Talking point

Operation Big Data: Challenge accepted (Fintech #3)

August 17, 2015

Big data is a hot topic. The large digital platform operators in particular have long recognised the economic potential of algorithm-based data analysis. They demonstrate this to billions of customers professionally every day. With their analytical technologies they generate high revenues and tie us loyal customers ever more firmly to their platforms via convenient and, above all, individualised services. A steadily growing number of companies want to imitate this lucrative lock-in effect so they can also capitalise on the benefits of big data. Nonetheless, in many sectors the implementation of modern data analysis tools is proceeding only sluggishly. Contrary to the expectations of some market participants, big data is not a simple add-on.

The professional handling of data on digital platforms shows, in particular, the potential of big data under almost fully digital conditions, that is, in an ideal environment. For companies like Google, data records and data analyses constitute core activities. To improve the results delivered by a search engine every search request entered has to be stored, enriched with metadata such as the IP address and then evaluated using algorithms specifically devised for this purpose. Furthermore, platforms promise optimum utilisation of their products and services when users register with them in order to open a customer account, for instance. With this approach, the platform can create an individual search profile for every user. This mass generation of search profiles enables the platform operators to monetarise various services by, for example, using these data for advertising purposes.

Advantages of a fully digital infrastructure

Such data analysis is made possible by the fact that every interaction between the user and the platform operator is executed via diverse digital channels. This provides the platform with all the (personal) data in digital form from the outset. An IT infrastructure developed specifically for this purpose and consisting of a network of powerful, state-of-the-art computer centres naturally enables the platforms to a) store these data in a structured format and b) evaluate them in real time if necessary, so c) the customer can be offered personalised services. Furthermore, this infrastructure enables data derived from other sources to be seamlessly integrated into already existing databases. For many companies, these ideal, virtually fully digital conditions are more or less a pipe dream. Correspondingly, they face new challenges when the idea is to replicate similar data analyses or implement algorithm-based solutions in short order.

Companies whose core businesses are not transacted (only) on the internet and whose value-added networks are not 100% digital have a much more difficult job of reorganising their business processes in such a way that these generate as much data as possible. In a hospital, many processes focus, for example, on nursing individual patients back to health. Depending on the medical condition, a substantial amount of information is collected to this end; this takes place via sensors which cover an array of activities ranging from measuring the pulse right through to recording conversations between doctors and patients. This information is in most cases not (permanently) stored on a digital or machine-readable basis, since up until recently it would neither have been possible nor useful to keep cost-efficient records of all these data. The pattern is similar in other sectors in which production steps are performed manually or customers interact largely offline.

Company architecture has to be adapted to the digital age

Evidently, the challenges of implementing modern data analysis tools go hand in hand with a company's progress in adapting its own corporate architecture to the digital age. Companies that succeed early on in digitising their upstream and downstream value-added networks to the greatest possible extent will form the basis required for the future use of algorithm-based data analyses. For the front-end this means, for example, that the development of digital distribution platforms and the channel-overarching standardisation of customer communication will noticeably simplify the creation of data profiles, assuming applicable data protection rules are in place. This can

help to boost the benefits to the customer. At the back-end, by contrast, the increased use of automation technologies, e.g. robotics at production sites, will enable every stage of production to be documented down to the last detail in order to optimise internal processes. Ever smaller and increasingly powerful sensors, which for example enable goods to be traced along the logistics chain, will give rise to further valuable data sets that can later also unleash internal and external optimisation potential (the "internet of things").

Although digital structural change has barely begun it is progressively forcing a rethink in all sectors. There are many companies founded back in the analogue era that are sitting on data goldmines which could start to yield rewards also on partly digitised processes and structures. However, they have dragged their feet on implementing the reforms required to do so, and technological progress is being impeded by the fact that IT infrastructures used up to now have not been optimised for modern digital data analyses. The existing data are usually stored in fragmented systems that are not linked in a common network. Moreover, different divisions of a company may perhaps use differing data formats and standards. The more complex a company's structures and the more they resemble the structures of a corporate group, the higher is the probability that the different business divisions do not collaborate sufficiently to enable potential synergies to be tapped. These incompatibilities have either grown over time or been created in-house due to a silo mentality that prevents the merging of complementary data. As long as such incompatibilities exist big data solutions can only be deployed at a few isolated data collection points. However, the algorithms used for data analysis stand out in comparison with conventional methods especially because they can discover unexpected correlations in very complex data sets in particular. This suggests that these algorithms can only develop their full potential once all of a company's data have been stored in an integrated, networked, compatible infrastructure whose features can be clearly grasped by various internal interest groups.

Big data also has its limits

In order to maximise the benefits of big data solutions it is vitally essential to structure data and enrich them with metadata. A database consisting of recordings of all of a firm's consultations – to cite an example of particularly unstructured data – can only be evaluated sensibly by an algorithm if it understands what type of advisory service is being offered, who the customer is and what problem needs to be solved. True, artificial intelligence (AI), that is the self-learning algorithm, can increasingly select the relevant details from complex types of data on its own. However, the algorithm still approaches its limits at present when it comes to understanding nested language constructs, comprehending ambiguities (e.g. irony) and interpreting differences in pitch. Existing AI systems (still) lack the human intuition and experience needed to address these weaknesses. This calls for data scientists and algorithm operatives who collaborate closely with the respective experts of a company in order to define the context of individual data sets, convert them into machine-readable format and, at the end of the day, interpret the results of the analysis. Moreover, the expertise required for such an undertaking is also needed to correct unrealistic results triggered by programming errors or poor data. In addition, there is a need for "common sense" in order to differentiate between causal relationships and purely correlative ones, something that algorithms can only deliver to a limited degree at present. Finally, an understanding of the day-to-day business will be required in order to translate the results of the algorithms into concrete recommendations for action.

To push ahead with the deployment of big data technologies many companies would have to start redefining and/or restructuring their processes and business models now. In practice, however, this proves to be a Herculean challenge especially for large companies in view of their ongoing day-to-day business. Especially companies from those sectors in which the digital transition process is in full swing would therefore be well advised to start off by hiring data scientists and algorithm operatives who, in conjunction with experts from the operating businesses, can gain an unprejudiced overview of all the databases and data-generating processes at hand. The resulting ideas and analyses should then be implemented in short order. Moreover, a new infrastructure should be built up in parallel in order to gradually replace outdated systems with new ones. By contrast, relatively small companies that prefer to seize on the opportunities of digital structural change more for their own use could find an attractive alternative in outsourcing. That is, instead of implementing proprietary infrastructures that are geared to big data, smallish companies could use cloud-based services. Collaborating with a firm specialised in this field will enable them, in compliance with applicable data protection rules, to store their data, enrich them with external data, structure them and evaluate the content.

Further commentaries from the Fintech series may be found here:
Fintech #4

Fintech #2
Fintech #1

Details about the opportunities and risks of "Big Data" can be found here.
Click here to find out what the Fintech movement holds in store.
Here's where you can discover why traditional banks should transform themselves into a digital platform (Fintech reloaded).
What is blockchain technologyall about?

Authors:
Thomas F. Dapp (+49) 69 910-31752
Alexander Karollus

...more information on **Technology & Innovation**
Talking Point - Archive