AI Frontiers: Where We are Today and the Risks and Benefits of an AI Enabled Future

Author

Mike Durland - Special Advisor, Global Risk Institue
Matthew Killi - Vice President, DeepLearni.ng

Mr. Mike Durland and Mr. Matthew Killi are independent contributors to the Global Risk Institute. They are solely responsible for the content of the article. Mike and Matthew's biographies are at the end of the article.

GLOBAL RISK INSTITUTE

EDITOR'S NOTE

"AI Frontiers: Where We are Today and the Risks and Benefits of an AI Enabled Future" is the next in a series of Global Risk Institute expert papers on the evolving world of machine learning. The authors Michael Durland and Matthew Killi are also preparing a GRI paper on the topic of the impact of AI on financial services for the fall of 2017.

EXECUTIVE SUMMARY

"Artificial Intelligence" is a very powerful narrative. In fact, today, many leading thinkers envision a future where machines surpass humans in intelligence. Many of those individuals worry about the abuses of AI, and although they don't dispute the potential good, they dwell more on the potential bad. Others have a more constructive imagination of the future. They see AI more as a powerful set of tools, with the potential to significantly augment human productivity. They see the risk of "singularity" as over hyped and distracting. In Part One of our two part series, we assess of the potential near term risks and benefits of Artificial Intelligence. Later, in Part Two, we explore how AI is expected to impact Financial Services and what specific use-cases we expect to see over the next 1-3 years.

The paper begins by differentiating between the concept of **automation** and **innovation**. Here we define **automation** as the use of technology as a **substitute** for an existing process, for example one that is carried out using human labour, and **innovation**, as the use of technology to **augment** human productivity, enabling humans to do things they could not do before. AI has the potential to both automate and innovate. Although subtle, this distinction is important. Automation and innovation have very different potential implications for both the future of labour and human progress.

We discuss the **semantics** behind AI, and how the phrase "**Artificial Intelligence**" creates a powerful fictional image that serves to both inspire innovation and evoke fear. The inspiration is important. Fictional narratives such as "Artificial Intelligence" are a vital component of driving forward human progress. Yet, at times this particular narrative acts more as a negative, evoking fears that to date seem mostly unfounded.

We then discuss the current optimism surrounding AI today. We provide a brief primer on the most important technology underlying AI today, machine learning. We contrast the various forms of learning paradigms: **supervised, unsupervised** and **reinforcement learning**. These concepts are important because they are at the center of an emerging issue in AI, namely, who is accountable for the actions of an AI and how the developer must take a deeper role in curation.

Following this section, we introduce the concept of **Narrow AI, General AI** and **Artificial Super Intelligence**. In Narrow AI, machines perform a

narrow set of tasks applied to a narrowly defined problem. Narrow AI's can be integrated to produce highly powerful applications. An example of this is the autonomous vehicle. General AI refers to a machine that is capable of performing the broad array of intellectual tasks of a human. In General AI, machines have human-like cognitive abilities and are capable of reasoning, making decisions, learning and communicating in natural language, and are able to operate in an open system. Creating General AI is a much different and more difficult challenge than creating Narrow AI. Artificial Super Intelligence, refers to a computer that is "smarter than a human", a machine that is capable of performing more than the broad array of intellectual tasks of a human. In this fictional form of machine intelligence, the computer would have the cognitive ability to outperform human brains across a large number of disciplines, even possessing scientific creativity and social skills. Today, all forms of artificial intelligence are instances of Narrow AI. In a world of Narrow AI, we can eliminate from our concern the notion of AI as an existential threat and instead focus on the impact that Narrow AI is likely to have on the world in which we live in today.

We begin our assessment of the risks and benefits of AI by introducing four key factors that are likely to shape the future of AI in the near term: **1) the identification and application of suitable use cases, 2) the access to large data sets, 3) the scarcity of talent,** and **4) the lack of platform technologies**. In other words, in order to successfully create AI today you must identify a suitable problem, have access to the data required to train the AI to solve the problem you identify, and have access to the talent and the tools required to developed the AI.

We build three broad scenarios that help us think about the future of AI: "**AI Winter**", "**Winner Takes All**", and "**Collaborative AI**". These scenarios are used to assess the potential benefits and risks associated with AI in the near future. We consider two potential benefits: **an increase in human productivity and efficiency,** and **an increase in our ability to drive future innovation**. The later benefit is a broad category but is meant to capture the tremendous potential for AI to drive future scientific innovation. We consider six potential risks: scope erosion, unemployment, wealth inequality, **the exploitation of data, black box vulnerability and the creation of new system risk**. The results of this scenario analysis are summarized in Table 1.

The "AI Winter" scenario, in which AI does not live up to its hype, provides the least benefits. Although a "full" AI winter deemed not likely, we do believe, that given the considerable hype surrounding AI, the probability of some type of cooling off period for AI is likely in the near future.

The "Winner Takes All" scenario, in which a number of companies exploit the potential of AI to achieve an early monopoly position, provides moderate benefits and material risks. We believe these risks are tolerable, and indeed likely necessary for society to further the development of important innovations in AI. We believe that such innovations will increase the potential for material long term benefits. However, we do believe this scenario should grab our attention. The current discourse of disruption and creative destruction must be understood in the context of a tenuous balance. The outcome we want for society is not disruption but progress. One way we can achieve this objective is the democratization of AI.

The "Collaborative AI" attempts to assess what needs to occur in order for society to maximize the future benefit of AI while minimizing its future risks. To achieve this objective it is helpful for us to perceive artificial intelligence, not as the automation of human cognition, but rather as an innovation capable of augmenting human productivity and efficiency. In this scenario, AI is not perceived as a substitute for human intelligence, in the possession of a concentrated set of large corporations but as a complement to human intelligence available to the masses. This scenario can be defined as the state in which we have successfully begun the democratization of the benefits of AI.

Automation vs Innovation

From the dawn of the industrial revolution, machines have improved human productivity and efficiency by automating existing processes, and innovating new ones. Automation and innovation in this context are different concepts. Automation can be thought of as a substitute for an existing process, for example one that is carried out using human labour. Innovation, on the other hand, can be thought of as the creation of new technologies that augment human productivity, enabling humans to do things they could not do before. During the industrial revolution automation and innovation were constrained

to physical processes. More recently, the digital revolution has brought exciting new potential. Advances in computer and information technology have enabled society to imagine the potential of automation and innovation in the realm of *cognitive processes*.

The Semantics

In the mid 1950's, John McCarthy, an American computer scientist, invited leading researchers from a broad array of disciplines to Dartmouth, New Hampshire to assess the rapid evolution of computer science, and whether someday computers would have the resources to be as intelligent as human beings [1]. This concept was so new to the human imagination that McCarthy felt it deserved a powerful name: "Artificial Intelligence".

The label Artificial Intelligence (AI) is powerful. Semantics are important. The science fiction fantasy of AI has served to shape our aspirations and inspire our efforts. Talent has been drawn into the discipline of AI because it captures their imagination. These individuals are inspired to be part of the construction of the future, one in which AI is a central feature.

Words help create images, and from the very onset of this label, these fantasies, and the fears associated with them, have greatly influenced our perceptions of artificial intelligence. For many, there is a lot to fear about AI. If we automate human cognition, do we not, as humans, now compete with machines for employment? And who controls the machine? Who benefits from the creation of AI? This lack of controllability and perceived participation in benefits reinforces the notion that AI is a threat, rather than an innovation for augmenting human productivity.

Why Now?

More than 60 years after that inaugural gathering in Dartmouth, attention grabbing headlines like the *New York Times* "The Great AI Awakening" [2] suggest that great progress has been made and that today we at the cusp of a watershed moment in Artificial Intelligence. The combination of vast volumes of data, unprecedented processing power, and increasingly sophisticated algorithms that enable machines to perceive images and sounds, and to discern complex patterns [3], leads many proponents to believe that AI is now poised to fundamentally transform human life.

Not everyone shares this view. Skeptics like to point out that during the past 60 years computer scientists have regularly claimed to be on the cusp of a breakthrough in AI, only to disappoint. In fact, these cycles of boom and bust became so consistent in the past, that they were known as "AI summers" and "AI winters" [4]. So what makes us think that the current hype will not fall to the same fate?

Historically, there have been three broad impediments that have limited the success of AI: computational power, access to large data sets, and the development of effective computational algorithms. Today, we have an abundance of both data and economical high-speed processing power – both of which are growing at an extraordinary rate. In fact, some 90% of all the data in the world today was collected in the past 2 years [5]. Now in the era of big data and cheap processing power, two of AI's biggest constraints are no longer binding.

While the removal of these two constraints is clearly a necessary condition for AI to evolve, their removal alone is not sufficient for AI to thrive. The promise of AI, and the hope that it will someday transform human progress, now rests of the evolution of algorithms, specifically machine learning algorithms.

Primer on machine learning

Artificial Intelligence is often used to describe the appearance of human intelligence exhibited by computers. In the "early years", AI was primarily based on rules-based programs that delivered rudimentary displays of 'intelligence' in a very narrow or specific context. Early progress was limited because real world problems are far too complex to program using a rules-based approach. As a result, these so-called "expert systems" failed to achieve intelligence in any but the narrowest definition. To advance AI, we must be able "create" intelligence without the need to enumerate the complex rules or concepts that govern intelligent behaviour. This is precisely the goal of machine learning.

Machine learning is the study of computer programs that can be trained to learn patterns, rules, and concepts from data, yielding trained models that describe the domain from which the data originated. Breakthroughs in subfields of machine learning like deep learning have recently allowed extremely complex models to be constructed, including models of written and spoken language, of our visual and

acoustic world, and of rational and irrational human behaviour. Often times, the ultimate goal of these complex models is to help us make more accurate predictions about the world and to make decisions with the best chances of success. Such models will receive input data about a domain (say, the films a person has watched in the past) and weigh the inputs to make useful predictions (the probability of the person enjoying a different film in the future). By endowing computers with the ability to learn complex patterns from today's abundance of data, computers are beginning to mimic humans' ability to learn from our environment, propelling us towards a new era of artificial intelligence.

There are many machine learning models, and there is usually more than one way to train each type of model. Moreover, the number of models and training algorithms are continually being discovered and at an increasing rate. Needless to say, each pairing of a model and training algorithm has its advantages and disadvantages and the pair selected to solve a particular problem will depend on many factors:

- size and richness of the data sets
- available computing resources
- model's ability to scale with increasing size of data set and computing resources
- type of feedback signal used during training
- complexity of the task
- performance requirements of the trained model
- speed and responsiveness of the trained model
- risks associated with mistakes and failures
- stakeholder risk tolerance
- cost of the end-to-end development of the model
- requirements on interpretability of the model.

Out of the factors listed above, the type of training feedback signal is one of the first considerations to be made, since it can significantly focus the model search and inform the feasibility of the project. There are three basic categories of feedback signals [6]: In the case of *supervised learning*, the training data consists of "labeled" examples. During training, the computer ingests a large number of examples that contain both input data as well as the correct answer. Over time, the algorithm finds patterns in the input data to help it predict the correct answer. In the case of *unsupervised learning*, there is no prior knowledge of a correct answer. The learning algorithm must draw inferences from the data through techniques such as Clustering, and Hidden Markov Models. In the case of *reinforcement learning*, the machine interacts directly with an environment and is rewarded when it achieves a desirable outcome and punished when the outcome is negative. Over time, the machine adjusts its decision making to maximize the rewards it receives.

Detailing these feedback mechanisms highlights the shifting role of the AI developer. We no longer need to explicitly write down the rules and decisions that an AI needs follow. Instead, our primary role is to curate the data that the machine will learn from and decide on what model and learning algorithm is most optimal for the problem at hand.

Where we are today

So, where exactly are we today in the evolution of AI? To answer this we must first distinguish between three broad stages of development: Narrow AI, General AI and Artificial Super Intelligence. On doing so, we can have a better sense of not only where we are, but where we are going. This will inform our discussion on the imminent risks and benefits, and help avoid overly hypothetical scenarios.

Narrow AI, sometimes referred to as Weak AI, is AI that specializes in one area. These machines perform a very narrow set of tasks that apply to a narrowly defined problem. Narrow AI's can be integrated to produce highly powerful applications.

General AI, sometimes referred to as Strong AI, or Human-Level AI, refers to a machine that is capable of performing the same broad array of intellectual tasks as a human. With General AI, machines have human-like cognitive abilities and are capable of reasoning, making decisions, learning and communicating in natural language, and are able to operate in an open system. Obviously, creating General AI is a much different and more difficult challenge than creating Narrow AI.

Artificial Super Intelligence, refers to a computer that is "smarter than a human", a machine that is capable of performing more than the broad array of intellectual tasks of a human. In this imaginary form of machine intelligence, the computer would have

the cognitive ability to outperform human brains across a large number of disciplines, even possessing scientific creativity and social skills.

Today, we are still in the early stages of AI development – all forms of artificial intelligence are instances of Narrow AI. This may sound disappointing, but in reality, this is not the case. Narrow AI, in and of itself, presents tremendous potential. For example, the automobiles we drive today are full of Narrow AI systems, from the computer that figures out when the anti-lock brakes should kick into the computer that tunes the parameters of the fuel injection systems. Our mobile phones are filled with little use cases of Narrow AI in the form of our favorite apps. The Google search is another great example of the application of Narrow AI, enabling the ranking and matching of webpages to your particular need or interest.

Sophisticated Narrow AI systems are now widely used across a variety of sectors and industries. Many of these applications combine several Narrow AI algorithms into an integrated system. This approach has resulted in the creation of a number of increasingly sophisticated applications of Narrow AI that we might mistake for General AI. Perhaps the best example being the autonomous vehicle which contains a variety of Narrow AI systems that allow it to perceive and react to the world around it.

The recent visible progress made in AI technologies, such as the autonomous automobile, has spurred the emergence of narratives prophesizing the obsolescence of not only jobs, but also the human race in itself. This discourse is not limited to headline grabbing articles but includes industry leaders as well. Elon Musk stated, “With artificial intelligence we are summoning the demons” and “If I had to guess at what our biggest existential threat is, it’s probably [artificial intelligence]”. Other predominate figures, such as Stephen Hawking and Bill Gates, agree.

Although it is tempting to delve into the realm of the science fiction of AI, it is becoming increasingly important for decision makers, business people, and regulators to take a step back. In a world of Narrow AI, we can eliminate from our concern the notion of AI as an existential threat and instead focus on the impact that Narrow AI is likely to have on the world we live in today.

Key factors that will shape the future of AI

How fast AI will progress and become a commercial reality will depend on several key factors. For the purposes of this paper, we consider four of these factors: 1) the identification and application of suitable use cases, 2) the access to large data sets, 3) the scarcity of talent, and 4) the lack of platform technologies. This is to say, that in order to successfully create AI today, you must identify a suitable problem, have access to the data required to train the AI to solve the problem you identify, and have access to the talent and the tools required to developed the AI.

The identification and application of suitable use cases

Somewhat harshly, Box and Draper said, “Essentially, all models are wrong, but some are useful” [7]. Models, like humans, are imperfect. And indeed, many are useful. Models are abstractions of our complex, uncertain and ambiguous world. The more complex, uncertain or ambiguous the use case, the more our models will be imperfect. And this is true for AI. The simpler the problem, the better the training data, the more stable the training data, the better the performance of the AI.

Today’s highly inflated expectations for AI probably exceed its near-term potential. As a result, it is reasonable to expect a period of disillusionment regarding AI that could inhibit its development. The mitigation for this is awareness and honesty. We need to be honest about what AI can do today and be less focused on the fictional expectations of what it may be able to do many years hence. Taking a more pragmatic and critical lens will help identify suitable use-cases and help quell unrealistic expectations.

Access to large data sets

Data is the lifeblood of AI. Training data is a necessary condition for the development of AI. As a result, it is reasonable to expect that in the near term applications of AI will evolve where the data resides. Today, much of the data used for the development of AI is our own personal information, data collected from our own personal transactions and countless interactions with technology. This may or may not be the best use cases for AI from a societal perspective, but it is where the data resides, and it is where many of the near term applications are likely to evolve.

While concerns over access and use of personal information are not new, AI has the potential to raise this issue to an entirely new level. Building privacy and security into the design of AIs is vitally important. The current complacency in the way we think about our personal information is subject to change. Without a robust privacy and security strategy, there is a high risk that consumers and businesses will be less willing to share data, and reactive regulation changes could drastically limit the collection, use and retention of data.

To assuage privacy concerns, we must recognize and address the fact machine learning algorithms have the potential to expose individuals to a new set of privacy risks. One issue arises from the "black-box" nature of certain algorithms. When we first create an algorithm, it is a blank slate waiting to learn or be trained. We then feed the algorithm with volumes of "training data" where it begins to find patterns in the data that help it make accurate predictions. As an example, it could identify customer micro-segments that have a higher propensity to buy certain products – just as Target's analytics team determined that women who buy unscented lotion are likely to be pregnant, and are likely to buy baby products [8]. The difference here is that the machine learning algorithm does this on its own and the correlations it discovers are opaque. The user simply inputs raw data into the black-box, and the algorithm provides the answer. It is difficult to determine if the AI is exploiting unethical patterns, which results in unwanted actions.

A second and related risk comes from the training data itself. Machine learning algorithms are blind to the world at large, and only learn from the data they are provided. More often than not, the training data from which they learn is skewed by either direct or indirect biases. Perhaps the most alarming demonstration of the perils of biased data was in the claim that AI could infer guilt based on the facial patterns of a person [9]. Numerous other incidences of biased data – ranging from the flawed polls of the US election, to the racist rants of Microsoft's AI bot – demonstrate how important it is to abandon the notion that computers are objective.

Once again, we see the role of the AI developer shifting towards curation. They must be accountable for understanding the biases in the data, assessing the potential for exploitive correlations in the algorithms, ensuring the limits of the AI are understood, and informing the appropriate governance model to oversee the AI. Without prudent curation, AI has the potential to kill the golden goose that is "Big Data" [10] by instigating backlash. It could drive us down a path where technologies like blockchain are developed to restrict access to our private data and use its use.

Scarcity of talent and the potential effect of AI on employment

Data may be necessary as a starting point to the development of AI, but equally important is access to talent. To successfully develop AI, organizations must have the ability to recruit and retain highly educated technical talent (individuals with machine learning expertise) and deep subject matter experts (individuals that have significant mastery of the use case for which the AI is being designed). This dual requirement potentially creates a significant war for talent, and a large barrier for entry for many companies, who may have access to large data sets and an interest in developing AI applications, but lack the culture to attract and retain the requisite talent.

We believe that it is highly likely that there will be a significant shortage in the supply of technical talent required to satisfy the growing demand for high-skilled data scientists. Technology giants invested $10.8B in AI in 2015, which largely went to recruiting and aqui-hires [11,12]. So extreme are the talent wars, that aqui-hire go for between $5-10M per employee. Today, universities are struggling to retain their talented academics that have technical expertise. The McKinsey Global Institute predicts there will be an incremental need for up-to 500,000 data scientist by 2025 [13]. Other sources cite that number to be significantly higher.

It also seems likely that there will be a growing need for low-skilled labour to support AI technologies and Big Data. As the *Harvard Business Review* report "The Humans Working Behind the AI Curtain" [14], there is a largely unnoticed labour force working under the veil of AI platforms. This labour force will be critical for maintaining the integrity of data, and facilitating the effectiveness and efficiency of AI.

In addition, it is likely that we will see a growing demand for "Business Translators" who have both technical competency and functional expertise, ranging from specific verticals to risk, privacy and regulatory. Machine learning algorithms rarely provide value on their own, but need humans to

distill actionable insights. Competitive advantage will become more and more dependent on the ability of business to attract and grow this talent base. Without this important bridge to business value, investment in AI will deteriorate.

On the flip side of increased demand for specialized talented, there is the potential for loss of low skilled jobs. Reminiscent to the early age of robotics, there is fear and anxiety surrounding the impact of AI on employment. The popular belief is that once AI development reaches a critical mass, we will see colossal job losses driven by lower cost of automation and superior abilities of machines. According to McKinsey, up to 40% of *workplace tasks* can be automated through robotics and artificial intelligence [15].

This statistic brings into focus the important debate regarding the risk of labour disruption, job quality, and productivity. We have seen these shifts throughout history, and while still debated, the consensus is that new job creation, higher wages and lower costs will in the least mitigate temporary loss of jobs. For better or worse, the threat of technological unemployment could create policies that work to inhibit the growth of AI.

Lack of platform technologies that are required to enable the democratization of AI

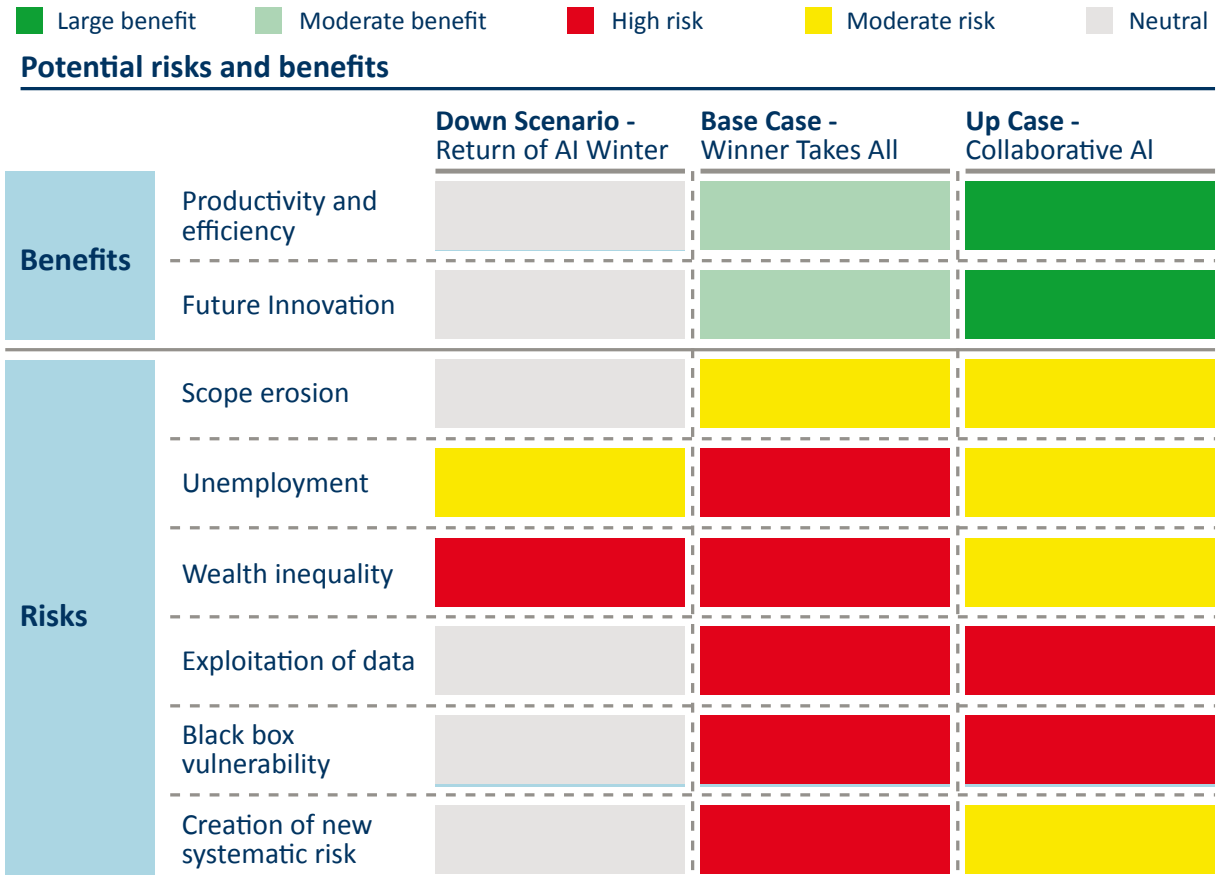
Often times when we think of AI, we think of a large-scale artificial intelligence system, for example the autonomous car. Yet, the reality is that entities, large and small, can and should benefit from the application of AI to their unique problems. For AI to reach its potential it must be democratized. A great example in our past is the creation of the spreadsheet. The spreadsheet democratized computer programming. It made programming accessible to the masses. Today, spreadsheets are used by hundreds of millions of people around the world, transforming productivity and efficiency for millions of unique use cases.

Although open source libraries such as TensorFlow are beginning to democratize AI, they remain complex. In addition, most open source algorithms are custom built for a very specific use case and data structure. And even then, these off-the-shelf solutions – across computer vision, voice recognition, sentiment analysis, search algorithms, natural language processing, and pattern recognition – rarely work within the specific business context. A massive amounts of effort are required to wrangle data, and radically customize the software. Even with a viable model in-hand, the process of deploying the solution into the production environment is far from trivial. We believe that the widespread adoption of AI is very much dependent on the existence of end-to-end platform technologies that streamline the development process.

The Risk and Reward of AI

The objective of the remainder of this paper is to provide a perspective on the future risk and reward of Narrow AI. We build three broad scenarios that help us think about the future of AI: “AI Winter”, “Winner Takes All”, and “Collaborative AI”. These scenarios are used to assess the potential benefits and risks associated with AI in the near future. We consider two potential benefits: an increase in human productivity and efficiency, and an increase in our ability to drive future innovation. The later benefit is a broad category but is meant to capture the tremendous potential for AI to drive future scientific innovation. We consider six potential risks: scope erosion, unemployment, wealth inequality, the exploitation of data, black box vulnerability and the creation of new system risk. The results of this scenario analysis are summarized in Table 1.

Table 1: Risk Assessment / Scenario Analysis



The Down Case - A Return of the AIWinter

In the first scenario, AI fails to live up to its excess hype and we have "a return of the AI winter". We believe a mild AI winter is moderately likely to eventuate in the near term, but that the winter will be shorted lived. We believe several forces make this scenario moderately likely.

First and foremost is the prevalence today of excess hype for AI. AI is an extraordinarily high potential technology but it will take time, stewardship and lots of misuses and misapplications for it to reach its potential.

Second, it is highly likely that in the next several years significant financial resources will be spent on unsuitable use cases for AI. Some of those use cases will be too ambiguous or uncertain in nature for AI to be suitable. Others will lack sufficient data to derive satisfactory results. Some will be applications that we would expect to achieve good results but as a result of still underdeveloped algorithms may not be primed for success.

Third, we believe that the use cases will follow the data. As a result, there will likely be a significant focus of personal transaction and engagement data sets. While there is tremendous potential for AI to provide valuable insights, drive loyalty and improve customer service, there is also significant potential for the creation of socially wasteful innovation, exploitation of data and user backlash.

In terms of risks and benefits, this scenario would likely still produce noticeable increases in human productivity and efficiency, and some important advances in human innovation, as a number of AI's become successfully commercialized. In addition, we

believe this scenario provides elevated risk of data exploitation as people seek for applications of AI and focus their time and effort on personal data to achieve commercial success. We also believe there is an elevated risk of black box vulnerability, especially in cases where AI's are built for non suitable use cases. We also think this potential creates an elevated risk of scope reduction, where misguided attempts at commercializing certain AI's would be shoe horned into working by accepting a material loss in scope. We believe that this scenario would have minimal effect on employment and wealth inequality. The creation of new systemic risks would be highly dependant on which use cases were successfully applied to AI.

The mitigation for this scenario is environment shaping. We need transparency around what the strengths and weaknesses of are of AI. As we mentioned earlier, the narrative of AI is inspiring. Significant interest has been drawn to AI because it captures our imagination. However, for AI to achieve its promise, we need a healthy dose of honesty, transparency and pragmatism. Today we are all inspired. However, at this stage in the cycle we might be better served by down playing the hype and focusing on the substance.

The Base Case – Winner Takes All

The second scenario, in which new concentrations of power are created in the race to develop and monetize the value of AI, is labeled "winner takes all". In this scenario, the near term development of AI is dominated by a handful of giants, who have unprecedented access to data and the ability to recruit, retain and indeed monopolize a large portion of the supply of technical talent in AI. These giants focus on a narrow set of use cases that support their own unique business interests.

Today, giants such as Google and Amazon, have the ability, the resources and the know how to make transformational advances in AI, producing substantial innovation with the potential to disrupt and destroy economic value across a variety of traditional sectors. We believe that the likelihood of this scenario happening is material in the near term.

This scenario is likely to produce benefits as new technologies make us more productive and efficient. Because the use cases will potentially be more restrictive in breadth for this scenario than in the up case scenario, we rank benefits as moderate. However, we believe that it is highly likely that the significant commitment of resources in this "winner take all" competition will be a driving force in the near term for much transformational innovation for AI. In this regard, this scenario has the potential for very significant innovation benefit which we rate high.

The "Winner Takes All" scenario has the potential to produce significant disruption. This potential disruption would be felt more by some sectors of the economy than others. Highly vulnerable areas include retail, transportation and financial services. We rate this scenario as having a high risk of negative employment impacts, increased wealth inequality, high probability of data exploitation and black box vulnerability and and the creation of new systemic risks.

We believe this risk is tolerable, and indeed likely necessary for society to further the development of important innovations in AI. Such innovations will hopefully increase the potential for material long term benefits. However, this scenario should grab our attention. The current discourse of disruption and creative destruction must be understood in the context of a tenuous balance. Obviously, the outcome we want for society is not disruption but progress. One way we can achieve this objective is through the democratization of AI.

The Up Case – Collaborative AI

In the third scenario, labelled as "Collaborative AI", we assess the democratization of AI. In this scenario, the industry adopts the mindset that power of AI is not solely in its ability to automate human cognition, but also in its ability to augment human productivity. Here, AI is not perceived as a substitute for human intelligence, controlled by a small number of large corporations but as a complement to human intelligence available to the masses. This scenario can be defined as the state in which the benefits of AI have been democratized.

In this scenario, not only will companies all around the world, large and small, be able to develop and deploy their own AI to help drive their own productivity and efficiency, but across organizations large and small, AI becomes a collaborative intelligence partner, specializing in certain tasks that it can perform well, while leaving the human partners to focus on tasks that

they perform well. Tasks will be divided based on areas of specialization. The advantage of having a collaborative intelligence partner cannot be overstated. AI's are better at certain tasks than humans. Humans are better at certain tasks than AI's [16]. Thus, AI is not a substitute concept but instead is seen as an important component of diverse high performance teams [17]. AI does the analytical heavy lifting, studying large swaths of data, keeping a real time pulse on the business by assessing customer sentiment and changes in organizational health, helping predict potential issues before they happen.

This benefit will not be constrained to corporations. In this scenario, a variety of Narrow AI applications become highly integrated with our lives, providing us with "partners" that enable us to share the burden of many tasks that would have once fallen solely on our own shoulders. Here "democratization" refers not to the development of AI, but instead to the use of AI. The use of an autonomous car is an excellent example of the democratization of AI.

In this scenario, the important factor is the evolution of platform technologies that ease the development of simple and highly customized AI. In the base case scenario, AI's are developed for us by someone else. Large entities control the talent, the know how and the data required to develop AI. As a result, they control the use cases. They control the data, which is their asset used for their benefit, not necessarily ours.

Platform technologies make AI development accessible to a broader set of users. They diminish the war for talent, by providing companies with simple to use tools that enable their own teams to develop their own AI applications. With platform technologies, AI's are trained by us for us. We choose the use case. We use data to enhance our own productivity.

In this scenario, we assess the benefits to human productivity, efficiency and innovation to be high. We see this scenario as much more favourable to employment, as AI does not cannibalize a current stock of employment opportunities but instead alters and enhances the roles of humans and machines within an organization. However, this is still likely to create an adjustment to employment that will be negative for some and positive for others. We therefore rate the risk of employment and wealth inequality as moderate.

We believe that this scenario has the potential to inspire a movement to democratize data, which would help minimize the risk of exploitation of data. We therefore assess the risk of exploitation of data as moderate.

Closing Thoughts

New advancements in machine learning have caught the attention and imagination of leading innovators and as a result, AI is rising to the top of the agenda for business leaders. Although it is still premature to assess whether or not we are on the cusp of the radical transformation foretold by science-fiction writers and visionaries, we do appear to be at a critical point in the evolution of AI. Today we have the opportunity to influence the discourse and narrative surrounding artificial intelligence, which will ultimately shape its development and integration into society. By taking a more critical and less fantastical view, we can improve the odds of fostering its benefits.

The adoption of a more balanced perspective will allow us to better address four key factors that could constrain AI's advancement. First, we must strive to be honest and transparent on what use cases are well suited to AI in order to avoid hype and disillusionment. Second, we must recognize that AI will create a paradigm shift in the use of personal information. As such, it is essential to keep privacy and security paramount in the design of AIs to ensure we do not kill the golden goose that is "Big Data". This will require more curation, governance and accountability over the actions and output AI applications. Third, we must anticipate and prepare for shifts in labour demand. New skills training, retraining and reorganization will mitigate the effects of the disruption, without slowing the productivity gains and other benefits of AI. Finally, considerable resources must be allocated to building and adopting more universal and accessible platform technologies in order to democratize AI.

In our risk assessment, we identified a variety of themes that have the potential to affect future scenarios for AI. The primary themes are: 1) the extent of the gap between hype and reality, 2) the extent to which AI technology is democratized, and 3) the extent to which AI is integrated with humans and in organizations. The potential benefits of AI technology are significant. The potential prize for early adopters will be gain significant competitive advantage.

Bibliography

[1] J. McCarthy, M. Minskey, N. Rochester, C. Shannon, Dartmouth (1955), "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence"

[2] G. Lewis-Kraus, New York Times (2016), "The Great AI Awakening"

[3] A. Ng, Harvard Business Review (2016), "What Artificial Intelligence Can and Can't Do Right Now"

[4] D. Crevier, Basic Books (1994), "Ai: The Tumultuous History of the Search for Artificial Intelligence"

[5] A. Dragland, Science Daily (2016) "Big Data, for Better or Worse: 90% of World's Data Generated over Last Two Years."

[6] V. N. Vapnik, Springer Verlag (2000), "The Nature of Statistical Learning"

[7] G. Box, N. Draper, John Wiley & Sons. (1987), "Empirical Model-Building and Response Surfaces"

[8] C. Duhigg, New York Times (2012), "How Companies Learn Your Secrets"

[9] S. Biddle, The Intercept (2016), "Troubling Study Says Artificial Intelligence can Predict who will be Criminals Based on Facial Features"

[10] N. Henke, J. Bughin, M. Chui, J Manyika, T. Saleh, B. Wiseman, and G. Sethupathy, McKinsey Global Institute 2016, "The Age of Analytics: Competing in a Data-driven World"

[11] Economist, (2016) "Million Dollar Babies"

[12] McKinsey & Company (2016), "Primer on Technology Superclusters and a Fact Base on Canada's Toronto-Waterloo Innovation Corridor"

[13] McKinsey Global Institute (2016), "The Age of Analytics: Competing in a Data Driven World"

[14] M. Gray, Siddarth Suri, Harvard Business Review (2017), "The Humans Working Behind the AI Curtain"

[15] M. Chui, J. Manyika, and M. Miremadi, McKinsey Global Institute (2016), "Where Machines Could Replace Humans—and Where They Can't (yet)"

[16] R. Smith, Economics (2015) "Idealizations of Uncertainty, and Lessons from Artificial Intelligence"

[17] C. Duhigg, New York Times (2016), "What Google Learned from its Quest to Build a Perfect Team"

Mike Durland

Mike is the former Group Head and CEO, Global Banking and Markets for Scotiabank. Mike retired from Scotiabank in 2016 to pursue a variety of business, philanthropic and academic interests. Mike is the CEO of Melancthon Capital, a Professor of Practice at the Munk School of Global Affairs, and a member of the Business Strategy Committee for the Global Risk Institute. Mike is a member of a number of corporate, academic, and philanthropic boards, and holds a B Comm degree from St. Mary's University, a PhD in Management from Queen's University and an honorary Doctorate from St. Mary's University.

Matthew Killi

Matthew is Vice President at DeepLearni.ng, a Toronto based company who creates bespoke AIs for large enterprises and helps organizations unlock value in their data assets with machine learning. Previously, he was a management consultant at McKinsey & Company. Before joining McKinsey, he researched theoretical physics at the Centre for Quantum Materials at the University of Toronto.