

# Smart Data

---

**REPORT 1: HOW TO MAKE DATA WORK  
FOR YOUR NEWS ORGANISATION**

---

APRIL 2017

# IMPRINT

**SMART DATA**  
**REPORT 1: HOW TO MAKE DATA WORK**  
**FOR YOUR NEWS ORGANISATION**

**PUBLISHED BY:**

WAN-IFRA  
Rotfeder-Ring 11  
60327 Frankfurt, Germany

**CEO:**

Vincent Peyrègne

**COO:**

Thomas Jacob

**DIRECTOR OF INSIGHTS:**

Dean Roper

**AUTHORS:**

Xavier van Leeuwe and Matthijs van de Peppel,  
NRC, The Netherlands  
Matt Lindsay, Mather Economics, USA

**COPY EDITING:**

Anton Jolkovski

**DESIGN/LAYOUT:**

Ivan Cotic, Plain&Hill

# CONTENTS

<b>ABOUT THE REPORT .....</b>	<b>4</b>
<b>TAKEAWAYS .....</b>	<b>6</b>
What you can learn from this report .....	6
<b>1. LET BUSINESSPEOPLE LEAD THE DATA TEAMS.....</b>	<b>7</b>
Centralised data department versus decentralised data team .....	8
<b>2. RETHINK KPIS .....</b>	<b>10</b>
The wrong KPIs make you do strange things .....	11
Find KPIs that represent company health .....	13
KPIs drive culture .....	14
Bite the bullet.....	16
<b>3. PUNCH THROUGH THE BIG DATA HYPE .....</b>	<b>18</b>
The power of big data.....	19
Case study: How customer engagement affects retention .....	21
Use common sense when approaching big data .....	22
Case study: What data should be captured?.....	26
<b>4. START WITH THE BASICS .....</b>	<b>27</b>
<b>ABOUT THE AUTHORS .....</b>	<b>30</b>

# About the report

**F**or obvious reasons, news publishers have been diligently, and in some cases desperately, re-evaluating their business approach over the last decade.

Where the aim once was to maximize the value of individual transactions, now it is to maintain and maximize the long-term value of customer relationships. Instinct and intuition, once the main influencers of business decisions, are being replaced by data-analysis technology.

The traditional approach was to grow an audience – in print and/or online – and sell that audience to advertisers. The larger the audience numbers, the higher the ad rates. Thus “gaming” audience measurements became an end in itself. The value of transactions with advertisers was maximized, but only by milking consumers for all they were worth. In many cases, the consumer was pushed to accept bundles of paper they did not want, view streams of meaningless photos on web pages, or worse.

Obviously, that is no way to treat someone if you are interested in maintaining a long-term, mutually beneficial relationship with them. And publishers urgently need to develop those relationships. It is no secret that print advertising revenue is falling – dramatically in some countries – and digital ad revenue is growing much too slowly to compensate for the drop.

So audiences must take center stage. In 2014, global newspaper circulation revenues exceeded advertising revenues for the first time this century, and the disparity grew in the following year, according to World Press Trends.

Keeping and growing audiences means treating them in a completely different way: finding out as much as possible about the people who comprise them (within legal and ethical bounds, of course) and giving each one what they need – even if some transactions sacrifice revenues in the short run. The new KPI is long-term revenue per subscriber.



**Anton Jolkovski**  
Managing Editor  
WAN-IFRA

That's where advances in data-collection and analysis technology, often lumped under the term "big data," come in. As technology enables publishers to keep tabs on individual customers' activities, a mass approach to statistics is giving way to what might be called a "granular" approach. Each customer or potential customer can be addressed as a unique entity – but can still be grouped with others with similar characteristics when it comes time to discern trends and formulate tactics.

In fact, it is possible to quantify almost every aspect of customers' activities and chart interrelations among them. Such mind-boggling possibilities may give rise to the temptation to invest large sums in systems and specialists, to engage in big data for big data's sake. As Xavier van Leeuwe of Netherlands-based NRC describes in the introduction to Chapter 1 of this report, such an engagement can easily become an expensive dead end.

This report is, among other things, a caveat: Technology is only a tool, and publishers should start by making better use of the data they already possess – and closely heed the human factors that can make all the difference. Big data is indeed a potentially valuable tool, but "smart data" means putting that tool to use as part of a strategic approach that increases revenues.

Van Leeuwe, his NRC colleague Matthijs van de Peppel, and Matt Lindsay of Mather Economics have gathered extensive experience with smart data – and can point to substantial success. At NRC Handelsblad, they used analytics and other customer-relationship tools to turn readership decline into sustainable growth and point revenue development back in the right direction.

The three have decided to share their experiences and knowledge by writing a book, "How to Succeed in the Relationship Economy: Make Data Work for You, Empathise with Customers, Grow Valuable Relationships." They generously agreed to let us publish excerpts from the book in this series of three reports. We and the entire WAN-IFRA community owe them a huge debt of gratitude. Thanks, guys!

# Takeaways

Big data is hot and full of opportunities. At the same time, there are many pitfalls to avoid when applying data and business analytics in your news organisation. This report explains, step by step, how news organisations successfully use data as an instrument to help them reach their goals: by building no-nonsense analytical teams, unleashing the power of KPIs, and bringing common sense to the big-data projects that in many cases have absorbed substantial investments in time and money.

## **What you can learn from this report:**

- When it comes to analytics, put businesspeople in the lead.
- Put data people directly in your business teams.
- Dare to question your KPIs: are these the drivers of real value?
- KPIs can change culture and drive success.
- Don't do big data just for the sake of big data.
- Always start data projects with a positive business case.
- There's an ethical threshold for storing data.
- Understand local privacy laws and regulations.
- The first millions in extra revenue will probably come from understanding the basics of your business.

# Chapter One: Let businesspeople lead the data teams



**Xavier van Leeuwe**

Director Marketing and Data  
NRC Media, The Netherlands

A few years back, I entered the office for my first day on a new job. The very first question my new boss shot at me was, “Do we keep the data warehouse?” It cost half-a-million euros per year in maintenance alone, and there were millions of euros in sunk costs from the isolated staff of analysts who had worked at it for years – but nobody seemed to know what to actually do with the data. I decided to start talking to the analyst who was operating the system. I quickly noticed he was an extremely intelligent man.

Nonetheless, our conversation did not go smoothly. He kept on telling me what data points were available. “Look at that connection with the retail information. We can visualise things,” he said. But every time I asked him how we could put this data to use, he told me that was an inappropriate question. That deeply puzzled me. I did not understand what he meant. Inappropriate?

So I went on to talk to the marketers. They had heard that the data warehouse could do great things, that the data warehouse had lots of connections. But they were not using the data, because they didn’t know what to do with it. I thought, how could this have happened? After a week, I decided to get rid of the entire data warehouse, since it was only costing us money.

I now understand what was going on back then, and I see the same problem in many organisations: There is small group of analysts who understand the power of data. They receive a budget, lock themselves up in a room with other specialists, and emerge a few months later, telling the businesspeople they have this cool thing that will improve the company. But the businesspeople are too busy selling stuff, visiting clients, and producing the product. The businesspeople are disconnected. And because these businesspeople have trusted their gut instincts for so many years, they will not change their inner compass just because some nerds say they know better about the business and clients.

So we threw out an entire data warehouse only to build a new one years later – this time run by businesspeople.

That example shows the classic gap between IT and business, between the analytical and the practical. If we take a step back, we can see what is happening here. Analysts like to talk about the data architecture, the variables, and the connections. They tend to focus on the *what*. Some analysts take it a step further and come closer to the *why*. They generate insights and pursue a singular truth: *Why is this happening?* Once they know why, they are satisfied. That is the reason many analysts are enthusiastic about what an organisation can learn from data, because they derive energy from the insight itself. But that is not enough. We want to know what data can *do* for the business.

### **Centralised data department vs. decentralised data team**

There are two ways to design a data-driven organisation. One is to build a centralised data department. This will bring great benefits. It will be very efficient, because it allows for specialised analysts: one for data architecture, one for database building, one for reporting, one for descriptive analysis, one for predictive analysis (which requires additional statistical skills), and another for visualisation. You can put the best of the best on each aspect of the work. You avoid political conflicts involving data ownership, and definitions are the same for the whole company.

Centralised data teams, however, tend to focus on the wrong things, because the Achilles heel of any central data department is its communication (or lack thereof) with the rest of the business. These teams of highly intelligent people can be so disconnected from real business issues that they produce things that are at best considered unsolicited advice. A good analyst can predict which customers will cancel their subscriptions by using a tool called a *predictive churn model*. But if there is no concrete action by the business to prevent those people from actually stopping – such as contacting them or sending a gift – nothing is improved. Some analysts seem to believe that having the analytical model alone will improve performance, but that is a common pitfall of data analysis. Many people forget that insights are worthless if you do not use them to make better products, more effective campaigns, and/or better experiences for your customers. Centralised data teams can produce analyses very efficiently, but if there is no action attached, there will be no impact.

There is another way to put data analysis at the heart of an organisation. At NRC, we found some innately motivated businesspeople with analytical minds and asked them to lead the data teams. They had been conditioned to aim for practical consequences for the company, because most of the time they were rewarded based on their results, and they received daily reports about sales, financials, conversions, or other business metrics. They mastered so-called “contextual knowledge” and knew the business issues that kept top management awake at night.<sup>1</sup>

To find the kinds of leaders you need, watch for analytical people among your project managers, marketers, and salespeople. What are they doing after work? What did they study? Once you’ve found a good candidate, we suggest you take that leap. Put an analytically-minded business person in charge of the data team, and keep him or her extremely close to – preferably inside – the business teams. If these business-led data teams are truly empowered and feel the freedom to follow their own course, they will start learning how they can implement their insights in the business, ask the right questions, and transform those questions into actionable insights. The last step is what adds value.

These decentralised data specialists are like one-man armies: “*data Rambos*,” one telecom analyst calls them. They are not experts in every field of analysis, but they can handle the basics of every aspect of the field and will make sure they all operate highly effectively. Their unique knowledge is developed within the company, which can provide a competitive edge.

The challenge inherent in this approach is persuading businesspeople to learn about data and analytics. Marketers and sales representatives are not used to working with relational databases or thinking about reporting, data visualisation, or predictive analytics. When you start working with data, you need to know how the data is generated and to learn the lingo of IT, the databases, the software solutions, and the people. If IT and business build a relationship, the company will reap the benefits of this mutual understanding for a long time to come.

---

<sup>1</sup> Thomas H. Davenport, one of the leading thinkers on incorporating business intelligence and analytics in organisations, asserts that this kind of knowledge is essential for effective use of data. (2001)

# Chapter Two: Rethink KPIs

On my first day at NRC, my manager introduced me to the company. After a brief explanation about the coffee machine and a chat about the weekend, he showed me my desk. It was settled in a nice corner in the marketing department, surrounded by my new colleagues.

And there it was, on the wall, just behind my new desk: a bold statement, printed on a piece of A3 paper, saying “HOI +1.” Because “hoi” means “hello” in Dutch, I thought it was some kind of office joke. So I asked my manager what it meant. First he started laughing, then looked a little bit puzzled. He looked at me as if he was wondering whether he had made the right choice hiring me.

On the same day, I found out that HOI is the name of the standard used by the Dutch audit bureau of newspaper circulation and that the marketing department at NRC had one - and only one - KPI: growth in circulation, as defined by the audit bureau statements. Because the overall market was declining, growth by even a single newspaper copy would satisfy the requirement. There had to be a “green” number in the quarterly audit bureau publications.

At that moment, it made complete sense to me. With newspapers, it’s always about circulation. When circulation increases, more people read the paper, the impact on society grows, reach grows, and advertisers will pay more for ads. Circulation was the main KPI for every newspaper in the Netherlands and for many around the world - and had been so for ages and ages.



**Matthijs van de Peppel**

Manager Data Intelligence & CRM

NRC Media, The Netherlands

## The wrong KPIs make you do strange things

Every industry has its own version of audience measurement. In television they go on and on about the Nielsen ratings; the online industry is obsessed with comScore. In the news industry it has always been audited circulation. But are those the KPIs that help us grow the business?

For many years, the audit bureau's numbers constituted the guide for daily working life at NRC. Everybody worked their hands to the bone for a little growth in the official audience measurement. But to be honest, the singular focus on that metric made us do strange things.

Because we were so focused on circulation growth, every single marketing promotion had to result in additional papers leaving the printing plant. That was the main goal of what we did. We did not really care if anybody actually wanted the newspaper. Sometimes we even pushed papers on people who didn't want them, just to grow the official numbers. Below are four examples of bad decisions we made because of the focus on audience measurement statistics.

At NRC it was possible to get a two-day subscription, with home delivery of just the Friday and Saturday papers. Subscribers who chose that product probably did so for a reason. But considering ourselves clever marketing guys, we gave them an extra day. We told these customers, "Your subscription will be extended. From now on, you will get the paper on Thursday as well. We are sure you will love it. In the first couple of months, you don't have to pay for it. Enjoy!" We did not get hauled into court, but it was a bad decision, because many customers were angered.

Another example was the reward for sales agents. The full six-day product<sup>2</sup> would bring in a very nice bonus – much higher than for a weekend subscription – because it meant more paper out of the print plant, which in turn was good for audit bureau measurements. You can imagine what happened. With the higher bonus in mind, sales agents were pushing the six-day subscriptions no matter what. Even if potential subscribers told salespeople that they wanted to read the paper only on weekends, the agent would tell them, "This is a killer deal – just throw those unread weekday papers in the dustbin."

---

2 In the Netherlands, newspapers are published six days a week; there are no Sunday newspapers.

Sometimes, we were hurting even our extremely valuable, loyal subscribers. We made it as difficult as possible to temporarily suspend delivery when on holiday. We hid the “suspend delivery” option on the web site and started charging an administrative fee of 10 euros. That discouraged lots of people from suspending delivery, which was good for the official audience numbers. But at the same time, many subscribers complained that they had to pay us *not* to deliver the newspaper.

Last but not least, we delivered our newspaper to people who were not really interested in it. We used the amazing power of *free*. You could get our newspaper for free, on trial. A lot of people will take anything when the word *free* is invoked,<sup>3</sup> so this appeared to be quite a successful way to push more paper out of the print plant.

With the benefit of hindsight, it seems clear that none of those actions make sense. The customer’s voice is completely ignored, and you don’t have to be a professor in statistics to predict that this kind of behaviour will backfire in the long term. And backfire it did. A third of the two-day subscribers canceled, because they didn’t want a third day pushed down their throat. The newly acquired six-day customers stopped as soon as they could, because they were tired of throwing away five newspapers a week. The free trials didn’t result in subscriptions and were actually costing us money because of the cost of acquisition and variable printing and distribution costs. Finally, loyal readers who didn’t want to pay for holiday delivery suspension simply stopped their subscriptions before they went on vacation. Many of them did not re-activate after their holidays and were permanently lost to us.

Maybe the weirdest thing of all is that it took us years to realise that those actions were not helping the company. Because circulation was going up, official audience measurement numbers were green, graphs on the wall were going the right way, and editors were cheering. Everybody was happy.

It was only after a couple of years, when we realised we were running out of tricks to inflate circulation numbers, that we started to question the objective on the wall: “HOI +1.”

---

<sup>3</sup> Dan Ariely, *Predictably Irrational* (2009).

## Find KPIs that represent company health

When that brutal realisation struck us, we had to rethink our one and only KPI. Was it really true that circulation was the one thing that drove our newspaper business? We came to the conclusion that that was not the case, because when we ship a lot of newspapers out of the printing plant, it doesn't necessarily mean we are earning a lot of money in the long term. It does not mean that people are reading those papers. The correlation between paper circulation and company health was true only in the old days, when everybody had the same six-day subscription and was paying the full price.

These days, circulation can consist of many different kinds of subscriptions, with widely varying margins. A full-price weekend subscription brings in a lot more money than a "five weeks for 10 euros" trial subscription, but in terms of circulation numbers, the trial for the full week is six times better than full price for the Saturday.

In the old days, more paper out of the plant meant more reach and more advertising money. Nowadays, it's sad but true: advertising income is down for newspapers both when circulation is down *and* when circulation is up. The correlation is weak.

So we waved goodbye to our one and only KPI, which had driven our company since 1828. Then we started the hardest part: defining new KPIs. We sat in a room for hours, asking ourselves, "What really drives the value of this company, if not circulation as defined by the audit bureau?" We tried to frame it into a golden rule that is simple and always works.<sup>4</sup>

We couldn't get away with just saying we need a healthy financial margin, mostly expressed in EBITDA (earnings before interest, taxes, depreciation, and amortisation). That makes perfect sense, but you can't tell a marketing department to go out and earn some EBITDA. So what's below the surface?

That's when we came up with *relationships*. The real value in our company comes from the relationships we are building with our readers. And those relationships have to be real. It doesn't count if you push the newspaper into a mailbox and nothing happens. It has to be a serious, two-sided relationship, just like the ones you cherish in your personal life.

---

4 After reading Jim Collins' book *Good to Great* (2001), we used his idea of the hedgehog principle.

From that perspective, the term *core relationship* was born. A core relationship has a positive value both for the customer and for the company. The calculation is simple: count each of your subscribers as a single, full-value customer. Don't divide a Saturday subscription by six, as many circulation audit bureaus do. And don't count subscriptions where costs exceed revenue, such as trials that stop automatically.

## KPIs drive culture

The change appears to be a subtle one. Core relationships or audit bureau standards – isn't it just another calculation of the same thing? But the impact on the company has been huge. The change in weekly reporting was the trigger that changed the organisational culture. From the moment we started to report on core relationships, the mindset started shifting. The marketing department, traditionally driven by the sales charts of the day, started to worry about building relationships with new and current subscribers.

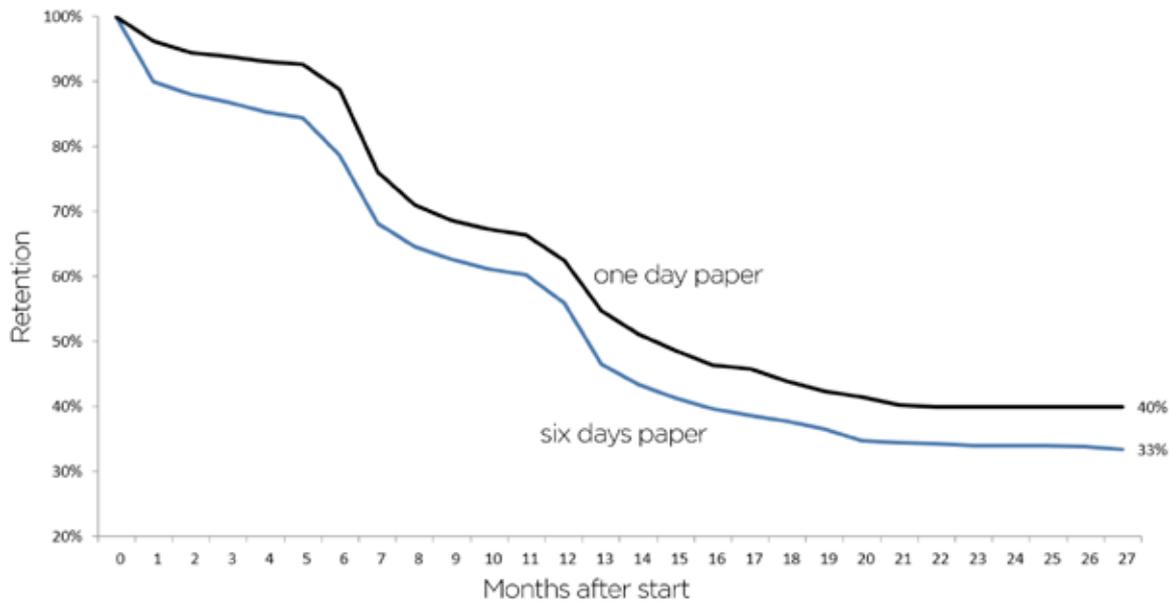
All of a sudden, our call centre was more interesting than any other department. We saw its true potential. The employees there were building relationships with our subscribers all day long. We wanted to get rid of stupid rules and stop hurting our customers with annoying fees or crappy service. Every individual subscriber counted – not just the six-day ones, who accounted for more in audit bureau measurement. The holiday suspension service became free of charge, and we started to promote it in the newspaper and with e-mails. Bonuses for sales agents became identical for all forms of core subscription. Like Starbucks, we introduced our own kind of LATTE system for sales reps.<sup>5</sup> We asked them to **listen** to customers, **acknowledge** their needs, **take** action by selling the subscription that fits their lives, **thank** them, and **explain** what the next steps are. That reduced cancellations and increased long-term retention.

We started to look for patterns in the data that indicated which products or promotions were building real relationships. For the first time, we analysed what happened after a subscriber came in. It turned out that 20,000 customers per year change their subscriptions because their current package doesn't fit their needs any more. We also found that the newly-acquired weekend subscribers stayed much longer than the full-week subscribers did, because a single newspaper per week was just better for their busy lives.

---

<sup>5</sup> Charles Duhigg, *The Power of Habit* (2012).

## RETENTION OF NEW SUBSCRIBERS AT NRC



Retention analysis shows more loyalty on the part of customers who chose a subscription for just one day per week. However, for the audit bureau, the six-day subscription counted for six times more circulation, attracting our focus but going against customers' needs.

Sales numbers also were not what they at first seemed to be. Because of the pushy behaviour of the sales agents, more than 10 percent of subscribers cancelled their orders within a month. We started to understand how we could listen to our customers by interpreting the data. If six-day subscriptions had a lower retention rate, then we had sold the wrong subscription in the first place. We hadn't listened properly to customers' needs.

In summary, the new KPI initiated a culture in which we listened to what our customers wanted, instead of trying to enforce what we needed. And from then on, we worked our hands to the bone to optimise customer experience and grow the number of core relationships. We started thinking of the longer term. How do we maximise the number of healthy relationships? Not by endlessly hunting for new relationships, but by making existing customers happy. To do that, we had to understand their needs. The table below illustrates how a new KPI drives a cultural change in many parts of an organisation.

## NEW KPI CREATES NEW CULTURE

	Circulation according to audit bureau	Relationships
Key metrics	Sales, revenue per transaction	Retention, lifetime value
Employees	Order takers, driven by senior management's needs	Customer success agents, driven by customer needs
Energy from	Salary, bonuses	Fulfilling a mission by meeting human needs
Interactions	Minimal: send the product	Critical: listen, touch base regularly
Way of working	Silo-oriented	Cooperation across silos
Focus	Self, vertical, internal, today	Other, horizontal, external, next year
Support	Minimise handling times, resolve incidents	Sincere attention, help use product more efficiently
Branding	Ad campaigns	Word of mouth
Salespeople rewarded for	Transaction, quantity	Lifetime value, quality
Acquisition	Always be closing, no thresholds, wide funnel, push what you want to sell	Attract the best customers, engage, motivate ambassadorship, intelligent threshold, chute, ask what customer wants

### Bite the bullet

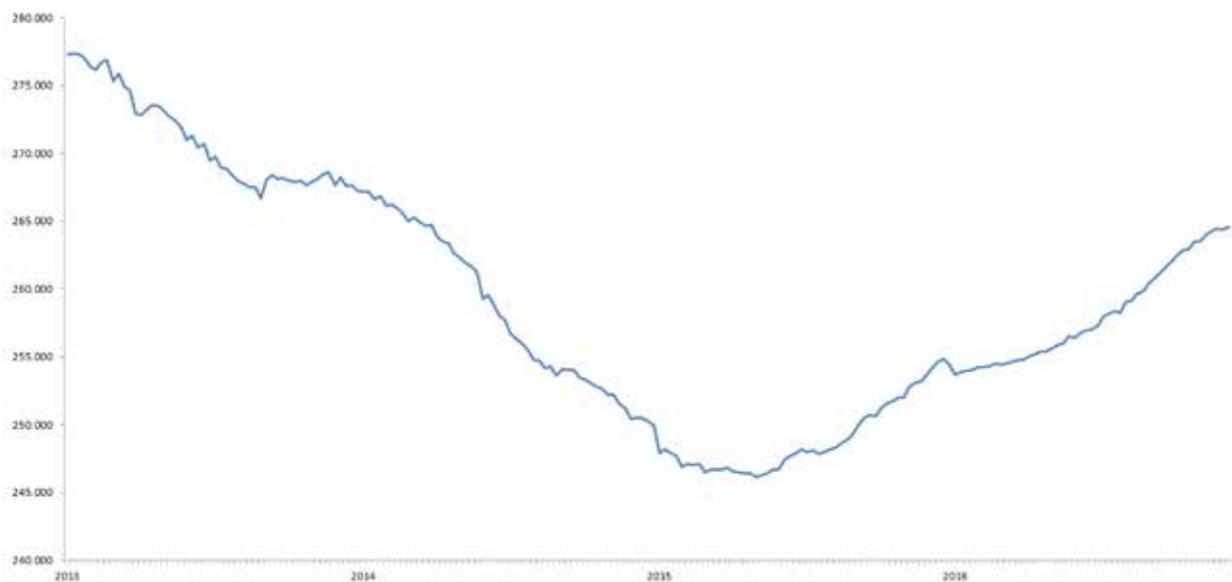
At the same time, some dramatic things happened to our circulation numbers in HOI. We stopped selling trial subscriptions that ended automatically, because our data told us that they were not building real relationships (much more on that later). So after a couple of weeks, a huge trial bubble burst, and we lost tens of thousands in the audience measurement standards.

We knew that was going to happen, and we were quite sure that it wouldn't really hurt our company financially, because those trial subscriptions had a negative value. But beyond a group of insiders, there were hundreds of employees and people outside the business who were not aware of the change in policy. And how could they have been? Even if we had told them with press releases and memos saying, "We changed our KPIs," it takes a while for every individual to say goodbye to a 189-year-old legacy – especially if your newspaper is the only one in the Netherlands to stop offering trial subscriptions and it is imploding in the quarterly industry reports.

So we had to bite a very big bullet. The editorial department questioned the sanity of the circulation department. How could the best newspaper in the country

have the worst circulation numbers? That is why you need top management support. In every market and at every company, the ride will be bumpy if you introduce new KPIs at the expense of the figures the rest of the company grew up with. You will need a very strong umbrella to protect you from the heavy rain pouring down.

### **NRC CORE RELATIONSHIPS 2013-2016**



# Chapter Three: Punch through the big data hype

**Matthijs van de Peppel**

Manager Data Intelligence & CRM

NRC Media, The Netherlands

In the past few years, there has been a lot of talk about “big data”.

At almost every conference and in many boards and marketing departments, the idea was put forth that the Holy Grail was to know everything about everyone at every moment. I was no exception. Developing holistic user profiles was on top of my priority list.

Then, one weekend, I visited my parents. Now and then, in their little village in the Dutch countryside, they make a brave attempt to understand how my brothers and I fill our working days. So I tried to explain my new project to collect as much individual data from subscribers and web site visitors as possible. We wanted to know much more than how to contact them and where to deliver the newspaper. What do they read? What is their income, their number of kids, their educational level, their birthday, their interests, the brand of computer they use?

We wanted all the personal information we could get our hands on to build profiles of them - everything from everyone, really.

At one point, my mother shyly asked, “Maybe I just don’t understand, but why does my newspaper have to know all that stuff about me? I don’t really like you to know everything about me. It feels a little intrusive.” By asking that simple question, my mother ensured that both my feet were firmly planted on the ground - as she had many times before.

## The power of big data

So what is *big data*? A common way to define the boundary between *regular data* and *big data* uses the three Vs:<sup>6</sup>

- **Volume:** the amount of data is large.
- **Variety:** the data are often not clean and tabular but rather messy, such as text or images.
- **Velocity:** new data arrive continuously.

In the publishing world, we usually speak of *big data* when combining online behaviour (such as clicks and scrolls) with offline data (such as names and addresses). There is some amazing software available to collect, combine, and report all that data. In the second report you will see what analytical tools, such as Mather Listener, can do.

Every company has the opportunity to collect huge amounts of data from its customers. That feeds the imagination, but what can you actually do with it? Here are some examples:

### Personalised products

By knowing your individual customers' interests, habits, and personal situations, you can match your digital products to their needs. Not everybody must see the same homepage on your web site, for example. Sports fans could be shown a different landing page than the one politics aficionados see.

### Targeted ads

Yes, Facebook has finally entered our story. It is earning huge amounts of money with targeted advertising. Facebook probably knows more about its customers than any other company in the history of mankind, because that's what it's primarily used for: sharing as much as possible about yourself. Facebook gives advertisers the opportunity to target campaigns at almost every imaginable segment of demographics and interests. Advertisers are investing large sums to reach exactly the audience they want to reach. Almost every traditional publisher dreams of reaping this kind of advertising money.

---

6 Doug Laney, 3D Data Management: Controlling Data Volume, Velocity and Variety (2001).

## **Data mining**

One of the most intriguing possibilities offered by big data is the chance to find hidden gems in your information. First, you store all the data you can get and drag it into the Amazon cloud – even before you have the slightest idea of what you will do with it. Then comes the magic part: unearthing value via data mining. Algorithms can find patterns in the data that you never could have imagined. This kind of analysis is used, for instance, to find correlations between DNA strings and the success rate of certain medicines. It can also be used to find factors driving churn (customers' cancellations), to cite just one example.

## **Predictive modelling and machine learning**

If your algorithms find correlations between certain variables, the next step is to make predictions. For example, foreseeing which TV series will be popular is a great feat on the part of Netflix's data team. Telecoms companies can effectively predict which subscribers are likely to churn. These predictions can be generated automatically, and the predictive model can improve itself via machine learning.

## Case study: How customer engagement affects retention

**Objective:** A publisher aimed to understand how engagement with its digital content affected retention of subscribers. (Subscribers got unlimited access to content on both desktop and mobile platforms in addition to print copies of the product.)

**Approach:** Data on site traffic was collected and matched to customer account information. Those customers who activated their digital access were identified, and their content consumption was measured in order to track their engagement. The site traffic data was captured using a tagging solution and stored in a cloud-based data warehouse. The raw site traffic data was aggregated by customer and matched to customer account data from the billing system. Customer activity was tracked for six months.

**Results:** Customers who had activated their digital access and consumed content digitally at least once a week were observed to have 91 percent retention. Customers who had activated their digital access but were infrequent consumers of digital content had 88 percent retention, and customers who had not activated their digital access had a retention rate of 84 percent. Controlling for other customer attributes using statistically valid samples for test groups, digital engagement was found to have a significant positive effect on customer retention. Based on those findings, the publisher started a content promotion campaign designed to encourage online registration and promote digital engagement.

### CONSUMPTION OF DIGITAL CONTENT VS. RETENTION



## Use common sense when approaching big data

There are lots of potential applications for big data, and storing data gets easier and cheaper every day. That combination presents a wealth of opportunities for new and existing companies. Consultancies pop up like flowers in springtime, software is developed at high speed, and no week goes by without an e-mail hitting our inboxes from some company boasting a state-of-the-art big data solution. There is an attractive new girl in town, and she keeps asking for a date. That's hard to resist.

But this girl is not the right partner for everyone. Big data is not equally relevant to every company, and building up holistic user profiles is probably not the thing to start with if you are turning into a data-driven company.

Before you start investing in Hadoop clusters, cloud storage, machine learning algorithms, and a lot of expensive and hard-to-find personnel, bring common sense into the equation. Ask yourself whether those investments will pay off at some point in the future. Are your customers interested in a personalised product? Are your advertisers interested in hypersegmentation? What is the likelihood that you will find patterns in your data that will earn you serious money? And last, but definitely not least: What will your family and friends think of you when you describe the kinds of information you are collecting about them? How do you feel about that?

If you have a solid business case for your big data project, the next question should be, what data is really needed here? Is it important to know “everything about everyone,” or will a couple of data points do the job?

## Find the million-dollar business case

Those were the questions we faced at NRC in the wake of my sobering visit to my parents. The first step we took was to visit companies in other industries that we had befriended and were far ahead of us on this path, to find out whether their extensive customer profiles had yielded golden insights.

We were greeted with openness and enthusiasm. Many companies were willing to share their experiences and show their systems. We saw amazing Hadoop installations, many terabytes of personal data, and numerous data scientists, analysts, and business developers who had invested serious time and money in big data. We asked only one question: Can you tell us about your million-dollar business case for big data? After all those investments, there must be at least one “golden application” for big data, right?

Despite the transparency, success cases were few and far between, and the connections between those successes and the big data projects were weak. One case involved a home improvement store that had an impressive loyalty programme that brought in huge amounts of data about their customers and their purchases. What you buy should say a lot about the stage of life you are in – how you live – so there are endless opportunities to analyse that data. But the store was unable to clearly attribute a revenue increase to having all that data about customers.

When we persisted, there turned out to be one thing that really paid off for the home improvement store, and it was much simpler than we expected. The store sent an e-mail to customers and asked what they were planning to renovate in the next couple of months. When customers responded that they planned, for instance, to remodel the bathroom, the company would send offers for bathroom tiles, showers, and taps. Response rates went through the roof.

That was the million-dollar business case. But it was not very complicated. It was one question and one answer – far from what you would call *big data*. The data volume is small, the data has low variety, and it does not come in continuously—it is just one data point through e-mail.

In fact, none of our contacts at those companies at the forefront of data usage could share examples showing that the holistic profiles of their customers really paid off. “But,” one analyst told us candidly, “it’s a great toy!”

## **Don’t do big data for the sake of big data**

Maybe that’s the risk of the big data hype. Most people in the boardroom don’t really know how these data projects work, but almost all agree they should “do something with big data,” because that’s what they hear at conferences and read in management books. For those who do know how big data works, this new era is a dream come true. There is money, fancy software, support from top management, and more data than ever. It’s like a planet-sized playground where you can play for ages. So if we lose track of common sense about data, chances are that companies are “doing big data” not for the business but just for the sake of big data.

As we wrote in chapter 1, it helps when businesspeople lead the data teams. They will make business cases and evaluate whether the investment in time and money is justified by a predicted revenue boost or an improved customer experience. If it is justified, go for it, keeping in mind that you have to start by doing three

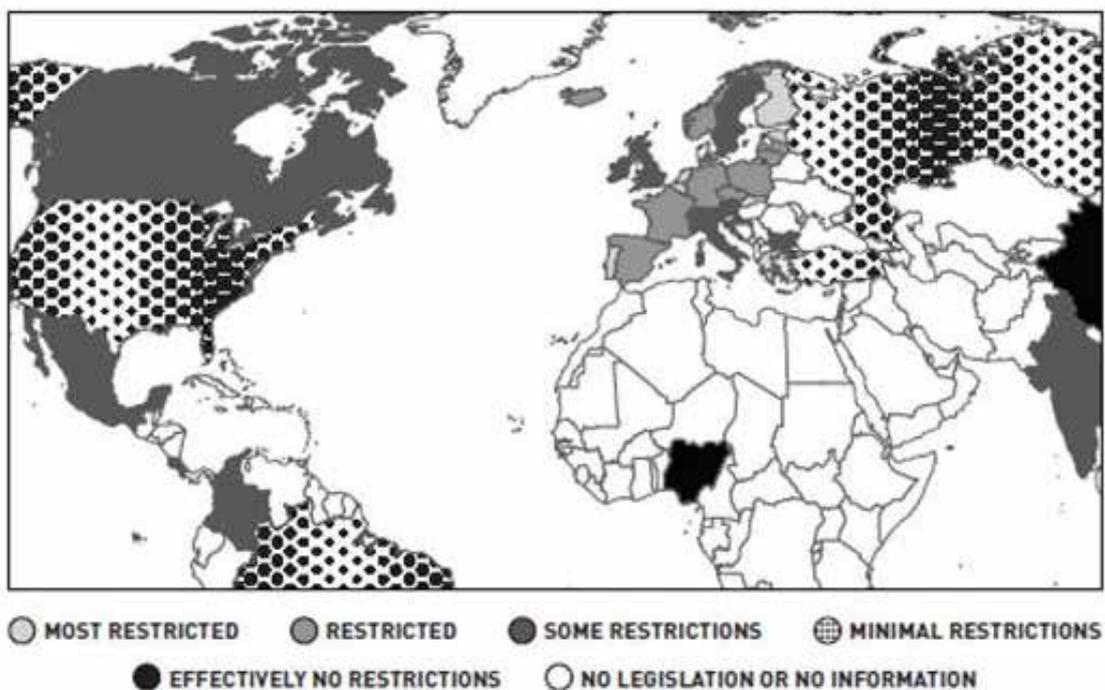
things: comply with privacy legislation, keep the relationship with your customers in mind, and be sure you understand your basic business processes.

## Privacy legislation

There are big differences in privacy laws and attitudes around the world. For example, people in the United States seem to have less concern about loss of personal privacy than Europeans do. Americans may more readily accept that data capture occurs and that companies use information on demographics and behaviours for a variety of purposes. There are Americans who actively try to thwart data collection attempts, but they are a very small part of the market. In fact, Americans may trust companies with their data to a greater degree than they trust their government.

Europeans appear to trust their governments with their data but keep a wary eye on data capture by the private sector. Privacy protection is a hot topic, legislation is strict, and regulatory agencies are active and powerful. Building a profile of customers without telling them exactly what information you store and what you do with it will soon be forbidden by the European Union.<sup>7</sup> More practical information about dealing with privacy in customer analytics is provided in the second report.

### DATA PRIVACY HEAT MAP (FORRESTER, 2015)



<sup>7</sup> By EU regulations on personal data that come into force in 2018.

## Step into your customer's shoes

In addition to determining whether a big data project will bring value to your business, there is also an ethical question to ask: To what extent do you store, connect, and analyse the personal data about your customers? If you are in the relationship business, you might want to be careful with the collection of knowledge about your partners.

If you decide that you want to know “everything about everyone,” ask yourself how you would feel if you found out that your wife or husband wanted to know everything about you at every moment. What if your partner tracked your phone, analysed your spending behaviour, and read your e-mails without letting you know? It would cause serious doubts about the trust and equality in the relationship.

At NRC, we have become very cautious in what we do with the personal data about our customers and website visitors. We always ask ourselves two questions before we start a big data project: Is there a positive business case? And can we explain to our mothers that we use this data?

In the future, another option will be to put the customer in the driver's seat. Tools will be developed to empower customers to control the communication and use of their personal data and dictate their own terms of service.<sup>8</sup>

---

8 Doc Searls, *The Intention Economy* (2012).

## Case study: What data should be captured?

---

**Objective:** A publisher wanted to develop a predictive model to support customer-acquisition campaign design and targeting. They sought to understand which product offers and customer attributes were important determinants of offer acceptance. The model findings would help the publisher design the acquisition offers and target the prospective customers most likely to accept the offers.

**Approach:** Mather Economics designed a model of customer acceptance rates using a field of econometrics called *discrete choice models*. To accurately predict outcomes, the model needed to include factors that were important to the customer's decision. If some important factors were omitted, the model's predictions could be biased and inaccurate.

The terms included in the model were product offer characteristics, customer attributes, seasonality patterns, and macroeconomic factors. Product offer characteristics included price per month, subscription term, automatic payment via credit card, and the length of the promotion period. Customer attributes included age group, income level, gender, presence of children in the household, and subscriptions to other periodicals. Seasonality and macroeconomic factors were represented by time variables such as the calendar month of offer and the Dow Jones stock-price average at the time of offer.

Mather Economics determined that those variables were the relevant data for the purpose of the analysis. Omitting certain customer attributes or other important data would make the model ineffective and unreliable, because those factors play important roles in predicting the acceptance rates of the offers. Sometimes the necessary data are not obvious until the modeling process begins and initial findings are obtained.

**Results:** A reliable, accurate model of customer acquisition was developed using data from prior acquisition campaigns. The model predicted acceptance rates for prospective customers, and the prospects were prioritised for campaign targeting. After testing the approach, the data feeds and list generation process were automated for ongoing operations. Using that targeting approach increased acceptance rates for direct mail campaigns by 40 percent.

# Chapter Four: Start with the basics

**Matthijs van de Peppel**

Manager Data Intelligence & CRM

NRC Media, The Netherlands

A couple of years ago we were in a room with three people - a marketer, a data warehouse developer, and an analyst - debating a question that at first looked very simple: "What is a subscription?"

We got three different definitions. The marketer was counting according to the audit bureau standard (a Saturday-only subscription is equivalent to one-sixth of a daily subscription, for instance), because the new KPI had not yet sunk in. The data warehouse developer counted the individual pieces of the subscription package (digital access as one, the printed newspaper as one), and the analyst wanted to count the bundles, regardless of frequency in delivery and number of pieces in the system.

Not being able to agree on the answer to this crucial question was confounding. We had to nail down this basic definition of our business for the brand-new data warehouse.

**S**o, now you have a data team led by business people, but you don't want to start with a big data project. What can you do?

We found that when trying to understand day-to-day business, there is a lot of work to do with the relatively small quantities of structured data already available in the current systems. It turned out there were a lot more basics than we thought, and we had to define many metrics before we could start to report and analyse them. What is a new subscriber? When is an ex-subscriber a prospect? When is a new subscription a “switcher”? What is the revenue per subscription? What are the variable costs? How many paying customers do we have, and at what price?<sup>9</sup>

Simply put, we didn't understand what was really happening with the 260,000 relationships we had – even though we declared that those relationships were the one thing bringing value to the company.

So we chose to invest time and money in developing definitions, data preparation in our data warehouse, and reports on the basics in our business intelligence tool. Those are not the sexiest subjects for data analysts and scientists. They're not big data toys. It's not Hadoop; it's just counting 260,000 records and their evolution. But at the same time, that is exactly what paid off. We started to see how many core relationships we had, as well as which promotions and products were driving the relationships and which were bad for relationships.

## STAGES IN DATA EXPERTISE FOR MARKETING PURPOSES

Basic	Advanced	Expert
Defining KPIs	Revenue forecasting	Customer profiles
Reporting	Volume forecasting	Propensity to buy
Financials	Retention analysis	Propensity to stop (churn)
Volume	Customer lifetime value (CLV)	Data mining
Upsell and downsell	Projection	Process mining
List management	Segmentation	Next best offer prediction
Funnel analysis	Identifying drivers of turnover	Online tracking of individuals
Return on marketing investment	Price elasticity analysis	Real-time dashboards
Online A/B-testing	Individual employee performance	Behaviour-based advertising

<sup>9</sup> In a 2013 article in the Harvard Business Review, Thomas H. Davenport calls this kind of descriptive analysis of basic business processes 'Analytics 1.0', as opposed to Analytics 2.0 and Analytics 3.0, which contain big data methods such as predictive and prescriptive modelling.

We began to understand what kind of service kept the relationships alive. After a while, we went from reporting the past to forecasting the future and started to understand the buttons we could push to improve results. More than a year after that confounding meeting at which we couldn't count our subscriptions, the improvements started to pay off, and we started to grow in terms of both relationships and our financial bottom line.

It is advisable to start with the basics. That is not a new or revolutionary recommendation, of course, but when you build a house, you begin with the foundation – not with a carport. For some reason, that's not how it usually goes in the field of data science. Many analysts will chose the sexy big data projects and skip the basics. It's like asking a mechanic if he would rather supercharge a car's engine or change the worn-out tires.

# About the Authors



**Xavier van Leeuwe** has over ten years of experience as a marketing executive at NRC, Telegraaf Media Groep, and de Persgroep. At the Amsterdam-based news organisation NRC, Xavier led the transformation of his team by building a customer-centric and data-rich culture at a news organisation with 189 years of heritage. Prior to being a media executive, he worked for several years as a financial and political journalist in the Netherlands. Xavier started his career as a researcher for the United Nations in Geneva. He has been a speaker at several conferences in Europe, the United States, and South America. Xavier is an active marketing tech blogger for [INMA.org](https://www.inma.org) and publishes a weekly free newsletter you can subscribe to at [Changemediaforgood.com](https://www.changemediaforgood.com). Xavier holds a cum laude master's in Business Administration and a degree in Journalism from Erasmus University Rotterdam.



**Matthijs van de Peppel** has been working in the newspaper industry for over ten years. Starting his career as a circulation marketer, he quickly developed as a online shop manager, online marketer, and project manager. Now manager of the data intelligence and customer relationship management team at NRC Media, a quality Dutch news organisation, Matthijs is responsible for bringing data insights into decision making and amplifying the voice of the customer. Matthijs holds a bachelor's degree in Dutch Language and Culture and a master's in Organisational Science from the University of Utrecht, and he attends a postgraduate curriculum for Business Analytics and Data Science at the Vrije Universiteit Amsterdam.



**NRC Media** is a Dutch news organisation that was founded in 1828. There are 7.5 million households in the Netherlands and 2.4 million newspaper copies distributed daily. NRC has over 265,000 active subscribers. Its news web site, [NRC.nl](https://www.nrc.nl), generates over 20 million monthly page views. The company has 360 employees, and its revenue ratio is 81 percent subscriber income and 19 percent advertising.



**Matt Lindsay** is president of the Atlanta-based Mather Economics, one of the fastest-growing private companies in the United States. Matt has over twenty years of experience helping businesses improve performance through pricing strategies and predictive models for clients including the Intercontinental Exchange, Gannett, The Home Depot, NRG Energy, IHG, McClatchy, the Everglades Foundation, Dow

Jones, and the New York Times. Matt began his consulting career with Arthur Andersen and has also worked with the the corporate Economics group of United Parcel Service. Matt is a sought-after expert and frequently speaks at media industry events including the News Media Alliance's MediaXchange, the INMA World Congress, and the WAN-IFRA World Newspaper Conference. Matt holds a doctorate in Economics from the University of Georgia, a master's in Applied Economics from Clemson University, and an undergraduate degree in Economics from the University of Georgia.



**Mather Economics** is a global consulting firm that applies a combination of proprietary analytical tools and operational expertise to help businesses better understand customers and, in turn, develop and implement pricing strategies that maximise operating margins, grow revenue, and improve customer loyalty. Mather was founded in 2002 and has forty employees. Mather assists about five hundred clients with their recurring

revenue yield management. These clients in turn serve thirty million households, with Mather analysing over \$4 billion a year in revenue generated by these households.

