Speech Nicole Stolk - The journey towards a data-driven organisation

In her speech at the Data Science Conference of the DNB Data Science Hub, Nicole Stolk addressed several preconditions for being a data-driven organisation.

Date: 12 May 2022
Speaker: Nicole Stolk
Location: Data Science Conference of the DNB Data Science Hub

Dear participants, good morning,

My name is Nicole Stolk. As one of the Executive Board Members of De Nederlandsche Bank, it is my pleasure to welcome you to Amsterdam and in particular to this conference. These two days you will be discussing new methods, new data, or combinations of the two. We have high hopes and feel this is an important and fruitful subject. This is why I am here today and why my fellow Executive Board Member Steven Maijoor will deliver his keynote speech tomorrow.

Me being here goes beyond just delivering a few standard remarks. Becoming a data-driven organisation cannot be achieved overnight. It is a journey, and I have been traveling with you for quite some time. I am not a data scientist myself, but I do see the potential value of basing our policy decisions and supervision more firmly on data-driven analyses. But this value can only be realised if we are comfortable with using data science in our daily work. Part of this sense of comfort will come if we can change the way we work. This involves a marked culture change which could be the topic for the next conference.

From a management perspective I am only comfortable relying on data-driven policy-making if it is underpinned by solid data governance. It should be clear who owns the data, what its quality is, where it is stored and who can access it for what purpose. Unfortunately, data governance is not nearly as sexy as data science. I don't know if you have ever lived with others – maybe as a student? – then you know how fights over who should keep communal areas clean can turn ugly. Data governance is no different. If it is well organised and properly implemented, everyone can almost effortlessly use the appropriate data; ownership and data lineage is clear and users know what can and cannot be done with the data. However, if governance is not properly implemented … Then data owners block access and shirk responsibility, individuals' privacy is violated and the same, or worse, almost the same data starts 'living' in many places. If this is what happens, I would be very uncomfortable trusting the analyses.

To tackle issues like these we initiated a bank-wide programme called "Mastering Data" in 2017. I chaired this effort and we made good progress on several fronts: for instance, we established the concepts of data ownership; every data set should have a single owner. We worked on ground rules for access. We set up a catalogue of available datasets and their owners. You probably have seen similar projects in your institutions, so you will know that this is a complicated issue. Currently we have a "Data Board" which is successfully pushing this agenda further.

Given the tremendous advantages of implementing, for example, the ideas you present in your papers, one would expect the required governance improvements to be implemented fairly quickly. Yet, implementation can be a rocky road. Why is that? Well, ideally, data owners try to encourage the use of data as widely as possible, within the restrictions, of course, that come with data that is often confidential. Only if data are actually used can the – often significant – costs of collecting them be justified. However, the incentives for data owners are often mixed. Allowing wider access might create value, which is a good thing for the organisation as a whole, but often benefits accrue to others than the data owner. Wider access inevitably also increases the probability of something going wrong. For example, data could leak, which reflects badly on the owner. These adverse

incentives are especially strong for data sets that have many users but no single natural home. A prime example of such a data set is the EMIR derivatives reporting. To counteract such perverse incentives, both data owners and data users should be appropriately incentivised – both financially and otherwise. The data strategy we formulated recently addresses these issues and will develop governance further.

Another fellow traveller in the journey towards a data driven organisation is the Data Science Hub. I remember it being a group of enthusiasts but without any formal role. They demonstrated the right type of energy but their efforts did not scale well, because some of the change required could not be achieved by just a group of volunteers. That's why the Executive Board decided two years ago to set up a unit that would act as a knowledge centre. This unit would foster a community of data scientists – who often work alone in their respective units – and thus encourage cross-fertilisation and code sharing. Similar to the incentive issues around ownership, the Data Science Hub could improve the individuals' incentives to cooperate and share code. We also envisioned a hub working together with 'spokes' in the other divisions of the bank delivering value-adding projects. In these projects the Hub would aim to implement new, scientifically proven methods: a Data *Science* Hub. I'm glad to see that – in organising this conference – the Data Science Hub is taking the connection with the scientific community seriously.

Let me come to a close. We are on a journey across often rocky terrain. But the progress is promising, and I hope to see many interesting findings and useful applications resulting from this conference. Enjoy being here in person. Make new friends. Visit Amsterdam. But watch out for the bicycle riders: they are all maniacs with no regard for human life!

Now, before I go, let me introduce the next speaker. We are honoured to welcome Mr Lukasz Kubicki. Lukasz, if I may, has extensive supervisory experience and is currently the Head of the SupTech, Technology and Innovation department at the European Central Bank. He will deliver the first keynote speech. Given the subtitle, *"applications of AI and ML for policy and prudential supervision"* I would have loved to stay and be inspired. Unfortunately, I have prior commitments and I must leave you in the capable hands of Ronald and Iman.

Thank you very much for your attention. Enjoy the conference.