

# Predicting the future of prediction

DOI: 10.29236/sistemas.n154a8

## Summary

Artificial Intelligence (AI) is, without a doubt, one of the disruptive technologies that will make an impact on the way we live, work and interact with each other now and in the future. This article highlights the current speed of AI development starting with some background on Computer Vision and then describes the latest research in the field of Natural Language Processing (NLP). Many of the traditional machine learning models have gained strength with the advent of Deep Learning and companies are embarking on their AI journeys; some are succeeding and some others are facing challenges when moving from proof-of-concept to production. An overview of the most common challenges for companies in Northern Europe is presented to bring a perspective of what is happening in other parts of the world. In the end, two examples of how face recognition is being applied in similar use cases both in China and Sweden are also described. The cases are used to compare the effects of those implementations as a framework of the importance of regulation in the future. All the topics are discussed at a high level to give a quick outline of how we can position ourselves today to best prepare for a paradigm change that has already started.

## Keywords

AI, Deep Learning, Computer Vision, Natural Language Processing, Operational AI

## Introduction

On the promise of creating intelligent machines, AI was born as an academic discipline in the '1950s to later discover that the challenge was much more difficult than expected. Machine Learning, a subset of AI, woke up the dormant field in the 1990s when teaching machines to become intelligent was a more feasible quest than creating them already intelligent. Machine Learning brought hope back to the AI research community with achievements as IBM's Deep Blue's win over Chess World Champion Garry Kasparov, in 1997<sup>1</sup>. However, it was not until Deep Learning came around, combined with the advances in computational processing power and the accelerated growth of data creation, that AI was back into the spotlight to prove its real potential.

Deep Learning, a technique within Machine Learning, learns from large amounts of samples to automatically extract features from data. Machine Learning, on the contrary, requires humans to extract those data representations (features). Computer Vision and Natural Language Processing or NLP are examples of Deep Learning. Computer Vision relates to the interdisciplinary field that focuses on how computers can understand, classify and extract information from images and videos. NLP aims to under-

stand how humans communicate with each other and how to build systems capable of replicating that behavior so we can interact with computers in a more natural way.

## Deep Dive into Deep Learning

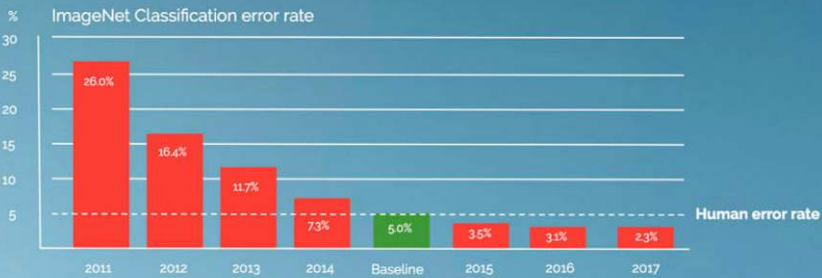
The cost of prediction is getting reduced by the technological paradigm shift that AI is bringing with its revolution. However, it is still very difficult to estimate accurately what will be its exact position in our society a decade from now when the velocity of its development is rolling faster than we have ever experienced before. To put things into perspective, let's look at the remarkable speed of development the areas of Computer Vision experienced in the mid-2010s and NLP has been experiencing in the past couple of years.

The year 2012 marked a significant breakthrough in the field of AI that brought back confidence to the six-decades-old field of science. By using Deep Learning, and more specifically deep neural networks, an algorithm called AlexNet trained on GPUs, reached an astonishing error rate for object classification of 16.4%<sup>2</sup> at the ImageNet Computer

<sup>1</sup> In Wikipedia (n.d.), Deep Blue versus Garry Kasparov. Retrieved January 25, 2020, from [https://en.wikipedia.org/wiki/Deep\\_Blue\\_versus\\_Garry\\_Kasparov](https://en.wikipedia.org/wiki/Deep_Blue_versus_Garry_Kasparov)

<sup>2</sup> Stanford Vision Lab (n.d.), IMAGENET Large Scale Visual Recognition Challenge 2012 (ILSVRC2012). Retrieved from <http://image-net.org/challenges/LSVRC/2012/results.html>

## Breakthroughs in Computer Vision



Source: Data from IMAGENET Large Scale Visual Recognition Challenge 2011-2017

Figure 1. Breakthroughs in Computer Vision (Peltarion, 2018)

Visual Recognition Challenge (ILSVRC). AlexNet dropped the error rate close to ten points from the best performing algorithm which in the previous year scored an error rate of 25.7%<sup>3</sup>. For humans this error rate is estimated at 5%<sup>4</sup>. This marked a breakthrough point in the field of Computer Vision that generated rapid improvements in the field by using more complex deep learning-based models. Three years later in 2015 error rates dropped to 3%<sup>5</sup> mark surpassing

the human baseline for the first time and giving us the possibility to create AI systems that can identify and classify images as well as us and sometimes better than us (Fig.1).

Similarly, the field of NLP has recently got unprecedented acceleration. In 2017 Google published a neural network architecture for language understanding called transformers. The new architecture surpassed its predecessors in accuracy and efficiency by using a self-attention mechanism that looks at the whole sentence and establishes connections between relevant words to build context. Previous architectures often treated words the same regardless of the context or word order. For example, in the sentences: “*opening a bank account*” and “*how is the bank of a river formed?*” traditional models give the word “*bank*” the same value representation<sup>6</sup>; In contrast, the

<sup>3</sup> Stanford Vision Lab (n.d.), IMAGENET Large Scale Visual Recognition Challenge 2011 (ILSVRC2011). Retrieved from <http://image-net.org/challenges/LSVRC/2011/results>

<sup>4</sup> Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S. et al (2015) IMAGENET Large Scale Visual Recognition Challenge, IJCV. Retrieved from <https://arxiv.org/pdf/1409.0575.pdf>

<sup>5</sup> Stanford Vision Lab (n.d.), IMAGENET Large Scale Visual Recognition Challenge 2015 (ILSVRC2015). Retrieved from <http://image-net.org/challenges/LSVRC/2015/results#loc>

<sup>6</sup> Uszkoreit, J. (August 31, 2017). Transformer: A Novel Neural Network Architecture for Language Understanding. [Blog]. Retrieved from <https://ai.googleblog.com/2017/08/transformer-novel-neural-network.html>

transformer-based model would give different values and a more accurate representation based on its context. The revolutionary architecture broke new ground on machine translation tasks.

In late 2018, Google opened sourced a pre-trained transformer-based model named BERT (Bidirectional Encoder Representations from Transformers)<sup>7</sup> a state-of-the-art model that re-shaped completely the NLP landscape. BERT was pre-trained on approximately 3.5 billion English words from books and Wikipedia. By releasing the model and its weights (meaning what it had learned from those texts) people can fine-tune it and make use of this massive network to create systems that understand English text better than ever seen before. It opened the possibilities to use NLP with better performance in tasks such as text similarity, question answering, content classification and language inference to name a few.

When BERT was evaluated against the General Language Understanding Evaluation (GLUE) benchmark, the results significantly outperformed across all the natural language understanding tasks<sup>8</sup>. Similarly, the results on the Stanford Question Answering Dataset surpassed the previous state-of-the-art scores including the human level<sup>9</sup>. These marked a main breakthrough within NLP that generated a wave of BERT-based models throughout all 2019 (e.g. 'Micro-

softs MT-DNN, 'Googles XLNet and 'Facebooks RoBERTa and XLM-R) pressing the researches to create more rigorous and complex tests to challenge language understanding.

### What does it mean in practice?

Computer Vision achieved in the lapse of three years an impressive development that is used broadly in areas where visual inspection is required. Examples of its applicability include:

- Detection of defective products for quality control in manufacturing lines;
- Image segmentation for product classification (e.g. clothing or furniture types) in the retail industry;
- Object localization as pedestrians, cars or traffic signs in self-driving cars, trucks or drones;
- Automatic sorting of produce in agriculture and food industries;
- Face recognition in images or videos for personal identification; and,

---

<sup>7</sup> Devlin, J., Chang, M., Lee, K., Toutanova, K., (2018) BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Computing Research Repository. Retrieved from <https://arxiv.org/pdf/1810.04805v2.pdf>

<sup>8</sup> Wang, A., Singh, A., Michael, J., Hill, F., Levy, O., Bowman, S., (2019) GLUE: A Multi-task Benchmark and Analysis Platform for Natural Language Understanding. Computing Research Repository. Retrieved from <https://arxiv.org/pdf/1804.07461.pdf>

<sup>9</sup> Devlin, J., Chang, M., (November 2, 2018). Open Sourcing BERT: State-of-the-Art Pre-training for Natural Language Processing. [Blog] Retrieved from <https://ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html>

- Identification of bone fractures or cancer from X-rays for medical diagnosis.

In the same way, NLP advanced enormously in a lapse of only one and a half years. Its applicability extends all the aspects where the meaning of text needs to be understood overpassing the keyword search. For instance:

- Question answering in customer service;
- Finding product features and descriptions in consumer support scenarios;
- Text classification that organizes messages in categories based on its content as in Help Desks, and then automatically redirects to the appropriate channels;
- Analyzing the sentiment in customer reviews, finding patterns in it and acting accordingly;
- Product classification based on similarities of its descriptions; and,
- Summarization of text that generates a short and accurate synopsis of a longer document.

The speed of development of new algorithms and deep learning architectures is certainly positive and very sophisticated models are being released and published more often than not. However, many or-

ganizations, companies and governments worldwide are only in the proof-of-concept phase or affronting challenges with AI operationalization.

### **Closing the gap between research and AI in production**

Business leaders are juggling many challenges, some might still be implementing the digital transformation strategies while others might be working intensively on the potential that big data, cloud and advanced analytics bring to their businesses. Those transformations might not have had time to settle down across the organization when new innovative development needs to be added to their company's strategies.

There are many challenges with AI adoption and the breach between moving from proof of concepts to operationalization needs to be closed. Below, the author briefly lists the most common reasons she encounters in her interactions with companies in Northern Europe when discussing and helping them on their AI journey:

1. Lack of talent. Attracting, acquiring and retaining the right talent is one of the biggest challenges organizations are facing. There are not enough professionals educated in the field with the right skill set. Big companies with large budgets attract the best talent and the availability in the

market has become very scarce and expensive. Companies that already have Data Scientists and ML Engineers in the organization, need to make sure to provide them with the tools and resources they require. Companies that do not, need to start investing in upskilling their teams and at the same time, considering hiring external expertise from organizations who specialize in the area. Always advisable to have in place an appropriate transfer knowledge from the experts that are outsourced to the people in the organization.

2. Executive sponsorship. An AI strategy needs to come from the top of the organization to the bottom, assuming an “AI-first approach” as Google's CEO Sundar Pichai mentioned in his keynote at I/O in 2017. Having executive sponsorship is key to successfully align everyone in the company towards the same goal. Even though AI applications get developed mostly by the technical teams, their output is a reflection of the company as a whole.
3. No Business Cases for AI or difficulties finding the right cases. The author advises not to start asking how AI can be applied to the business, instead look at the business and identify the areas that have the most impact if they are made more efficient and effective. Then, brainstorm at

ways to improve those areas and if AI becomes part of the solution then a good business case has been identified. AI cannot be imposed on any business and many times it might not be needed as digital transformation or advanced analytics can give the company those insights and improvements required. However, when identifying potential AI use cases, business unit stakeholders need to be invited to the table. The technical team is able to guide if AI is appropriate or not as part of the solution but the stakeholders have the domain expertise and can focus on those areas that bring concrete business value. If the use case does not increase revenue, reduce costs or bring tangible benefits that use case is not going to get the executive sponsorship either.

4. Data and more data: AI is very data-hungry and in order to implement successful applications organizations require a high level of digital maturity. On this aspect, the author has experienced on one hand, over-optimistic companies who believe they have data but unfortunately is not usable or accessible; the quantity or quality is not optimal, or it lacks consent for its use. On the other hand, some companies quickly realize their data is insufficient and abandon their developments too quickly. In both cases, it is key to do a data inven-

tory and establish a data strategy across the company. If the data is not available today it does not mean it cannot be available tomorrow. Leaders need to design and update their AI strategy with the expectation that they will gather more data year after year as it is better to start working on the data strategy sooner rather than later.

5. Right Infrastructure and tools: Deep Learning might seem just as another type of Software Engineering, especially when creating models involves a lot of coding. However, AI is very different and learning models are all about experimentation rather than product development. Data partly replace code in AI systems and a learning algorithm is used to automatically identify patterns in the data instead of writing hard-coded rules<sup>10</sup>. The data dependency in AI systems is much higher and as an experimental field, version control becomes

imperative and more rigorous than in software engineering. Models need to be reproducible with all its components as: operating system; installed packages, frameworks and libraries; model source code and model configuration; pre-processing settings; input signals on training data and target values among others. Patchwork technical solutions need to be avoided as much as possible as they do not allow for scalability, continuous change nor speed.

### China as early adopter

With a clear goal to become the world's leader in AI in 2030<sup>11</sup> China is applying AI at large-scale with massive surveillance one of the most predominant. According to an analysis in 2019 by Comparitech<sup>12</sup> eight of the top most-surveilled cities in the world are located in China. London (6th place) and Atlanta (10th place) are the only two cities outside China to make the top ten. The extensive network of computer-controlled cameras combined with computer vision and face recognition has facilitated the implementation of many AI systems for security and government's purposes. Some examples include: face scanners at entrances in office buildings and housing complexes; border control checkpoints; apprehension of suspects and most-wanted in public areas and catching and finding traffic violators.

---

<sup>10</sup> Artep, A., Brinne, B., Crnkovic-Friis, L., Bosh, J., (2018) Software Engineering Challenges of Deep Learning. 44th Euromicro Conference on Software Engineering and Advanced. Retrieved from

[https://peltarion.com/resources/software-engineering-challenges-of-deep-learning?gclid=EAAlaQobChMIhsOuv5Kx5wIVeAqaCh3RPQtEAAAYASAAEglqMvD\\_BwE](https://peltarion.com/resources/software-engineering-challenges-of-deep-learning?gclid=EAAlaQobChMIhsOuv5Kx5wIVeAqaCh3RPQtEAAAYASAAEglqMvD_BwE)

<sup>11</sup> Jing, M. (Nov 15, 2018) China to boost its 'national team' to meet goal of global AI leadership by 2030. South China Morning Post. Retrieved from

<https://www.scmp.com/tech/innovation/article/2173345/china-boost-its-national-team-meet-goal-global-ai-leader-ship-2030>

<sup>12</sup> Bischoff, P. Surveillance camera statistics: which cities have the most CCTV cameras? August 15, 2019. Comparitech. Retrieved Jan 31, 2020 from

[https://www.comparitech.com/vpn-privacy/the-worlds-most-surveilled-cities/#China\\_leads\\_the\\_world\\_in\\_CCTV\\_surveillance](https://www.comparitech.com/vpn-privacy/the-worlds-most-surveilled-cities/#China_leads_the_world_in_CCTV_surveillance)

Educational institutions from primary schools to universities across the country are not excluded from the AI adoption wave in China. Face controls are not only used to gain entry and secure school facilities but also to record attendance, measure student's attentiveness in class and analyze behaviors of teachers and students among others. In a primary school, children wear a headband with sensors that pick up electrical signals from the neurons in the brain (Figure 2). The data measures the attention levels throughout the class and is then sent in real-time to a computer on the teacher's desk and parents via an app<sup>13</sup>. BrainCo, the firm that created the headbands, has delivered more than 20,000 units to China

alone. The intensive use of monitoring systems is raising concerns among university students, teachers and parents about information security, privacy and effectiveness of the technology but in this case, the school was using the headbands without major restrictions.

## Regulation

The panorama looks very different in other parts of the world both in the adoption of the technology and its regulation. In Europe, for instance,

---

<sup>13</sup> Wang, Y., Hong, S., Tai, C., (October 24, 2019) China's Efforts to Lead the Way in AI Start in Its Classrooms. The Wall Street Journal, Technology. Retrieved from <https://www.wsj.com/articles/chinas-efforts-to-lead-the-way-in-ai-start-in-its-classrooms-11571958181>



Figure 2. School children wearing AI headbands in class. Source: Tai, C.,(2019) The Wall Street Journal



the General Data Protection Regulation (GDPR) is a privacy legislation law that regulates how businesses can collect, use and store personal data. A municipality in the north of Sweden was fined 20,000 Euros for running a pilot project under three weeks using a facial recognition system to track attendance in one class of 22 students. The Swedish Data Protection Authority found the use of surveillance cameras, in an everyday environment, an intrusion to students integrity and argued that attendance can be done in other ways that are less intrusive than face recognition<sup>14</sup>. Biometric data, like the one used in face recognition, is under GDPR legislation very sensitive personal data that requires exceptions to be able to use it.

How can the same technology have such different outcomes in two similar scenarios? And what would be the best way forward to not restrict adoption but at the same time respect people's privacy and integrity? Regularization is one of the critical topics that need to be developed extensively in the years to come and it is not going to be an easy task. In China, the use of face recognition has been in place already a few years back but in Europe, it is a nascent area with already controversial outcomes. Besides, the European Commission is considering a ban for up to 5 years on the use of facial recognition in public spaces until appropriate regulations for its use are in place<sup>15</sup>. In

China, the primary school was ordered to suspend the use of the head-tracker after the news and footage from The World Street Journal circulated online<sup>16</sup>, echoing the concerns that already were building up and after pressure from the local and international public. Regulators have on their shoulders the responsibility and challenge to assess the benefits and risks that AI brings with its use and consequently, dictate its recommendations. At the same time, balancing out the opportunities and drawbacks of this technology, that increasingly evolves and has broadness applicability, will not be an easier task either.

## Final reflections

Artificial Intelligence will redefine human lives, jobs and interactions with each other and the environment surrounding them. AI is becoming ubiquitous by augmenting the human cognitive capacity, making processes more efficient and operations more effective in every domain. However, within its potential

---

<sup>14</sup> European Data Protection Board (August 22, 2019) Facial recognition in school renders Sweden's first GDPR fine. National News. Retrieved from [https://edpb.europa.eu/news/national-news/2019/facial-recognition-school-renders-swedens-first-gdpr-fine\\_en](https://edpb.europa.eu/news/national-news/2019/facial-recognition-school-renders-swedens-first-gdpr-fine_en)

<sup>15</sup> Delcker, J., Smith-Meyer, B., (January 17, 2020) EU considers temporary ban on facial recognition in public spaces. PoliticoPro. Retrieved from <https://www.politico.eu/pro/eu-considers-temporary-ban-on-facial-recognition-in-public-spaces/>

<sup>16</sup> Pinghui, Z., (November 2, 2019) Chinese primary school stops using headbands to study pupils' concentration levels after public outcry. South China Morning Post. Retrieved from <https://www.scmp.com/news/china/society/article/3035983/chinese-primary-school-stops-using-headbands-study-pupils>


and myriad of possibilities also lies drawbacks. It is not only up to regulators or researchers at Universities and large tech companies to shape the way AI evolves, but it is also everyone's responsibility. From the analyst, data engineers and developers who build the algorithms; business leaders who embedded predictions into applications used by millions; media and journalism that inform people; governments that respect human rights and facilitates adoption; to finally end-users who need to be judgmental on its consumption.

Several topics were not discussed in this article as bias, discrimination, algorithms' explainability or ethics but that does not make them, in any aspect, less relevant or less fundamental on their critical role for responsible AI. Society needs to prepare for the paradigm shift that has already started and the best way to do it is to stay informed, learn more about what AI is and its implications for everyone independent of their social role. This exercise needs to start today because tomorrow is just hours away.

Suggested readings: The essential AI Handbook for Leaders, Free online course Elements of AI, Blog Machine Learning from Everyone.

## References

- Artep, A., Brinne, B., Crnkovic-Friis, L., Bosh, J., (2018) Software Engineering Challenges of Deep Learning. 44th Euromicro Conference on Software Engineering and Advanced. Retrieved from [https://peltarion.com/resources/software-engineering-challenges-of-deep-learning?gclid=EAIaIQobChMIhsOuv5Kx5wIVeAqaCh3RPQtfEAAYASAAEgIqMvD\\_BwE](https://peltarion.com/resources/software-engineering-challenges-of-deep-learning?gclid=EAIaIQobChMIhsOuv5Kx5wIVeAqaCh3RPQtfEAAYASAAEgIqMvD_BwE)
- Bischoff, P. (2019) Surveillance camera statistics: which cities have the most CCTV cameras? August 15, 2019. Comparitech. Retrieved Jan 31 2020 from [https://www.comparitech.com/vpn-privacy/the-worlds-most-surveilled-cities/#China\\_leads\\_the\\_world\\_in\\_CCTV\\_surveillance](https://www.comparitech.com/vpn-privacy/the-worlds-most-surveilled-cities/#China_leads_the_world_in_CCTV_surveillance)
- Delcker, J., Smith-Meyer, B., (January 17, 2020) EU considers temporary ban on facial recognition in public spaces. PoliticoPro. Retrieved from <https://www.politico.eu/pro/eu-considers-temporary-ban-on-facial-recognition-in-public-spaces/>
- Devlin, J., Chang, M., Lee, K., Toutanova, K., (2018) BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Computing Research Repository. Retrieved from <https://arxiv.org/pdf/1810.04805v2.pdf>
- Devlin, J., Chang, M., (November 2, 2018). Open Sourcing BERT: State-of-the-Art Pre-training for Natural Language Processing. [Blog] Retrieved from <https://ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html>
- European Data Protection Board (August 22, 2019) Facial recognition in school renders Sweden's first GDPR fine. National News. Retrieved from [https://edpb.europa.eu/news/national-news/2019/facial-recognition-school-renders-swedens-first-gdpr-fine\\_en](https://edpb.europa.eu/news/national-news/2019/facial-recognition-school-renders-swedens-first-gdpr-fine_en)

- In Wikipedia (n.d.), Deep Blue versus Garry Kasparov. Retrieved January 25, 2020, from [https://en.wikipedia.org/wiki/Deep\\_Blue\\_versus\\_Garry\\_Kasparov](https://en.wikipedia.org/wiki/Deep_Blue_versus_Garry_Kasparov)
- Jing, M. (Nov 15, 2018) China to boost its 'national team' to meet goal of global AI leadership by 2030. South China Morning Post. Retrieved from <https://www.scmp.com/tech/innovation/article/2173345/china-boost-its-national-team-meet-goal-global-ai-leadership-2030>
- Pinghui, Z., (November 2, 2019) Chinese primary school stops using headbands to study pupils' concentration levels after public outcry. South China Morning Post. Retrieved from <https://www.scmp.com/news/china/society/article/3035983/chinese-primary-school-stops-using-headbands-study-pupils>
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S. et al (2015) IMAGENET Large Scale Visual Recognition Challenge, IJCV. Retrieved from <https://arxiv.org/pdf/1409.0575.pdf>
- Stanford Vision Lab (n.d.), Imagenet Large Scale Visual Recognition Challenge 2011 (ILSVRC2011). Retrieved January 25, 2020, from <http://image-net.org/challenges/LSVRC/2011/results>
- Stanford Vision Lab (n.d.), Imagenet Large Scale Visual Recognition Challenge 2012 (ILSVRC2012). Retrieved January 25, 2020, from <http://image-net.org/challenges/LSVRC/2012/results.html>
- Stanford Vision Lab (n.d.), Imagenet Large Scale Visual Recognition Challenge 2015 (ILSVRC2015). Retrieved January 25, 2020, from <http://image-net.org/challenges/LSVRC/2015/results#loc>
- Tai, C., (October 24, 2019) . China's Efforts to Lead the Way in AI Start in Its Classrooms. The Wall Street Journal. [Video] Retrieved from <https://www.wsj.com/articles/chinas-efforts-to-lead-the-way-in-ai-start-in-its-classrooms-11571958181>
- Uszkoreit, J. (August 31, 2017). Transformer: A Novel Neural Network Architecture for Language Understanding. [Blog]. Retrieved from <https://ai.googleblog.com/2017/08/transformer-novel-neural-network.html>
- Wang, A., Singh, A., Michael, J., Hill, F., Levy, O., Bowman, S., (2019) GLUE: A Multi-task Benchmark and Analysis Platform for Natural Language Understanding. Computing Research Repository. Retrieved from <https://arxiv.org/pdf/1804.07461.pdf> 

**Liliana (Díaz) Lindberg** AI Solutions Engineer at the Swedish start-up Peltarion. She guides companies to solve business challenges by successfully using AI & Deep Learning. Liliana holds a BSc. Systems and Computing Engineering from Los Andes University, Colombia; an MSc. in Geographical Information Systems from the University of Calgary, Canada; and a Master's level Business Leadership Specialization from Duke University and Google where she worked for seven years with cutting edge technologies like Google Maps, Google Cloud, Big Data and AI.