

# Machine learning in risk models – brontobytes and Brontosauruses

Speech delivered at the Bundesbank-BaFin FinTech Festival

19.05.2022 | Berlin | Joachim Wuermeling

1 Introduction

2 Data: quality, quality, quality

3 Supervisory approaches: a question of characteristics and explainability

4 Conclusion

---

## 1 Introduction

Ms Roth,

Ladies and gentlemen,

The Brontosaurus, a long-necked 30-tonne herbivore, is one of the most well-known dinosaurs and surely familiar to many of you as well. Perhaps it was a fascination with dinosaurs that led to one of the largest units of data volume being called the brontobyte. And just like the Brontosaurus, the brontobyte is also awe-inspiring; it seems gigantic and somehow unmanageable and inconceivable.

The brontobyte comes after the gigabyte, terabyte, petabyte, exabyte, zettabyte and yottabyte. One brontobyte equals around 10 to the 27th power of bytes, or a 1 followed by 27 zeros. This volume of data would be equivalent to the US (United States) Library of Congress 100 trillion times over; every person in the world would have 12,500 copies of this library. This is big data in the truest sense of the word.

However, we have not yet entered the brontobyte era. Right now, we are in what is known as the zettabyte era: in 2016, annual global internet traffic exceeded one zettabyte. Around 1 million such zettabytes make up a brontobyte. We're still a fair way off that mark. Who is supposed to process all this information?

This is where artificial intelligence (AI) and machine learning (ML) come into play, because they are capable of managing these data volumes. And in the financial sector, banking supervisors – BaFin (Bundesanstalt für Finanzdienstleistungsaufsicht) and the Bundesbank – have a vested interest in this. You see, digital technologies can make individual banks more stable, through better analyses and data management or the use of advanced analytics in risk management. Thus, AI and ML can create added stability for the financial system as a whole.

We, as supervisors, therefore wish to act as enablers of the digital transformation. We support the use of AI, and encourage you all to consider and make use of its potential – though all the while we are keeping an eye on the risks, of course.

The extent to which financial institutions are already using ML was addressed in a joint BaFin (Bundesanstalt für Finanzdienstleistungsaufsicht)-Bundesbank consultation. Banks and insurers are already employing machine learning in a number of areas, such as money laundering and fraud detection and in analyses used in lending processes, so mainly in risk management. Companies are also using ML methods in sales and product pricing.

ML is used only occasionally in internal Pillar 1 risk models, but some banks and insurers see great promise in using it. Already, ML is being used as an aid to validate internal models.

The consultation shows that the areas in which ML methods can be applied are many and varied, and the potential is huge – extending far beyond the current area of application. The wide range of possible applications shows that “big data” also has “big potential”. For every institution, and for the stability of the financial sector as a whole. It is worth exploring and investing in that potential.

But even though the term “big data” almost glosses over it, we’re not just talking about the quantity but also, at a very basic level, about the quality of the data.

## **2 Data: quality, quality, quality**

The results of machine learning processes very much depend on the data they are fed with – as they say in IT (Informationstechnologie), “garbage in, garbage out”. If you feed a machine learning engine with data that are not varied enough, or are insufficient or fake, you are going to get incorrect results – and then run the risk of making the wrong business decisions.

A problem arises when data are not sufficiently representative. Take speech recognition, for example: even though this software is usually given a female name, such as Alexa, Siri and so on, it has a harder time recognising female voices than male ones. The reason is that Alexa and Siri are trained with databases in which male voices are overrepresented.

Another example, this one from the world of football, which again shows how dependent ML methods are on data: while matches were being held without spectators due to the coronavirus pandemic, one Scottish football team used an AI-based camera system to livestream its matches. This camera system is intended to follow the ball by means of a built-in AI – an AI that has been trained to identify the ball. During one live transmission, the camera was supposed to track developments on the pitch but kept focusing instead on one of the assistant referees. The AI had mistaken his head for the ball. Apparently, the procedure had not been fed with enough data to allow it to learn the difference between a head and a ball. Although this ML error did not lead to a flawed business decision, the spectators may well have seen more of the assistant referee than of the match ...

As humorous as this example may seem, it does illustrate a fundamental requirement for the application of machine learning methods: it shows just how important high-quality data are – whether on women's and men's voices or on balls and heads.

AI and ML are only as intelligent as the data they are fed with. Data quantity does not automatically translate to system quality. The dataset is the be-all and end-all for AI analyses. A deep understanding of the data used is crucial; data competence has become the key competence in the financial sector!

Incidentally, palaeontologists also have to deal with incomplete datasets: although the Brontosaurus is one of the best known dinosaurs, a skull has never been found. So the picture we have today of the Brontosaurus is based on assumptions, meaning we can only make the educated guess that the sauropod had a short, high skull.

Such conclusions are arguably acceptable in palaeontology – I would rather not recommend them as a basis for a business policy decision, and would instead suggest looking for a better dataset.

### **3 Supervisory approaches: a question of characteristics and explainability**

So how do the supervisory authorities view ML? First and foremost: pragmatically. We look at those elements that are risk-sensitive. We stick to our technology-neutral and risk-oriented approach. What we do not need is a definition of ML that is as universally applicable as possible, in which we may easily become entangled, and which may smother innovation or fail to cover every situation.

Instead, we are focusing our supervisory practices, inspection techniques and inspection intensity on which ML characteristics, if any, are present in a particular methodology and how pronounced they are. This approach – focused on individual characteristics rather than a rigid definition – helps us to identify ML innovations and the risks they pose, to address them appropriately, and to avoid lumping all new applications at banks together – that is the underlying principle of risk-oriented supervision.

The responses to the consultation reaffirm the approach we are taking. There is a broad consensus that no explicit definition should be prescribed. It would in any case be virtually impossible to capture the diversity of procedures and their continuous development in a rigid definition. The focus should not be on generalised requirements, but on specific use cases.

And the results of the consultation also make it clear that ML does not need its own new set of rules. Many supervisory requirements, such as those covering model development, operation and validation or data management, come about because they address complex statistical models. I believe that the current technology-neutral and risk-oriented regulatory approach should also stake out a general framework for ML.

Our supervisory approach also focuses on ensuring that data are of the right quality. That was a point that elicited a unanimous response in the consultation, although good and meaningful data have always been a key factor in the success of models, whether ML or not. Procedures which process huge amounts of data do not alter this principle.

Another key criterion is explainability. ML is often referred to as a “black box”, an opaque system in which data are fed in at one end and results are produced at the other without us being able to understand how one relates to the other.

In this sense, dinosaurs, too, are a black box that we cannot fully comprehend, never having seen them. Did they have fur or feathers? How large was their brain? And what colour were they even? With ML, too, we often can't tell straight away which relationships serve as a basis for decision-making.

However, if decisions made by ML systems have far-reaching effects on us as individuals or on a bank's risk situation, the decision-making process is significant. Consider, for example, AI-based assessments of creditworthiness. We want to understand decisions, we want to shed light into the dark.

The more complex an ML model's design, the more difficult it is to describe the relationship between input and output verbally or using mathematical formulae. It is then often difficult for modellers, users, validators and also for us supervisors to verify the results in detail. And every stakeholder also has a different understanding of and need for explainability.

Checking that a model's behaviour overall is explainable and plausible is therefore more important than for every detail to be verifiable. Is there a logical relationship between input and output? Does the model behave as we would expect where decisions are verifiable? In short, does it make sense?

It will often be a matter of treading a golden middle way of the "grey box", between the need for explanation on the one hand and performance on the other. What is important is understanding and critically assessing the methodology in question; it is not a question of every single process having to be signed off by a human being.

The use of artificial intelligence is monitored as part of existing banking regulation. I am therefore fairly critical of the idea of special approval requirements, as proposed by the European Commission for credit assessments. Where there is double regulation and there are dual supervisory processes, there is a risk of innovation being smothered and of supervisors losing the sense of what can and should be done that they draw from their daily work.

## **4 Conclusion**

Ladies and gentlemen, innovative technologies always offer enormous opportunities, not only for the individual bank but also for the financial system as a whole. We can and should think of innovation and stability together!

Supervisors remain technology-neutral in their work and operate within the framework of the existing rules. This gives you, as representatives of banks and fintech companies, planning certainty when it comes to investments in ML methods. And the faster you invest in AI, the greater your benefits will be in an intensely competitive market; think process automation, speed, operational quality and cost efficiency.

Big data and AI are not only a powerful duo in terms of technology, they are also of central strategic importance for the financial sector. If Brontosaurus is the past, then brontobyte is the future. AI and ML must be deployed in a way that leverages the potential offered by the vast amounts of data without ignoring the risks associated with their use. Here, data competence is vital.

Ladies and gentlemen, the dialogue on opportunities and risks between supervisors and you – the banks, insurers and fintech firms – will certainly continue, and I look forward to being part of it.

Thank you very much for your attention.