



What it takes to deliver advanced analytics

The six essentials for a decision-ready business

Analytics: the promise and the pain

Businesses are buzzing with the promise of advanced analytics. And rightly so. When tied to a business strategy, a well-planned analytics initiative can make a big difference.

Analysis from 2012 found that companies that relied on data-driven decision-making were, on average, 5 percent more productive and 6 percent more profitable than their competitors¹.

And a more recent survey by Information Week identified the most compelling drivers for business analytics²:

Accurate forecasts of financial and operational results.

Optimized business operations to make sales faster, pricing more flexible and processes more efficient.

Identification of risk factors like customer churn, fraud, and defaulting.

The ability to predict new opportunities for upsell, cross-sell and whole new products.

The ability to engage intelligently based on smarter segments, and personalized interactions.

The expectations couldn't be higher. The trouble is, the reality doesn't usually match up:

Answers take far too long, with data scientists spending up to 80 percent of their time just 'wrangling' data³. The result? Decision makers are left with little time to actually make adjustments.

Confidence in the analysis is often low, thereby inhibiting decision-making. Just 10 percent of respondents to a TDWI survey considered their data to be of the highest quality⁴.

Incomplete data sets make comprehensive analysis difficult at best. If one region's CRM is better managed than another's, comparative analysis just isn't possible.

Different applications give different versions of the truth. The same customer could be registered as a twenty-year-old in one application and a thirty-year-old in another.

This much is clear: even after significant investment in next-generation analytics tools, teams, software, and warehouses, a large proportion of analytics projects usually underwhelm, disappoint, or outright fail⁵.

This eBook is about understanding why.

We'll outline six essential principles that keep your teams decision-ready — and what you can do to put them in practice.

Let's dive in.

¹ *Big Data: The Management Revolution*, Harvard Business Review

² Information Week: Analytics and BI Survey 2015, page 24

³ For big data scientists, 'janitor work' is the biggest obstacle, New York Times

⁴ *Customer analytics in the age of social media*, TDWI

⁵ Almost a third of BI projects fail to deliver on business objectives, ComputerWeekly

“Technologies such as predictive analytics, geospatial analytics, text analytics, and in-stream analysis are all poised to double in use over the next three years if users stick to their plans.”

Fern Halper
TDWI Best Practices Report
Next Generation Analytics and Platforms
1Q 2015

“A new spirit of experimentation is in the air: for the first time in years, more companies are trying out multiple analytical tools rather than standardizing on a few favored ones. Data quality and ease of use remain the toughest challenges.”

Doug Henschen
Information Week
2015 Analytics and BI Survey

The next generation of analytics challenges



The next generation of analytics challenges

No trend has played a bigger role in shaping the next generation of analytics than big data. But the next generation of analytics and business intelligence systems will be defined by a number of new factors:

Multiple data sources — some internal, some external.

New goals — predictive, prescriptive, operational intelligence.

Complex data models — with disparate data including text, social, mobile, big data, third party data, video, voice, and geospatial.

Dynamic analyses — real-time streams as well as batches.

Massive volumes — large enough to affect budgetary decisions.

Distributed processing and storage — with new tools, platforms, and languages.

With all these changes, next-generation analytics call for an urgent re-think of the way analytics stacks and teams are set up.

Without the right architecture and management, there's a very real danger your efforts to understand your business and customers could end up fragmenting your analytics operations into technological silos.

When confronted by these challenges, most organizations make two crucial mistakes:

- 1. They DIY themselves into a hole** — with hand coding taking so much effort, analysts and data scientists end up with little time to actually analyze. Worse still, the errors and limitations of all that hand coding make their analytics operations hard to scale.
- 2. They buy another application** — while the disappointment of previous analytics projects is an important motivator, acquiring another tool for visualization or analytics is rarely the answer. The result is often just more confusion.

If next-generation analytics aren't living up to all they're cracked up to be — and the applications and the people working on them aren't the problem — then what is?

*Your data isn't
decision-ready*



Your data isn't decision-ready

Any way you slice it, the success or failure of your next analytics project depends on the data being analyzed. No matter how well it's visualized, no matter how sophisticated your models, no matter how many data scientists you hire, without great data your analytics project is doomed.

If the data going into your analytics stack is dirty, noisy, duplicated, incomplete, poorly integrated, or delivered too late, the insights coming out of it can't be trusted.

Of course, this isn't exactly breaking news. Data quality problems and integration challenges are the two most commonly cited barriers⁶ to analytics success among people who work on analytics tools.

But the fact remains — until your data's ready, your people and processes won't be decision-ready.

We define great data as:

Clean — accurate, de-duplicated, timely, and complete data that matches pre-determined standards.

Safe — data that's compliant and secured so only the right people have access to it.

Connected — data that reflects the whole truth consistently.

As many data teams are discovering, having great data that's clean, safe, and connected is the key to successful next-generation analytics and BI projects.

Treating intelligent data management as a strategic first step delivers a solid foundation for all analytics programs. Great data has some important qualities. It's:

Trusted — ready for reliable analysis, giving your decision makers confidence in the facts they're basing their judgments on.

Timely — ready for new questions, so your decision makers can get the insights they need in time to take action.

Inclusive — ready for new data sources, like partner data, social media, or geospatial data.

Accessible — ready for non-IT users like business analysts who don't have the time or skills needed to clean and integrate data.

Scalable — ready to take on more projects and data sources.

Put simply, intelligent data management ensures you're working with great data. And great data ensures your analytics projects, tools, and teams are decision-ready.

⁶ Information Week: Analytics and BI Survey 2015, page 7

Your data isn't decision-ready

Eight symptoms of bad data

Most organizations don't realize just how poor their data really is. Here are eight signs that messy data is lurking beneath your shiny new applications:

Dueling spreadsheets — clashing reports that confuse decision-makers.

Data reticence — management doesn't trust the data is worth acting on.

Attribution blindness — you can't tell what works and what doesn't.

Insight lag — it takes too long to get the answers you need.

Application apathy — your people lose enthusiasm for key applications.

Talent drag — your rare, expensive data analysts aren't having fun.

Market myopia — no one has a complete picture of anything.

Compliance chaos — your people are scrambling to rise to even the most basic data stewardship and privacy challenges.

The six essentials of great data



The six essentials of great data

If the key to next-generation analytics success is great data, how do you ensure your approach to data management yields the right results?

We believe intelligent data management comes down to six essential principles:

- 1. Repeatability**
- 2. Abstraction**
- 3. Alignment**
- 4. Collaboration**
- 5. Automation**
- 6. Governance**

We'll explain what we mean by these principles and show you what you can do to put them into practice. Deliver on these six essentials and you'll be ready for next generation analytics.

The six essentials of great data

One Repeatability

Hand-coded data integrations and manual data cleansing solve your problems once. The trouble is these problems will come up over and over again. Every new dataset will need the same granular cleansing as the last one. And every new data source will need to be extracted, transformed, and loaded.

If the efficiency and scalability of your next-generation analytics matter, repeatability is an essential target.

That means finding a way to comfortably reuse the logic you've created to cleanse, integrate, and transform your data so you aren't reinventing the wheel over and over again.

Action points

1. Start by documenting your current processes for crucial steps that come up every time. Things like data cleansing and data integrations.
2. Analyze the processes to find out how long they take and whether or not your team follows the same steps every time.
3. Identify opportunities to standardize these steps and create a common roadmap and logic for teams to follow.
4. Once you know these details, look for tools that can help you standardize these steps in an efficient and scalable way (more on this in principle #5).

Why this matters

The top three impediments to information management success are accessing, integrating, and cleansing data⁷.

Make these steps repeatable (and the logic you rely on reusable), and you'll have overcome the most important obstacles to analytics success.

Because the real value of your analytics projects is in its ability to deliver insights when decision makers still have time to adjust their plans. Lower the time spent managing your data and you'll lower your time to insight.

⁷ Information Week: Analytics and BI Survey 2015, page 27

The six essentials of great data

Two Abstraction

Analytics is a journey. You may start with a data mart, add a data warehouse, add real-time streams and then eventually decide to move to a data lake.

So it pays to build a data management architecture that works across any type of data, any storage technology, and any BI tool. In particular, it should be able to run on top of any type of storage.

If introducing Hadoop or NoSQL means you need entirely new data management tools then you're at serious risk of creating yet another disconnected 'silo of data.'

By separating your integration modules and data management processes from these systems (the factories), you future-proof your raw materials (the data).

Layered architectures give you the flexibility to easily add new data sources and target applications.

Action points

1. Choose data management tools that abstract your data from the underlying storage technology.
2. Make sure your integrations and cleansing processes take place in the data management layer rather than in applications.

Why this matters

It's clear that the technology used for data storage in analytics is changing rapidly.

But when your data management platform is architected to abstract you from that change, it gives you the flexibility you need to make sure your teams and processes are always decision-ready — no matter what.

The six essentials of great data

Three Alignment

Vocabulary is a particularly contentious part of every analytics project. A 'client' to one team may mean the 'client company' to another. When you zoom in to even broader terms like 'solutions' or 'deals', these discrepancies can cause some serious problems.

Successful analytics projects combine the domain expertise of people on the business side with the technical proficiency of people on the IT side. So a key aspect of decision-ready data is that it's all based on a common business vocabulary that aligns the expectations of the business with IT's execution.

Alignment ensures your integrations and deployments deliver even if someone miscommunicated what the data actually means.

Action points

1. Create a business glossary that's easy to share and work with.
2. Actively manage and share your terms, definitions, term owners, and rules across both the business and IT. For instance, when analyzing revenue, ensure it's clear whether it pertains to gross revenue, net revenue, local revenue, or global revenue.
3. Make sure your business terms are linked to the actual data. By physically linking the two you'll create an environment where there is no room for ambiguity between business (term) and IT (physical location) of the data.

Why this matters

A healthcare organization was experiencing conflicting reports. The problem turned out to be in the understanding of the meaning of the term 'claims paid date.'

A different analyst interpreted this to mean the date the claim was approved, the date the check was cut, and the date the check cleared. The resulting misunderstandings were worth millions of dollars.

It may sound simplistic, but the first step to IT and business alignment is getting both sides to speak the same language.

The six essentials of great data

Four Collaboration

The consumerization of IT has changed the way enterprises interact with technology. For better or for worse, non-technical employees, such as business analysts and business professionals are the primary users of analytics tools.

So it's essential that your approach be one that empowers non-technical users to self-service and work more naturally with both IT and the data.

An IT bottleneck negatively affects motivation and kills the initiative of the domain experts who'll invariably be the ones extracting ROI from your analytics tools.

Action points

1. Listen closely to the needs of specific user types and give them access to tools they can use themselves.
2. Specify rules, terms, and requirements from IT — from workflows to vocabulary.
3. Open up the lines of communication so that IT and business users can collaborate directly with one another.

Why this matters

The top three users of advanced analytics are business analysts, statisticians, and business users⁸. That's a lot of non-technical expertise being directed at fairly complex technology.

The democratization of data analysis is happening, regardless of how difficult it is. The most logical way forward is to make sure it isn't difficult for your technical experts to work closely with your domain experts.

⁸ TDWI Best Practices Report, Next Generation Analytics and Platforms, 1Q 2015, Page 14

The six essentials of great data

Five Automation

As mentioned in the first principle on repeatability, the efficiency and scalability of your analytics operations is crucial to its long-term success.

But as long your data management depends on manual efforts to integrate and clean your data, your analytics projects just won't be sustainable. The pragmatic thing to do is to automate those parts of your processes that take up the most time and require the least manual intervention.

Not only does this mean your expensive big data talent won't spend the bulk of their time on menial data wrangling tasks. It also means they'll be doing more analyses and more innovation.

Action points

1. Identify the most time-consuming data management processes and make the business case for automating them.
2. Look for tools to help you automate the data management processes your organization has defined. Processes like address correction, relationship mapping, and format correction.

Why this matters

Southwest Power Pool benefited in non-trivial ways when it adopted automation. It reduced typical analysis time from one day to 20 minutes by automating its processes for data integration, replication, and delivery⁹.

That kind of uplift is important when you're testing hundreds of scenarios and looking to derive answers based on complex models.

But whatever the scope of your analytics operation, its effectiveness shouldn't come at the cost of its efficiency.

⁹ Southwest Power Pool's customer success story, Informatica

The six essentials of great data

Six Governance

In order to ensure intelligent data management happens across the enterprise, data governance is an essential process.

Not only does it signal that data will be managed as the crucial business asset it is, it also ensures everyone involved in your analytics projects gets the support and shared perspective they need.

Depending on your organization's objectives, you can use a broad or narrow data governance program to oversee data quality, manage security and compliance, or standardize processes and technologies.

Action points

1. Identify any data risks and the policies needed to manage them.
2. Outline the policies, standards, and rules your teams need to follow.
3. Identify the most important stakeholders who would benefit from next-generation analytics and start the process of discussing a potential data governance strategy.
4. Look for points of conflict and issues that relate to a lack of accountability. These are the kinds of pain points an effective data governance strategy can help with.

Why this matters

Effective data governance programs — no matter how narrow or broad their focus — resolve conflicts, define decision rights and frame analytics projects in the context of the enterprise's wider needs.

They're an integral part of the process, making sure different teams don't duplicate their work, technology purchases, and compliance controls.

So an effective data governance strategy is a sure fire way to make analytics success predictable and sustainable.

Getting decision-ready: making next-generation analytics work

In a recent survey, more than 50 percent of respondents said next-generation analytics were extremely important to strategic decision-making, with just under 50 percent saying it was extremely important to improve business processes and performance¹⁰.

If you're reading this eBook, you're likely an equally staunch believer in the potential of next-generation analytics. We believe — and we hope we've convinced you — that the most important gap between the promise of the new analytics and the reality you work in today is the quality of your data.

No matter how many great applications and visualization tools you throw at the problem of poorly managed data, the results will always be the same —unreliable analysis and stunted decision-making.

Make no mistake — the ultimate value of analytics solutions doesn't lie in their ability to deduce insights from large amounts of data. It lies in their ability to improve critical business outcomes.

So even if the answers coming out of them are the answers your business needs, they'll count for nothing if they come too late, too infrequently, or at too high a cost.

The goal of next-generation analytics may be to fuel better decisions. But only intelligent data management gives you the tools to become a decision-ready enterprise.

¹⁰ TDWI Best Practices Report, Next Generation Analytics and Platforms, 1Q 2015, Page 7