

Independent information
management consultant
Ronald Damhof developed the
Data Quadrant Model

**“Make
data management
a live issue for
discussion
throughout the
organization”**

**Ronald
Damhof**

INDEPENDENT
CONSULTANT
INFORMATION
MANAGEMENT

The data management field is awash with jargon. Most business managers have no idea what all those terms mean, let alone how to use them in understanding the precise value of particular data and how to handle it. To allow an enterprise-wide discussion on data, Ronald Damhof developed the Data Quadrant Model.

Damhof works as an independent information management consultant for major organizations such as Ahold, De Nederlandsche Bank, the Dutch tax authorities, Alliander, and organizations in the financial and healthcare sectors. These are data-intensive organizations which share a growing realization that the quality of their work is increasingly determined by the quality of their data. But how do you move from that realization to a good data strategy? A strategy which everyone in the organization understands, from the director in the boardroom to the engineer in IT? Damhof developed a quadrant model to make data management a live issue for discussion.

To push or pull?

Damhof starts by explaining a concept which everyone will have encountered in high school: the 'Push Pull Point'. This concerns the extent to which demand impacts the production process. He takes as an example the building of a luxury yacht, a process that does not start until the customer's order is known. The decoupling point is at the start of the production process. We can take matches as an opposite example. If a customer wants matches, he or she goes to the supermarket and buys them. Unless he wants black matches, then he is out of luck. The decoupling point is right at the end of the production process. The production of a car, however, comprises standard parts and customized parts. Customers can still state that they want a specific colour, leather upholstery or different wheel rims. The decoupling point lies somewhere in the middle of the production process. "Similarly, in the production of a report, dashboard, or analytical environment, the decoupling point lies somewhere in that middle area," Damhof explains.

The decoupling point divides the production process into two parts: a push and a pull side, also referred to as a supply-driven and a demand-driven part. Push systems are aimed at achieving eco-

"A quote I have stolen from Gartner analyst Frank Buytendijk: in an average organization the car park or art collection is better managed than data"

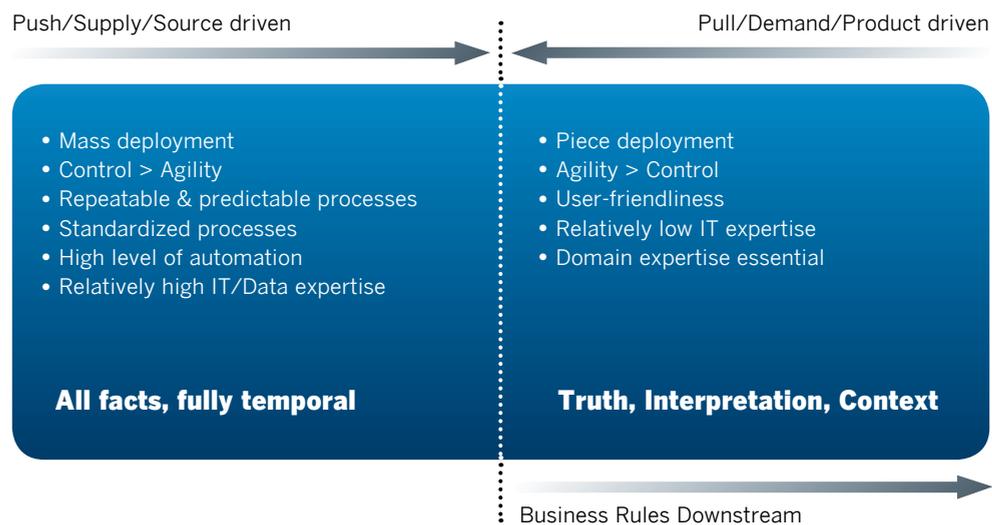
Ronald Damhof

nomies of scale as volume and demand increase, while the quality of the product and the associated data remains guaranteed. On the other hand there are pull systems which are demand-driven. Different types of users want to work the data to produce 'their' product, their truth, on the basis of their own expertise and context.

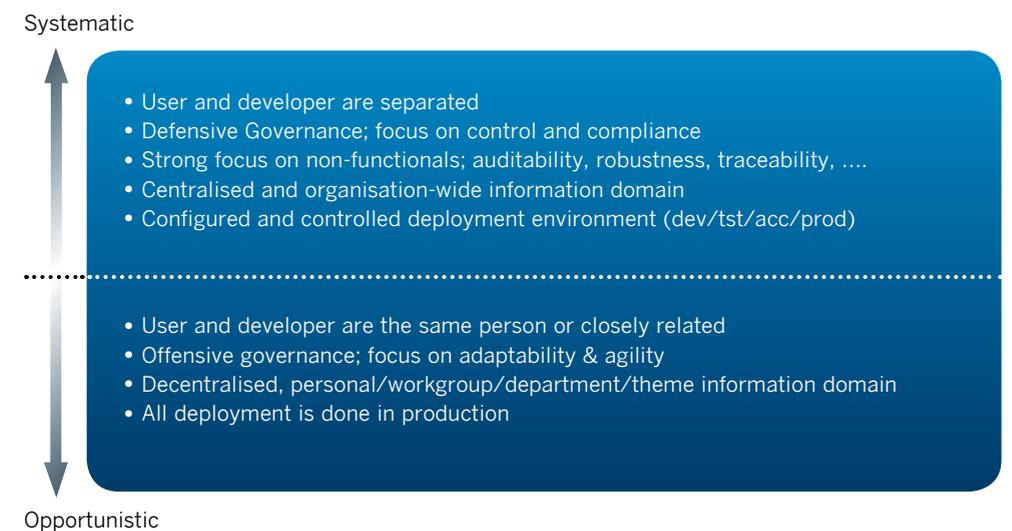
Opportunistic or systematic development?

On the y-axis Damhof projects the development style dimension. "By that I mean: how do you develop an information product? You can do so systematically; the user and the developer are then two different people and you apply defensive governance, aimed at control and compliance. This puts into practice everything that engineers have learned in order to create software on a sound basis. You often see this in centralized, enterprise-wide data, such as financial data and data which is reported to regulators." You can also use an opportunistic development style. "In that case the developer and the user are often one and the same person. Take for example the data scientist who wants to innovate with data, who wants to produce and test analytical models. Or situations in which speed of delivery is essential. The governance in these cases is offensive, which means the focus is on flexibility and adaptability."

The Data Push Pull Point



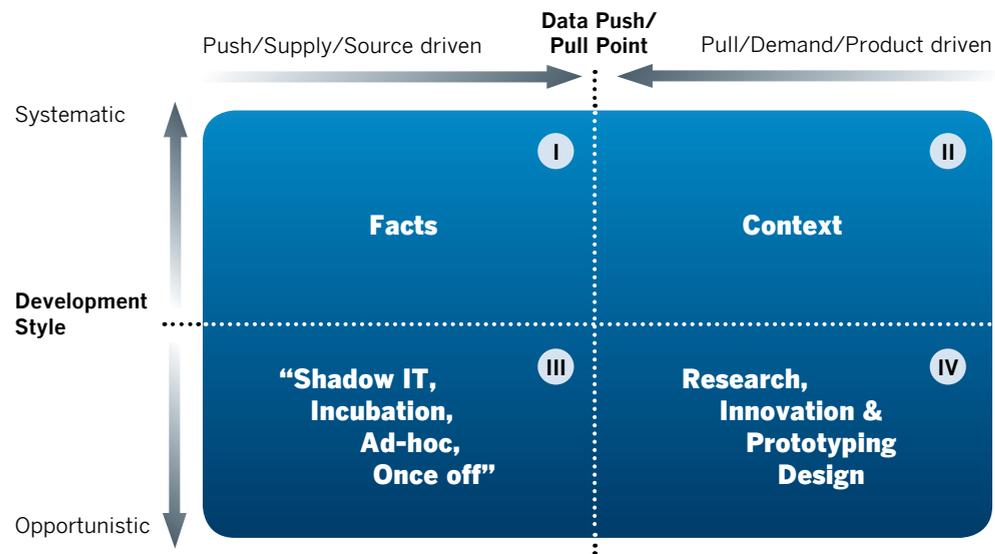
The Development Style



Data Quadrant Model

The combination of these two dimensions produces the following picture.

A Data Deployment Quadrant



“Quadrant I is where you find the hard facts,” Damhof explains. “This data can be supplied intelligibly to quadrants II and IV in its full, raw volume. Data in quadrant I is produced by highly standardized systems and processes, so it is entirely predictable and repeatable.” Diagonally opposite, in quadrant IV, is data that is characterized by innovation and prototyping. “This is the quadrant in which the data scientists work, who actually have only three demands: data, computer power, and cool software.” Increasingly, separate departments are set up as innovation labs giving data scientists free rein to use the data for experimentation and analysis, with the aim of innovation. “You need this type of data management to discover and test good ideas. When a concept works, it needs to be raised from the fourth to the second quadrant, because you can only achieve economies of scale with data if you can generate and analyse it systematically. You can then use it enterprise-wide.

“I often talk to data scientists who obtain very sound insights in a kind of sandbox environment,” Damhof continues. “But they forget or are unable to monetize those insights in a production situation. They cannot bring their insights from quadrant IV to quadrant II. This is where governance comes into play.” And therein lies the major challenge for many organizations, as Damhof knows only too well. “If you explain this model to managers and ask where their priority lies, they will all say they first have to get their foundations in order, the first quadrant. But if you ask what they are investing their



“With organizations generating ever greater volumes of data, they can no longer be so slapdash in the way they handle it. Now is the time to make sure your data management and the associated governance are properly set up. The Data Quadrant Model helps you to achieve this”

Ronald Damhof

money in right now, where they are innovating, it is often in the fourth quadrant. It is great that they are engaged in this more experimental and exploratory form of data management, but that is only possible if your foundations are right. Otherwise it is like having a hypermodern toilet that is not connected to the sewer system, so it turns into a total mess.” Ask the average data scientist what takes up most of his or her time and he or she will answer getting the data to the right qualitative level: the aim of quadrant 1. “Only a data scientist with powerful analytical software, a lot of computer power, and high-quality data will genuinely make a difference.”

Reliability versus flexibility

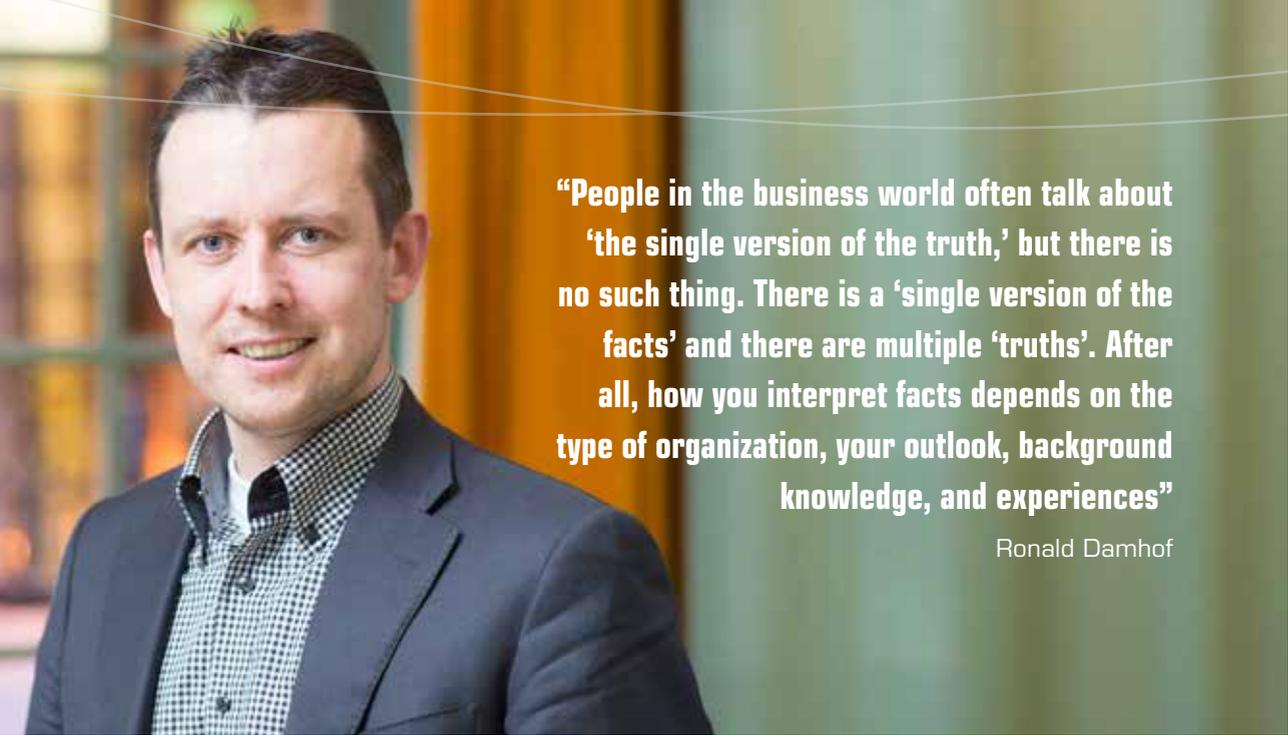
“Managers insist that systems must be reliable and flexible, but these qualities are inversely related. A highly reliable and robust system is less flexible. And in an extremely flexible system it is necessary to lower the requirements with regard to reliability,” says Damhof. “The Data Quadrant Model makes this clear to managers. In quadrant I reliability takes precedence over flexibility and in quadrants II and IV flexibility takes precedence over reliability.” Quite a few different types of expertise and competence are therefore required in order to make optimum use of data.

Expertise and competences

You often find that organizations require a single person to supply expertise and competences which cover the entire quadrant. Such people do not exist. Employees in quadrant I have an engineering profile. They are information and data engineers, trained in data architecture and data modelling. “Note that this is not the classic IT profile. These are engineers who can carry out model-driven development and have a solid understanding of the need for conceptual and logical modelling.” This expertise is very scarce. Quadrants II and IV on the opposite side require people with expertise in the respective business domain supplemented by Business Intelligence and/or analytical competences.

Facts and truth

Damhof also calls quadrant I of the model ‘the single version of the facts’. Those facts are then made available to employees in quadrants II and IV. That enables them to create their own truths. Since the same facts are used to create multiple truths in the right-hand half of the model – depending on the



“People in the business world often talk about ‘the single version of the truth,’ but there is no such thing. There is a ‘single version of the facts’ and there are multiple ‘truths’. After all, how you interpret facts depends on the type of organization, your outlook, background knowledge, and experiences”

Ronald Damhof

context and background of the data user – Damhof calls this half ‘the multiple version of the truth’. You should bear in mind that the ‘truth’ quite often changes over time. “You often hear companies talking about ‘the single version of the truth,’ but there is no such thing. After all, how you interpret particular facts depends on the context, your outlook, background knowledge, and experiences.”

Quadrant III

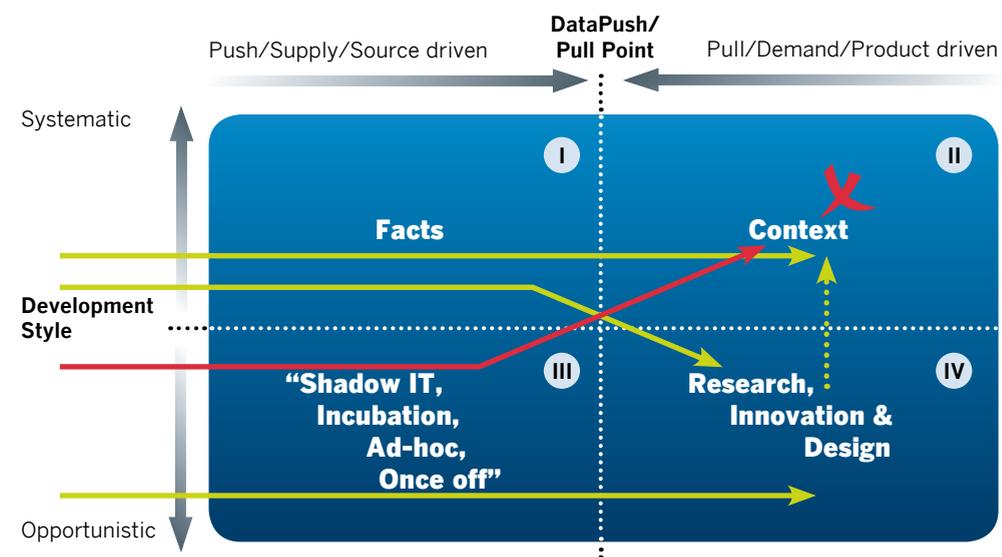
So far, Quadrant III has received little mention, even though it is incredibly important. It is the quadrant of data sources which are not under governance, like an ad hoc download which you obtain from an open data provider, a list in Excel that you want to use, or a set of verification data which you have received on a CD. “You may even want to combine governed data from quadrant I with your own dataset in quadrant IV, that’s fine,” says Damhof.

The journey through the quadrants

In order to get value from data, you can make various movements in the model. You can move from fact-based data management towards a model in which the context is also important (from quadrant I to II). “This actually is the classic journey of ‘unlock data and produce an information product,’” says Damhof. This is often inefficient, however, because this process is based on known requirements and wishes on the part of the user. “And the user does not really have that knowledge in advance.” Many organizations opt for a more agile-driven form, such as from quadrant I to quadrant IV to quadrant II. Have the employees in quadrant IV produce an information product in an iterative way using the data in quadrant I/III. You then promote the product to quadrant II only if it is important to bring this under management.

It is also possible to move from quadrant III to quadrant IV. “You have your own datasets and you want to try something? Great,” says Damhof. The only movement an organization must never make is from quadrant III to quadrant II. “Because in that case you use data that you are not entirely sure of, as it has not been subjected to good governance in the required way. An example is a compliance report for the regulator which you want to produce using data which is not under governance. You should not seek to do that.”

How we produce, process variants



Make data management a live issue for discussion

In his day-to-day work Damhof finds that his Data Quadrant Model helps organizations to talk about data management. “From my current customer, De Nederlandsche Bank, I regularly hear statements such as, ‘I want to move this data product from quadrant IV to quadrant II;’ or, ‘We must put the data in quadrant I first, but the submitter is really responsible for the data in quadrant I;’ or, ‘I want some space to store data temporarily in quadrant III.’ Everyone understands what it means. That is new; the organization has never thought about data in that way before. And that actually applies to almost every data-intensive company. Organizations have long spoken of data as an ‘asset,’ but in practice they handle data in a very unstructured way. As a result they never monetize that asset. With organizations generating ever greater volumes of data, they can no longer be so slapdash in the way they handle it. Now is the time to make sure your data management is properly set up. The Data Quadrant Model will help you to achieve this.” ■