're not really using the wisdom of AI/ML to help enhance life." Overall, ethical standards tended to be viewed as an issue of potential compliance rather than a proactive strategy for the responsible use of AI.

5. AI/ML Ethics in the Public vs. Private vs. Government Sectors

It was found that there is a considerable gap between public and private companies when it comes to managing risk, understanding AI/ML, respecting the importance of systematizing and reporting on ethics and ethical behaviour, awareness of looming legislation and regulation on certain aspects of ethical sustainability both within a company, and also within the supply chain of that company.

Public Companies tend to have a strict reporting structure to shareholders and regulators, whereas private companies (although aware of the importance of ethical standards) are not held accountable in the same way. One publicly held company interviewed had an entire department with a staff of 25 people devoted to data privacy and trust. The same company also developed a "Bias Model" algorithm that is incorporated into the design processes of every data analytics project, and an internal audit process to ensure it falls within their ethical guidelines. "These questions have come out of the boardroom and entered the minds of the general public and once that happens, it's only a matter of time before the boardrooms will have to take a serious look at things," said the interviewee, who is a former employee of the company.

> For many of the small and midsize companies, education about the technology itself, ethical use and regulation appears to be the primary need.

Similarly, governments also seem to be leading the way when it comes to both the use and ethical considerations around AI/ML, obviously due to greater standards of trust and accountability to taxpayers. For example, one large public sector organization recently created a new advisory position to oversee the responsible use of data and AI/ML tools and systems. The position "came about when they realized that the FOIP [Freedom of Information and Protection of Privacy Act] department wasn't really equipped to deal with AI/ML and data governance issues, and also a growing understanding of the need for ethical oversight as malfeasance such as the Cambridge Analytica scandal and others came to light," said the advisor. That

individual has also been approached by another large public institution to help create a similar position as part of its organization-wide AI strategy.

6. Small/Medium vs. Large Companies on Ethics Services/Tools
The level of sophistication in understanding, awareness, current and/or future use of AI/ML and the implications of ethically aligning systems and people to maintain a strong reputation in the market is noticeably different depending on the size of the company. Therefore, the opportunities presented vary.

For many of the small and midsize companies, education about the technology itself, ethical use and regulation appears to be the primary need. Some midsize companies also appear open to outsourcing consulting services to build ethics guidelines, and meeting compliance requirements for existing and upcoming regulations around data governance and AI. Many of the larger organizations are already addressing many of these needs with internal teams or big consulting firms. However, there may be niche opportunities for employee education and training. Another opportunity with some large companies and governmental organizations may be in consulting around AI specifically with respect to bias and inclusion in both hiring and data collection, as much of the current focus is around privacy and cybersecurity. But as one individual pointed out, "Because of high levels of public scrutiny and accountability, ethics consulting services would have to be credible and sustainable."

7. Regulation & Legislation (Proactive vs. Reactive)
It was interesting to discover that there is a tendency to be more reactive than proactive when it comes to ethics and compliance to standards, with regard to AI/ML specifically. One could conclude that where there is no standard, there is wiggle room to react as required. But it is important to note that many of the individuals we spoke with felt that existing legislation that protects data and privacy went some way towards regulating the use of this technology from a cyber and information security perspective.

For example, interviewees in healthcare, charitable and financial sectors all pointed to "rigorous" provincial and federal regulations, including the need for Privacy Impact Assessments (PIA). In healthcare, it was unanimously agreed that most human research and development projects must undergo academic Research Ethics Board (REB) reviews. So while almost all the interview subjects agreed that ethical and responsible use of AI/ML was important, many of the small to medium enterprises (SMEs) seemed to be waiting for legislative standards to be put in place before addressing the ethical issues specific to AI/ML within their own organizations.

8. Market Sectors with High Potential
Based on the market research interviews, the sectors with the highest potential for ethical AI/ML services and/or tools appear to be healthcare, financial, telecommunications, and the government sector. This conclusion is drawn from the more widespread and advanced usage of the technology in those sectors, and the higher levels of data governance already in place. There is some indication of interest in preparing for eventual regulation of AI/ML technologies through education and consulting. There even appears to be some potential for some sort of certification process if and when there are standards to be measured against,

Sectors with the highest potential for ethical AI/ML services and/or tools appear to be healthcare, financial, telecommunications, and government.

similar to International Organization for Standardization (ISO), Inter-Control Center Communications Protocol (ICCP) certifications. The caveat here is that the certification would need to be widely recognized as credible and sustainable.

Organizations involved in charitable fundraising also show potential, although AI/ML technologies seem to be at a fairly early stage in their use. An interviewee who fundraises for a university said they would be interested in 3rd party services (education, consulting, tools) "because as a university we want to be at the forefront of any new innovation that comes along that helps serve our students better…[we're] always looking for certifications in order to be transparent and to do business effectively - if it's something being used by other institutions in Alberta and Canada."

If, and when, regulatory compliance specific to automated systems becomes a requirement, there also seems to be interest among SMEs for an off-the-shelf testing tool that can be brought in house and deployed by the individual companies. This seems of far greater interest than an external auditing process. As one interviewee said, "Audits would be useful, but nobody likes to be audited."

9. Market Sectors with Low Potential

Private sector SMEs appear to show the lowest potential at this time in market maturity since compliance appears to be the driving force that would move them toward resource investment in addressing the ethical use of AI. "Ethics services and tools are probably needed but cost may be prohibitive unless compliance is required," said an executive from a company with 5 employees.

Sectors that focus on industrial data rather than collecting human information are the least likely to need or want 3rd party AI/ML ethics services or tools. Due to early indications of this we chose not to continue surveying sectors such as manufacturing, agriculture, and oil and gas. As a participant from a company in the infrastructure sector noted, "because we collect machine data rather than PII, we don't really see the need for ethics testing services at this time. But that may change if regulatory compliance is needed."

Limitations of the Study

Limitations of this study include, but are not limited to the fairly small sample size and geographic focus. We conducted interviews with 28 individuals who were mostly located in Edmonton, but there was some representation from Calgary, Ontario and British Columbia. Also, given the globalized nature of the technology sector, it's worth noting that opportunities internationally, but scanning those marketplaces were beyond the scope of our research at this time.

Additionally, the purposeful nature of our outreach limited our sample and findings to only companies and sectors that are already invested in, or considering using AI/ML in the near future. There could be sectors and companies who are interested in the technology but not yet in a position to discuss implementation and its associated risks and benefits from an ethical perspective. But growth in the use of automated systems in most sectors is expected. As one technology consultant put it, "I don't think everyone fully understands the power of artificial intelligence yet, so hopefully we can do some things to make it easier to understand what we can unlock here, and hopefully the legislation and policy isn't too far behind."

Conclusion & Opportunities

The use of artificial intelligence is growing rapidly across most industry and government sectors, as is the evidence and awareness of both its positive and negative impacts. This, along with the ongoing development of regulatory frameworks, and the growing interest in socially responsible business practices, suggests that organizations should pay attention to the ethical dimensions of AI.

Education

As a first step, organizations need education and training about ethical considerations such as bias, diversity and inclusion, privacy, safety and security in the context of AI design, development and deployment. Organizations who market their services internationally, particularly those who wish to do business in Europe, need to understand regulatory regimes related to responsible data collection, its use in training automated systems and data governance. Organizations can also benefit from understanding how ethical practices built into the design of AI can provide a competitive advantage and help attract and retain socially responsible investors, employees and consumers.

As organizations move forward to embrace the benefits of using AI-enabled technologies, they should also be aware of, and proactively, addressing the risks these technologies bring.

AI Ethics Readiness

The market also indicates some appetite to start building and creating ethical guidelines and governance structures, as well as policies, processes and procedures. Companies who are planning to or currently developing AI and/or deploying such systems throughout their supply chain will at some point, need to address these ethical challenges. Third party, independent advice is key to providing meaningful consultation with respect to ethical guidance. Independence ensures that the work is credible and objective, not subject to misaligned incentives by companies whose primary concern is to develop and deploy AI solutions.

As organizations move forward to embrace the benefits of using AI-enabled technologies, they should also be aware of, and proactively, addressing the risks these technologies bring. There is an opportunity for progressive organizations to start the work of AI ethics now, to take a leadership position and to set their own path rather than waiting to be told “what not to do” by regulators. However, even those who wait, will eventually be called upon to address these issues, and that day may be sooner than some expect.

Acknowledgements

This report builds on work conducted by Katrina Ingram and her masters research on applied ethics in AI. Thank you to the following contributors who made this work possible: Kaly Vittala, Remington Irwin, Hero Laird and the Digital Law & Innovation Society, Anna Bubel of Another Way Consulting, Bruce Alton of A-Partners Ventures Inc. and Vi Becker Consulting. Special thanks to the individuals who participated in our research study. We greatly appreciate your time and perspectives.

About Ethically Aligned AI

Ethically Aligned AI help organizations build better AI systems and help consumers to make sound choices in selecting trustworthy AI-enabled products or services.

We are on a mission to help companies understand the risks and benefits of using AI in their business practices; and to advise them on how to improve these systems to ensure responsible use, thus reducing potential risk harm to their reputations and harm to their finances. We offer a range of educational and consulting services as well as technical tools to assist our clients.

Download our FREE ethics toolkit, read about current issues in AI ethics or find out more about our education and consulting services at ethicallyalignedai.com or contact us at hello@ethicallyalignedai.com.