informatica

decision ready.

Laying the foundations for next-generation analytics.

A workbook for building an intelligent data layer for a decision-ready enterprise.

01
02
03
04
05
06

Contents

Tip: Click on parts to jump to the particular section you want

The rise of next-generation analytics

Interest in data-driven decision making is at an all-time high.[1]

But it's never been harder to make analytics projects work. More data, more sources, more structures, more users, more use cases — next generation analytics initiatives are incredibly complex.

It's why so many projects have failed.[2] And it's why we've written this workbook — to help you make sure your project doesn't.

The challenge is twofold. On the one hand, you need to deliver clean, complete, and timely data in support of actionable business insights that will improve your organization's effectiveness and competitive position.

But at the same time, you need to be able to deliver "good enough" data for fast-paced, innovative projects that are using analytics to ask new questions. The goal is to be able to quickly see if the answers have value before investing the resources to operationalize the analytics that provide those insights.

What's needed is a common approach to intelligent data management that works across both these worlds.

We've written this workbook to help you develop this approach and plan a successful analytics project.

Because the success of analytics projects comes down to a variety of different factors ranging from issues with planning (whether or not the goals are clearly defined), to issues with technology (whether or not your architecture is flexible enough to meet shorter timeframes comfortably).

By the end of it, you'll be ready to lay the foundations for intelligent data management so that your initiative can deliver clean, complete, and timely data rapidly and reliably.

Let's start.

Introduction

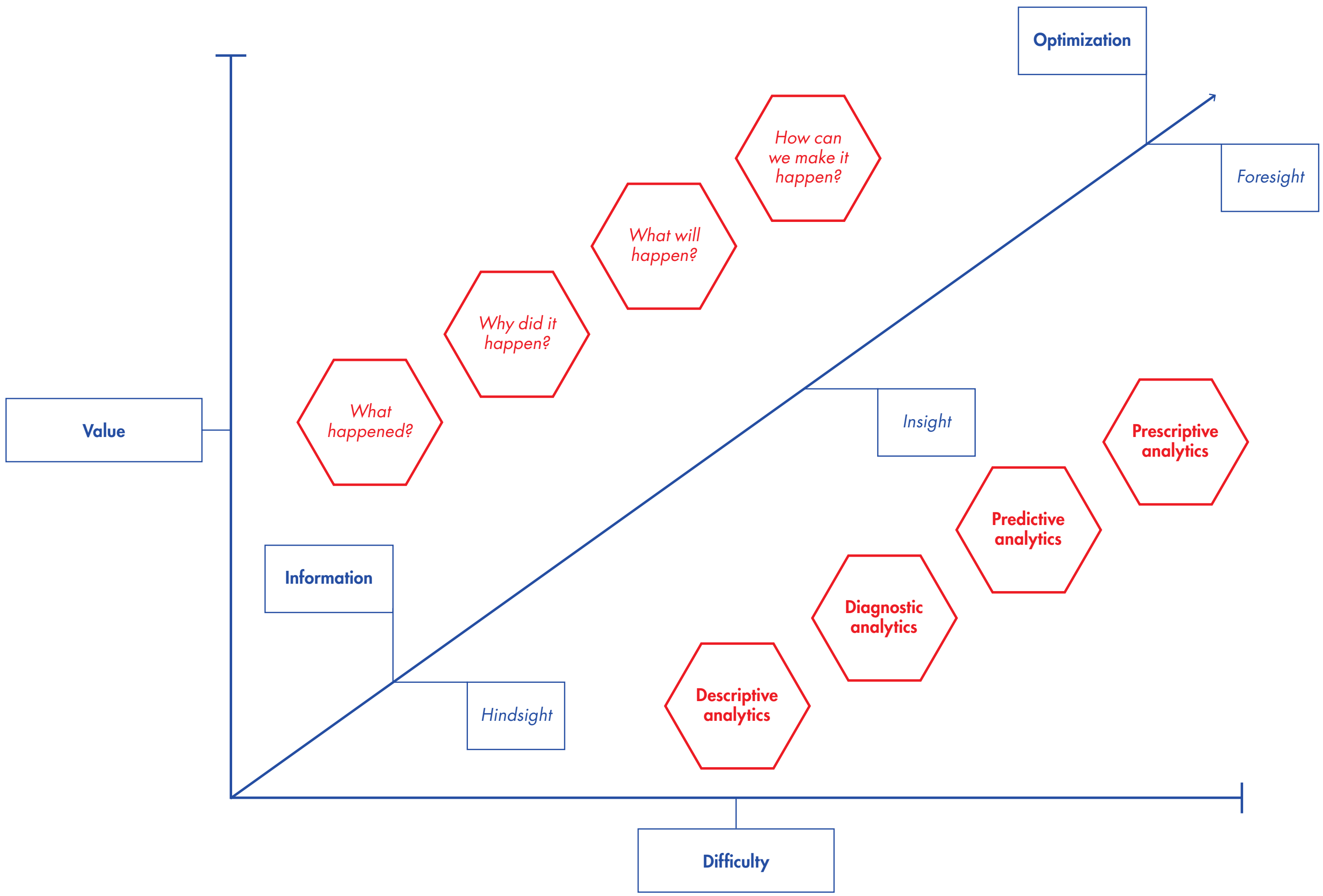
Why analytics initiatives are hard

Advanced analytics tools and projects are evolving rapidly.

But as soon as an innovative new tool yields useful business insights, it has to be operationalized. Which is great for the people analyzing the data.

But it's a challenge for the people managing the data. Because data that is good enough for ad-hoc analyses usually isn't good enough for making significant business decisions.

And since lines of business are seeking competitive advantage from analytics, IT is under ever more pressure to improve its ability to provision the data rapidly.



Source: Gartner IT Glossary: Predictive Analytics

Why analytics initiatives are hard

The need for speed

When you struggle to meet the business' tight timeframe requirements, the business doesn't just sit and wait. They turn to new-wave data visualization tools like Tableau and QlikView to solve their immediate challenges.

But while these tools have exceptional capabilities when it comes to analysis, visualization, and reporting, they aren't built to solve underlying data management issues like data quality, integration, security, and governance.

Which means your colleagues in the business are effectively forcing a trade-off between the speed with which they can stand up a solution and the quality of data being fed into it.

So it's fair to say that the challenge for your analytics initiative is about delivering analytic insights rapidly — but also ensuring the data going in is of the appropriate quality level for the purpose intended.

The need for agility

When it comes to analytics initiatives, change is both constant and inevitable.

- *The nature of projects is changing:* From traditional, IT-led production-quality business intelligence to ad-hoc business-led analytics.
- *The use cases are changing:* From straightforward historical reporting to real-time decision support and, predictive analytics.
- *The tools are changing:* From the old paradigm of business intelligence and data warehouses to the new wave of analytics and data visualization tools like Tableau and QlikView.
- *The sources of data are changing:* Data has been captured in application-specific data silos that weren't architected to share their data. Now, we are adding big data, and increasing volumes of external data, all of which will have different or unknown structures, data context, and quality.
- *The analytics data storage technology is changing:* For years, business intelligence was done with traditional data warehouse technology. Now we're seeing rapid technology change to newer, more specialized forms of analytics data storage such as Hadoop, NoSQL, data warehouse appliances, and data virtualization — most of which are evolving rapidly.
- *The data itself is changing:* From highly structured data to huge volumes of data with little or no structure.

Most current enterprise data lives in a set of application-specific silos that make it difficult to discover, access, and share data. So if every new analytics project isn't designed from the outset to be a shared resource, it runs the risk of becoming just another data silo. And even the new wave of analytics technologies for persistence and analysis will end up raising the danger of slowing down business insight delivery.

In short, the success and failure of your analytics initiatives will depend on your ability to manage change rapidly and comfortably.

Why analytics initiatives are hard

The business is going to be forced to choose between traditional BI tools that are too slow to adapt and shiny, new analytics tools that can't deliver production-quality data.

If you're reading this workbook, you know all these challenges all too well. Our aim is to guide you to make foundational decisions about what you need to build intelligent data management into the very fabric of your analytics initiative.

Intelligent data management

We define intelligent data management as the collection of capabilities you need to produce great data in a scalable, flexible, and reliable way.

We define this great data as:

Trusted – Ready for reliable analysis, giving your decision makers confidence that the data fueling their decisions is clean, complete, and without duplicates.

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, blogs, machine data, or geospatial data.

Accessible – Ready for non-IT users like business analysts to easily discover, assess, understand, and use.

In short, by the end of this workbook, you'll know what you need to build out the foundations for intelligent data management so that all your analytics projects — present and future — are fueled with great data.

Part 1

Aiming for the right initiative

The importance of focus

No two analytics initiatives are the same. But every analytics initiative — big or small — should start with a clearly defined vision of the foundations and future analytics architecture being built.

Without a clearly defined vision of the analytics data architecture you need, you'll be investing precious time, energy, and resources into ad-hoc projects that may deliver a project quickly, but at the cost of slowing down future development by creating another silo of data.

They won't make your ability to provision great data any faster. They won't make your underlying architecture any more agile. And they won't make your life any easier. A clearly defined vision of your future state architecture is necessary.

In our experience, there are two perfectly valid approaches to building towards your vision, once you've defined it.

First, you can go big and build a completely new, enterprise-wide data architecture for analytics. These kinds of initiatives tend to take longer and require buy-in from the highest levels of the organization.

Second, you can build towards the future analytics data architecture your enterprise needs one project at a time. This approach is far more common and relative to the first approach, it's easier to show business results quickly while growing incrementally.

This would be a great time to talk to your Enterprise Architect about your strategy for a future state architecture and how it should aim to support the business' goals and strategy. Bear in mind, your data architecture should be designed for interoperability: all data should be easily discoverable, accessible, and shareable by any authorized user.

Depending on the executives backing you, the use cases you're solving for and the amount of time and budget allocated to your initiative, you'll know which of these two approaches makes the most sense for you.

But knowing the difference between the two is essential.

Because one of the most commonly cited reasons for analytics failure is a lack of clear goals and expectations.[3] So getting to grips with the precise expectations of your most important stakeholders is fundamental to the success of your initiative.

So before we go any further, we're going to help you get clear and specific about the scope and goals of your initiative.

Three big factors that shape analytics initiatives

The way we see it, the scope and goals of your initiative should be determined by three factors:

1. The expectations of the executives backing you:

If they're expecting a tactical project to solve near-term issues, everything from your team to your architecture should aim to deliver on those goals. At the same time, you'll need to be careful not to create new data silos that will slow down future projects.

But equally, if they're expecting a strategic initiative to fuel multiple analytical use cases for the whole enterprise, you need to think deeply about your analytics and data architecture and design accordingly.

2. The analytical use cases you're aiming for:

In other words, the specific problems you're attempting to solve with next-generation analytics. If you're building churn analytics for customer service, your priority must be to build a reliable, scalable, and flexible architecture for them.

But you will also need to think in terms of an enterprise data management architecture that will fit the needs of future analytics use cases and allow you to adapt to analytics technology changes without disruption to your business.

3. The time and budget allocated to your initiative:

Project delays and cost overruns are often a result of poorly communicated goals and mismanaged expectations.

There are a million reasons for goals and expectations to become misaligned, but you have to ensure your project's goals are aligned with the amount of time and money made available to achieve them.

If the business lacks clarity in terms of its expectations from your initiative, don't take that as license to try and 'boil the ocean'. Probe for their business goals and objectives, and then manage your analytics initiative accordingly.

To sum up: manage expectations, get consensus on the business problems that need to be solved and stick to clearly defined budget and time frame requirements.

Part 1

Defining the scope of your project

In this section, we're going to help you align your expectations with those of the stakeholders on the business side. So grab a pencil to fill this section out.

Once you're done, print some copies and ask your colleagues on the business side to fill it out too.

That way, after you and your colleagues have filled this out, you'll be able to compare their answers to yours and analyze the difference (if any) so you can re-align before you move forward.

Tip: If you have multiple analytics initiatives in the planning stages, print multiple copies out and fill them in for each one.

Part 1

Defining the scope of your project

Project Name:

1. The business problem

Remember, specificity is key. So rather than saying "we need customer analytics" use this format to outline the core business problem you're trying to solve with analytics.

We lack insights into…

So far, we've relied on…

To make decisions about…

This isn't sustainable because…

Defining the scope of your project

2. The stakeholders

Here we're trying to list out the people and/or teams who'll be relying on your analytics initiative.

Pencil in your answers and if any surprises crop up when you compare it to your colleagues' answers, it'll probably be worth your time to talk to the people you've missed and find out what their expectations are.

Decision makers and budget owners

The following teams and/or individuals will be funding this project and relying on reports and/or insights from it:

....

....

Analysts, scientists, and end-users

The following teams and/or individuals will be using the front-end technology from this project to analyze the data and produce reports and/or insights:

....

....

Implementers (IT)

Who are the IT managers, architects and key implementers of these initiatives?

....

....

Defining the scope of your project

3. The time frame

Write down (in hours, days, weeks, months, or years) how much time you have to make this analytics initiative go 'live'.

(In the case of broader, enterprise-wide initiatives, list out the end date as well as the time frame for any strategic milestones you may have in mind):

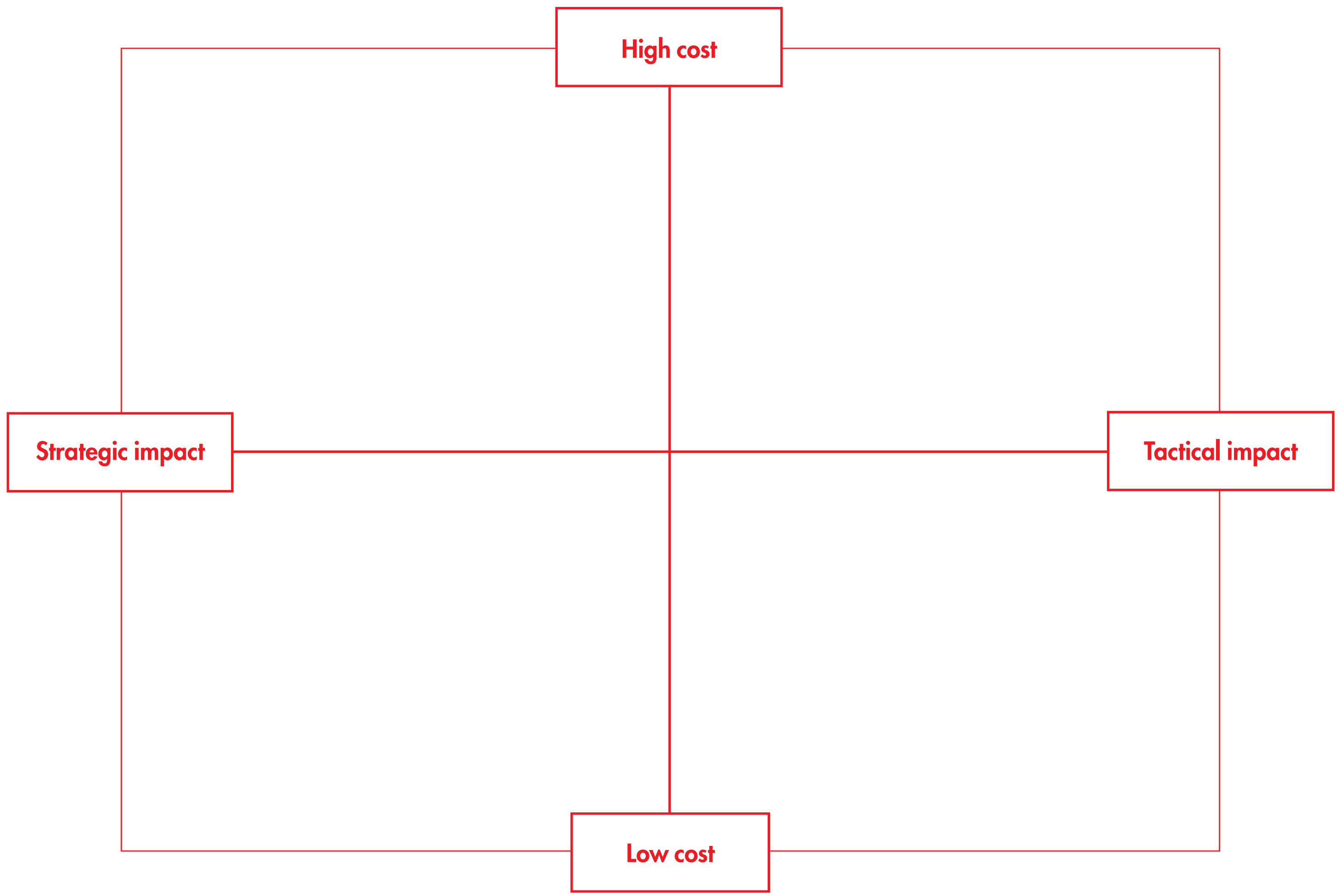
Part 1

Defining the scope of your project

4. The business impact

Use this two-by-two matrix to align your understanding of all the planned analytics projects' perceived impact with key stakeholders in both the business side and IT. Based on the cost of your project (your budget) and the nature of its impact on the business, where does this project fit?

Tip: We've defined 'impact' on a spectrum from 'strategic' to 'tactical' to keep this general. But based on your specific circumstances, you may want to define the impact based on different criteria such as time (immediate impact or long term impact).



Part 2

Leveraging intelligent data management

Part 2

Seven essential criteria for intelligent data management

Whether you're standing up a new data warehouse for BI, or implementing a data lake for analytics innovation, the success or failure of your initiative will invariably be tied to the quality of the data everyone's using.

So it is crucial that you choose a set of tools and platform that can enforce the standards for the intelligent data management you set. That is, you need a platform that's capable of delivering on the following seven criteria:

1. Flexibility

So you can deliver the right data to a whole range of different analytics projects, tools, use cases, and users.

Technology change will only increase for the foreseeable future in the analytics space. And the reason for that technology innovation is to support better, faster analytics innovation.

It's essential that you establish the foundations for data management that's capable of meeting diverse analytics needs — for everything from traditional IT-led production analytics to more complex predictive analytics in the cloud.

Rapid change will remain a constant. So it makes sense to optimize for agility when you're setting up intelligent data management.

2. Repeatability

So that you aren't constantly re-inventing the wheel. Hand-coded data integrations and manual data cleansing may solve your problems — once. But they may very well slow you down on future projects.

Accessing, integrating, and cleansing data are three of the biggest data management challenges you'll face.[4] And without reusable logic and standards to support your initiative, your analysts will end up spending 50-80 percent of their time managing the data,[5] and just 20 percent analyzing it.

The easier it is to discover, share, and re-use assets that have already been created, the quicker your teams will be.

Seven essential criteria for intelligent data management

3. Abstraction

So that changes in analytics data persistence technology don't break your enterprise data management process and delivery.

Data persistence technology for analytics is changing rapidly and will continue to do so. It is critical that your enterprise data management tools and platform can insulate you from this change so that your people, tools, and skills are not impacted by the choice of a specific persistence technology to meet a technical need.

Intelligent data management is about ensuring that clean, complete, and timely data is always available for analytics — no matter where the data lives or where it's needed.

4. Alignment

So that both IT and business users can communicate effectively. This starts with a common understanding around business terms, meaning, and context that is shared across all of business and IT.

For instance, in one healthcare organization there were three conflicting meanings of the term "claims paid date."

Some analysts took it to mean the date the claim was approved, some took it to mean the date the check was cut, and others took it to mean the date the check cleared. The resulting misunderstandings were worth millions of dollars.

If the people managing the data aren't on the same page as the people using the data, everyone loses. It becomes a matter of "dueling spreadsheets."

5. IT and business collaboration

So that different stakeholders can work together to define what the data means, what great data looks like, and who can use it.

Once you have determined common business terms and meaning, it becomes important to have tools that specifically enable business-IT collaboration, both in the management of data and the policies around the data. This will avoid errors, but even more importantly, it'll be a key factor in accelerating project delivery.

So it's crucial that you ensure the data doesn't just meet your own standards — but theirs too.

Seven essential criteria for intelligent data management

6. Shared metadata

So that your data management architecture can manage data complexity and provide a basis for increasing automation going forward.

Specifically this refers to:

- **Shared business metadata:** Across business terms, definitions, owners and policies.
- **Shared technical metadata:** To provide data lineage views of how your data is flowing through your environment so that you can understand where the data is coming from, where it's being used, and how it's being transformed along the way.
- **Automation:** Metadata is increasingly important for automating data management processes. It can be used to suggest a next step in the data prep process based on previous behavior of other users. And it can also provide completely automated ways to do things like routing, data cleansing, and data table joins.

7. Data governance

So you can manage data as an asset.

This includes the rules, policies, and best practices that ensure data is entered, stored, and managed appropriately across the enterprise.

It's often hard to ensure that good habits stick. So it's essential that both the tools and platform you choose support your efforts to guide the way data is managed across the enterprise.

Of course, depending on the scope of your initiative, you'll know exactly how much data governance is necessary. Too much and it's unsustainable, too little and it's chaos.

For instance, if it's to enable a decision to invest $5 billion in a new plant, the data had better be perfect.

It's why we're big fans of what we call 'just enough' data governance.

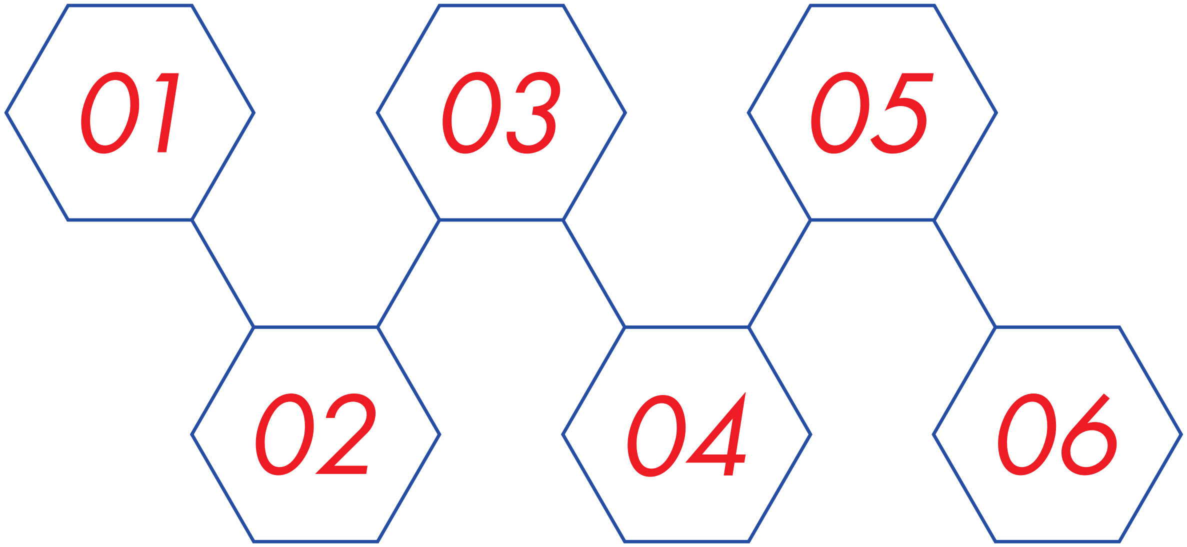
We'll be covering this notion in more detail later. For now, suffice it to say that if your tools and platform don't build governance practices and processes into your architecture, the long-term viability of your initiative is likely to suffer.

Part 2

Six essential capabilities

In practice, intelligent data management comes down to six essential processes. On the surface, they may seem simple.

But as long as your analytics initiatives rely on inefficient, ineffective, or manual ways to do these six things, you'll always struggle.



Six essential capabilities

01

Accessing the data

"The whole manual metric collection process took about three weeks, and by then the data was stale."

— Joery Carty, Manager of Employee Information Delivery Group, FPL Group.[6]

Between application-specific data silos and regulatory requirements, the challenges you face when trying to access data within your enterprise are non-trivial. So if you're relying on the manual collection of data from multiple other systems, departments, and databases, you're going to slow down the whole initiative.

What to do about it

First you need to find what data you have available. Next, you need tools to automate the profiling of data for your analytics projects.

This will help you to understand what the data looks like, what data quality issues are present, and what data relationships exist.

Without this, there's no way the project can become as scalable, flexible, and reliable as you need it to be. Of course, even with automating the access of data there are two variables you need to consider.

1. Real time or batch?

You need to be clear about whether or not you need real-time access to the data. If it's quarterly data from the enterprise's general ledger, it's possible you'll only ever need to analyze it as a batch. On the other hand, if it's transactional data from your ERP, or social media streams, real-time might be essential.

2. Is it regulated?

Another variable that can't be ignored is whether or not the data is sensitive and regulated. If it is, you'll need to control who can view the data and what they're allowed to do with it.

Six essential capabilities

Use this table to list out the data sources you need and what you need to know about accessing each one.

The tools you'll need:

Use these analyst reports to pick the right tools for your initiative:

To profile and parse your data:
Gartner 2014 Magic Quadrant for Data Quality Tools, November 26, 2014

To secure access based on user roles:
Gartner 2014 Magic Quadrant for Data Masking Technology, December 10, 2014[7]

Data source	Where it lives	How you access it today	Do you need it in real time or batch?	Is it regulated?
E.g. Sales data	E.g. Salesforce	E.g. SQL/RDBMS	E.g. Real time	E.g. No

Six essential capabilities

02

Cleaning the data

Every analytics initiative is built on an assumption that the data being analyzed is reliable. But data quality erodes if it is not actively managed. And most efforts to make data clean and accurate are manual and sporadic at best.

Which means the quality of your data is decreasing a lot faster than your ability to fix it.

For instance, customer data changes by 1 percent to 1.5 percent per month. Compounded, that's 27 percent in a single year. So if you're going to wait a year before you assess and improve the quality of your customer data, any analysis of that data during the year can't be trusted.

Data quality must be a continuing and systematic effort.

And the more data sources you rely on, the harder it is to maintain the quality of data your analytics depend on. So rather than relying on manual transformations and corrections of all the data, you need tools to:

- Discover data quality issues
- To correct those data issues
- To monitor data quality on an ongoing basis
- To enable IT-business collaboration

Moreover, this is where effective IT and business collaboration comes in. It's essential that you're able to work with your line of business counterparts to understand what the data's being used for. It's the only way to understand the business impact of having bad data in a critical situation.

Why automated data quality matters

Consider the case of the Queensland Police Service.[8] It found that the quality of its policing was being undermined by data that was inaccurate, duplicated, or missing altogether.

Invalid business names, addresses, and telephone numbers — it found that a single suburban department store had up to 120 variations in its name — were making it incredibly hard to rely on the data.

But if the police service hadn't automated its approach to data quality, the issue would have amounted to the equivalent of 79 years of manual data management and more than $4.4 million in staff salaries.

The tools you'll need:

Use these analyst reports to pick the right tools for your initiative: Gartner 2014 Magic Quadrant for Data Quality Tools, November 26, 2014

Six essential capabilities

03

Integrating the data

The ability to connect different data sources is central to the success and failure of your analytics initiative. But the more data sources you need to integrate, the harder it is to do manually.

That's because the more sources you need to integrate, the more data structures you're trying to reconcile. The thing to consider here is that the most interesting analytics insights being produced today are coming from analyses based on data from widely different data sources.

Hand-coded integrations are one-off 'works of art' that just take up too much precious developer time. They're also filled with the errors that come from doing such complex integrations manually.

Without re-usable logic and documentation, it becomes extremely costly and difficult to change, scale or even maintain hand-coded integrations

It's a waste of your analysts' and scientists' time and hurts the long-term viability and scalability of your initiative.

Why automated data integration matters

What's needed is an integrated approach that enables automation of data management processes.

Let's face it, IT will not get the budget or headcount to keep up with the explosion in data volume and complexity, the explosion in analytics technology, the requirement for business self-service, and actually accelerate the delivery of these projects. The only way to be successful is to look for ways to automate the data management process.

Automation will allow you to leverage intelligent suggestions around highly complex tasks like joining two tables or applying data de-duplication rules.

The impact is non-trivial. For instance, analysts at the US Air Force Knowledge Services needed to integrate data from as many as 26 different source systems and databases.[9] By automating this part of the process, they were able to reduce reporting time from six months to less than an hour.

And more importantly, it freed them up to generate approximately 20,000 reports a week.

The tools you'll need:

Use these analyst reports to pick the right tools for your initiative: Gartner 2015 Magic Quadrant for Data Integration, July 29, 2015

Six essential capabilities

04

Mastering the data

By now it should be clear that the more data and data sources your analytics project depends on, the harder it becomes to manage all that data in a way that speeds up data delivery for traditional BI and new and innovative analytics users.

A key technology that can help is Master Data Management (MDM), which enables you to gather and manage all the data relating to key business entities such as a customer, part number, or partner. You can share that "golden record" across any applications and analytics use cases that may require that data.

For instance, you could create a 'master' profile of your customer 'Emily Browne.' MDM would gather data about Emily (and all your other customers) from all relevant sources, cleanse that data, resolve duplicate data, and deliver a single version of the truth about the entity 'customer.'

Consider MDM for more complex environments with many internal and external data sources where trustworthy data about key business entities will need to be shared across multiple applications and uses.

What to do next

MDM manages all those profiles in the context of 'domains'. These are the business entities you care about like 'products' and 'customers'.

Of course, depending on the focus of your analytics project, you may care about multiple 'domains'.

While 'customers' and 'products' are the two most-common domains our clients have built master profiles for, we've implemented MDM for domains like channel partners, suppliers, menu items, and even wellheads (for an oil company).

Part 2

Six essential capabilities

Go through the list of data sources you created in step 1 (on page 21) and list out the different domains you could use 'master' profiles of.

The tools you'll need:

Use these analyst reports to pick the right tools for your initiative: Gartner 2014 Magic Quadrant for Master Data Management of Customer Data Solutions, October 30, 2014

The Forrester Wave™: Master Data Management (MDM) Solutions, Q1 2014

Six essential capabilities

05

Managing external data

Analysts estimate that large organizations are sourcing more than 50 percent of their data from outside sources. This data will often have little structure or an unknown structure and will come with varying degrees of data quality.

You will need to manage these sources to deliver clean, complete, and timely data that is appropriate for the type of project that the data is being used for.

What to do next

List out the external data sources your analytics initiative depends on:

For instance: Data providers like Experian, IMS Health, Bloomberg. Social data sources like Foursquare, Twitter, and Facebook

Six essential capabilities

06

Governing the data

Data governance is managing data as an asset. This includes rules, standards, policies, and processes so business and IT can collaborate to control the way data is sourced, managed, cleaned, accessed, and secured within your enterprise.

Most importantly, it is about governance strategy, people, processes, and technology, in that order.

'Just enough' data governance

Depending on the scope of your analytics initiative, this might mean an enterprise-wide program that touches every department. But it could just as well mean a handful of people implementing a small number of rules and policies.

The important thing is to manage the data that matters the most.

You may never be able to keep up with all the data. So figure out what is important and what will have the most business impact.

We're big advocates of what we call 'just enough' data governance where the scale of the program you implement reflects precisely what your initiative needs. But we also recognize the value of enterprise-wide data governance. It all depends on your goals.

We've written about Holistic Data Governance in more detail elsewhere, in case you're looking to learn more about it.

Get the whitepaper.

Six essential capabilities

What to do next

List out the other people and departments you'd need to work with to get the information you'd need for your analytics and governance initiative.

Picking the right tools:

Data governance means managing a lot of moving parts — people, processes, and tools. You need tools that give you the ability to enforce your rules, policies, and standards across the enterprise.

Use this analyst report to pick the right tools for your initiative: The Forrester Wave™: Data Governance Tools, Q2 2014

Taking stock

By now, you'll have filled out a list of the internal data sources, the external data sources, the domains and departments your analytics initiative is going to have to rely on to succeed.

Use this information to flesh out your vision for the long-term foundations of your analytics initiatives and engage other stakeholders from IT and the executive team.

Assessing your own data management chops

If you're looking for a way to socialize the amount of change your enterprise needs to undertake to become truly decision ready, we recommend using this Data Management Proficiency Assessment Model from SiriusDecisions.

Now we'll look at what the ideal architecture for your analytics initiative needs to look like, in the context of the six steps we just described.

Part 3

Aiming for the right architecture

Part 3

A general architecture

On the next page, we've illustrated a highly generalized architecture for an enterprise-wide analytics initiative. We've done this for a few reasons.

First, to show you where intelligent data management fits in the context of your data sources, your data storage, and your analytics applications.

And second, so you can start to define:

- Your current state data management architecture
- Your own vision of the appropriate future analytics architecture you're going to be building towards.

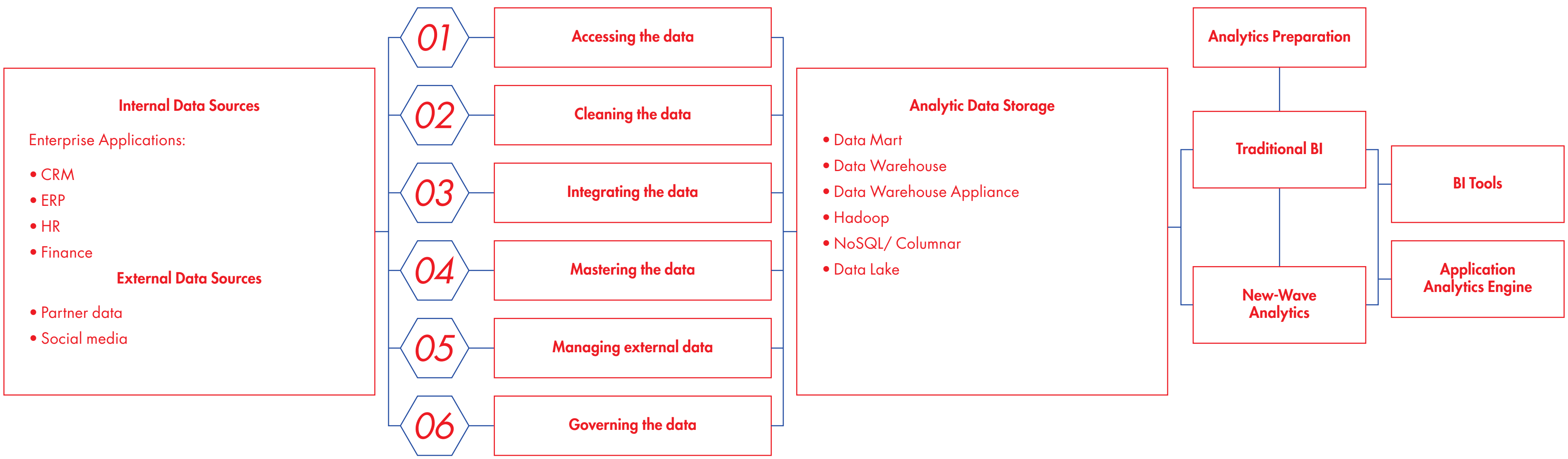
It's important to remember that you will want a single data management architecture that works across your existing BI/data warehouse environment as well as any new or emerging analytics use cases that your organization may be considering.

The important thing is that your architecture should be flexible enough to deal with change while improving overall productivity.

Feel free to use this as a starting point and modify this diagram to match your environment.

Part 3

A general architecture



Part 3

Nine common analytics use cases

Different projects rely on different architectures. So the right architecture for your initiative will be based on the nature of your initiative, the amount of time and money available to you, and of course, your current architecture.

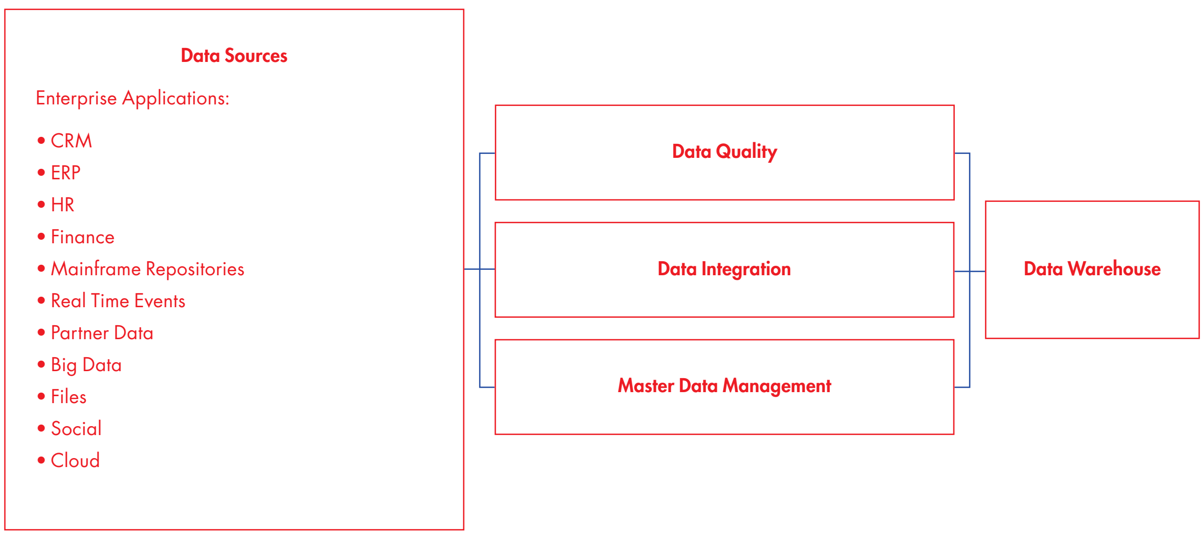
Over the following pages, we'll describe nine common analytics architectures (generalized for this workbook) representing the kinds of real-world use cases our customers approach us with.

Look at each one to see how, even though they're serving a single project use case, they actually lay the foundations for a future analytics architecture, capable of spreading intelligent data management across other projects.

Part 3

Nine common analytics use cases

1. Creating or modernizing your data warehouse



About this architecture

This generalized architecture was built for a company that had multiple data sources it needed to fit into a fairly complex data model.

The idea was to analyze correlations between patient treatments, outcomes and costs, but it only had a short amount of time within which to start delivering quality data to the data warehouse.

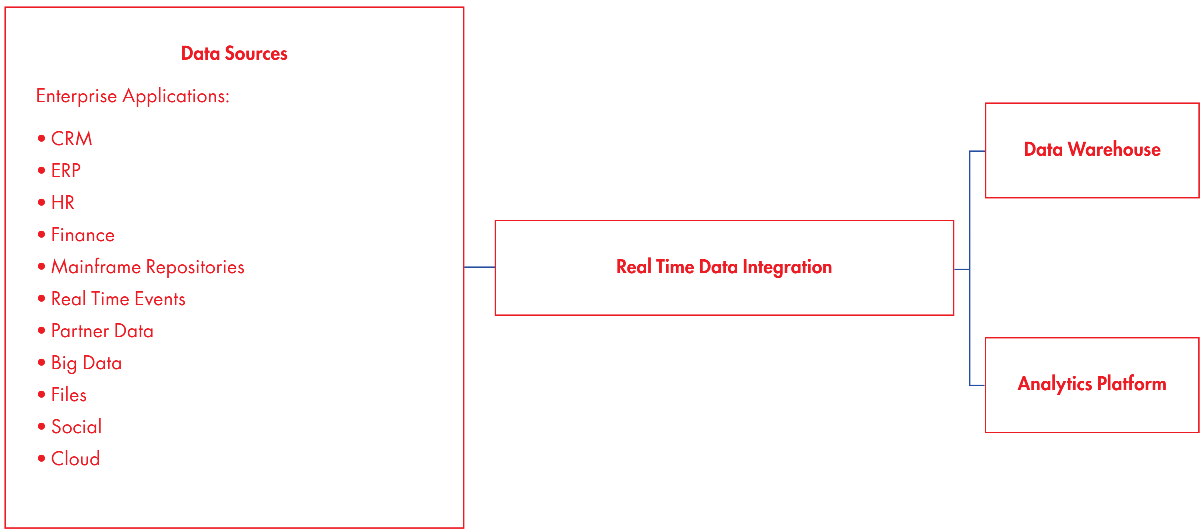
At the core of this use case are tools for data integration and data quality so you can ensure you're loading the right information, of the right quality, and the right format into your data warehouse.

If you have MDM in-house already you save time because you're able to re-use a golden record of data about your key business entities. For more complex projects, MDM is a big help because you only have to master your key entities once, and then share that trusted data.

Part 3

Nine common analytics use cases

2. Upgrading your data warehouse to handle real-time data



About this architecture

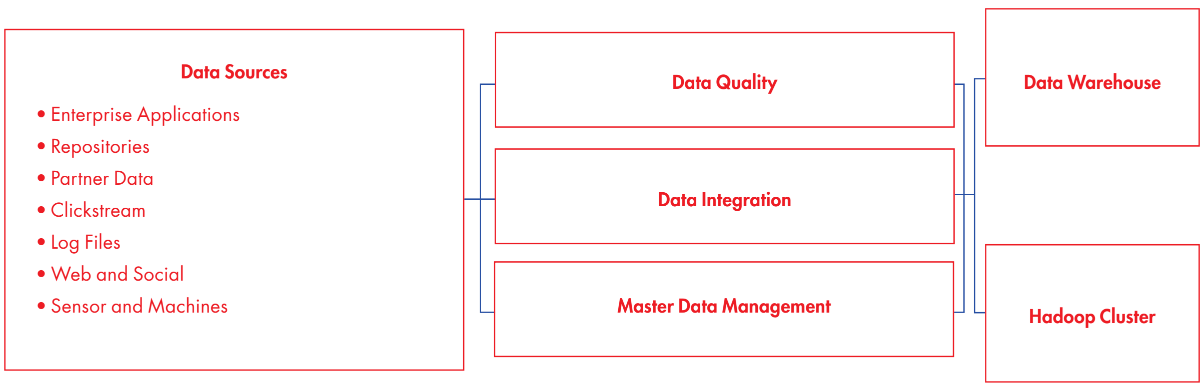
Increasingly, IT is being asked to support more rapid business decision making. This means moving from a paradigm of weekly or monthly reporting to one in which business decisions are based on data with a freshness window of 10-15 minutes.

Solving for this can be as straightforward as upgrading your current data integration tool to support real-time data updates.

Part 3

Nine common analytics use cases

3. Extending your data warehouse with Hadoop



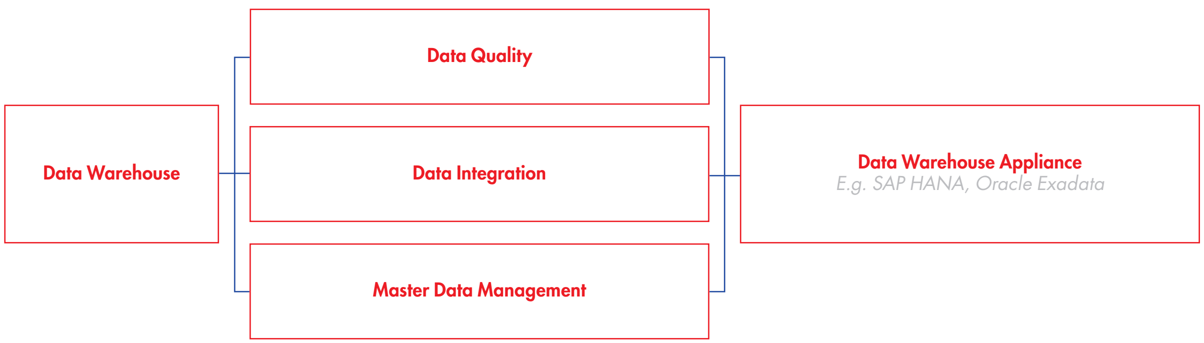
About this architecture

The use case focused primarily on reducing costs. Because constantly expanding your data warehouse is an expensive exercise.

By leveraging Hadoop, you can use commodity hardware to offload your data warehouse and keep only the most business-critical data in the data warehouse itself. At the same time, you can also provide a pool of less-frequently used data for exploration.

Nine common analytics use cases

4. Making the move to a data warehouse appliance



About this architecture

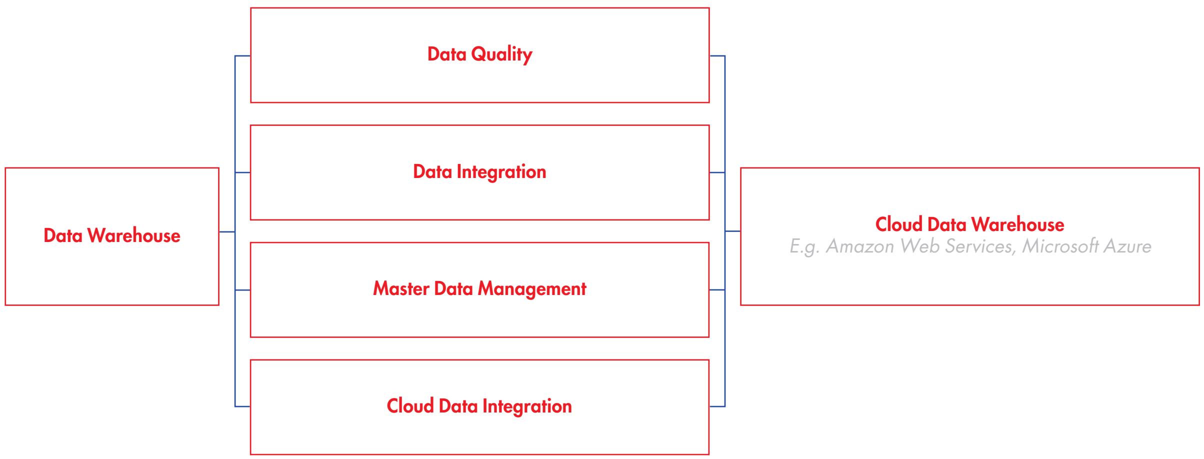
This use case was all about speed. When organizations move to specialized data warehouse appliances like SAP HANA or Oracle Exadata, they're able to dramatically increase the speed of their analytics processing by leveraging specialized hardware and software for high performance.

It means they can generate results a whole lot faster than they would with their existing data warehouse technology.

Part 3

Nine common analytics use cases

5. Leveraging a cloud data warehouse for quick project starts



About this architecture

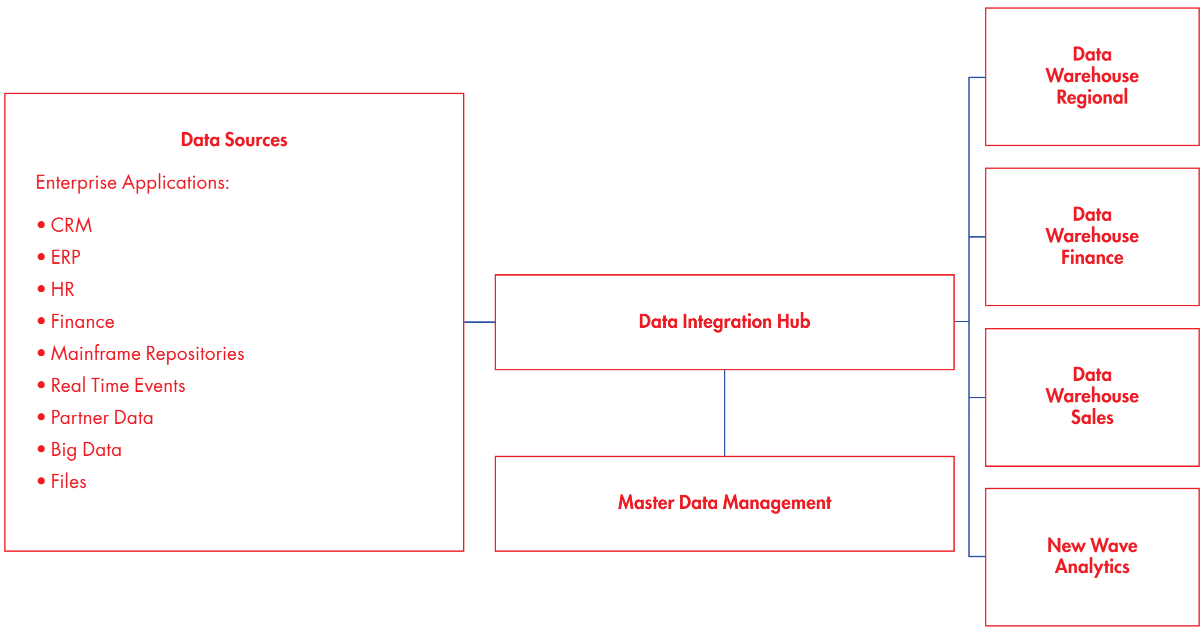
Here, the company found it had the need to start up new analytics projects frequently. For instance, to analyze the impact of a product launch on social media, in real time. Only the launch was in two weeks' time.

That's a big challenge. And in this case, it was clear that the team didn't have time to buy new hardware and software when these requests came in.

In cases like this, it's important to have data integration and data quality tools that can work on-premise, in the cloud, or in a hybrid environment. If available, it always makes sense to leverage MDM for more complex environments.

Nine common analytics use cases

6. Simplifying the environment with a data broker



About this architecture

In some companies, different business analysts produce conflicting reports because they're using different data sources to conduct their analyses.

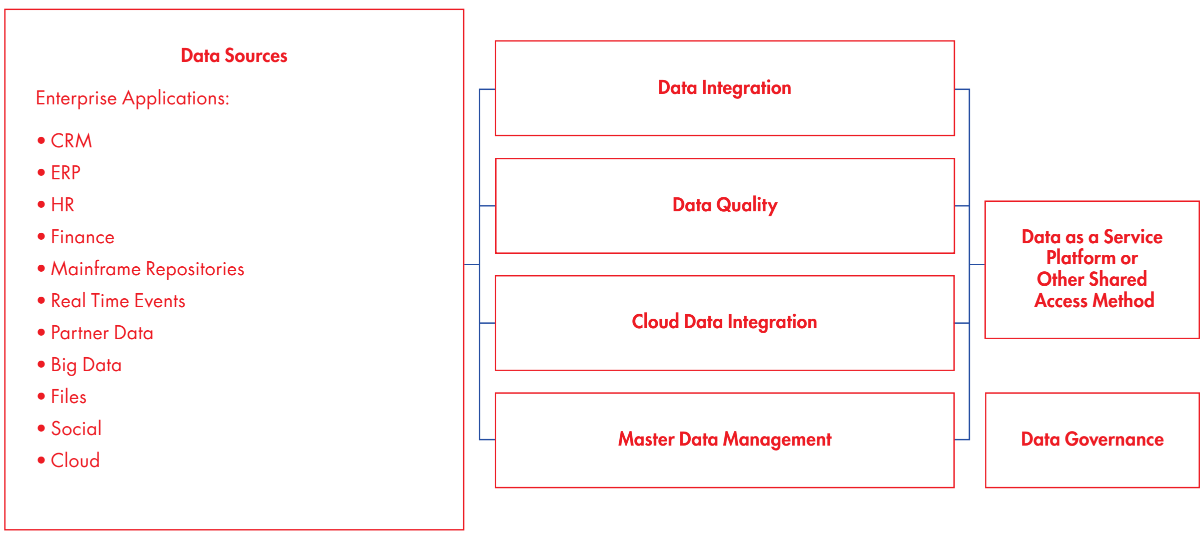
In such cases, it helps to build a central integration hub that can replace the many (often hundreds) of batch processes needed to integrate analytical applications and silos.

That way you can ensure everyone's making decisions based on a single, central source of truth with master profiles everyone can trust to be clean, complete, and up-to-date.

Moreover, the simplified environment creates some crucial benefits for IT. It's easier to understand, manage, and change. And it also costs a lot less to maintain.

Nine common analytics use cases

7. Building a safe place for self-service analytics



About this architecture

This architecture aimed to provide business analysts with a single, shared source of managed, secured, and governed data. This could be a simple common repository, a data lake, or a data virtualization layer for example.

This necessitates the need for the deployment of tools for intelligent data management so that central IT could create a safe place where analysts could discover and gain self-service access to data sources from all over the enterprise.

By using tools like master data management and cloud integration, they were able to ensure they could serve a variety of data domains and analytics tools.

Nine common analytics use cases

8. Big data for innovation

About this architecture

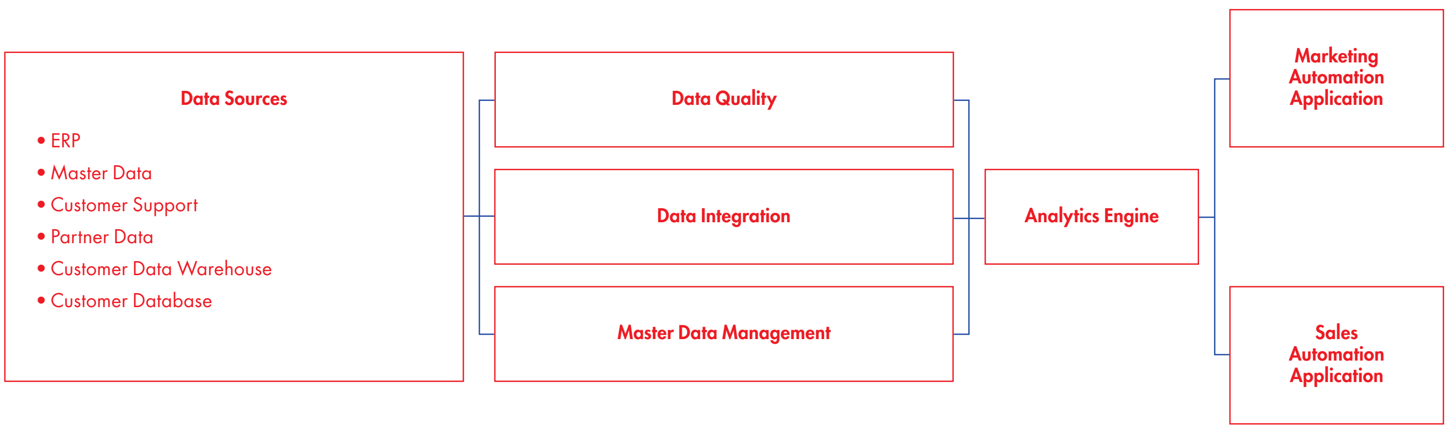
Here, the big data challenges were clear. It was taking too long to onboard, prep, and make the growing volume and variety of data available to business analysts. Building a data lake was essential.

At the same time the team wanted to avoid turning it into a 'data swamp'. So they deployed data management tools to ensure the data being fed into the various target applications and analytics tools was clean, complete, and timely.

The point here is that data could be quickly ingested into the data lake. Later as data was determined to be important, structure could be added, and the data could be cleansed, transformed, and mastered as appropriate for business needs.

Nine common analytics use cases

9. Embedding sales and marketing applications with predictive analytics



About this architecture

In this case, the enterprise already had an analytics engine in place.

This engine provides recommendations as to the next best step to take when marketing to a customer or when a sales representative speaks to a customer.

But the company found that by adding data from the right sources, it could significantly improve the quality of the recommendations being generated.

The aim was to establish the foundational processes of intelligent data management and then feed great data into the recommendations engine so it could fuel the company's sales and marketing applications.

Mapping your current architecture

Your current architecture

Now that you've seen a range of other architectures, map out your current and future architectures. Then you can identify what you're going to need for your analytics initiatives to succeed.

Internal Data Sources

External Data Sources

01 **To access the data:** *E.g. Data profiling, Data masking*

02 **To clean the data:** *E.g. Data quality*

03 **To integrate the data:** *E.g. Data integration*

04 **To master the data:** *E.g. Master data management*

05 **Managing external data:** *E.g. 3rd party data*

06 **To govern the data:** *E.g. Business glossary*

Analytic Data Storage

Analytics Preparation

Traditional BI

New-Wave Analytics

BI Tools

Application Analytics Engine

The substance of next-generation analytics

Next-generation analytics hold a lot of promise. The ability to consolidate institutional knowledge and turn it into something that can be analyzed and leveraged is going to be fundamental to the success of modern enterprises.

What's clear is that it's going to take a whole lot more than a couple of dashboards to make that happen. You will need to leverage both the 'traditional' BI/data warehouse worlds and the next-generation analytics worlds to deliver the best and most productive results.

It's going to take a commitment to clearly defined goals, a robust foundation of intelligent data management, and an architecture that makes analysis scalable, reliable, and flexible. This architecture must work with any data, any analytics use case, any analytics technology, and any analytics tool to deliver the data needed for actionable business insights.

But most importantly, it's going to take the focus and guidance of people like you to ensure your enterprise has everything it needs to become decision ready.

By now, we hope to have armed you with the advice and answers you'll need to engage your colleagues and start down your journey towards next-generation analytics.

Go forth and deliver great, actionable business insights.

Further reading

TDWI Checklist Report: Trusted Information for Analytics.

In a comprehensive and insightful report, data management and business intelligence expert David Loshin looks at the key tasks and activities that bolster an organization's trust in information for analytics.

Get the report.

About Informatica

We're Informatica and we help the world's biggest enterprises with the intelligent data management they need to become decision ready. If you're looking to deliver advanced analytics to your organization, you're going to need to become decision ready too.

We should talk.

Sources

1. Data analytics dominates enterprises' spending plans, Forbes. *http://www.forbes.com/sites/louiscolumbus/2015/03/15/data-analytics-dominates-enterprises-spending-plans-for-2015/4*

2. Why do so many analytics projects fail, Analytics Magazine. *http://www.analytics-magazine.org/july-august-2014/1074-the-data-economy-why-do-so-many-analytics-projects-fail*

3. Vague goals seed big data failure, InformationWeek. *http://www.informationweek.com/software/information-management/vague-goals-seed-big-data-failures/d/d-id/1108384?*

4. Information Week: Analytics and BI Survey 2015, page 27 *http://reports.informationweek.com/abstract/81/12544/Business-Intelligence-and-Information-Management/2015-Analytics---BI-Survey.html*

5. For big data scientists, 'janitor work' is key hurdle to insights, New York Times, August 2014 *http://www.nytimes.com/2014/08/18/technology/for-big-data-scientists-hurdle-to-insights-is-janitor-work.html*

6. FPL Group's customer success story, Informatica *https://www.informatica.com/about-us/customers/customer-success-stories/fpl-group.html*

7. Gartner does not endorse any vendor, product or service depicted in its research publications, and does not advise technology users to select only those vendors with the highest ratings or other designation. Gartner research publications consist of the opinions of Gartner's research organization and should not be construed as statements of fact. Gartner disclaims all warranties, expressed or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.

8. Queensland Police Service's customer success story, Informatica *https://www.informatica.com/about-us/customers/customer-success-stories/queensland-police-service.html*

9. US Air Force Knowledge Services' customer success story, Informatica *https://www.informatica.com/about-us/customers/customer-success-stories/air-force-knowledge-services.html*