

# Data challenges associated with AI

An official noted that the panel would be discussing the advance of artificial intelligence (AI) and technological innovation in the financial sector in connection with data and data protection. AI is about empowering machines to make decisions by using large amounts of data. This subject is not theoretical, however. AI already plays a very important role not only in the financial sector but also more generally in society.

## 1. AI creates both challenges and opportunities in the financial sector

### 1.1. The use of AI and other technologies offers perspectives and challenges to the financial industry

An industry representative feels this is an extremely interesting topic. SMEs are dealing with many of the realities of AI in the financial industry, but there are two main challenges for the industry. First, there are challenges around customer data. There is a certain amount of modesty about what the financial services industry feels comfortable doing with customer data. AI is often considered as an opportunity for things such as know your customer (KYC) and anti-money laundering (AML), but in the financial industry there is an opportunity to create new products, be more innovative and access more customers. AI does not just have to help prevent fines; AI can assist the development of new products and give access to more customers. Additionally, there is an opportunity to be more transparent with customer data than Google or Facebook. It is time to 'brainstorm' the kinds of products that customers feel comfortable buying after providing their data. There is a potential for revenue here but accessing this revenue would require substantial investment.

Secondly, the industry representative observed that many market participants see poor returns from investing in AI, which is partially caused by a lack of data standards. There is a requirement for better data standards that are not merely legislative but also machine-readable, i.e. they are accessible so that developers can use them to help make better technology. It will become necessary to use much of the detailed data that is currently being used to create AI. This suggests that there is a huge opportunity for the industry to collaborate and to collate data sets which can teach machines to work much more effectively.

Another industry speaker praised the fact that ideas like 'machine learning' and 'AI' are now household terms. This serves to remove the fear around AI. Turning to a concrete example the industry representative described, which is a flight-delay insurance product developed by Swiss Re's, which uses similar methods to Amazon's dynamic pricing.

A third industry representative agreed that technology offers huge opportunities, noting however that Europe is somewhat side-lined. The companies most actively using data in their businesses are the platform companies like Google and Amazon or WeChat, Tencent and Alibaba in China. 70% of the global market capitalisation of these companies is owned by seven out of eight of these companies. Europe owns around 2% of it. Europe has enough ambition, but they are in competition with institutions which are 'ahead of the game'. The industry representative does not fear competition, but these companies are significantly ahead.

A Central Bank official noted that AI is also of use to supervisors. De Nederlandsche Bank is experimenting with machine learning, because it possesses a substantial amount of

data. The Central Bank official feels that regulators could also improve the data being put into AI models. AI has become possible because of the huge amounts of available data, but the quality of this data is extremely important in ensuring reliable outcomes.

### 1.2. There is a desire for explicability, but this creates several difficulties

An industry representative suggested that the four levers to serving customers well, are transparency, explicability, independence and ethics. Transparency and explicability are positive traits, but it is very difficult to explain how AI is being used. This activity is very complex and algorithm-based. It is impossible to release algorithms on the market to explain what a company does, and the company itself might not fully understand itself what it is doing. Additionally, there are limitations to transparency. Transparency is always good, because companies should explain to their customers how they are using their data, but it is impossible for regulators to rule flatly that everything must be transparent. In particular, it would be inadvisable to be transparent about AML procedures, because criminals will take advantage.

A Central Bank official felt it is important to consider what explicability means. Regulators should not set out exactly what is being sought. Defining the criteria of explicability will be very complicated in certain cases, due to the nature of AI. The Central Bank official noted that only 1% of money laundering is identified and wondered whether Europe would accept an algorithm which increased this number by a factor of 10 but was not entirely explicable. Explicability must also depend on the purpose of the model.

An industry representative agreed on the importance of explicability. Airlines have used 'dynamic pricing' since the 1950s. Passengers pay a different price for the same service, and often the only difference is when a passenger bought his or her ticket. In the insurance space, the methodologies used for pricing and rating insurance products have more or less stuck to generalised linear models. The reason for this is explicability. Insurers seek to explain to both regulators and consumers how factors such as age effect changes to the ultimate rate for a target cohort. Some firms are using approaches involving less explicable methodologies such as deep learning, but these methods only form part of the process. If there is a methodology that is much more accurate for assessing underwriting risk but not explicable, the industry will have to decide whether or not to use it.

### 1.3. Regulation and practices are fragmented between different industries and across jurisdictions

An industry representative described that in principle, many products in insurance and banking have been data-driven for a long time. However, the technology behind them is not as developed as the technology used in bigtechs like Google and Amazon. The challenge for banking and insurance concerns how to access cutting edge technologies notably because their activity is much more highly regulated than other industries. For example, Amazon engages in 'dynamic pricing' by predicting demand and adapting its prices accordingly. Consumers and regulation in the insurance and banking industries would not be happy to see these practices being used.

A Central Bank official suggested that the purpose of introducing a regulation was to achieve a Europe-wide

implementation of these principles. There were good intentions, but this is clearly insufficient. The point of regulation was to ensure a level playing-field, but the rules are not applied identically across member states.

#### **1.4. There is a clear need for more education on the topic of AI and technology across the industry and among regulatory and supervisory authorities**

An industry representative highlighted the importance of education. Poor results from investments in AI are often caused by huge misunderstandings about how to use data scientists. AI is often mistaken for machine learning, which is somewhat simpler and more accessible. The industry representative wants Europe to take charge of educating and upskilling its workforce – both in the financial industry and at the regulatory level – with a detailed understanding of technologies such as AI, algorithms and the cloud. This will facilitate the development of a legislative agenda that makes sense for innovation, developers and technology.

Otherwise, firms will continue to implement things in their own way and at some point, the legislation will catch up, which is when fines and legislation happen. Europe must anticipate this and install an efficient system of welfare around its legislation on the implementation of AI and algorithms.

Another industry representative agreed on the need to demystify technology. This is not only a task for regulators; it applies to workforces in companies of all sizes. The industry must demystify machine learning and AI. When there is better understanding, fear will subside and innovation will be enabled. A Central Bank official noted that their number one objective is to increase knowledge about these subjects.

## **2. Regulatory intelligence: how regulation interacts with the development of new technologies**

A Central Bank official is 'tickled' by the expression that AI is a highly regulated area. The technology exists now and it is already shaping the preferences of the public. One of the problems with AI is that it is supposed to exist outside the regulatory remit.

A regulator described how several global multinational tech companies are seeing the consequences of misusing AI. These companies are the 'guinea pigs' in AI regulation: they have huge amounts of data, vast resource centres, large resources in relation to employees and finance, and they are innovators and disruptors. They want to be part of the financial sector not only by creating cryptocurrencies but by developing technology for KYC and ID verification purposes. The regulator summarised the recent regulatory fines imposed on Google, Facebook and Amazon. Google were fined \$170 million by the FTC (Federal Trade Commission) for their failure to protect children from an excess of advertising when using YouTube. Similarly, Facebook knowingly kept insecure the details of 50 million customers, and did not declare it for some time. They were also recently fined \$5 billion due to a similar issue. Beyond the issue that most companies do not have an annual budget of \$5 billion for the payment of fines, there is a wider problem. In rectifying these issues, the industry has lost consumer trust. Finally, Apple was fined for its Siri devices, which use AI processing, having been found collecting confidential and highly sensitive information from their customers. Apple used these data to develop how Siri interacts with customers, but this data was not anonymised. While Apple have realised their mistake, the regulator feels that they will have 'a hell of a job' to rectify related consequences, beyond the fact that the case will likely become a matter for the European Data Protection Board (EDPB).

### **2.1. Existing principle-based regulation is an enabler of innovation**

A regulator suggested that the EU data protection regulation (GDPR) does not prevent the use of technology or AI; it simply requires firms to think about what they are doing, the purpose for

which they are using AI and the outcomes they wish to achieve. Data protection also means that firms must consider the benefits a technology will have for the firm, its regulatory implications and the benefits that will accrue to consumers.

A regulator considered that data-protection legislation enables innovative development by providing a framework for balancing data subjects' rights with innovative developments. When developments occur in a regulatory framework, it is assumed that there are sufficient safeguards for data subjects' rights.

A Central Bank official observed that regarding AI, De Nederlandsche Bank has recently published a discussion paper which defined six important principles, spelling 'SAFEST': Soundness, Accountability, Fairness, Ethics, Skills and Transparency. The Central Bank official suggested that there is no need for further regulation in the field of AI. In addition to existing high-level regulation (GDPR), as in other areas, the ESAs could provide further advice in the form of guidance like the SAFEST principles.

However, an industry representative noted that GDPR is still being interpreted in many different ways across different member states. This creates difficulties. For instance, the privacy authority in the Netherlands considers the use of a TIN or a social-security number as a matter of privacy, and this then cannot be used by a bank. This is one of the only ways to make sure somebody exists. While the month of a person's birth is usable, their day of birth is not. A more Europe-wide approach to explaining these principles would be welcome.

### **2.2. Regulation cannot offer the industry complete ex ante guidance**

A regulator opined that regulators 'would love' to give further guidance to the industry, but they must be informed about problems first. Firms must identify what their risks are. If these issues cannot be solved, supervisory authorities can discuss this with firms in a 'friendly manner' and determine what can be done. Only data controllers can identify the problem. The regulator reminded the panel of the mandatory requirement to have a data-protection officer.

A Central Bank official considered that regulators and supervisors should not provide firms with guidance on ethics. Companies should decide these things themselves. While the early involvement of supervisors can be very positive, there is a tendency for the industry to seek explicit guidance on exactly what is permitted and what is not permitted. This is not the task of a regulator or supervisor, because it is a box-ticking approach.

An industry representative considered that regulation is based on laws that everyone accepts, but it can eventually create dilemmas. Financial institutions struggle with fairness and bias in their application of AI. In consumer lending, some use instant-lending techniques. If a consumer files an application after midnight in a bar there can be a bias against that person, because the system concludes something about the liability connected with that person. Although the system looks at all kinds of other data to make this decision, the decision may ultimately be completely unjustified. In the past, there has been postcode and race-based underwriting in certain jurisdictions, for which there was a technical explanation but which was also completely unfair. It was rightly regulated away and punished. As a bank, there are challenges here.

### **2.3. The industry view: regulation stifles the development of innovation**

An industry representative suggested that much of the legislation that has been introduced has had an impact on the level of innovation being brought into the industry by AI. A Central Bank official felt there is a dilemma around fostering innovation in the financial sector. If regulators prohibit something too quickly,

the potential for innovation is destroyed. The pace of change in AI is very fast. It is important not to take a definitive stance, because technologies continue to evolve. Cloud computing is an instructive example here. Five or 10 years ago, supervisors were firmly against cloud computing, but now they are in favour of it because of developments in the technology. The technology is a moving target, which means that supervisors should keep their minds open.

Another industry representative felt that guidance from supervisors is also capable of preventing innovation, considering what has happened to the fintechs and SMEs which have implemented the General Data Protection Regulation (GDPR). Conversely, although many large tech companies have been eventually fined, in practice they have largely benefited from years of managing data and building models with limited regulation.

#### 2.4. Regulators do not understand the technology to a sufficient degree

An industry representative felt there is a need to educate legislators and regulators. They must understand in much more detail how AI is implemented. While algorithms are supposed to be 'explainable', this very quickly becomes extremely complicated. There will be large amounts of documentation that is inexplicable to anyone unless they are skilled enough to understand how algorithms work.

A Central Bank official noted that the people who understand new technologies are the people working in these companies, not regulators or supervisors. The industry should accept responsibility for ongoing issues and work on them. Supervisors have their own perspectives but work in these areas should not begin with regulations and detailed guidelines, especially with regard to issues such as AI.

An industry representative stressed the importance of understanding the detail of these technologies. This is why education is so important in this area. If very high-level regulations are implemented, they will have a serious impact on smaller companies, fintechs and innovators rather than the larger companies.

### 3. Understanding the impact of GDPR

#### 3.1. GDPR is a solid framework for balancing data subjects' rights and technological innovation

A regulator outlined how one part of GDPR is an AI toolkit, which allows public or private sector entities to carry out a Data-Protection Impact Assessment (DPIA) when they are processing personal data using a huge customer database and big-data processing. The toolkit enables a firm to assess what it seeks to achieve and determine how it is going to manage the problems which arise from this. If the problems identified are high risk and cannot be mitigated, the firm can consult the supervisory authority in any member state. There can be a problem if a data subject objects to the automatic processing of his or her data. In DPIAs touching on this, it is essential to demonstrate very clearly the evidential requirements being used to justify the use of AI and how it will benefit consumers.

Another regulator noted that the importance of the tools provided by GDPR, became clear to the wider public, notably following the Cambridge Analytica scandal. Indeed, the DPIA is supposed to be conducted at a very early point in the development in order to take account of the rights of data subjects. GDPR provides a good toolkit for managing these challenges. One of the main challenges of AI is ensuring clarity and transparency for data subjects on how data is collected and when and how decisions are taken based on that data. However, while in some cases – such as 'cookie boxes' on the internet – this is relatively clear, this is not always the case. Sometimes it is challenging for a data controller

to identify the situation in which personal data is being collected, which means it is also not clear for the data subject.

#### 3.2. The GDPR is not sufficiently detailed

An industry representative cited the Swiss Re's internal digital governance framework, which is used to evaluate everything in the data and analytics space. Such a framework could prove to be too burdensome to use for smaller players, however. However, it is essential to ensure that the regulation of AI is better in Europe than in other jurisdictions and that it does not prevent innovation. The only way to get this balance right is to go into detail. The industry representative considered finally that any further regulation of AI is unnecessary.

Another industry representative noted that GDPR is a Europe-wide regulation, which should represent an opportunity for collaboration and the sharing of data sets in a context where Europe is competing with China and the US in AI. These issues should be at the top of the agenda.

A regulator suggested that a code of conduct could be implemented. If the industry has a problem with AML or any other issue, there could be a rulebook with 'dos' and 'do nots' which could indicate how personal data can be processed in specific situations. Both the European Data Protection Board and member states' national authorities would be open to such discussions. An industry representative felt that a lack of guidance and a lack of detail prevents machine learning from taking place. There is a lack of detailed understanding about aggregated, combined or statistical data.

#### 3.3. Is the GDPR sufficient? There are different views on the need for further regulation or further guidance

An industry representative felt the intentions behind GDPR are very positive. GDPR is a great example of collaboration across different countries to develop a piece of legislation across Europe. However, it is somewhat of a missed opportunity in terms of what it could have been. Hopefully, there will be better and more pragmatic guidance on how to implement it. The industry representative felt that the GDPR is very high level and there is poor guidance on implementation. Another industry representative noted in particular that it is sometimes difficult to understand which data is owned by whom, even when it is anonymised. In principle, GDPR is well regarded in the insurance industry, but there are questions about data lineage and what happens to data in the value chain where data is processed. The industry representative noted that in particular there are cases where it is unclear whether an institution is a data processor or not and hence must take responsibility for customer data. The guidance on GDPR does not take account of the fact that rogue developers in financial institutions could overwrite a log to place blame for a problem on an external company rather than an internal process. GDPR is not bad per se, but it is not sufficiently detailed to avoid having an impact on innovation.

A regulator felt that GDPR established the basic framework for the development and application of AI. Given that GDPR itself is comprehensive and creates a multi-layered legal framework, it is important to assess carefully whether there is any need for additional legislation.

An official considered that financial regulation in Europe is better equipped to manage innovation than GDPR. This is not due to the substance of GDPR but the way it is structured, because it becomes excessively detailed. The official pleaded against any further regulation. Another official observed that this is an area where the regulators consider further regulation is necessary, while the industry seeks further detail. Each side should consider the other's views. ●