Artificial Intelligence/Machine Learning: The CRO's Agenda

Author: LOIS TULLO
EXECUTIVE-IN-RESIDENCE, Global Risk Institute

Artificial Intelligence[1] (AI) was googled for 3.5 billion searches per day in 2018, over 40,000 search queries every second on average, and 1.2 trillion searches per year worldwide as listed by Internet Live Stats. With all the hype, what is important for CROs and financial executives to cull from the noise? I spoke with John Hull[2] and Ryan Riordan[3] at GRI's 2019 RISK SUMMIT about how CRO's are approaching AI opportunities and risks.

BUSINESS STRATEGIES THAT DRIVE AI INVESTMENT

To be successful AI must be embedded in a business's strategy, deployed either to address a problem that the business is trying to solve, or to capture an opportunity that the business has identified. Financial services organizations have in recent years identified applications like:

- *KYC/AML[4]: accounted for 92% of all operational risk losses in 2018*
- *Cybersecurity: the 2018 Cyber Incident & Breach Trends Report reported an estimated two million cyber-attacks in 2018, resulting in more than $45 billion in losses worldwide, while 2019 is seeing increases in overall attacks with the greatest increase in malware style attacks[5]*
- *Fraud detection: the proliferation of misinformation through text and video has the power to potentially create a trust barrier to relying on information provided digitally*
- *Portfolio Management: increased momentum towards fully automated trading, monitoring (risk management and compliance) and clearing and settlement (P2P, Centralized, Blockchain)*
- *Customer Service: under pressure to become more efficient through automation, deploying AI across financial services has a $1 trillion potential. Predictions have been articulated of potential cost savings of $490 billion in front office (distribution), $350billion in middle office, and $200 billion in back office (manufacturing) functions.[6]*

---

1 Machine learning is an analytic technique that "learns" patterns in datasets without being guided by a human analyst. AI refers to the broader application of specific kinds of analytics to accomplish tasks like driving a car or identifying a fraudulent transaction.

2 John Hull, Maple Financial Professor of Derivatives & Risk Management at the University of Toronto

3 Ryan Riordan, Associate Professor & Distinguished Professor of Finance, Queens University

4 OXR 2018 largest losses study for Know Your Client/Anti Money Laundering

5 https://www.internetsociety.org/resources/ota/2019/2018-cyber-incident-breach-trends-report/

6 https://next.autonomous.com/augmented-finance-machine-intelligence

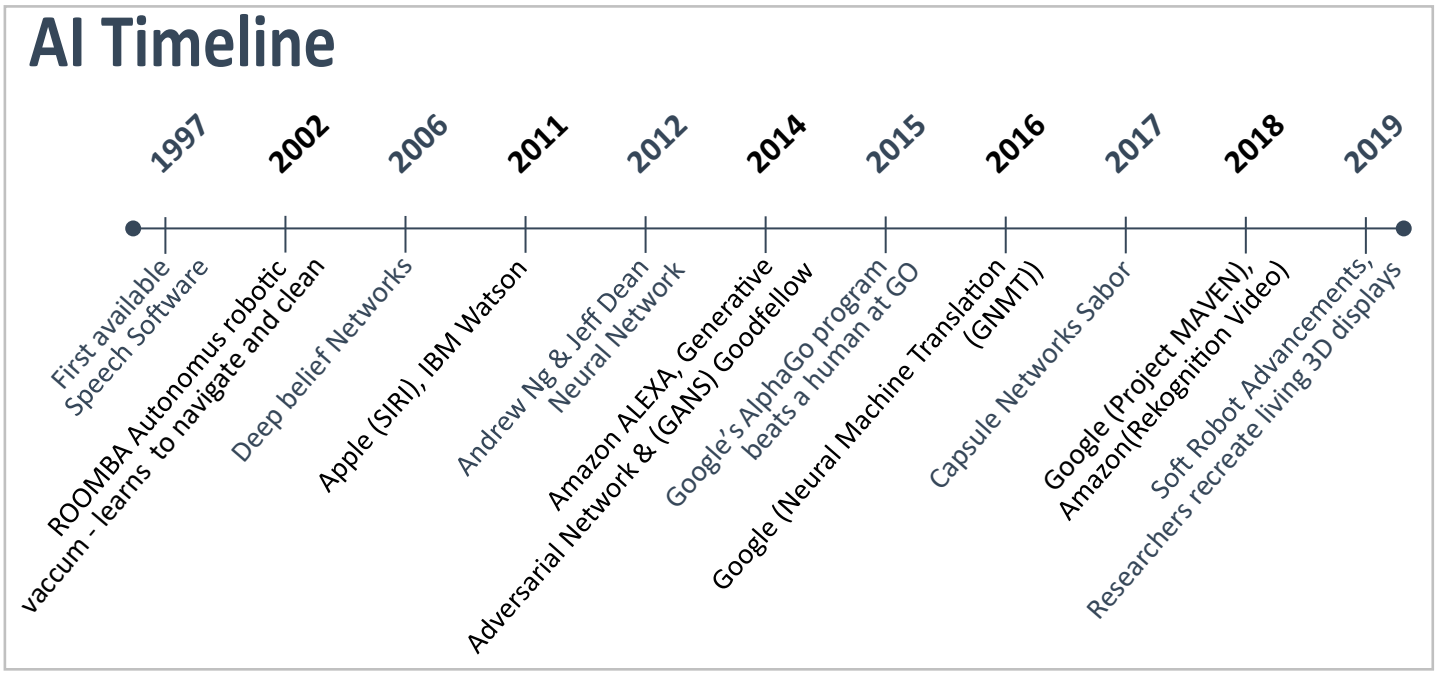
At the SUMMIT we asked our members "Where do you see the most value for AI in the next two years?" They ranked customer-facing applications, operations & compliance, and trading & risk management as their highest AI priorities over the next 2 years.

Where do you see the most added value for AI in the next two years?
Polls Results:

Customer-facing applications	33%
Operations & compliance	26%
Trading & risk management	26%
IT, Calculations optimizations, distribution of work	12%
Other	2%

WHY NOW? THE AI TIMELINE

A recent timeline of ML/AI starts to tell the story of, why now:



Driving today's accelerating momentum is the ongoing improvement in the data, tools and techniques available to enable AI/ML.

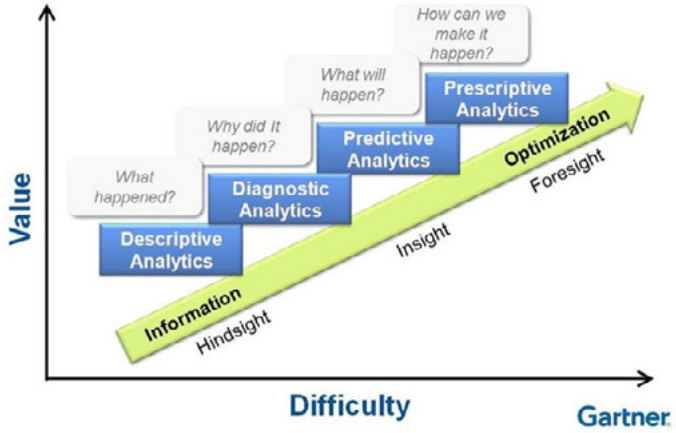
Big Data: It's estimated that by 2020, every person on earth will generate 1.7 MB of data every second, according to DOMO[7]. Databases are becoming increasingly versatile and powerful. In addition to traditional relational databases, we now have powerful graph databases that are more capable of connecting data points and uncovering relationships, as well as databases that specialize in document management.

Progress in Data Analytics: There is far more data being generated today than humans can analyze in any meaningful way. Techniques like machine learning, predictive analytics[8], and data visualization can help us find meaning by digging deeper into large data sets and improving the speed and accuracy of decision-making. **See Figure 1 for Progress in Data Analytic**s

[7] https://www.domo.com/assets/downloads/18_domo_data-never-sleeps-6+verticals.pdf

[8] https://altviz.co/articles/retail-analytics

FIGURE 1:PROGRESS IN DATA ANALYTICS



Structured learning: Data is the fuel that powers AI, and large data sets make it possible for machine learning applications to learn independently and rapidly. The abundance of data we collect supplies our AI models with the examples they need to identify differences, increase their pattern recognition capabilities, and see the fine details within the patterns, for both text and facial recognition.

A supervised model, the most common form of machine learning across all disciplines, is a model that is trained on a rich set of properly "tagged" transactions. Each transaction is tagged as either fraud or non-fraud. The models are trained by ingesting massive amounts of tagged transaction details in order to learn patterns that best reflect legitimate behaviors. When developing a supervised model, the amount of clean, relevant training data is directly correlated with model accuracy.

[9]**Unstructured learning:** AI enables us to make sense of massive data sets, as well as unstructured data that doesn't fit neatly into database rows and columns. AI is helping organizations create new insights from data that was formerly locked away in emails, presentations, videos, and images.

The development of unsupervised models designed to spot anomalous behavior expands the use cases where tagged transaction data is relatively thin or non-existent. In these cases, a form of self-learning is employed to surface patterns in the data that are invisible to other forms of analytics.

More powerful computing: Increasingly powerful hardware is required to do extensive computations very quickly as a model may be required to calculate and update millions of parameters in run-time for a single iterative model as used by deep neural networks.

Examples of increases in computing power of hardware include the transition from the use of CPUs[10] to GPU[11] clusters, TPUs[12] and faster FPGAs[13] – which are designed specifically for AI/ML. GPU clusters can perform operations on a batch of 128 or 256 images at once in just a few milliseconds. There is a downside to that increased capability: the power consumption increases to around ~250 W and requires a full PC that additionally requires 150W of power, which leads to a total of 400W per machine.

In the future quantum computing will infinitely speed up processing capacity. Currently, quantum computers are processing up to 15 seconds of uninterrupted gating processing time and over 3 minutes of annealing. The D-Wave machine is a quantum annealer running

9 https://su.org/blog/artificial-intelligence-and-big-data-a-powerful-combination-for-future-growth/

10 Central Processing Unit

11 Graphics Processing Unit- A programmable logic chip (processor) specialized for display functions. The GPU renders images, animations and video for the computer's screen.

12 A tensor processing unit (TPU) is an AI accelerator application-specific integrated circuit (ASIC) specifically for neural network machine learning.

13 A field-programmable gate array (FPGA) is an integrated circuit designed to be configured by a customer or a designer after manufacturing.

adiabatic quantum computing algorithms. This is great for optimizing solutions to problems by quickly searching over a space and finding a minimum (or "solution"). The latest announcement from Google states that the D-Wave machine is more than 10^8 times faster than simulated annealing running on a single core.[14]

We are following developments in quantum computing closely, as its arrival will mean that the RSA encryption underpinning all of today's digital commerce can be cheaply and quickly broken, necessitating a major rebuild of digital commerce infrastructure. Currently, quantum computing remains on the medium-term horizon. A major milestone on the path to quantum computing being able to break the current economic infrastructure that relies on conventional encryption is the development of stable quantum gating technology. While universal gating quantum computing systems rely on building reliable qubits, basic quantum circuit operations, similar to the classical operations, can be put together to create any sequence, to run increasingly complex algorithms. Algorithms like Shor's (to break RSA cryptography) and Grover's (faster search) require this quantum gating technology.[15]

Converging technologies: AI will not be developed in isolation. The potential benefits and capabilities which AI can offer are firmly interlinked with the development of other technologies such as blockchain and cloud computing. Emerging technology concepts, such as federated learning[16], differential privacy[17] and homomorphic encryption[18], suggest some potential paths forward.

REALIZING BUSINESS POTENTIAL IN FINANCIAL INSTITUTIONS

There are many opportunities for organization to use ML/AI, we are focusing on AI that is specifically targeted on the financial industry. Data from one AI vendor[19] show the Top 3 vendor applications in banking so far are fraud and Cyber security – 20.2%, Risk Management 14.4.%, and Compliance – 11.5%.

AI is being used in **fraud prevention** to identify irregularities in patterns, while tools like machine learning and data analytics are being used to scrutinize the vast amounts of data available. AI will take client and transactions data and compares these against publicly available data to recognize then flag suspicious activity for teams of security staff to investigate further. Current and anticipated security measures are requiring biometric data such as facial recognition, voice recognition, etc., to provide identity protection to customers. However, like so much in this field, this comes at some cost, notably of raising ethical and regulatory questions about data privacy and governance, and concerns about the decision-making biases current AI are learning from the bias-infused historical data with which they are operating.AI is currently being used in Cyber security protection for:

- *Spam filter applications (spamassassin)*
- *Network intrusion detection and prevention*
- *Fraud detection*
- *Botnet detection*
- *Secure user authentication*
- *Cyber security ratings*
- *Hacking incident forecasting*

[14] What's the difference between quantum annealing and universal gate quantum computers? Anastasia Marchenkova, https://medium.com/quantum-bits/what-s-the-difference-between-quantum-annealing-and-universal-gate-quantum-computers-c5e5099175a1

[15] Quantum Cryptanalysis: Shor, Grover, and Beyond, September/October 2018, pp. 14-21, vol. 16, https://www.computer.org/csdl/magazine/sp/2018/05/msp2018050014/17D45Xq6dCs

[16] Federated Learning is a machine learning setting where the goal is to train a high-quality centralized model with training data distributed over a large number of clients each with unreliable and relatively slow network connections

[17] Differential privacy is a system for publicly sharing information about a dataset by describing the patterns of groups within the dataset while withholding information about individuals in the dataset.

[18] Homomorphic encryption is a method of encryption that allows any data to remain encrypted while it's being processed and manipulated.

[19] https://emerj.com/ai-sector-overviews/ai-in-banking-analysis/

For instance, in order to develop an artificial intelligence application that can detect **malware**, the AI needs some definitions of distinctive features of malware. Comparing harmless software with known malware is one way to provide the system its training ground.

Features to used in analyzing software for malware comparison may include:

- *Accessed APIs,*
- *Accessed fields on the disk,*
- *Accessed environmental products (camera, keyboard, etc.),*
- *Consumed processor power.*
- *Consumed bandwidth.*
- *Amount of data transmitted over the internet.*

A new malware detection system is then built around the identified distinguishing features. DarkTrace Ltd. is currently using AI to detect cyber attacks, with claims of a 99% accuracy rate.[20]

From an **anti-money laundering perspective**, AI can intelligently extract risk-relevant facts and pattern recognition from a huge volume of data, holding out the promise of making the process of identifying high-risk clients easier in the fight against financial crime. It is also able to track the changes in regulations around the world, identify gaps in customer information stored by the financial institution and provide know your customer (KYC) alerts to perform regulatory outreach to customers to collect the outstanding information. However, the track record of operational costs and effectiveness of running AI for AML purposes remains thin. AI KYC/AML applications are currently being used for:

- *Accurate client risk profile and enhanced due diligence*
- *Ultimate beneficial ownership*
- *AML screening and investigation*
- *Improved client onboarding and document management automation*
- *Managing regulatory change and compliance*

AI could become the primary way that financial institutions **interact with their customers**, at cost points orders of magnitude lower than offshore service centres. Widespread implementation of chatbots that enable bank customers to chat with their financial institution using text-based natural language that facilitates finance-specific interactions. Providers such as Livechat, Ada, Hubspot, and Livechat all provide ready to use AI platforms for customer interface on both mobile and stationary devices.

AI Portfolio Management and client enablement: Steven Yadegari presented recent uses of AI implementations at the Summit for Asset Management in New York:[21]

- *Automated insight: reading earnings transcripts to assess management sentiment*
- *Relationship mapping: identifying nonintuitive relationships between securities and market indicators*
- *Alternative datasets: analyzing alternative data such as weather forecasts and container ship movements, monitoring search engines for words on specific topics to structure hedging strategies*
- *Growth opportunities: using corporate website traffic to gauge future growth along with clients' behavioral patterns*
- *Client outreach: smart client outreach and demand generation via analytics, using alternative data sources such as social media data.*

AI is also being used in the front, middle, and back-offices for **Efficiency Improvements** in:

[20] https://medium.com/@akshay.kurhade/how-can-ai-play-a-huge-role-in-cyber-security-e8bcc6b15636

[21] https://www.tsamtoronto.com/wp-content/uploads/Steven-Yadegari.pdf

- *Operations intelligence: using machine learning to automate functions*
- *Risk performance: AI-based algorithms and machine learning to monitor for suspicious transactions and trigger response protocols*
- *Reporting and servicing: generating reporting for clients, portfolio and risk commentary and marketing material using natural language processing*
- *On-demand reporting: chatbots and machine learning used to respond to employee or investor queries, generating management reporting on-demand*
- *Employee insights: monitoring employee conduct risk and employee morale.*

While not addressed in this article, it is also important for members of the financial industry to be aware of how their customers are also being transformed by AI as this could potentially increase credit, market and 3rd party risk.

ADDRESSING THE RISKS OF AI

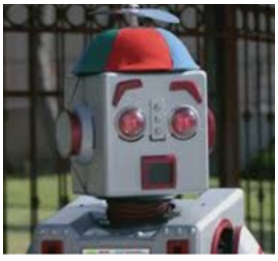
While considering the question, "why now?", we must also consider the flip side of the coin, why not now? What are the risks that CROs are concerned about? Our discussion pointed to:

1. ML/AI is still in its infancy and organizations are likely overstating the usefulness of AI for their customer service applications, including chatbots because:
    - *Chatbots deliver easily measurable cost savings, however customer relationship value and satisfaction is often difficult to evaluate*
    - *The natural language processing technology behind them is still relatively primitive*
    - *Even the best banking chatbots, such as Bank of America's Erica, are still only able to handle rudimentary requests from customers*

2. AI presents a skills gap where only 1 in 4 employees[22] is ready to work with AI. And while most cite the growing skills gap as the number-one factor influencing their workforce strategy, only 3 percent plan to significantly increase their investment in reskilling programs in the next three years[23].

Machine Learning works using statistical algorithms allowed to "learn" models of future decision making and action from historical data without being explicitly programmed. Serious ethical concerns surround the **bias or error** of either the baseline data or the programming of the algorithms used for decision making. Several high-profile examples where ML/AI have reinforced racist or biased human behavior have been publicised. Other **ethical concerns** captured public attention when on March 18, 2018, Elaine Herzberg[24] was struck and killed by a self-driving car.

MIT[25] outlined several methods where bias can slip into the AI/ML process. First, framing the problem may be a rather nebulous concept (for example: "creditworthiness"). Second, the data you collects may be unrepresentative of reality, or reflect existing prejudices (for example: the hiring algorithm at Amazon that discounted female applicants based on historical male hiring data). Third, preparing the data and the 'art' of deep learning: choosing which attributes to consider or ignore can significantly influence the model's prediction accuracy.



Example of a controversial of AI/ML projects:

facial recognition software that can predict the sexual orientation of people based on their facial characteristics

To answer the question: "Should New Privacy Legislation Regulate Artificial Intelligence Research and Applications?"

[22] https://aibusiness.com/bridging-ai-skills-gap-2018-long-read/

[23] https://www.accenture.com/ca-en/company-reworking-the-revolution-future-workforce

[24] https://www.forbes.com/sites/cognitiveworld/2019/09/26/what-happens-with-self-driving-cars-kill-people/#56be6ef6405c

[25] https://www.technologyreview.com/s/612876/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/

we need to look at how Big data allows for "privacy externality" where information others disclose about themselves also implicates you[26]. Machine learning increases the capacity to make inferences. The patterns found by machine learning analysis of your online behavior disclose your political beliefs, religious affiliation, race, ethnicity, health conditions, gender and sexual orientation, even if you have never revealed this information to anyone online. The presence of AI/ML in virtually all online spaces making deductions about you means that giving consumers control over their own information will not protect them from indirectly disclosing even their most sensitive information. The real ethical problem is the downstream use of the technology to harm other individuals in vulnerable groups.

In a financial services context, organizations must ask their own ethical questions[27]:

i. Does automation lead to financial surveillance, who owns the data, and do we have consent?

ii. Does automation reduce the ethical awareness and responsibility of financial professionals, does the use of AI pass off the responsibility from the person to the model?

iii. Does AI reduce accountability to financial customers, can they explain how the AI model made its decision?

iv. Does AI reduce the customer awareness of ethics, due to the enhanced speed reducing the ethical pause time?

Machine learning models can be both inscrutable[28] and nonintuitive[29]. Existing **laws** like the Fair Credit Reporting Act (FCRA), the Equal Credit Opportunity Act (ECOA), and the General Data Protection Regulation (GDPR), as well as techniques within machine learning, are focused almost entirely on the problem of inscrutability[30]. While such techniques could allow a machine learning system to comply with existing law, doing so may not help if the goal is to assess whether the basis for decision-making is normatively defensible. To know why the rules are what they are, one must seek explanations of the process behind a model's development, not just explanations of the model itself.

HOW SHOULD WE ADDRESS AI RISKS?

Current non-discrimination laws cover specific sectors such as housing, employment, credit, and insurance and specific groups of people who might be the victims of discrimination because of their race, gender, religion, national origin, age or disability. There is no exemption from these rules simply because a new advanced analytic technique such as AI or machine learning is being used.

Good privacy legislation in the age of AI, should include the following components:[31]

- *The use of AI must have a "deeply rooted" right to the information it is collecting.*
- *Consumers must be able to opt out of the system.*
- *The data collected and the purpose of the AI must be limited by design.*
- *Data must be deleted upon consumer request.*

[26] https://www.brookings.edu/blog/techtank/2019/04/01/how-to-address-new-privacy-issues-raised-by-artificial-intelligence-and-machine-learning/

[27] Ethics of using AI in the Financial/Banking industry Swapna Malekar, https://www.finn.ai/article/ethics-of-ai-in-banking/

[28] Dealing with inscrutability requires providing a sensible description of the rules.

[29] Addressing nonintuitiveness requires providing a satisfying explanation for why the rules are what they are.

[30] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3126971

[31] Principles for AI: A SourceBook, Roger Clarke http://www.rogerclarke.com/EC/GAIP.html

LOOKING FORWARD

We see promise in AI's application in areas like fraud reduction, money laundering, cyber security, portfolio management, and customer service. However, in the identification of irregularities or patterns in transactions, AI anomaly detection apps[32] increasingly rely on the comparison of biometric data such as facial and voice recognition, client and transaction data, publicly available data and data of unclear legality available only on the dark web. The reliance on such data, along with the challenge of transparency and repeatability of neural net ML models, highlights the risk of potential bias and other ethical concerns. To address these concerns, we must educate ourselves about the technology and its governance issues before we can effectively and responsibly link AI to organizational strategy and enterprise risk management while beginning to take advantage of the opportunities that AI presents.

[32] Artificial intelligence and the future of banking and finance, Wall Street, May 7, 2019 https://wall-street.com/artificial-intelligence-and-the-future-of-banking-and-finance/