OUTCOMES
BEYOND
PERFORMANCE

INVESTORS' EXPECTATIONS ON RESPONSIBLE ARTIFICIAL INTELLIGENCE AND DATA GOVERNANCE

April 2019

For professional investors only

www.hermes-investment.com

HERMES
INVESTMENT MANAGEMENT

TABLE OF CONTENTS

FOREWORD

“Increasingly as AI systems have become the focus of innovation investment, standards of expected ethical governance where AI is deployed have risen.

We need a clear engagement framework to cater for the impact of the introduction of AI applications, which now constitutes significant business risks.

It follows that clarity in accountability and the ethical design of AI systems to ensure data security, non-bias and transparency are essential.

This excellent Hermes paper points the way to the key requirements which I hope investors will treat as the new yardstick for corporate stewardship where AI is introduced and applied.”

– Tim, Lord Clement-Jones
Former Chair of the House of Lords AI Select Committee
Co-Chairman of the UK All-Party Parliamentary Group on Artificial Intelligence (APPG AI)

AUTHORS

Dr Christine Chow, Katherine Frame, Sonya Likhtman, Nicholas Spooner, Janet Wong.

BRYAN CAVE LEIGHTON PAISNER BCLP

A Special thanks to **Mark Lewis, partner in the Technology & Commercial practice of Bryan Cave Leighton Paisner LLP (BCLP),** who authored the section on the legal and regulatory challenges of this framework.

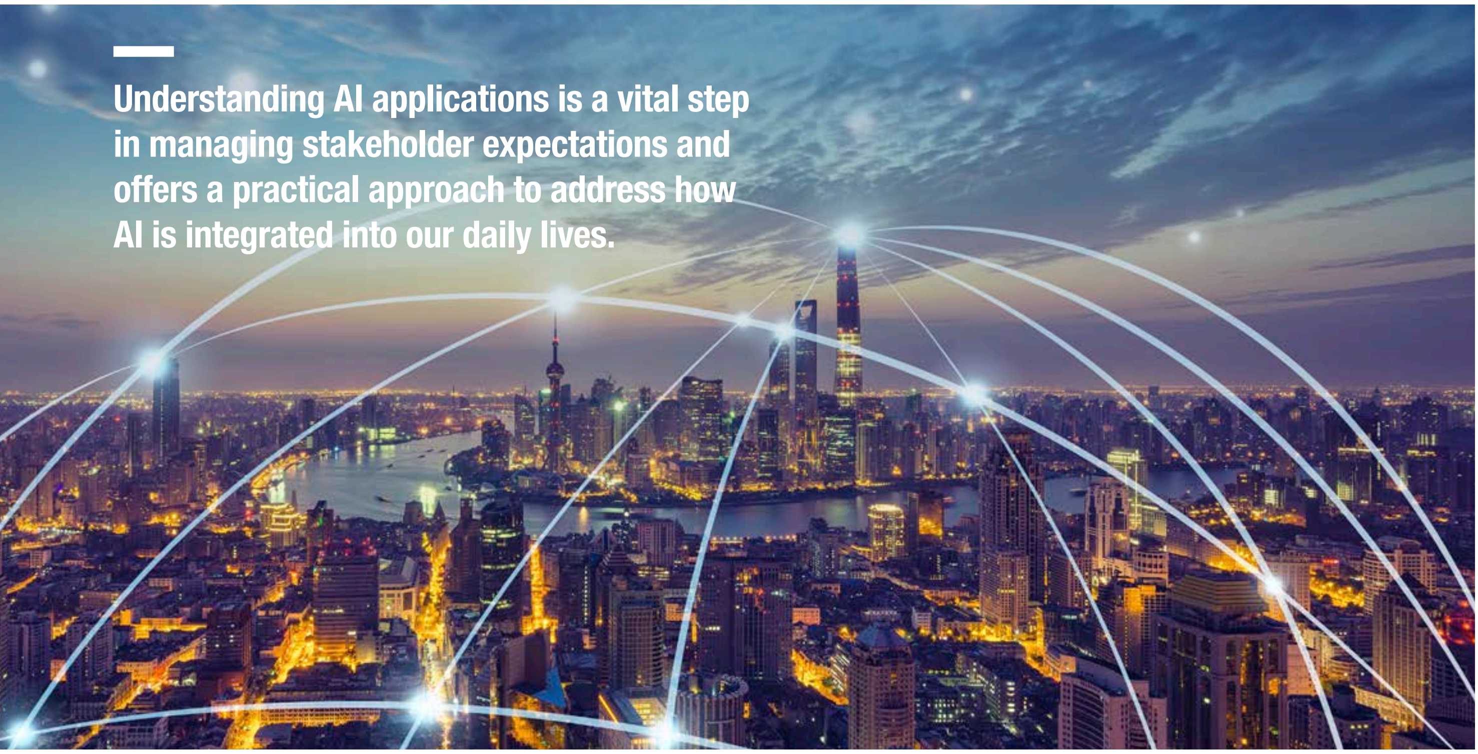
Thanks to Claire Milhench and Charlotte Judge for editing and formatting this report.

CONTEXT

> Now, we are seeing society beginning to wonder – how would I know if I'm being discriminated against by an algorithm, showing me things that an algorithm believes that I should be looking at? Does that create echo chambers? Does that create more polarised views? What is the societal impact of having people living in more of their echo chambers and silos? These questions are very deep questions because they are related to societal peace, and the very fabric that makes us who we are as a functioning society.[1]

We welcome input to further develop the principles and engagement framework included in this paper, as our collective experience and thinking on this important topic grows.

- We made these comments in December 2018 whilst participating in a technology podcast hosted by law firm DLA Piper.
- This paper outlines the importance of artificial intelligence (AI) as an ESG consideration for investors. We explain that any inherent biases from AI can be identified in one of three areas – input data bias, process bias, and outcomes bias – due to the nature of how data is applied in specific contexts. We discuss the evolving regulatory framework that different sectors may be exposed to, when companies deploy AI as part of their business strategy or in their operations.
- We encourage corporate boards to be accountable for the responsible use of AI and establish internal data governance mechanisms.
- We propose a structured approach for investors to engage on AI and data governance based on six core principles and have set out a common framework across sectors. This framework is separated into two related strands. The first, taking a risk factor approach, assesses the materiality of issues for companies based on legal, regulatory and financial outcomes. The second, taking a process-based approach, evaluates the impact of biases that may arise from data input and processes. We then suggest two main ways to address the unavoidable biases that may arise due to applied human values and judgement or systems architecture, based on the explainability and oversight.
- Technology companies have been the initial targets of engagement on AI and data governance. However, big data and AI are used in every sector. There are some common concerns, notwithstanding the benefits, challenges and applications.
- We conclude by stating how important it is for investors to take the time to understand how machine learning works. AI is built on 1950s and 1960s mathematics supported by more recently available increased computing power. It is unlikely to be a 'silver bullet' or panacea to solve problems. Neither does it need to be a 'black box' that is beyond good corporate governance or investor capacity to appraise, once it is understood properly. Understanding AI applications is a vital step in managing stakeholder expectations and offers a practical approach to address how AI is integrated into our daily lives.



Understanding AI applications is a vital step in managing stakeholder expectations and offers a practical approach to address how AI is integrated into our daily lives.

[1] **From DLA Tech podcast series December 2018. Soundbite video:** https://dms.licdn.com/playback/C4E05AQEabJ5oYNOvHQ/bcc2136ecb0a4557a2a2cfab062accde/feedshare-mp4_3300/1488578169071-zmy00q?e=1552478400&v=beta&t=vVCvccwrhoeAHYkLqAPepd3n_-hLwmswzPPO9-AtJIw

EXECUTIVE SUMMARY

Human beings generate data for machine learning, which is used in artificial intelligence (AI) applications on a daily basis. This could happen offline, when we do our grocery shopping at the weekend and accumulate loyalty points; but increasingly it could happen online, such as when we search for local cinema listings, make bookings for dinner through the internet, shop online and use free maps for navigation.

The mathematics that underpins what we call AI nowadays dates back to the 1950s and 1960s. There were two 'winters' in the early 1970s and late 1980s where research funding reduced due to criticisms, as machine learning was deemed overhyped. Machine learning advanced only recently because of enhancement in computing power, optimisation techniques and data availability. In the 21st century, machine learning as AI captures our imagination.

According to an Economist Intelligence Unit survey of more than 600 senior executives worldwide[2], 36% said that AI and machine learning had played a significant role in their organisation's digital strategy. Some 45% saw AI and machine learning as the most important technology to play a significant role in their organisation's digital strategy in the next three years. A survey[3] by Gartner, a global technology consulting firm, showed that 37% of organisations had already implemented AI in some form; the number of organisations implementing AI grew 270% in the past four years and tripled in the past year.

37%

of organisations had already implemented AI in some form

270%

– amount that implementation grew over the last 4 years

Scientific breakthroughs using AI in the fields of healthcare, autonomous driving, agriculture, manufacturing, and climate change solutions demonstrate tangible business and social benefits. AI may enable better predictions to be made in healthcare and improve the process for drug discovery and applications. This will help reduce wasted research and clinical trial failures. AI adoption also promises to improve product quality and make manufacturing processes more efficient. Targeted advertising improves marketing, cross- and up-selling. Automated compliance and operations management reduces the cost burden on institutions and helps compliance or engineering professionals to focus on more value-adding tasks. However, despite the benefits, AI applications could also become a double-edged sword, through carelessness in data and analysis management, user manipulation, invasion of privacy and social segmentation. It could also have profound impacts on wider society, as it accelerates the process of automation, affecting jobs and livelihoods. Together these raise a series of significant social and ethical questions that will require increasing governance and management.

Our responsible AI and data governance principles

Although at the early stages of our engagement on AI, we have gained insights from companies which have helped us to establish the following principles for best practice AI and data governance, which then links to our framework for stewardship. Hermes encourages investors to apply these governance principles to the responsible use of AI.

Trust | Transparency | Action

Integrity | Accountability | Safety

Trust

Companies should earn trust by educating users on their rights to data privacy and give users control and the right to consent to the use of their data by providing fully free choices.

Transparency

Companies should be transparent about tracking methods in the full value chain and disclose how they measure the robustness of data governance and the fair and safe use of AI. Companies should inform users when their data is being used for scoring and screening purposes.

Action

Companies should thoroughly explore and make all reasonable efforts, in good faith, to avoid unintended consequences such as data and process bias, which may lead to discrimination.

Integrity

Companies should demonstrate integrity in the treatment of customers, suppliers and users. They should avoid user manipulation, including approaches that encourage addiction, such as shopping, gaming and device addiction that goes beyond the limits of targeted advertising. Companies should have risk disclaimers about addiction and consider providing users with an opt out option from targeted advertising.

Accountability

Companies should establish a clear accountability system internally and externally within their AI development and application ecosystems. There should be an appropriate due diligence process for supply chains and third-party access. Companies should build systems that allow for auditability and put in place appropriate insurance where possible.

Safety

Human safety is of paramount importance, especially when it comes to access to critical services, such as water, electricity and healthcare or control of transportation such as autonomous vehicles. Companies should demonstrate that their AI applications put human safety as a priority over profit and revenue.

[2] https://assets1.dxc.technology/digital_transformation/downloads/Digital_Decisions_Survey_Report.pdf

[3] https://www.gartner.com/en/newsroom/press-releases/2019-01-21-gartner-survey-shows-37-percent-of-organizations-have

Our AI and data governance engagement framework

Hermes has been engaging with companies on data privacy and AI since April 2018. Building on our experience, we have created a framework for engagement based on the six principles. These principles can be used for opening a dialogue with a company on AI applications. We have also developed a further supporting engagement framework with two related strands. The first of these focuses on identifying material issues due to the regulatory, legal and financial impacts on companies. It follows a risk-factor assessment logic.

The second strand is a process-driven approach. This is designed for circumstances in which engagement is with more technically trained specialists and focuses on identifying salient social issues. These salient issues include potential breaches of human rights or societal norms that risk creating the most severe negative impacts through a company's activities or business relationships. Generally, companies do not intentionally set out to create negative impacts, and therefore we expect these to arise due to biases in the input, processes and network architecture that give rise to unintended outcomes that a company fails to explain or be held accountable for, hence the emphasis on these in our framework.

A risk factor assessment
Material
legal and financial outcomes

Regulatory

Counterparty

Cybersecurity

Exploitation

Operational

A process-driven approach
Salient
social impacts

Input bias

Process bias

Outcomes bias

Explainability

Oversight

Output versus outcomes

Output is the result of input and processes; outcomes are differences made and impact generated by the outputs. In an article by Hurley and Adebayo (2016)[4], the authors highlighted that by using zip codes in the US as a data input to help determine credit worthiness of individuals, the inherent bias in input data may lead to unfair lending based on ethnicity as zip codes often reflect segregated communities. The output is lending recommendations and the outcomes are potentially unfair lending practices.

Material legal and financial outcomes – a risk factor approach

- **Regulatory risks:** A company should be aware of the emerging legislation, regulation, codes of practice and governance standards, such as the emerging European Union *Draft Guidelines for Trustworthy AI*. Data privacy related issues should reference the General Data Protection Regulation (GDPR) or applicable data protection acts, e.g. California's recently enacted privacy law.
- **Counterparty risks:** A company should undertake proper due diligence of its supply and contractual counterparties, including data brokers and AI analytics providers, to avoid violation of GDPR, breach of contract or allegations of negligence.
- **Cybersecurity risks:** A company should have risk classification and cybersecurity architecture that are fit for purpose to avoid unsolicited third-party access and data theft. In the UK and EU, companies should be aware of specific cyber resilience and breach reporting laws, e.g. the UK's Network and Information Systems Regulations 2018.
- **Exploitation risks:** A company should have secured the necessary intellectual property rights and contractual licences in AI development to deliver business benefits as expected.
- **Operational risks:** A company should have thoroughly investigated and addressed execution risks that may incur liabilities, custom claims and reputational damages.

Salient social impacts – a process-driven approach

- **Input bias:** A company should have a clear rationale for each of its machine learning objectives. It should have a strong ability to identify the sources, nature and ownership of its input data, including generative or simulated data, especially if third-party data brokers or providers are involved.
- **Process bias:** We must expect some bias in any process as human beings are inherently biased. A company should be aware of the biases in each AI function with a systematic approach to demonstrating that such biases can be tested. There should be diversity, governance and challenge processes within the modelling team to identify unintended biases where possible.
- **Outcomes bias:** Data input and process generates output, but the consequences of output is what generates outcomes. Some data inputs created by human systems are inherently biased (see output versus outcomes). This may include religious and political views, cultural norms, gender and ethnic perceptions or simply the use of languages. A company should be aware of unintended outcomes biases based on data inputs affected by context and address such unavoidable biases by applying responsible AI principles, summarised in the following two areas – explainability and oversight.
- **Explainability:** Explainability breaks down the AI applications – input/process/output – into understandable steps. It is vital to building trust with users especially when outcomes are perceived as biased. Transparency reduces information asymmetry between a company and its stakeholders, and hence potential mistrust.
- **Oversight:** A company should thoroughly understand the risks and opportunities presented by AI and data governance in its emerging and transitioning business models, with the capability to identify different parties that should be responsible for the impact of AI, putting human safety as a priority over revenue and profit.

[4] https://digitalcommons.law.yale.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1122&context=yjolt

INTRODUCTION

"Why is it, that when my partner looked up holidays in France, advertisements of these holidays popped up on my Facebook news feed?" a colleague asked.

In fact, many friends have asked the same or similar questions. "When I clicked on the trailer of the movie Captain Marvel, my husband started getting updates from his Google news feed too. We don't even share the same devices," I answered. In an attempt to address these puzzling, everyday questions, we decided to map out how we can be affected by products and services that are now an integral part of our daily lives.

Privacy International's report on "How Apps on Android Share Data with Facebook (even if you do not have a Facebook account)"[5] put in simple terms how the above scenarios have been made possible. The digital footprint of technology giants can be summarised in Figure 1 below.

The report suggests that 61% of the apps tested automatically transfer data to certain social media providers the moment a user opens the app, together with a unique identifier, such as the Google advertising ID. If combined, data from different apps can paint a detailed and intimate picture of people's activities, behaviours, (hidden) interests and routines. In other words, the surveillance mechanism that powers the generation of 'big data' for deep learning and artificial intelligence (AI) has, in some ways, taken away a user's privacy without their explicit consent. This is happening every time we are active online, such as when we search for local cinema listings, make bookings for dinner through the internet, shop online and use 'free' maps for navigation. By understanding the impact of AI and data governance, we can begin to understand our wider human rights in the digital era, and hopefully reinstate them.

Figure 1: Activities of the digital giants (adapted from source[6])

	Google	Amazon	Apple	Microsoft	Snap Inc.	Facebook	Tencent	Alibaba Group	Baidu
Digital media and content development	✓	✓	✓	✓	✓	✓	✓	✓	✓
Artificial intelligence and machine learning	✓	✓	✓	✓		✓	✓	✓	✓
E-commerce marketplaces	✓	✓	✓	✓		✓	✓	✓	
Bricks and mortar	✓	✓	✓	✓					
Hardware devices	✓	✓	✓	✓	✓	✓	✓	✓	✓
IT and cloud services	✓	✓	✓	✓			✓	✓	✓
Messaging and communication	✓	✓	✓	✓	✓	✓	✓	✓	
Navigation	✓	✓	✓	✓			✓	✓	✓
Other business services	✓	✓		✓		✓		✓	✓
Payment and financial services	✓	✓	✓	✓	✓	✓	✓	✓	✓
Advertising	✓	✓	✓	✓	✓	✓	✓		✓
Operating systems	✓	✓	✓	✓			✓	✓	✓
Search engine / browsers	✓	✓	✓	✓			✓		✓
Healthcare	✓		✓	✓			✓	✓	

[5] Murgia M (2018) Popular apps share data with Facebook without user consent Financial Times 30 Dec 2018 and https://privacyinternational.org/report/2647/how-apps-android-share-data-facebook-report

[6] https://www.imd.org/contentassets/1bd6c626f9934fc4a472fa1ec0a06366/digital-giants-compared.pdf

BACKGROUND

Over the last two years, we have seen growing evidence of how targeted advertising and the application of artificial intelligence (AI) can have political and religious ramifications. This prompted our review on the appropriate use of information identifiers that aim to improve decision-making capabilities through a continuous learn and adapt process, commonly referred to as big data for machine learning, the bedrock of AI applications.

The mathematics that underpins what we call AI dates back to the 1950s and 1960s. There were two 'winters' in the early 1970s and late 1980s where research funding reduced due to criticisms, as machine learning was deemed 'overhyped'. Machine learning advanced only recently because of enhancement in computing power, optimisation techniques and data availability. In the 21st century, machine learning as AI captures our imagination.

There are three main types of learning styles – supervised, unsupervised and reinforcement[7]. In supervised learning, data is pre-labelled and classified. Methods include random forest[8] or support vector machines[9]. In the case of semi-supervised learning, desired labels may be missing in the output. For example, an algorithm may examine a photo, and identify a human being and a dog, but miss a tub of ice-cream. This is a missing label. Some mobile phones now automatically tag items as food when the user points the device camera to a food item, such as an apple. However, if the camera has not been trained on data that has a classifier of marshmallow, it will not be able to identify it. Worse still, if an unexpected item is introduced, the AI may fail altogether.[10]

Unsupervised learning uses pattern analysis and clustering methods such as principal component analysis (PCA). It draws inferences from data sets without pre-determined labels or classification. Netflix uses PCA to customise advertisements.

Reinforcement learning is an optimisation process to find the best solution for a situation. It is designed to maximise future reward over many steps, over time. For example, Deepmind's AlphaGo, which managed to beat the world Go champion Le Sedol in 2016, used the Monte Carlo Tree Search method of reinforcement learning. This type of learning has a policy that awards desirable behaviour and is often compared to the method of training an animal where behaviour favoured by the trainer is rewarded and hence reinforced. This method tends to drive the algorithm towards the criteria defined by the modeller.

Computer algorithms that learn using these methods rely on massive input data to generate relevant outcomes. Such data is composed of the personal information of users...

Deep learning is not a learning style, but it uses a many-layered neural network in an attempt to mimic human learning through the concept of gradient descent and is often used in image classification and natural language processing.[11] Using this method, an algorithm can 'see' by learning features of an image step by step, or layer by layer, called forward propagation and then back propagate to find the best fit of an image based on the features learned.

Computer algorithms that learn using these methods rely on massive input data to generate relevant output. Such data is composed of the personal information of users, from hard facts such as names, addresses and telephone numbers, to behavioural information such as habits, perceptions, opinions, likes and dislikes, daily routine preferences etc. It is unclear to the general public how data is sourced, used, and how computer algorithms work, as their 'secret sauce' provides core competitive advantages to businesses in the form of user or customer insights. The output of such algorithms often becomes data input again in a feedback loop in the case of reinforcement learning. If bias exists, such feedback loops are likely to amplify biases and create errors without detection until the algorithms are debugged or made redundant. Over the years, companies that deploy AI in their search functions or social media platforms have provided limited transparency as to why the outputs of algorithms have changed overnight, what prompted such changes, and what triggered the redundancy of an algorithm.

In general, we are concerned that companies lack the ability and determination to decipher and explain the impact of 'black box' operations (i.e. complex unarticulated or explained operations) in AI systems. Although Google published a white paper on AI governance in January 2019 to outline a plan of work[12], it is also evident from this same paper that not all the issues identified will have straightforward solutions. This is particularly true when it comes to managing social media content where one person's freedom of speech could become a provocative and discriminatory statement in the eyes of another.

[7] http://ceur-ws.org/Vol-2301/paper_2.pdf

[8] The creation of a 'forest' of possible decision trees to optimise randomness that aims to reduce data overfitting problems.

[9] The method of segregating two or more clusters of data by determining how definitive their segregation is.

[10] https://www.quantamagazine.org/machine-learning-confronts-the-elephant-in-the-room-20180920/

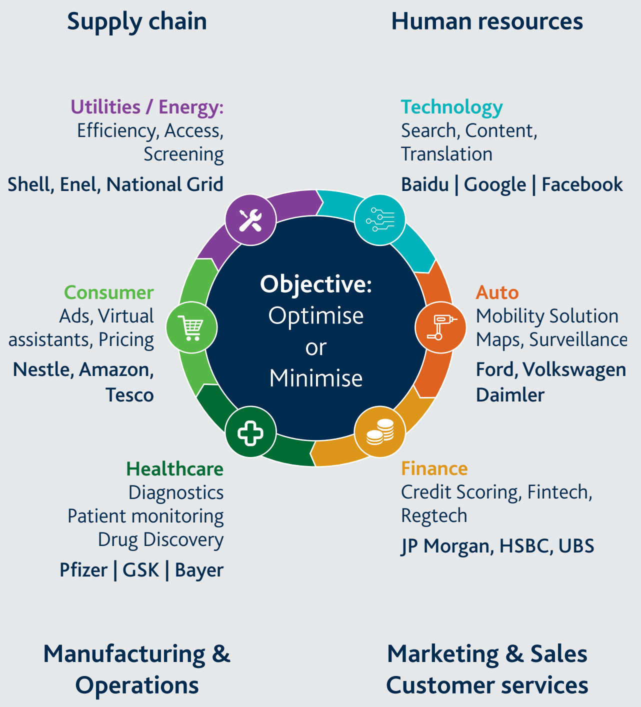
[11] We find the best analogy of describing how deep learning works is to imagine the experience of a mountain climber looking for a way out in the fog. For further information, please read: https://ml4a.github.io/ml4a/how_neural_networks_are_trained/

[12] https://ai.google/static/documents/perspectives-on-issues-in-ai-governance.pdf

RELEVANCE TO LONG-TERM INVESTORS

The examples we have shared so far have mainly focused on technology companies, but artificial intelligence (AI) applications are not restricted to this sector. Companies in other sectors use AI in their supply chains, business operations, marketing and sales, and service delivery, providing inputs to key decision-making (Figure 2 & 3).

Figure 2 Key applications of AI in different sectors and example companies


Supply chain
Human resources
Utilities / Energy:
Efficiency, Access, Screening
Shell, Enel, National Grid
Technology
Search, Content, Translation
Baidu | Google | Facebook
Consumer
Ads, Virtual assistants, Pricing
Nestle, Amazon, Tesco
Objective: Optimise or Minimise
Auto
Mobility Solution Maps, Surveillance
Ford, Volkswagen Daimler
Healthcare
Diagnostics
Patient monitoring
Drug Discovery
Pfizer | GSK | Bayer
Finance
Credit Scoring, Fintech, Regtech
JP Morgan, HSBC, UBS
Manufacturing & Operations
Marketing & Sales Customer services

The growth of AI applications in businesses is accelerating and will continue to do so as quantum and DNA computing technologies mature.

Figure 3 A summary of key applications[13]

Sector	Companies	Examples of applications
Technology	Baidu	Baidu Brain powers open AI Apollo platform
	Google	Google Search, Google Translate, YouTube
	Facebook	Content management on social media
Autos	Ford	Invested in Argo start up
	Volkswagen	Co-invests with Ford created its own AI Driving unit
	Daimler	Invested in Nvidia electric control units
Finance	JP Morgan	Treasury services and corporate payments
	HSBC	Credit scoring assistance, anti-money laundering
	UBS	Wealth advisory, IT platform enhancement
Healthcare	Pfizer	Imaging and diagnostics
	GSK	Drug discovery
	Bayer	Patient monitoring
Consumer	Nestle	Personalised nutrition
	Amazon	Hiring, targeted marketing
	Tesco	Mobile coupon app
Utilities/ Energy	National Grid	Process drone data to spot leaks and need for repairs
	Shell	Reduce cost of exploration, enhance precision drilling
	Enel	Identify faulty solar infrastructure

The growth of AI applications in businesses is accelerating and will continue to do so as quantum and DNA computing technologies mature. According to an Economist Intelligence Unit[14] survey of more than 600 senior executives worldwide, 36% said that AI and machine learning had played a significant role in their organisation's digital strategy. Some 45% saw AI and machine learning as the most important technology to play a significant role in their organisation's digital strategy in the next three years. A survey by Gartner, a global technology consulting firm, showed that 37% of organisations had already implemented AI in some form; the number of organisations implementing AI grew 270% in the past four years and tripled in the past year. The social impact of AI and data governance is therefore intricately linked to the long-term value creation of companies. As a recent article stated:

> **A central promise of AI is that it enables large-scale automated categorisation. Machine learning, for instance, can be used to tell a cancerous mole from a benign one. This 'promise' becomes a menace when directed at the complexities of everyday life. Careless labels can oppress and do harm when they assert false authority.**[15]

[13] Hermes EOS 2019 analysis

[14] http://www.dxc.technology/digital_transformation/insights/146023-2019_the_year_of_digital_decisions

[15] Penn J (2018) AI thinks like a corporation—and that's worrying The Economist 26 Nov 2018

With growing tensions between government and businesses on data privacy, and an erosion of trust between consumers and businesses, we expect the long-term value of companies to be impacted by regulation and societal expectations of responsible business behaviour.

This takes us back to the animal training analogy. For example, if a dog is trained to bite humans, and is rewarded for it consistently, this will become its normalised behaviour.

Despite the risks, scientific breakthroughs using AI in the fields of healthcare, autonomous driving, agriculture, manufacturing, and climate change solutions (such as increased energy efficiency and electricity grid optimisation) demonstrate tangible social benefits. We have entered a period where academic and applied research are co-developing alongside each other at an unprecedented speed. AI may enable better predictions to be made in healthcare and improve the processes for drug discovery and application. This will help to reduce wasted research and clinical trial failures. AI adoption also promises to improve product quality and make manufacturing processes more efficient. Targeted advertising improves marketing, cross- and up-selling. Automated compliance and operations management reduces the cost burden on institutions and helps compliance or engineering professionals to focus on more value-adding tasks. The relevant research and applications should continue, albeit in a more transparent, accountable, fair and regulated manner.

On the other hand, allowing data and AI to be controlled by a small group of firms is likely to pose a long-term threat to society, triggering calls for more regulation.[16] In February 2019, the UK Parliament published the conclusion of an inquiry on disinformation. It announced plans to regulate social media platforms, mandating a code of ethics overseen by an independent regulator. The measures are targeted at monitoring any anti-competitive practices, politically-motivated voter manipulation and user privacy.[17] The public consultation of the Online Harms White Paper was published in April 2019. In the US, there are stronger voices over breaking up the big technology incumbents.[18]

With growing tensions between government and businesses on data privacy, and an erosion of trust between consumers and businesses, we expect the long-term value of companies to be impacted by regulation and societal expectations of responsible business behaviour. The responsible use of AI will, in our view, become the new social licence to operate. This is supported by academic and industry collaborative research that has identified at least six scenarios that show AI has human rights impacts, using the United Nations Guiding Principles for Business and Human Rights as a reference framework.[19]

OUR RESPONSIBLE AI AND DATA GOVERNANCE PRINCIPLES

Institutional investors are well placed to engage on this topic through the lens of stewardship and long-term sustainable development. When establishing engagement objectives, we test companies against the following principles:

Trust

Companies should earn trust by educating users on their rights to data privacy and give users control and the right to consent to the use of their data by providing fully free choices.

Transparency

Companies should be transparent about tracking methods in the full value chain and disclose how they measure the robustness of data governance and the fair and safe use of AI. Companies should inform users when their data is being used for scoring and screening purposes.

Action

Companies should thoroughly explore and make all reasonable efforts, in good faith, to avoid unintended consequences such as data and process bias, which may lead to discrimination.

Integrity

Companies should demonstrate integrity in the treatment of customers, suppliers and users. They should avoid user manipulation, including approaches that encourage addiction, such as shopping, gaming and device addiction that goes beyond the limits of targeted advertising. Companies should have risk disclaimers about addiction and consider providing users with an opt out option from targeted advertising.

Accountability

Companies should establish a clear accountability system internally and externally within their AI development and application ecosystems. There should be an appropriate due diligence process for supply chains and third-party access. Companies should build systems that allow for auditability and put in place appropriate insurance where possible.

Safety

Human safety is of paramount importance, especially when it comes to access to critical services, such as water, electricity and healthcare or control of transportation such as autonomous vehicles. Companies should demonstrate that their AI applications put human safety as a priority over profit and revenue.

We have entered a period where academic and applied research are co-developing alongside each other at an unprecedented speed.

[16] Kelion L (2018) UK PM seeks 'safe and ethical' artificial intelligence BBC News 25 Jan 2018; Lucas L (2018) Singapore to develop code for ethical use of AI and personal data Financial Times 13 June 2018.

[17] https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/179102.htm

[18] https://www.theguardian.com/commentisfree/2019/mar/09/elizabeth-warren-break-up-facebook-google-amazon

[19] https://cyber.harvard.edu/publication/2018/artificial-intelligence-human-rights

RESPONSIBLE AI AND DATA GOVERNANCE ANALYTICAL FRAMEWORK

Hermes has been engaging with companies on data privacy and artificial intelligence (AI) since April 2018, based on the six principles, which can be used for opening a dialogue with a company on AI applications. Building on this experience, we have created a further framework, based on two strands, which can be used in deeper engagement conversations (Figure 4). The first strand focuses on materiality of issues due to the regulatory, legal and financial impacts on companies. It follows a risk-factor assessment logic.

Generally, companies do not intentionally set out to create negative impacts, and therefore we expect these to arise due to biases in the input, processes, network architecture which gave rise to unintended outcomes.

The second strand of the framework is a process-driven approach. This is designed for circumstances in which engagement is with more technically trained specialists responsible for creating and monitoring AI systems. This strand focuses on identifying salient social issues. Salient issues are potential breaches of human rights or societal norms that risk creating the most severe negative impacts through a company's activities or business relationships. Generally, companies do not intentionally set out to create negative impacts, and therefore we expect these to arise due to biases in the input, processes, network architecture that give rise to unintended outcomes that a company fails to explain or be held accountable for, hence the emphasis on these in our framework.

As an introduction to the material legal and financial outcomes section below and elsewhere in this paper, **BCLP partner Mark Lewis explains as follows**.

> For me, the starting point in risk assessing an outcome resulting from the deployment of any technology, relatively new and fast developing (like AI) or otherwise, is to ask: what is its functionality or application? Without functionality in context, for example, AI in driverless vehicles, assessing specific outcome risks is challenging. Nevertheless, we need a generic framework as a starting point. This is what the framework in Figure 4 is designed to achieve.
>
> Certain sectors, including financial services, healthcare, technology and utilities, will of course raise particular regulatory considerations that apply specifically to those sectors, such as rules and guidance on the use of AI and algorithms in the financial services and health sectors, or cyber resilience and cyber breach handling regulations in the essential and digital services sectors (see Appendix 2 for an outline of the UK Network and Information Systems Regulations). Companies operating in those sectors will therefore contend with heightened legal risk and severity profiles, just for being there. Throughout, I refer to those considerations and profiles in what I hope is sufficient, through necessarily high-level, detail.
>
> I have tried to be as practical as possible, focusing on the "here and now", but also on the realistic and foreseeable regulatory and legal horizons. Essentially, this means putting oneself into the position of thinking and prospectively acting as good businesses might when contemplating or deploying AI, either alone or with others as counterparties or within supply chains. This has meant avoiding – for now at least – fascinating philosophical considerations, such as how cognitive AI might be programmed or teach itself to reach decisions based on human or other ethical computations (see http://moralmachine.mit.edu for conundrums in driverless car AI). So what follows also outlines the right normative corporate behaviours in the face of current and foreseeable legal and regulatory risks.
>
> I have considered and applied law and regulation as each applies to businesses based and/or operating in the UK, with references to EU laws and regulations where applicable. For now, it is beyond the scope of this exercise to consider and refer to the many other laws, regulations, ethical codes and governance frameworks for AI that may impact businesses based or operating in the UK, providing goods and/or services to international markets.
>
> Mark Lewis, BCLP, partner

Figure 4 AI governance analytical framework[20]

1. Material legal and financial outcomes	Risks	Details
1.1 Current and emerging regulatory, code of conduct and governance frameworks	Regulatory risks	1. There are various initiatives and proposals around the world for legislation, regulation, ethics, codes of practice and governance standards applying to AI at supranational (e.g. EU High-Level Expert Group on AI, Draft Ethics Guidelines for Trustworthy AI), governmental, parliamentary and even at corporate level. While these specific AI initiatives do not yet have the force of law, any company deploying AI directly or through counterparties and/or supply chains should be aware of the prospect and progress of legislation, regulation, ethics and governance standards applying to AI in its sector and to its operations. Failure to do so would be a risk in itself. 2. As corporate governance and best industry practice emerge, failure to adhere to what may be found to be good industry practice or good corporate governance standards in the use of AI could result in legal claims from regulators, consumers and others. 3. In the context of AI deployment, General Data Protection Regulation (GDPR) in the EU and the UK Data Protection Act will be relevant and apply to companies, both in their bricks-and-mortar and online operations. The risks and penalties of failure to comply with GDPR are, or should be, well-enough known. 4. In the broader context of AI deployment, there are other related regulations that will apply to companies operating in certain sectors, e.g. the UK Network and Information Systems Regulations 2018 (NIS) that apply to, amongst others, companies providing specified essential services. 5. All companies should understand the impact and outcomes of deploying AI in their businesses – and before doing so. Failure to do so must be considered a risk, with the possibility of employee, customer and supplier claims with resulting liability, e.g. at an operational level for unfairness and/or discrimination in recruitment processes under gender, ethnicity, race or similar equalities legislation. This last example is cited to illustrate how the deployment of AI even at a relatively low operational level, i.e. in recruitment processes in any corporate, could lead to claims that in turn might result in material liability. 6. In summary, a failure by any company and its senior management to adequately take account of the above risks and impacts and to act to prevent them could expose it and senior management to regulatory enforcement action and liability and ultimately reputational loss and damage which may be considerable, if not existentially threatening.
1.2 Counterparty and supply chain due diligence, including data brokers, cloud brokers and data analytics providers	Counterparty risks	1. If a company fails to undertake proper due diligence of its supply chain/contractual counterparty, it could find itself inadvertently to be in breach of regulation (e.g. GDPR) and/or be liable to customers or other parties negligence and/or breaches of contract. 2. Similarly, if a company fails to protect itself contractually in its arrangements with its supply chain or counterparties, it may be unable to manage the allocation of risk and liability to those counterparties, customers and third parties, including GDPR and other regulatory risks and claims
1.3 Third party access / data theft / cyber resilience	Cyber risks	1. Investors will expect companies to have risk classification and cybersecurity architecture that are fit for purpose to avoid unsolicited third party access and data theft. 2. If a company suffers a cyber breach (third party gaining access to personal data used in the course of AI processing and/or theft of personal data in those processes) and was found not to have taken all steps required under GDPR to secure its networks or systems and therefore cyber resilience, it would be in breach of its GDPR obligations and could face enforcement measures (fines, etc) by the data protection authorities, as well as compensation claims by data subjects. Such penalties could be draconian, and the quantum of compensation claims material. The company would also, of course, suffer reputational damage. Companies providing essential services (e.g. drinking water and distribution, power and transportation) as well as certain digital services providers are subject to additional (i.e. additional to GDPR) network and information systems resilience and reporting obligations under the NIS Regulations 2018. 3. In addition, under GDPR and (where applicable) NIS, in the event of a cyber breach, companies must generally report it to the relevant authorities within 72 hours and may be required also to inform data subjects. Breach of these reporting obligations, or the mishandling of them, could subject the company to further liability and reputational damage.
1.4 The protection of intellectual property (IP) and IP ownership	Exploitation risks	1. Failure by a company to secure by contractual assignment or other means the necessary IP ownership rights in AI development (e.g. algorithms) could leave it without the commercial rights to exploit intellectual property rights in the AI in which it had invested. 2. Failure by a company to obtain by contractual licence (e.g. a software licence, as opposed to ownership) suitable (e.g. as to scope, permitted users) rights to deploy AI technologies would deprive it of the right to use the AI, face supplier/third party IP claims (including for breach of copyright) and be prevented from using the AI in necessary processes, so affecting its operations and creating an additional business risk, e.g. the inability to process customer transactions because of illegality of software/algorithm use. 3. Failure by a company to undertake adequate investigations/due diligence into permitted usage of its existing software estate in conjunction with AI technologies could deprive it of the right to use the AI, face supplier/third party IP claims (including for breach of copyright) and be prevented from using the AI in necessary processes, so affecting its operations and creating an additional business risk, i.e. inability to process own-account or customer transactions because of illegality of software/algorithm use.

[20] The following framework is provided by Mark Lewis of BCLP and is valid as of April 2019. The legal and regulatory frameworks throughout this paper as provided by Mark contain only a general summary of the law and/or regulation, and only general ideas and/or guidelines. It does not contain, and should not be relied on in any manner as, legal or professional advice. No statement of law or regulation in this document is intended to be, or should be relied on as, comprehensive. You should always seek specific legal and/or compliance advice on the particular facts of any given situation.

1.5 Best/safe standards; securing commercial benefits; liability, remedy, penalty, reputational risks and insurance availability	**Operational risks**	Failure by a company adequately to contract with a customer/provider/supplier: 1. To ensure compliance by the customer/provider/supplier with applicable laws), could result in regulatory liability for the company and significant third-party liability, including corporate and (possibly, depending on where it is in the supply chain) consumer customer claims, also (depending as mentioned) under consumer protection legislation, in contract and for negligence or other torts. 2. To capture contractually the commercial benefits and outcomes (e.g. through KPIs, detailed product/service specifications or specified business outcomes) could leave the company liable to pay contractual charges without achieving the mandated/desired business outcomes, including loss of market opportunity and possible loss of reputation. 3. To allocate properly execution risk and liability as between the company and its customers/providers/suppliers could leave it without adequate legal recourse against the customer/provider/supplier while facing the risk of failure of the project and exposure to third party customer and even (depending on where it is in the supply chain) consumer protection claims, actions and liabilities. 4. To investigate the insurability of the contract risks as allocated to the company and to put in place, where available, suitable cover or require the customer/provider/supplier to do so, could leave the company's balance sheet exposed to uninsured claims and settlements. All of the above failures by a company could expose it to regulatory enforcement action (including under GDPR, but also (depending as mentioned) consumer protection and product liability claims) and other liability and ultimately reputational loss and damage which may be irrecoverable.

2. Salient social impacts[21]	Risks	Details
2.1 Input bias	**Data set bias**	1. A company should have a clear rationale for its machine learning hypotheses or objective functions[22] for each of the AI analytical tasks it seeks to undertake. The Model card documentation framework developed by Google is a good practice example.[23] 2. A company should have a strong ability to identify the sources, nature and ownership of its input data. If a company uses third party data brokers, there should be sufficient due diligence before the beginning of a contractual relationship, with an ongoing monitoring mechanism to ensure that there is transparency in the way the data is collected, cleaned and inputted into any AI systems. 3. A company should have records of viable reasons when changing data attributes, as input bias is the most significant step in creating intended and unintended biases in machine learning and AI applications. When generative or simulated data is used, any bias that could be amplified should be taken into accounting for testing.
2.2 Process bias	**Classification / clustering bias**	1. For all three types of learning, there must be clear machine learning hypotheses or objective functions. 2. The algorithm modeller should be fully aware of the biases (but are not necessarily required to demonstrate that their models are unbiased), with a systematic approach to demonstrate that such biases can be tested. 3. Note that more complicated controls for algorithm testing may introduce more biases and may create more moral hazards as the modeller(s) may no longer be fully responsible for the outputs of the algorithms. There should be diversity, governance and challenge processes within the modelling team to identify unintended biases where possible. 4. For unsupervised learning, the challenges of cluster analysis apply, which include overfitting, how to determine outliers, and the shape and the number of clusters in a data group. Companies should be prepared to address these challenges with reasonable efforts.
2.3 Outcomes bias	**Discrimination exclusion allegations**	1. Some data inputs created by human systems are inherently biased. This may include religious and political views, cultural norms, gender and ethnic perceptions. 2. In natural language processing, algorithms could integrate biases in the human use of languages. Modellers should be aware of these unintended outcomes based on data inputs affected by context and any process biases. A company can address such unavoidable biases by applying responsible AI principles that focus on ensuring explainability and oversight.
2.4 Explainability	**Trust** **Transparency** **Action**	1. Explainability breaks down the AI applications – input/process/output – into understandable steps. It is vital to building trust with users especially when outcomes are perceived as biased. Transparency reduces information asymmetry between a company and its stakeholders, and hence potential mistrust. 2. Aside from GDPR obligations, a company should make all reasonable efforts (take action) to provide users with genuine choices in relation to personal data access and applications. Any tracking methods should be adequately disclosed with proper consent mechanisms in place.
2.5 Oversight	**Integrity** **Accountability** **Safety**	1. A company should oversee development of a thorough understanding of the risks and opportunities presented by AI and ensure data governance in its emerging and transitioning business models. The insurability and auditability of AI systems should be assessed. We expect board level oversight for the responsible use of AI and that internal governance mechanisms are established to support that. 2. A company should have an appropriate due diligence process in place for its supply chain and third-party access. A company should make explicit statements with demonstrable efforts to avoid user manipulation that puts revenue and business objectives ahead of human safety. 3. Safety also includes ensuring that AI systems generate outputs and outcomes as intended, as per the machine learning hypotheses or objective functions under all circumstances.

[21] The following part of the framework was created by Hermes EOS and is considered work in progress to account for ongoing development in consultation with stakeholders and our engagement work.

[22] An objective function is a function that seeks to maximise or minimise an output. For example, in a production function, resources and labour are two inputs that will combine to produce outputs. An objective function may be to maximise production output or to minimise labour costs.

[23] Page 14 Box 6 on Model and Data Cards, Google's AI Perspectives White Paper Jan 2019: https://ai.google/static/documents/perspectives-on-issues-in-ai-governance.pdf

SECTOR APPLICATIONS

Technology

The EU's General Data Protection Regulation (GDPR), which came into effect in May 2018, has significantly enhanced public awareness of data privacy, amongst other things, as a human right.

Facebook has developed a range of analytics tools for third-party businesses enabling them to track the personal information of their users or customers, even if they are not Facebook users.[24] At this point, it is important to understand the roles played by different service providers in the data governance chain. In general, a data controller is the person who, alone or jointly with others, determines the purposes and the means of the processing of personal data. If an app sends personal data elsewhere for processing, then it is likely that the business that provides this app service may be the data controller. A data processor is the person who processes personal data on behalf of a data controller. Alphabet's Google classifies itself as a processor for users of tools like Google Analytics, which has integrated previously developed and acquired products such as Ads Data Hub and DoubleClick Bid Manager. Data processing tools such as Facebook Analytics and Facebook Insights enable businesses to create targeted advertisements based on users' 'likes', social media posts and other in-group interactions and measure the performance of different outreach channels.

App store operators such as Apple and Alphabet's Google do not require apps to disclose to their partners all the information of their users, and users can choose not to disclose contacts and location, but information submitted to apps directly, such as health information for tracking devices, is shared by the data controllers with the data processors. The analytics functions of technology firms are in general only interested in collecting information to train their algorithms, as societal behaviour, values and 'likes' continue to evolve. During this process, however, a company could become powerful in influencing and pre-empting users' behaviour if information is presented in a way that triggers an emotional response by users. For example, a company could become manipulative in encouraging excessive paid-for gaming behaviour when a user has established a clear pattern in entertainment preferences.

Recently reported technology events have demonstrated the impact of artificial intelligence (AI) outcomes on the two strands that we have identified – legal materiality and social salience. In one case, Cambridge Analytica, the political consultancy, was revealed to have harvested personal data from millions of people's Facebook profiles without their consent.[25] In December 2018, Facebook was fined €10 million (£8.9 million) by Italian authorities for misleading users over its data practices,[26] which dwarfed the £500,000 fine levied by the UK Information Commissioner's Office in September over the company's failure to protect user data.[27]

In November 2018, a coalition of seven consumer organisations filed complaints with local data protection regulators over Google's tracking system.[28] In January 2019, the French data protection authority, CNIL, fined Google €50 million for failing to provide users with transparent and understandable information on its data use policies, and processing information for personalised advertisements accessible by third-party businesses.[29] Clearly, such incidents are not conducive to building public trust in how personal data is being used and possibly exploited. Many businesses that are part of the social media and data analytics ecosystem are likely to need to raise their game to ensure proper governance is in place to manage third-party access, as regulations have started to catch up with reality. After an 18-month investigation, the Digital, Culture, Media and Sport Select Committee of the UK Parliament concluded in its final report that social media companies should be required to remove inappropriate content on their platforms and should operate according to a code of ethics, overseen by an independent regulator.[30] The public consultation of the Online Harms White Paper was published in April 2019.

€10M

fine to Facebook from by Italian authorities for misleading users over its data practices

€50M

fine to Google from the French data protection authority, CNIL

The social impact of AI is more complicated than its legal, regulatory and financial impact. Regulators have started to get tough on technology companies over the uses and abuses of personal data, and failures by social media platforms to police the way that third parties, use them to manipulate others. As social media companies employ more human content reviewers to screen for violence and other inappropriate content, it is reasonable to be concerned about the mental health impact on its workforce, in addition to the high turnover of contract staff, consistency of performance, working conditions and personal safety.[31] A few questions immediately come to mind:

- Have the reviewers been adequately trained on conscious and unconscious bias?
- Does the company expect reviewers to understand and be able to adapt to daily amended community standards and remain consistent in their judgement?
- What is the benchmark for judging whether a comment is discriminatory or offensive?

These should be part of the investor engagement conversation.

[24] https://www.wsj.com/articles/you-give-apps-sensitive-personal-information-then-they-tell-facebook-11550851636?emailToken=80a84ecb53d239afbbdc3668e642e71eCq/XWYSaoG3Bb/23L+pDMvqf41UzmEDEX2

[25] https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

[26] https://www.theguardian.com/technology/2018/dec/07/italian-regulator-fines-facebook-89m-for-misleading-users

[27] https://www.theguardian.com/technology/2018/oct/25/facebook-fined-uk-privacy-access-user-data-cambridge-analytica

[28] https://www.reuters.com/article/us-eu-google-privacy/european-consumer-groups-want-regulators-to-act-against-google-tracking-idUSKCN1NW0BS

[29] https://www.theguardian.com/technology/2019/jan/21/google-fined-record-44m-by-french-data-protection-watchdog

[30] https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/179102.htm

[31] https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona

Companies must have appropriate governance and accountability frameworks in place to reassure consumers that they take such risks seriously.

The accidental deletion of YouTubers playing Pokemon Go due to, allegedly, the algorithms having mistaken "CP" as child pornography rather than combat points has prompted criticisms from users.[32] This follows growing discontent amongst users and viewers about the way YouTube is run. Google has a tradition of creating a remix of all the most popular YouTube videos every year in a short clip named the Rewind video. The 2018 version was the most disliked video in YouTube's history as voted for by viewers, who criticised it for failing to genuinely represent the best creators, many of whom are amateurs. They also felt it did not respect the tradition and values of YouTube as a grassroots community, instead favouring large companies with access to production budgets and expertise.[33] Again, a few questions came to mind:

- Is there a transparent process for selecting YouTubers that should appear in the annual celebration video?
- What is the accountability system involving AI-human interactions for content review?
- Has the company provided a timely and adequate explanation of the incident and what process and procedures are put in place to avoid such 'errors' in the future?
- Under what circumstances would the company be liable to penalties? Is this waived or covered already in the YouTuber agreement?

Using the analytical framework in Figure 4, we conducted a preliminary assessment of AI's impact for technology companies (Appendix 1).

Consumer and retail

In the age of AI, consumer and retail companies see retaining customer trust as a priority.[34]

Their reputation and brand rely on customers trusting that the company will operate with integrity and be accountable for its actions. The use of predictive analytics to assess how consumers shop provides a significant opportunity for consumer and retail companies to improve their operational efficiency and strategic approach. However, without careful application, consumers may question companies' motives. To maintain trust whilst reaping the benefits of AI, companies must respect consumer privacy and demonstrate how AI can deliver benefits for their customers, rather than being a tool for manipulation and personal data capture.

A reliable way to maintain trust is by being transparent. Companies must do more to disclose how they use AI in their operations and how it may affect customers.[35] For instance, many retail companies use customised emails to prompt customers to buy more products. When used excessively or intrusively, the method can lead to customer frustration and dissatisfaction. Instead, companies can gain insights into how AI can be used safely, fairly and effectively by engaging with a broad range of stakeholders, including customers. Companies must assure customers that their personal data will be treated with care, rather than sold on to unsolicited parties. Furthermore, companies can educate and empower their consumers to opt out of AI and targeting schemes.

Another way to maintain trust is by demonstrating a commitment to fairness and equality. Dataset bias is a major problem facing consumer and retail companies. Having an inaccurate view of the consumer base may limit product development to particular demographics and risk categorising consumers based on inaccurate stereotypes.[36] In online employment advertisements, there have been cases of higher-paid jobs being shared disproportionately with men, as well as gendered assumptions about suitability to roles being exacerbated.[37] Companies must gain greater understanding and control over which datasets are fed into algorithms to avoid exacerbating historic biases and social injustices. According to the Institute of Business Ethics, doing so involves conducting due diligence on third party data and algorithm providers. Equally, companies must have appropriate governance and accountability frameworks in place to reassure consumers that they take such risks seriously. There should be a clear escalation procedure for assigning responsibility for when something goes wrong during AI applications.

Financial services

Financial institutions increasingly use AI in third-party market research, process automation, payment services, trading and devising investment strategies, customer profiling and credit analysis, robo-advisers for wealth management services, and recruitment.

We need to be cautious of the potential risks and implications of AI applications in financial services. In November 2017, the Financial Stability Board published a paper entitled Artificial intelligence and machine learning in financial services, highlighting concerns over AI's opaque model and the importance of having expertise in AI oversight and audit.

[32] https://www.techspot.com/news/78814-youtube-bans-several-pokmon-go-channels-over-mistaken.html

[33] https://qz.com/1495042/youtube-2018-rewind-youtube-made-the-worst-youtube-video-ever/

[34] https://www.forbes.com/sites/kenkrogue/2017/09/11/artificial-intelligence-is-here-to-stay-but-consumer-trust-is-a-must-for-ai-in-business/#c7a6bf5776e7

[35] https://sloanreview.mit.edu/article/every-leaders-guide-to-the-ethics-of-ai/

[36] https://www.ibe.org.uk/userassets/briefings/ibe_briefing_58_business_ethics_and_artificial_intelligence.pdf

[37] Raso, Filippo and Hilligoss, Hannah and Krishnamurthy, Vivek and Bavitz, Christopher and Kim, Levin Yerin, Artificial Intelligence & Human Rights: Opportunities & Risks (September 25, 2018).

Customer profiling and credit scoring

There are challenges associated with big data credit-scoring tools in banking and insurance.

In the Yale Journal of Law and Technology, Hurley and Adebayo (2016) highlight some of these. They include: insufficient transparency about the data collection and transformation process, as credit-scoring companies treat their sources as proprietary trade secrets; input data that is potentially inaccurate, as companies collect alternative data from various sources; the potential for biased and discriminatory scoring, as some inputs involve human interpretation; and; the risk that these tools will be used to target vulnerable consumers, which can further entrench discriminatory lending patterns. These challenges can lead to an inaccurate assessment of an individual's repayment ability, resulting in unfair lending.

Opaque algorithms can also result in a potential breach of the existing regulations relating to fair lending. Federal laws in the US prohibit lenders from directly taking sensitive characteristics such as race or gender into account when making lending decisions. However, it is difficult to guarantee that machine learning processes do not indirectly consider such characteristics, even when they are not directly designated as input values. For example, research findings show that zip codes in the US can be a close proxy of ethnicity.[38]

The use of big data in customer profiling and credit scoring raises concerns about compliance with data privacy regulations. GDPR in the EU empowers individuals with the right of information and access, the right of rectification, the right of portability, the right to be forgotten, the right to restriction of processing and the right to restriction of profiling. It stipulates that consent must be given in an intelligible and easily accessible form, with the purpose for data processing attached to that consent and with provisions to rescind that consent. It is unclear how easily individuals can now opt out of sharing data for customer profiling. It is also unclear whether opting out will affect individuals' credit scoring, which will affect their eligibility for credit-based products, such as loans, and the pricing of insurance products sold to them. In human capital management, banks such as Goldman Sachs and DBS are increasingly using AI to screen job applicants.[39] They may be exposed to the same issues faced by companies as described in the consumer and retail section.

More asset managers[40] are attempting to deploy AI techniques in asset selection. However, in doing so, they also have to uphold their fiduciary duty to ensure the appropriateness of their investment strategy, and that this is not based on a random pattern identified from big data. In addition to the above, the use of third party fintech (financial technology such as mobile payments) and regtech (regulatory technology that might be deployed for legal and compliance functions) should both be subject to stringent due diligence and testing before implementation.

Healthcare and pharmaceuticals

AI has massive potential to improve healthcare outcomes and pharmaceutical applications.

Perhaps the most highly anticipated application of machine learning is the potential to revolutionise drug discovery and development. Many have criticised the pharmaceutical industry recently for a lack of new innovation[41] and the average capitalised research and development cost per new molecular entity has increased 148% from US$1.1 billion in the 1990s to US$2.6 billion today.[42] This remains one of the largest barriers to success in the industry.[43] However, the promise of AI is that it will advance the development of high-value precision medicine, whilst at the same time reducing research and development costs. This is because it is able to analyse massive datasets in much less time than that taken by human researchers, with Intel[44] estimating that AI could cut drug development costs by up to 50%.

50%
AI could cut drug development costs in half

148%
research and development cost increased

The biopharmaceutical industry has seen a rapid growth in AI-based start-ups focused on drug discovery, largely due to the availability of big data in life sciences and the rapid progression in deep neural networks. In 2016-2017 a number of significant AI-big pharmaceutical collaborations were announced, including Pfizer and IBM Watson, Sanofi Genzyme and Recursion Pharmaceuticals, and GSK and Exscientia. However, so far as we are aware, there are still no AI-inspired US Food and Drug Administration-approved drugs on the market. Despite the opportunities AI may offer, we must be cautious and cognisant of the potential negative social impacts.[45]

Perhaps the biggest social risk stemming from AI applications in healthcare and pharmaceutical research is the potential to incorporate, entrench and amplify existing economic and social biases in healthcare. If data shows that poorer patients do worse after receiving chemotherapy for end-stage cancer, and are less likely to benefit from further treatment, should AI recommend against it[46]? This will lead to further disparities in our already unequal healthcare systems.

Additionally, AI might work less well where data is scarce or more difficult to collect or render digitally, which could affect people with rare conditions, or underrepresented groups in clinical trials and research data.[47] For example, historical data shows that women experience different symptoms from men when having a heart attack and may be disproportionately underrepresented in AI input data. This may affect diagnostic outcomes and hence the effectiveness of treatment for female patient groups.[48]

[38] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2892349

[39] https://www.computerweekly.com/news/252443238/Bank-uses-AI-to-select-job-candidates

[40] https://www.scmp.com/business/money/money-news/article/2182593/fidelity-looks-ai-future-portfolio-management-asset

[41] Huseyin N, et al. (2015). Why the drug development pipeline is not delivering better medicines BMJ 2015; 351 :h5542

[42] Tufts CSDD Briefing, 2014

[43] https://www.jpmorgan.com/commercial-banking/insights/ai-revolution-drug-discovery

[44] Chamraj H. and Ambert K. How AI will revolutionize precision medicine, Intel Artificial Intelligence Products Group (March, 2018)

[45] https://www.forbes.com/sites/forbestechcouncil/2017/08/03/artificial-intelligence-in-drug-discovery-a-bubble-or-a-revolutionary-transformation/#1dba1da24494

[46] https://www.nytimes.com/2019/01/31/opinion/ai-bias-healthcare.html?mod=djemAIPro

[47] http://nuffieldbioethics.org/project/briefing-notes/artificial-intelligence-ai-healthcare-research

[48] https://www.nytimes.com/2019/01/31/opinion/ai-bias-healthcare.html?mod=djemAIPro

Whilst there may not be immediate solutions to systemic issues of data bias within the healthcare system, we could ask companies and stakeholders to regularly monitor the output of the algorithms and downstream consequences for bias, and to have appropriate governance and accountability frameworks in place to begin addressing the identifiable issues.

Utilities

As the utility sector evolves and becomes more complex, AI will play an important role in ensuring the reliability of critical infrastructure.

The transition to a low-carbon economy, decentralisation of the energy system and the greater need for customer engagement are key challenges for the sector, all of which AI will play a central role in addressing.

The incremental approach to AI integration, the perceived lack of sensitive data, and the heavily regulated nature of the sector may not trigger the same alarms as other sectors, where the application of AI may seem more disruptive. However, with so many key services reliant on the stable provision of energy, the scale of the potential risk makes the responsible application of AI an area of focus. The prevention of unintended consequences will be crucial to maintaining this stability. With so many potential applications within the value chain, it may be the case that there are competing interests, which occurring in unison, pose risks to reliability.

The intermittency of renewables creates challenges for the management and distribution of energy. AI, such as that being trialled by Google's DeepMind with National Grid[49], will assist in load forecasting and the prediction of supply and demand peaks, and enable greater efficiencies to be realised. Customers becoming energy producers, otherwise known as prosumers, will need AI to control their own energy management, but also to participate in smart contracts that sell off surplus energy at desirable price points. These, and the many other application examples of the nexus between AI and the internet of things, offer significant opportunities for the optimisation of the system. Nevertheless, this infrastructure also creates vulnerabilities for the grid. And an over-reliance on AI may expose the system to cyberattacks. As mentioned earlier, as providers of essential services, utilities in the UK will be subject to the NIS and have to comply with these cyber regulations.

At the retail end of the value chain, utilities may play a much greater role within the homes of consumers, providing enhanced services for customers through demand-side management. This could include shutting down unused appliances, tracking energy demand and optimising use to lower prices. All these applications will use increasingly sensitive sources of personal data, exposing utilities to previously unencountered risks such as how the AI is using the data and/or who it may be distributed to. Where utilities are publicly owned, the increased data collection and interpretation may pose further concerns to the privacy rights of customers.

CONCLUSION AND NEXT STEPS

This paper aims to provide the rationale and evidence that artificial intelligence (AI) governance is an emerging yet critical ESG consideration in responsible investment. We decided to spend a reasonable amount of time explaining the different types of machine learning and their characteristics because we believe this will strengthen the readers' ability to formulate specific and relevant engagement questions.

As regulations evolve, and the legal framework for AI strengthens, more companies applying machine learning in different components of their businesses will become more open to engagement on AI. As business models are transformed by AI, our engagement with companies on these issues will only intensify.

As our knowledge accumulates through positive interactions and open dialogues with companies, we hope to amplify the positive outcomes of AI and support the efforts of regulators, companies and civil society to minimise the negative consequences. This, we believe, shall be the commitment of all responsible investors and shareholder representatives.

The transition to a low-carbon economy, decentralisation of the energy system and the greater need for customer engagement are key challenges for the sector, all of which AI will play a central role in addressing.

[49] https://www.ft.com/content/27c8aea0-06a9-11e7-97d1-5e720a26771b

APPENDIX 1 – AN EMERGING POLICY ENVIRONMENT

Governments and regulators have shown an increased interest in the governance of artificial intelligence (AI) applications especially in 2018 (Figure 5), but the policy framework governing the responsible use of AI is only emerging (Figure 6).

Figure 5 An overview of national AI strategies (adapted from source[50])

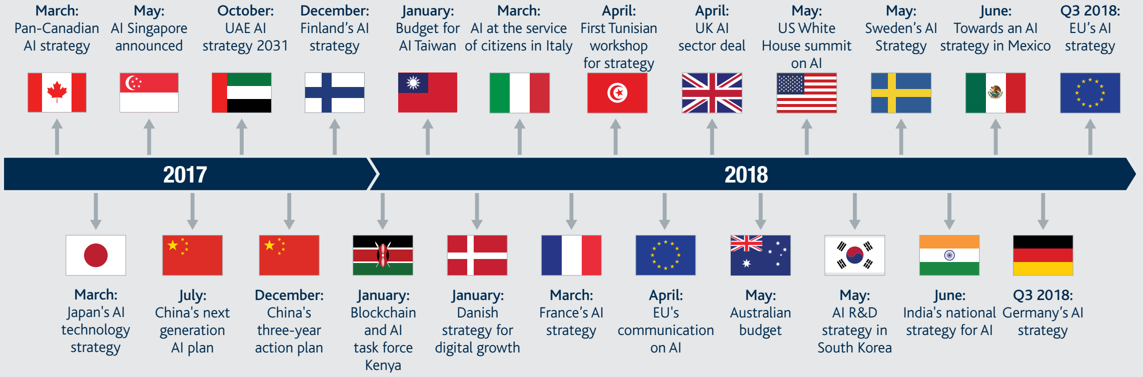

March: Pan-Canadian AI strategy
May: AI Singapore announced
October: UAE AI strategy 2031
December: Finland's AI strategy
January: Budget for AI Taiwan
March: AI at the service of citizens in Italy
April: First Tunisian workshop for strategy
April: UK AI sector deal
May: US White House summit on AI
May: Sweden's AI Strategy
June: Towards an AI strategy in Mexico
Q3 2018: EU's AI strategy
2017
2018
March: Japan's AI technology strategy
July: China's next generation AI plan
December: China's three-year action plan
January: Blockchain and AI task force Kenya
January: Danish strategy for digital growth
March: France's AI strategy
April: EU's communication on AI
May: Australian budget
May: AI R&D strategy in South Korea
June: India's national strategy for AI
Q3 2018: Germany's AI strategy

Figure 6 An emerging policy environment on the responsible use of AI

Year	Country	Reference
2014	**UK**	Information Commissioner's Office published a white paper on governance implications of data control and processing.[51]
2017	**China**	China State Council announced AI Development Plan with plans to establish AI laws and regulations.[52] Ministry of Industry and Information Technology published three-year action plan.[53]
2018	**UK**	Financial Conduct Authority chair voiced concerns about businesses putting customers in a disadvantageous position by using big data and AI.[54] Government established Centre for Data Ethics and Innovation Consultation.[55]
2018	**US**	The Executive Office of the President published a paper on AI for American Industry[55] and a memo on AI leadership[57] followed by an executive order expressing its interest in having a seat at the table of AI governance.[58]
2018	**Singapore**	The new Advisory Council on the Ethical Use of AI and Data will help the government develop standards and governance frameworks for the ethics of AI.[59] Monetary Authority of Singapore introduced new principles to promote responsible use of AI.[60]
2018	**European Commission**	The Ethics Guidelines for Trustworthy Artificial Intelligence (AI) were revised and republished in April 2019. [61]
2019	**Singapore**	Info-Communication Media Development Authority announced Model Artificial Intelligence (AI) Governance Framework at World Economic Forum.[62]
2019	**Australia**	Criminal Code Amendment Bill passed in April 2019 which means that social media executives would face up to three years in jail or be fined up to 10% of their company's revenue if they fail to take down violent content expeditiously.[63]
2019	**UK**	The Department for Digital, Culture, Media and Sport (DCMS) published Online Harms White Paper for public consultation for 12 weeks. It proposed the establishment of an independent regulator for social networks and internet companies.[64]
2019	**US**	In April, Algorithmic Accountability Act introduced as a new bill (draft legislation) by the US Federal Trade Commission. It requires companies to audit their machine learning systems for bias and discrimination and take corrective actions in a timely manner. A separate bill, the DETOUR (Deceptive Experiences to Online Users Reduction) Act, aims to ban companies from designing, modifying or manipulating user interface in a way that impairs users from making educated decisions before consenting and giving up access to their personal data.[65]

[50] https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd

[51] https://ico.org.uk/media/for-organisations/documents/1546/data-controllers-and-data-processors-dp-guidance.pdf

[52] https://www.natlawreview.com/article/china-s-vision-next-generation-artificial-intelligence

[53] http://www.miit.gov.cn/n1146295/n1652858/n1652930/n3757016/c5960820/content.html

[54] https://www.fca.org.uk/news/speeches/how-can-we-ensure-big-data-does-not-make-us-prisoners-technology

[55] https://www.gov.uk/government/consultations/consultation-on-the-centre-for-data-ethics-and-innovation/centre-for-data-ethics-and-innovation-consultation

[56] https://www.whitehouse.gov/wp-content/uploads/2018/05/Summary-Report-of-White-House-AI-Summit.pdf

[57] https://www.whitehouse.gov/wp-content/uploads/2018/07/M-18-22.pdf

[58] https://www.whitehouse.gov/articles/accelerating-americas-leadership-in-artificial-intelligence/

[59] https://www.opengovasia.com/singapore-announces-initiatives-on-ai-governance-and-ethics/

[60] http://www.mas.gov.sg/News-and-Publications/Media-Releases/2018/MAS-introduces-new-FEAT-Principles-to-promote-responsible-use-of-AI-and-data-analytics.aspx

[61] https://ec.europa.eu/futurium/en/ai-alliance-consultation

[62] https://www.imda.gov.sg/about/newsroom/media-releases/2019/singapore-releases-asias-first-model-ai-governance-framework

[63] https://www.nytimes.com/2019/04/03/world/australia/social-media-law.html

[64] https://www.bbc.co.uk/news/technology-47826946

[65] https://www.theverge.com/2019/4/9/18302199/big-tech-dark-patterns-senate-bill-detour-act-facebook-google-amazon-twitter

The Ethical Guidelines for Trustworthy AI (Draft) published in April 2019 (prior draft for public consultation was published in December 2018) by the European Commission's High-Level Expert Group on AI encouraged AI to be grounded in, and reflective of, fundamental rights and societal values.[66] Trustworthy AI should be lawful, ethical and robust, from a technical and social perspective.

There are seven requirements: human agency and oversight – AI should not trample on human autonomy; technical robustness and safety – AI should be secure and accurate; privacy and data governance – personal data collected by AI systems should be secure and private; transparency – data and algorithms used to create an AI system should be accessible, and the decisions made by the software should be understood and traced by human beings; diversity, non-discrimination and fairness – services provided by AI should be available to all without bias and discrimination; environmental and societal well-being – AI systems should be sustainable and enhance positive social change; accountability – AI systems should be auditable and covered by existing protections for corporate whistleblowers.[67]

To achieve that, information needs to be provided to stakeholders (customers, employers, investors) in a clear and proactive manner, ensuring traceability and explainability. There needs to be diversity of views, beliefs and perspectives when setting up teams that develop, implement and test AI products. Trustworthy AI values should be integrated into organisational culture through design and use of AI systems, with training and continuous education.

Some companies have created internal AI guidelines and begun to socialise the ideas and seek feedback amongst staff, but we are still some way from integrating AI impact and governance awareness into our corporate culture, even though business models have already been impacted by them.

APPENDIX 2 – TECHNOLOGY COMPANIES AND THEIR AI IMPACT[68]

1. Material legal and financial outcomes	Risks	Technology
1.1 Regulatory framework	**Regulatory risks**	**1.** The UK Network and Information Systems Regulations 2018 (in effect, the UK's first cyber laws, "NIS") apply to certain technology companies if they are "relevant digital service providers" ("RDSPs"). Digital services regulated by the NIS are the provision of: **(a)** online marketplaces; **(b)** online search engines; and **(c)** cloud computing services, including PaaS, IaaS and certain SaaS services, as well as certain cloud broking services. Under the NIS, a technology product or services supplier, provider or distributor (a "Tech") that is an RDSP must take appropriate and proportionate measures to ensure the security of network and information systems, and to notify its supervising regulator (for RDSPs, the UK Information Commissioner's Office, "ICO") of cyber breaches having substantial impact without undue delay and in any event within 72 hours of becoming aware of the incident. The ICO has various powers of enforcement, including the power to impose fines at a maximum of £17 million for the most serious of "material contraventions". **2.** In addition, Techs may need to comply with separate regulatory rules and guidance from their sector regulator(s). The UK proposes to devolve to sector-specific regulators decisions about the need to regulate AI, and the extent of regulation required, for AI deployed in their sectors. **3.** There are various initiatives and proposals – including, in the case of the technology sector, threats – around the world for legislation, regulation, ethics, codes of conduct and governance standards applying to the use of AI at supranational (e.g. EU High-Level Expert Group on AI, Draft Ethics Guidelines for Trustworthy AI), governmental, parliamentary and even at a corporate level (e.g. Google's Perspectives on Issues in AI Governance). While we are not aware of these specific initiatives and proposals yet having the force of law as applied to AI, the technology sector is often at the forefront of governmental, parliamentary, regulatory and other supervisory concerns in areas of its operations that: **(a)** impact privacy and data protection and the massive and covert use of personal and big data, including the involvement of data brokerage supply chains; **(b)** result in unwanted and malevolent online targeting, manipulation of information ("fake news") and discriminatory and/or other offensive outcomes; **(c)** result in failure by certain Techs adequately to monitor criminal and/or other damaging or offensive online material and activity and take down offending websites and/or human/robotic actors in due time; **(d)** facilitate or at least don't protect adequately against cyber-crime; **(e)** facilitate or at least support distributed and other denial of service attacks, cyber warfare and espionage by states and parastatal actors; **(f)** support the creation and deployment of AI in military applications, including autonomous weapons; and **(g)** involve collaboration with politically exposed persons and companies generally and specifically for some or all the issues covered in sub-paragraphs (a) to (f). The above list is by no means exhaustive. Accordingly, it follows that any Tech deploying AI directly or through counterparties and/or supply chains must be acutely aware of the prospect and progress of legislation, regulation, ethics, codes of conduct or practice and governance standards applying to AI in its sector. Failure to do so would be a material risk.

[66] https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

[67] https://www.theverge.com/2019/4/8/18300149/eu-artificial-intelligence-ai-ethical-guidelines-recommendations

[68] The legal and regulatory statement in this table covering sections 1.1 to 1.5 is provided by Mark Lewis of BCLP: see footnote 20.

1. Material legal and financial outcomes	Risks	Technology
		4. As corporate governance and best industry practice emerge, failure to adhere to what may be found to be good or universally applicable industry practice or corporate governance standards in the use of AI could result in regulatory claims, as well as legal claims from consumers and others. 5. GDPR in the EU will of course apply to Techs, both in their bricks-and-mortar and online operations. The risks and penalties of failure to comply with GDPR are, or should be, well-enough known – especially in this sector. 6. All companies, including Techs, should understand the impact and outcomes of deploying AI in their businesses – and before doing so. Failure to do so must be considered a risk, with the possibility of employee, customer and supplier claims with resulting liability, e.g. at an operational level for unfairness and/or discrimination in recruitment processes under gender, ethnicity, race or similar equalities legislation. This last example is cited to illustrate how the deployment of AI even at a relatively low operational level, i.e. in recruitment processes in any company, could lead to claims that in turn might result in material liability. In this regard, Amazon has withdrawn the use of AI in its recruitment processes that has discriminated against non-male job applicants. In summary, a failure by any Tech and its senior management adequately to take account of the above risks and impacts and to act to prevent them could expose it and senior management to regulatory enforcement action and liability and ultimately, of course, reputational loss and damage which may be considerable, if not existentially threatening.
1.2 Supply chain due diligence e.g. data brokers, cloud brokers	Counterparty risks	1. If a Tech fails to undertake proper due diligence (DD) of its supply chain/contractual counterparties, it could be in breach of regulation (including GDPR and the NIS) and, depending on its role in the supply chain, consumer protection and product liability, and/or be liable to customers or other parties for negligence or breaches of contract. 2. Similarly, if a Tech failed to protect itself contractually in its arrangements with its supply chain or counterparties, it may be unable to manage the allocation of risk and liability to customers and third parties, including GDPR and NIS, consumer protection and product liability risk and claims.
1.3 Third party access / theft / cyber resilience e.g. Cambridge Analytica and Facebook	Cyber risks	1. If a Tech suffers a cyber breach (a third party gaining access to personal data used in the course of AI processing) and/or theft of personal data) and it is found not to have taken all steps required under GDPR and NIS to secure its networks or systems and therefore cyber resilience, it would likely be in breach of its GDPR and NIS obligations and could face enforcement measures (fines, etc) by the data protection authorities and (for NIS) the UK ICO (see 1.1), and face compensation claims by data subjects and others. Such penalties could be draconian, and the quantum of compensation claims material. The Tech would also, of course, suffer reputational damage. 2. In addition, under GDPR and NIS, in the event of a cyber breach, the Tech must generally report it to the ICO within 72 hours, and may be required to inform data subjects. Breach of these reporting obligations, or the mishandling of them, could subject the Tech to further liability and reputational damage.
1.4 Intellectual Property ("IP") / ownership / exploitation	Exploitation risks	1. Failure by a Tech adequately to secure by contractual assignment or other means the necessary IP ownership rights in AI developments (e.g. algorithms) could leave it without the commercial rights to exploit IP in the AI in which it had invested. In this sector, this failure could cause acute problems, including unlimited liability for infringement of third party IP rights and severe reputational loss and damage. 2. Failure by a Tech adequately to obtain by contractual licence (e.g. a software licence, as opposed to ownership) suitable (e.g. as to scope, permitted users) rights to deploy AI technologies would deprive it of the right to use the AI, face supplier/third party IP rights claims (including for breach of copyright) and be prevented from using the AI in necessary processes, so affecting its operations and creating an additional business risk, e.g. the inability to process customer transactions because of illegality of software/algorithm use. Again, this could cause acute problems in this sector, as noted in 1. 3. Failure by a Tech to undertake adequate investigations/due diligence into permitted usage of its existing software estate in conjunction with AI technologies could deprive it of the right to use the AI, face supplier/third party IP rights claims (including for breach of copyright) and be prevented from using the AI in necessary processes, so affecting its operations and creating an additional business risk, i.e. inability to process own-account and customer transactions because of illegality of software/algorithm use.

1.5 Best/safe standards; securing commercial benefits; liability, remedy, penalty, reputational risks and insurance availability	**Operational risks**	1. Failure by a Tech adequately to contract with a customer, provider or supplier: **(a)** to ensure compliance by customer/provider/supplier with applicable laws), could result in regulatory liability for the Tech (including under GDPR and NIS) and significant third party liability, including corporate and (possibly, depending on where it is in the supply chain) consumer customer claims, also (depending as mentioned) under consumer protection legislation, in contract and for negligence; **(b)** to capture contractually the commercial benefits and outcomes (e.g. through KPIs, detailed product/ service specifications or specified business outcomes) could leave the Tech liable to pay contractual charges without achieving the mandated/desired business outcomes, including loss of market opportunity and probable loss of reputation; **(c)** to allocate properly execution risk and liability as between the Tech and its customers/providers/suppliers could leave it without adequate legal recourse against the customer/ provider/supplier while facing the risk of failure of the project and exposure to regulatory, other third party and (depending on where it is in the supply chain) consumer protection claims, actions and liabilities; and **(d)** to investigate the insurability of the contract risks as allocated to the Tech, and to put in place, where available, suitable cover or require the provider/supplier to do so, could leave the Tech's balance sheet exposed to uninsured claims and settlements. All of the above failures by a Tech could expose it to regulatory enforcement action (under GDPR and NIS) and also (depending as mentioned) consumer protection and product liability claims and other liability and ultimately reputational loss and damage which may be irrecoverable.

2. Salient social impacts[69]	Risks	Details
2.1 Input bias	**Data set bias**	1. A Tech should be able to give examples of successful and adaptive AI systems. For each AI system it tries to test, there should be a clear learning objective, issues identified and how input data, processes and network architecture can be better managed and controlled for training and testing purposes. 2. A Tech should be able to demonstrate clear data ownership, an understanding of why a certain type of data is introduced or discarded. It should be aware of input biases coming from human labelling for supervised learning and have processes in place to limit and manage that.
2.2 Process bias	**Classification / clustering bias**	1. A Tech should be able to answer questions on how challenges of pattern recognition or cluster analysis have been identified, addressed, and at least partially overcome to improve its AI systems. If sequential data is used, such as in text mining, it should be able to explain methods used to mitigate translation out of context and overfitting. If deep vision techniques are used, it should be able to explain the network architecture applied. 2. A Tech should be able to explain to investors how AI modelling talents are acquired to limit structural bias in relation to its overall human capital management approach. 3. A Tech should have the capability to monitor data bias amplification risks when generative or simulated data is used, or when output data is feedback into the process as input data.
2.3 Outcomes bias	**Discrimination exclusion allegations**	1. A Tech should be able to provide examples of how it explains unexpected outcomes from AI algorithms. If there is a risk of population subgroups, such as groups by gender or ethnicity, may be discriminated against in the AI function, the Tech should be able to demonstrate that its has reviewed and tested error rates across the different sub-groups, and has made all reasonable efforts to minimise differences. 2. A Tech should be able to explain the accountability system for AI-human interaction. When human reviewers are involved, their mental health and physical safety should be considered as part of the human capital management process due to the impact that may be imposed by the outcomes of AI systems. Adequate training should be provided on conscious and unconscious bias. If content management standards are subject to frequent changes that require human input, a Tech should be able to explain how the consistency of performance is maintained.
2.4 Explainability	**Trust** **Transparency** **Action**	1. A Tech should be able to disclose and explain the issues in 2.1 to 2.3 above, especially when input factors are changed, and when algorithms are made redundant. 2. A Tech should be able to explain to investors its chosen methods of AI systems; how it reproduces mechanisms in which AI generates outputs; and how AI learns to adapt as new input and algorithms attributes are updated. 3. A Tech should demonstrate that it has made all reasonable efforts to provide users with genuine choices in controlling personal data access and applications. Any tracking methods should be adequately disclosed with proper consent mechanisms in place.
2.5 Oversight	**Integrity** **Accountability** **Safety**	1. A Tech should have board level oversight for responsible use of AI. A Tech should be able to demonstrate that AI systems generate outputs and outcomes as intended. A Tech should have put in place systematic remedial actions as per UN Guiding Principles for business and human rights. 2. A Tech should thoroughly assess and be prepared for the unintended consequences of its AI applications.

[69] The following part of the framework was created by Hermes EOS and is considered work in progress to account for ongoing development in consultation with stakeholders and our engagement work

HERMES
INVESTMENT MANAGEMENT

OUTCOMES
BEYOND
PERFORMANCE

HERMES INVESTMENT MANAGEMENT

We are an asset manager with a difference. We believe that, while our primary purpose is to help savers and beneficiaries by providing world class active investment management and stewardship services, our role goes further. We believe we have a duty to deliver holistic returns – outcomes for our clients that go far beyond the financial – and consider the impact our decisions have on society, the environment and the wider world.

Our goal is to help people invest better, retire better and create a better society for all.

Our investment solutions include:

Private markets
Infrastructure, private debt, private equity, commercial and residential real estate

High active share equities
Asia, global emerging markets, Europe, US, global, small and mid-cap and impact

Credit
Absolute return, global high yield, multi strategy, global investment grade, unconstrained, real estate debt and direct lending

Stewardship
Active engagement, advocacy, intelligent voting and sustainable development

Offices

London | Denmark | Dublin | Frankfurt | New York | Singapore

Why Hermes EOS?

Hermes EOS enables institutional shareholders around the world to meet their fiduciary responsibilities and become active owners of public companies. Hermes EOS is based on the premise that companies with informed and involved shareholders are more likely to achieve superior long-term performance than those without.

For more information, visit **www.hermes-investment.com** or connect with us on social media:

This communication is directed at professional recipients only. The activities referred to in this document are not regulated activities under the Financial Services and Markets Act. This document is for information purposes only. It pays no regard to any specific investment objectives, financial situation or particular needs of any specific recipient. Hermes Equity Ownership Services Limited ("HEOS") does not provide investment advice and no action should be taken or omitted to be taken in reliance upon information in this document. Any opinions expressed may change. This document may include a list of HEOS clients. Please note that inclusion on this list should not be construed as an endorsement of HEOS' services. This document is not investment research and is available to any investment firm wishing to receive it. HEOS has its registered office at Sixth Floor, 150 Cheapside, London, EC2V 6ET. EOS00381 0005914 04/19