Nederlandse Vereniging van Banken

Response to the DNB discussion paper General Principles for the use of AI in the financial sector – 18 October 2019

The NVB welcomes the on-going discussion on the use of Artificial Intelligence (AI) to deepen our common understanding of AI. AI offers interesting opportunities for the financial services industry as AI is envisioned to bring innovation, development of new propositions for all customers and healthy competition to the sector. As with any new technological advancements there are also new challenges to consider.

Banks are carefully considering AI, having a close eye on both the specific possibilities and the challenges and risks AI might bring for the banking sector. Although the application of AI in the financial services industry is relatively sparse at the moment, the role of the financial services industry in society and its potential societal impact with the use of AI justifies additional prudence. We believe that collaboration and dialogue with financial supervisors, data protection and competition authorities on the impact of this technology on the banking sector is the best way forward to develop a well-balanced and clearly defined supervisory framework. Therefore dialogue and alignment across all regulators are critical in the process.

We welcome DNB's approach of initiating the dialogue with a discussion paper on the specific opportunities and challenges for financial services. The principles are supportive in the further development of AI in a responsible way. DNB's guidelines create the right pre-conditions and criteria for the use of AI in financial institutions. Due to the early stage of this technology and the diversity in use cases (and the lack of a 'one size fits all' standard), we believe a risk-based and principle-based approach should be at the core of any future governance model.

The NVB calls for specific attention to the overall unclarity around the legal and regulatory framework that can lead to fragmentation throughout Europe, which can jeopardise the development of AI goods and services across the single market. This should be first assessed before any new measures specifically aimed at AI are considered or introduced. In order to take DNB's guidance on the application of AI a step further, we would like to recommend DNB to include a point of view on how to foster innovation in the application of AI in the financial sector.

We welcome further dialogue on existing principles, good practices and guidelines on the deployment of AI and call for support in creating an innovative environment with possibilities for further experimentation and development of financial products and services. In our opinion the application of AI fits within the existing risk management framework. The focus should be on operationalisation of the use of AI within this framework, giving careful consideration to specific aspects of AI. We would also see value in further discussing AI and DNB's expectations in practice once this paper is finalised to align expectations and to avoid misunderstanding in practice

In line with the ambition of the Dutch government and the European Commission's agenda, confidence and trust in AI-based products and services should be promoted whilst industry efforts are focused on releasing the technologies' potential through EU-wide cooperation, increasing public-private partnerships, balancing the flow and wide use of data while preserving privacy, security, safety and high ethical standards.

AI in the banking sector

We believe AI presents strong opportunities for prosperity and growth, both for society, and specifically for financial services. The application of AI will be in the interest of consumers and businesses, providing better, faster products and services, providing relevant or right information at the right time. It is expected that AI will lead to easier and broader access to financial services. As such AI has the

potential to democratise financial services for more customers. One of the opportunities we observe already, for example, is how AI helps financial inclusion by creating access to credit for people and businesses - that are currently shut out of the market - at the same or lower risk costs, made possible by considering new data sources. Also, robo-advice provides portfolio investments to a larger group of clients who previously had no access to these services.

The use of AI also allows banks to develop new propositions, business and risk management models. One of the many key advantages is better risk management as advanced data analytics contributes to a better internal understanding of bank activities (e.g. provisioning and capital models), operational risks and improved monitoring of compliancy. Also in the field of financial crime and transaction monitoring AI may greatly improve the effectiveness of combatting crime.

The wide range of opportunities for the use of AI and its current use cases have been adequately addressed in the discussion document, but it should be noted that in the foreseeable future new use cases will emerge. Given the need for innovation in the financial sector and fostering legal clarity, a future-proof regulatory framework requires a technology-neutral approach.

The use of Artificial Intelligence spans a broad range of techniques and technologies. Application of AI can be found in a wide range of use cases, each with specific goals and context. As the technology is constantly evolving there is no precise definition which adequately covers the field of AI. The definition by the Financial Stability Board (FSB), and used by DNB, which defines AI as ‘the theory and development of computer systems able perform tasks that traditionally required human intelligence’ is deemed too broad and too generic. It is in fact so broad that anything that is done using machines is considered to be an application of AI. We urge DNB to limit the scope of the term “AI”. For any discussion on the use of AI we are of the opinion that one should primarily focus on the activity in which AI is involved, rather than creating an overarching view on this technology.

The NVB acknowledges that AI-based technologies can also deepen existing challenges and create new risks if not developed and deployed in thoughtful ways with appropriate governance safeguards and data protection mechanisms in place. Any principles and guidelines for its use should be closely aligned with the existing risk management framework already in use by financial institutions. Any policies directly or indirectly aimed at the use of AI should be assessed with great care, thereby maintaining financial stability as well as flexibility in the application of AI.

Most challenges are more closely linked to the type of activity and its context rather than the technology used. For banking these challenges include consumer protection & inclusion, prudential risk, ethics, data & privacy, operational and fraud risk, reputational risk, equal and fair treatment, but also cybercrime, industrial espionage and unfair competition. For the financial sector as a whole these risks are broader, and also include e.g. solidarity and risk selection in the insurance industry. While many risks involved with the use of AI are not industry specific, some risks that are applicable to the banking sector require more safeguards than other industries given its role in society, its potential impact on the financial system, as well as the impact on the life of individuals.

Several pieces of EU and national legislation aim at providing the highest levels of consumer protection, while also ensuring financial stability. These also include eligibility for certain products and services, and financial inclusion. As with any automated system, outcomes with the support of AI systems needs to be fair and equal to customers. We also like to note that for legal requirements, e.g. combatting financial crime, exceptions are made for these particular applications.

Customers need adequate protection regardless of where they access services and who provides those services. As such any guidelines for the use of AI should be set on an international level. Due to fragmentation of regulation and the existence of different conflicting regulatory frameworks this is currently a major challenge for Dutch banks.

The use of personal data in relation to use of AI in the financial sector should also be a key part of the discussion between financial services providers and supervisors. As personal data is important in the development of AI, the increased collection and use of data also requires strong safeguards for data protection and privacy across all actors active in financial services. Simultaneously, the GDPR affects automated decision-making as such that the GDPR requirements on transparency, data minimisation, purpose limitation and automated decision-making could potentially limit the innovative nature of AI. Alignment and legal clarity between competent authorities in the use of data and the application of AI is therefore necessary.

The effort towards formulating general AI principles and guidelines should give consideration to creating a level playing field in the use of AI, both geographically and across industries. As the use of AI transcends our national borders we also need to engage in this discussion in an international and European context. Currently the United States and China dominate AI development, so ideally AI will be governed by global principles, but should at minimum be set at a European level to create a level playing field between the member states. Given the European Commission's approach to lead with its own human centric approach it could possibly guide global principles and guidelines. In some countries outside of the EU the principles of the GDPR are currently being implemented as well, so the same could be true for AI principles in the future.

As discussions on appropriate guidelines and regulations already take place at an European level, DNB should strive for European harmonisation across member states, but also across European Competent Authorities (ECA's). We believe the effort for guidelines and principles should not be a siloed Dutch discussion but one that builds on the discussions in Brussels, Frankfurt and the EU member states as well as global discussions.

As financial services become increasingly cross-industry, cross-border and internationally focused we would like to stress the principle of 'same services, same activities, same risk, same rules and same supervision in different geographies' to create a level playing field. While European banks have a very strict regulatory framework, other companies and non-banking players (including BigTech) increasingly offer similar activities. Therefore, the approach through the Financial Supervision Act alone might not be sufficient for maintaining a level playing field, and thus fragmented and unbalanced customer protection will follow.

Going forward, a well-balanced approach in the application of AI should promote digital business models and innovation whilst managing risks associated with new technologies. In general, we currently identify five main challenges for the use of AI in the banking sector:

- Low customer confidence: ethics concerns and lack of transparency & explainability.
- Asymmetric regulation: between industries and geographies, outdated regulatory framework for digital strategies, multi-layered legal & regulatory environment increasing the complexity of application and interpretation of the applicable legal & regulatory frameworks, including data protection.
- Shortage of AI skilled talent and experts: development of skills and knowledge, competition with other sectors.
- Ambivalence of policy makers: insufficient knowledge and understanding on the impact of this technology on the banking sector
- Data availability: Lack of data availability itself is a concern as data availability is a requirement to build any solution involving AI. While good data sets are a prerequisite, there are challenges due to limitations (or restrictions) of the use of personal data according to GDPR.

In continuous discussion, these challenges should be part of the dialogue on the use of AI in the financial sector. DNB could be a linking pin between the financial sector and take a leading role towards Dutch regulators and ECA's in promoting the use of AI and innovation within the financial sector.

The SAFEST framework

AI is not a fixed tool or process but is by definition “in development”. DNB’s SAFEST framework provides a comprehensive overview of the most relevant aspects of use of AI in financial institutions, or any industry for that matter. The framework is beneficial for the purpose of starting a broader public discussion, while providing a good starting point for further discussion within the financial sector.

Operationalising these principles is a major challenge and should be carefully considered in dialogue between supervisors and financial institutions. As the majority of the guidelines are covered by existing policies and/or legislation, regulatory clarity on any new additions is critical. Also, as no clear agreed AI definition currently exists it is hard to define the precise scope of AI systems to which SAFEST should apply. The framework should therefore allow for reasonable flexibility given the early development phase of AI. For that reason, we support a risk-based approach along the lines of DNB’s purpose vs. materiality matrix, for model classification, validation and monitoring. The matrix provides a valuable starting point for AI models categorization.

As mentioned earlier, principles and guidelines on the use of AI should be considered in an international context, across competent authorities, industries and geographies to maintain a level playing field. Therefore, the SAFEST framework should foremost be used for discussion purposes, and not become an operational regulatory framework by itself, which would only be relevant in the Dutch financial sector and limited to prudential supervision.

Going forward, we would like to encourage DNB to expand and align this dialogue with AFM, AP and ACM to formulate a coherent approach to the use of AI in general, and within financial services specifically, in its interactions with ECA’s and the financial services industry.

We would also like to encourage DNB to review the current suitability of existing requirements on governance and risk management regarding the use of AI, in particular (existing) model risk management frameworks. Future-proof regulation should be technology-neutral and digital first, meaning policymakers and supervisors should take fully digital business models as a leading service model.

Principles of the SAFEST framework

Soundness (Section 4.1: Principles 1 – 5)

Principle 1: Ensure general compliance with regulatory obligations regarding AI applications
As soundness is a broad and loosely defined concept and lacks a clear definition in this specific context of prudential risks associated with AI, further dialogue on its interpretation is welcomed. We interpret specific soundness requirements in the use of AI as (1) AI models are ‘fit for purpose’ and (2) are of known quality.

Compliance with regulatory obligations should be ensured with the use of any automated system, which is also true for the use of AI. The use of AI involves multiple laws and regulations and involves aspects such as data protection, liability, cybersecurity, copyright and intellectual property, non-discrimination and duty-of-care. We support the general design principle of compliance-by-design if applicable, appropriate and feasible.

While AI might introduce specific complexity, the same requirements for prudential risks apply to all IT systems that are part of the existing banking infrastructure. The requirements of sound and controlled business operations (WFT) also cover many other obligations. We agree the specific aspects of AI are to be included in the existing risk management frameworks.

This also includes the necessary precautions for business continuity. As such we do not see any new compliance and business continuity issues that are specific for the use of AI, but existing policies should incorporate the use of AI as part of the technology stack.

Principle 2: Mitigate financial (and other relevant prudential) risk in the development and use of AI applications
NVB proposes to distinguish between applications that could yield systemic risk (for instance systems to determine the RWA or to perform liquidity management) that do not (directly) interact and affect customers and AI systems that are used for client-interaction purposes. Principles like *fairness* and *ethics* seem to be important for the latter category, while less relevant, or sometimes even irrelevant, for the first category.

It should be noted that the choice of *targets* and *evaluation metrics* is not only relevant to the soundness of a system, but could also have significant impact on the fairness and ethics of the system. Furthermore, the principle pays specific attention to the need for periodic retraining of models and/or systems. Many AI systems can be used in an "online learning" mode already (each new datapoint yields a partial recalibration of the system), which could yield significant benefits (e.g. less maintenance). At the same time, we acknowledge that substantial systemic risk could arise in case such mechanics are not fully understood, implemented and properly monitored.

Accuracy, as an important part of soundness, should be further elaborated on. This can relate to accuracy of underlying data and to accuracy of modules. An advantage of applying accuracy as a starting point is that it is more suitable for quantification than other principles.

Principle 3: Pay special attention to the mitigation of model risk for material AI applications
Machine learning could be very helpful to improve the quality of risk management models (e.g. provisioning and capital models); for example to detect nonlinear relationships between risk factors (which are not easily discovered with `traditional' techniques). Moreover, the use of (online) machine learning could also be useful to reduce the cost of developing and maintaining such risk management models. Currently, most institutions use `traditional' models, typically regression-based. An important reason is that such models are considered to be very `transparent' and `explainable'. To facilitate the use of machine learning models in this context, regulators should be willing to i) invest in the required (technical) skills for assessing such systems, and ii) accept that such approaches cannot offer the same degree of transparency and explainability as traditional models do.

Principle 4: Safeguard and improve the quality of data used by AI applications
Algorithms used in AI systems can only be as good as the data used for their development. There is not a standard definition available for what is actually high quality data as it depends on multiple factors amongst financial institutions, but also across sectors. Often available and suitable data sets take time to acquire or are not available at all (i.e. due to data privacy restrictions), or are otherwise difficult to clean or transform for an intended (new) business.

While we fully agree with the statement the "Original data sets used to (re)train and (re)calibrate models are systematically archived.", it should be noted that this could yield tensions with the GDPR. It could be helpful if the DNB and AP agree on a compliant strategy (for instance, provide guidance when a dataset is properly pseudonymized or anonymized or can be used for AI purposes and still be in line with the requirement of data-minimalisation).

Principle 5: Be in control of (the correct functioning of) procured and/or outsourced AI applications
While AI outsourcing should be treated as any other outsourcing of IT systems and other services for which the financial institution is accountable, the fragmentation of the value chain should be taken into account. The DNB research report "Unchained, supervision in an open banking sector"[1] identifies the

[1] https://www.dnb.nl/nieuws/nieuwsoverzicht-en-archief/DNBulletin2018/dnb380421.jsp

complexity of new value chains in financial services. New entrants might become the ultimate responsible party as they become the primary contact for customers, while other third parties might make use of outsourcing themselves. Any requirements on soundness should equally apply for such non-banking players. Although we are sympathetic to this principle, it could seriously affect the level playing field if incorporated into regulation.

Moreover, the dependence on large IT companies should be taken into account. Excessive concentration of the IT market in the hands of few key Tech players might also lead to concentration of risks. Assuming accountability in unbundling value chains is only possible if financial institutions can set commonly (EU) contractual clauses, set policies and have access to perform audits. As we have seen with dominant larger IT suppliers, e.g. with the use of cloud services, the 'right-to-audit' on the entire value chain, while already a challenge, might also pose a challenge for the use of AI.

Accountability (Section 4.2: Principles 6 – 8)

The obligation to be in control of procured and/or outsourced applications is not new and is already incorporated into EBA guidelines. While AI is not necessarily a new technology, it continues to evolve at a rapid pace and has the potential to change the way the sector operates. The use of AI adds significant complexity to systems used in the organisation, but this does not change accountability for the organisation. The requirements for accountability and auditing should be tailored based on the use of the AI use case and their potential impact and risks. For instance, requirements for marketing models can be different than those for operational risk management.

Principle 6: Assign final accountability for AI applications and the management of associated risks clearly at the board of directors level
As a business enabler, we believe the responsibility for AI should not confined to the technology functions of the organisation. For AI systems to be effective, accurate, precise and specific, the development of AI systems draws on multi-disciplinary teams of data scientists and researchers, engineers and product owners. Next to involving domain experts, also second line of defence roles should be and are closely involved in any development and deployment of an AI system.

Principle 7: Integrate accountability in the organisation's risk management framework
The existing three-lines-of-defence model already sets high standards in effective risk management and control. Banks have the structure to guarantee the appropriate accountability, risk management and auditability for the use of AI models.

As AI technology is used in various part of the banks value chain and in different use cases a specific AI policy, as is suggested under principle 5, could be less effective. Each financial institution should have the flexibility to embedded the use of AI in their policy framework as they see fit. Application of AI should be embedded in relevant policies based on its use cases and context, which also includes outsourcing to third parties.

Fairness (Section 4.3: Principles 9 – 10)

We subscribe to the need for fairness in the application of AI. What is meant by "fairness" may differ across legal, political, social, historical, and cultural considerations. We therefore need to strive for an appropriate degree of fairness and will need further dialogue to operationalise fairness within the context of AI. For instance, what precise constraints and/or requirements would the need to 'not inadvertently disadvantage certain groups of customers' yield? The main purpose of many machine learning and statistical models is to differentiate between objects/customers on basis of their properties. Therefore it could be useful to adhere to the principle of equal treatment in equal situations

Finally, addressing fairness and inclusion in AI covers all aspects of the use-case life cycle: setting concrete goals, the use of representative suitable datasets to train and verify the model, and the continuous testing of the final system for unfair outcomes. Therefore fairness is broader than a

conduct risk issue only. After all, recent examples show that even the brightest data scientists and AI engineers could not foresee unintentional biases in their AI built systems and models.

Principle 9: Define and operationalise the concept of fairness in relation to your AI application
We urge DNB to reconsider the chosen approach in principle 9 and refrain from technical specifications as currently stated, given its potential impact and possible unintended consequences. While fairness in AI is obviously important for trust in the use of AI, more research needs to be done on which approach is best suited to operationalise compliance with this principle.

Imposing fairness will yield a trade-off with the performance of the AI system. As it is still unclear what the impact of imposing fairness would yield, and there are no relevant empirical studies available, we feel it is too early to include `fairness-by-design' as suggested.

Principle 10: Review (the outcomes of) AI applications for unintentional bias
The NVB recognises the importance of Human-in-the-loop (HITL) and Human-on-the-loop (HOTL) within the context of AI applications. In practice, its application will differ amongst AI applications depending on the methods used and the desired outcomes. In the field of Machine Learning in particular, leveraging both human and machine intelligence can lead to increased accuracy and higher quality of results and predictions. Nevertheless, "*AI should not increase unintentional bias*". In that sense, we need to be mindful and aware of any unintentional biases in historical data as well.

Also it should be noted that under the provisions of article 22 of the GDPR individuals already have the right not to be subject to a decision based solely on automated processing which produces legal effects concerning him/her or has similarly significant impact. This provision includes after-the-fact reviews, the right to obtain human intervention and redress.

Ethics (Section 4.4: Principles 11 – 12)

Ethical considerations should be ensured in the use and development of AI. For that reason the European banks, within the European Banking Federation (EBF), have contributed to the work of the High Level Expert Group on AI on ethics principles for trustworthy AI[2]. At the same time ethics is a shifting, amorphous concept that can rapidly change among different cultures, societies, and values. Ethics in general is concerned with human behaviour that is acceptable or "right" and that is not acceptable or "wrong" based on conventional morality. General ethical norms encompass truthfulness, honesty, integrity, respect for others, fairness, and justice. Ethical guidelines therefore are best accomplished through a set of abstract and high-level principles which would leave enough flexibility for financial institutions for embedment of ethics in practice.

Principle 11: Specify objectives, standards, and requirements in an ethical code, to guide the adoption and application of AI.
Ethical considerations are applicable to all activities of the financial institution, and are not only attached to the use of AI. Also AI is a multi-layered concept. There is no commonly agreed definition of what AI is, and we are of the opinion that it is thus important for ethics standards to be technology neutral: to apply to all technologies alike and not set different standards for different solutions.
In section 4.4 it is stated that "financial firms should ensure that the outcomes of these systems should not violate the firm's ethical standards". Approaches to dealing with ethical problems are already established in ethical codes, such as the Bankers Oath and requirements on duty of care, on which AFM is the competent authority. These ethical codes include the use of any systems within that organisation including the use of AI. Financial institutions can choose different approaches to implement ethics in the organisation, either in the form of an ethics committee and/or via ethics officers. Current ethics governance structures should also include ethical considerations on the use of AI within that organisation. Therefore an specific AI ethical code is not preferred, but should rather be incorporated in general ethical guidelines used by the financial institution

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

Principle 12: Align the (outcome of) AI applications with your organisation's legal obligations, values and principles
We agree on the alignment with the institution's values and principles, and our legal obligations including duty of care. The principle brings up financial well-being as consideration for AI applications. Unless duty of care is meant with financial well-being, we suggest to refrain from the use of financial well-being in this context as it is a undefined legal concept and is already covered in general ethical considerations and regulations.

Skills (Section 4.5: Principles 13 -15)

Principle 13 : Ensure that senior management has a suitable understanding of AI (in relation to their roles and responsibilities)
Principle 14: Train risk management and compliance personnel in AI
As financial institutions increasingly come to rely on automated systems sufficient knowledge and understanding of its workings and potential risks is already part of current risk management requirements. This includes any new technologies being used, and applies also for AI and its specific challenges.

Many financial institutions have set up training programs to increase AI awareness for senior management and relevant personnel in first, second and third line roles. We recognize the need for broad training efforts across all senior management, but also consider the emphasis on Board of Directors too strong in the document.

As a result of multi-disciplinary teams developing AI systems there is hands-on involvement by employees, both from specific domains and second line functions, in discussing and assessing the risks and challenges involved. Relevant training can either be formal training, on-the-job training or other forms of training & development appropriate for the various aspects involved with the use of AI, which include technical skills, risk management, business skills, ethics etc. Discriminatory effects in AI should be avoided and/or mitigated at the early stages of AI development process. Hence training AI experts involved in AI development on how to make the right (ethical) decisions is key. As the developments in the field of AI are cross-sectoral, we also see the need for cross-sectoral knowledge exchange to improve insights in the use of AI.

It suffices to say that the same skillset is required for policymakers and supervisors. They should have the necessary knowledge and understanding of the technology and its impact in the banking industry. As the field of AI is constantly developing more frequent knowledge exchanges and alignment between the regulators and the sector will increase a better understanding.

Transparency (Section 4.6: Principles 16 – 17)

Transparency and explainability are key to building and maintaining citizen's trust in AI assisted systems. Organisations should be transparent in the way they use AI in their business processes and how they impact automated decision-making. Ex-ante transparency and ex-post explainability (within limitations of what can be disclosed) are of importance to customers and, therefore, we assume this principle also applies to customers and to other external stakeholders as well.
As AI will become an instrumental way to provide more personalisation in a broad(er) range of IT services, including financial services, AI awareness will be a necessity for the general public as well. Some banks are already involving customers and end-users in the product lifecycle, thereby increasing transparency and AI knowledge of end-users too.

Nederlandse Vereniging van Banken

Principle 16: Be transparent about your policy and decisions regarding the adoption and use of AI internally
Principle 17: Advance traceability an explainability of AI driven decisions and model outcomes

Transparency: In AI, we would highlight a risk-based approach and different degrees of transparency required depending on the audience and the stage along the model lifecycle. We recognise that the degree of transparency is different towards model developers, internal (business) users, second line officers, external stakeholders such as competent authorities, jurisprudence and customers.
We emphasise the need for general and practical principles to ensure sufficient flexibility for situations where full transparency cannot be provided or even desired (e.g., in fraud).

Explainability: Similarly, different use cases of AI call for different degrees of explainability. AI models model a complex reality and cannot be – by design – be explained and understood in simple terms. In parallel, explanability as a form of transparency could be distinguished from transparency in a strict sense as described in the GDPR. Explainability of algorithms is closely related to the techniques applied and different techniques presume a different level of explainability. For instance, simple machine learning algorithms are inherently explainable, whereas more complicated machine learning models sacrifice explainability for accuracy and performance. Ultimately, in their AI decisions financial institutions are dealing with multiple trade-offs between transparency and explainability versus accuracy and performance. In the light of this complicated environment , the levels of transparency and explainability will vary amongst stakeholders, product lifecycles and geographies.

In this regard, the NVB welcomes the proposed ‘heatmap’ approach by the DNB, based on principles such as appropriateness and proportionality, taking into account multiple factors as described. The NVB favours a risk-based approach based on the impact of the outcomes of the system as ensuring transparency and explainability. While firms should have a good understanding of their own data processing, their models and obtained results, as well as the appropriate level of detail should be based on relevance and the impact of the outcomes of the system. Also transparency and explainability should be technology-neutral. The use of AI should not increase requirements on transparency and explainability by default.