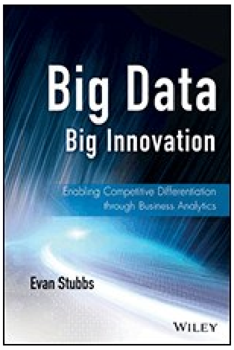

Big Data, Big Innovation: Enabling Competitive Differentiation through Business Analytics

by Evan Stubbs

Chapter 6: Operating Models

Overview

Insight is easy; execution is hard. When it comes to big data and analytics, confusing the two is probably the single biggest reason organizations fail to see the returns they expect. At first, it seems counterintuitive. After all, the point of analyzing information *is* insight, isn't it?

The problem with insight is that in isolation, it's worthless. It's what you do with it that matters, not whether you have it. Markets are hard and, to coin a phrase, never bring a knife to a gunfight.

Everyone has *that* friend, the one who has an answer to everything, the serial entrepreneur, the one who would be rich if only someone would bankroll his great ideas. We love being around these people, but late in the night, after more than a few drinks, they're usually a bit of a bore. It's not that they're wrong; they're just missing the point.

Ideas are cheap. It's doing something with them that's the hard part. Given enough information, there's no end of interesting ideas a reasonably motivated person can come up with—ideas that, if fostered for long enough, sometimes germinate into potential innovations. Insight is the road that never ends; if you're not careful, the journey sometimes becomes more important than the destination.

There's always one more fact to find, one more way of slicing the data, one more information source, one more report, one more mashup. It's addictive. Discovery can be a dangerous siren; more than one explorer has become wrecked upon her shores.

The smartest group in the world is just another cost if they don't add value. And, while insight or efficiency is still a form of value, it's not the best type of value. The best teams excel in producing both internal *and* external value, as shown in Figure 6.1.

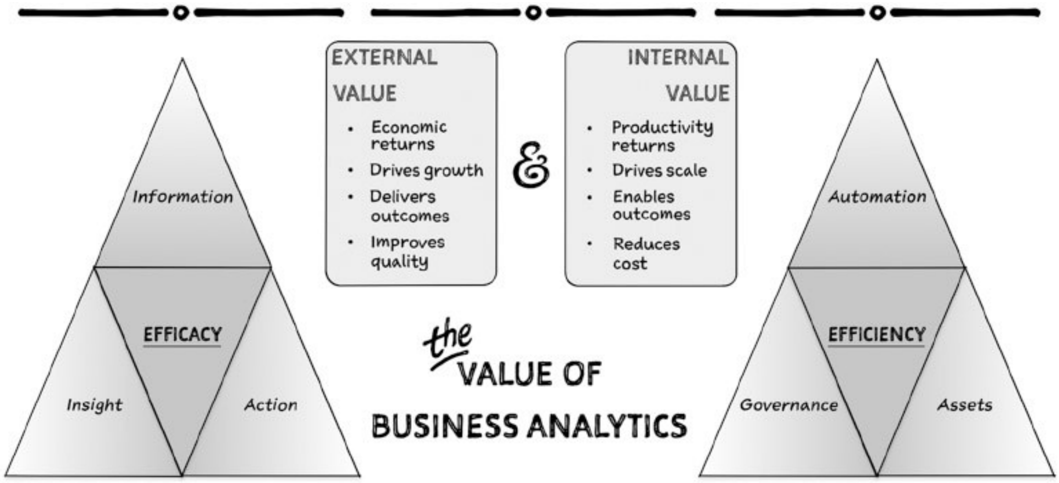

Figure 6.1: The Value of Business Analytics

The true goal is *external value*. These are the outcomes that everyone *outside* the analytics group is happy to recognize. Usually tangible, they're normally closely linked to growth, improvement, or efficacy in some form. Common measures include revenue, profitability, or cost reductions. They're normally created through using analytics or big data to identify and deliver quality improvements in some form. It's created through sourcing information, generating insight, and acting on that insight to realize value.

On the way, those same teams need to create *internal value*. These are outcomes that people *inside* the analytics group see as valuable. Importantly, not everyone outside the group will always recognize these outcomes as being important. Sometimes intangible, they're normally closely linked to new capabilities, productivity, or efficiency. It's created through improving automation, managing intellectual property as assets, and ensuring governance processes are aligned against requirements.

While internal value *does* normally reduce structural cost in some form, the linkage between the outcome and the cost reduction is often unclear. For example, the use of prebuilt modeling processes might help teams do more work.

Unfortunately, while productivity might help the organization avoid hiring people at some indeterminate point in the future, it's often hard to translate the opportunity cost savings into something an accountant will recognize. It's not impossible, but when it does happen it becomes external value.

The real value of business analytics comes through balancing these two forms of value. External value provides the return from big data. However, it's impossible without creating internal value on the way.[1] Doing this is impossible without having an operating model that aligns investment to outcomes, balances risk against reward, and gives each activity a "home."

This chapter tends toward the technical; it's intended for people who want to have a framework to map responsibilities between different groups in an organization. It covers three things:

1. What's the goal?
2. What's the enabler?
3. How does it create value?

The rest of this chapter will answer these questions and lay the foundation for effective delivery.

Readers interested in "getting to the meat" and reviewing the operating model are free to skip ahead to the section titled, "What Does It Look Like?"

[1]For a far more comprehensive review and explanation of these concepts along with how to quantify them, see Chapters 4 and 6 of Evan Stubbs, *Delivering Business Analytics: Practical Guidelines for Best Practice* (Hoboken, NJ: John Wiley & Sons, 2013) and Evan Stubbs, *The Value of Business Analytics: Identifying the Path to Profitability* (Hoboken, NJ: John Wiley & Sons, 2011).

What's The Goal?

Business analytics teams exist to create value. Like the alchemists of old, they are the modern-day magicians that are tasked with transforming data into value. Unlike the alchemists, though, their task is doable. Rather than transmuting lead into gold, it simply requires the ability to uncover advantageous patterns and act on them.

This is the core of *external value*. To create it, the team needs to get the right information, generate the right insights, and help the organization act on that insight to drive tangible returns.

Practical examples include:

- Using changes in calling patterns to identify people at risk of canceling their telephone services
- Mapping social-relationship information to identify and quarantine people at risk from known pedophiles
- Analyzing retail purchasing information to identify products that tend to sell well together and by doing so, change stocking patterns to maximize cross-selling products
- Modeling price sensitivity patterns to identify the ideal price to offer every customer to maximize margin while maintaining sales conversion rates

Despite how obvious this may be, surprisingly few organizations are any good at using big data to create external value. Rather than cultivating a forest, they focus on the trees and believe (falsely) that the forest is a natural outcome. Instead, to stretch the analogy, they usually end up with a series of disconnected hedges.

At a minimum, external value involves three activities:

1. Ensuring quality information
2. Generating knowledge through insight
3. Realizing value through action

Quality Information

Managing analytical information is mainly concerned with transforming source data into forms that are fit for other uses. There are four major activities that occur in this space. Of these, most organizations are only good at one. Developing an understanding of the other three activities is a key step in driving true economies of scale.

These four activities are:

1. Operational data preparation and delivery
2. Operational data quality
3. Analytical data preparation and delivery
4. Analytical data quality

The operational side of information management is usually well understood. Running a business requires many systems. Some provide transactional support—common examples include order management, case management, or customer relationship management. They provide the operational support that an organization needs to run its day-to-day operations. There are also normally a variety of systems that facilitate functional, business, and organizational planning.

While these use the information contained in the transactional systems, they require the information to be aggregated and transformed; knowing that a small can of beans was sold last Tuesday at 2:15 P.M. in store 31 is less useful in planning than knowing that over the last three months, total sales of beans in a particular geography has been increasing by 2 percent compound. Getting from one view to the other involves having a warehouse designer aggregate transactional sales by category, geography, and time period.

Sitting between all these systems is usually a warehouse that attempts to centralize all the organization's information in one location. Operational data preparation and delivery involves pulling all this information together and delivering it in the right form to the right system in the right order to make sure everything gets what it needs at the right time. This can be surprisingly complex, especially when one considers that different systems update at different times and, if the updates are not cascaded through the right systems in the right order, data can quickly get out of date.

Data modelers do this using a variety of extract, transform, and load (ETL) or extract, load, and transform (ELT) jobs, so named because they describe the major activities that need to occur. These are usually strongly governed and relatively inflexible—once defined, they will usually remain as-is until their source or destination data structures change. Every change carries cost; in practice, this happens as infrequently as possible.

Even unsophisticated organizations are usually still competent at operational data preparation and delivery, largely by necessity. Without the ability to manage data, it is usually extremely hard for decision makers to get any visibility over how the business is performing. There is an important caveat that goes along with this, however: simply getting the data into the right form has little relationship to whether the data is trustworthy or accurate. Over time, the organization starts to realize that despite having lots of data, most of it is relatively untrustworthy. This may be because of duplicate customer records (often because people use different addresses or change names) or it might be because front-of-house staff take shortcuts when entering information to speed up order processing (using all zeros is a common way of avoiding entering codes).

As organizations mature, they increasingly understand the importance of operational data quality and have usually established parallel processes to ensure the information used by the organization is correct. Common focus areas include data profiling and data cleansing. Again, these activities are ideally transformed into a variety of assets that have the potential to be deployed operationally.

This is a critical part of ensuring continuous data quality—when cleansing is treated as a one-off activity, information quality resumes its gradual decay over time once cleansing is finished. By operationally deploying these assets into ETL or ELT jobs, organizations can ensure that information is always correct and cleansed before it hits the warehouse or other destination systems. Organizations that forget this critical step and assume that cleansing is a one-off activity usually find that their information sources regress back to their original state.

At this point, organizations have a good grasp on operational data management as well as a set of high-quality and trustworthy information. However, there are still two other activities that, while similar, require a slightly different approach. Analytical data preparation and delivery shares many core requirements with its operational counterpart but extends these to include the need for a variety of statistical and temporal transformations.

A common example is the creation of "RFM" data that, for each customer, describes their most *recent* transactions (on a rolling basis), the *frequency* with which they transact over a certain time period, and a variety of measures describing their *monetary* spend (including their mean expenditure, maximum expenditure, and so on). This represents a fairly simple example—because the resulting tables are designed to be fed into a variety of models for training or scoring purposes, these additional fields can end up being highly complex mathematical derivations.

Analytical data quality is similar in the sense that it represents a superset of the requirements behind operational data quality. In addition to the need for profiling, cleansing, and matching, analytical data quality is also concerned with statistical characteristics such as completeness and importance. Because missing values can severely restrict one's choice of algorithms, increasing the "completeness" of data (even when it doesn't exist) is a major driver behind analytical data quality. Imputation is focused on generating replacement values without statistically biasing the original dataset or losing the importance of clearly distinguishing between "real" data and imputed data. Not all data is necessarily important or relevant when it comes to developing models. Identifying outliers and isolating the truly "important" information is another major source of analytical data quality.

Much like analytical data preparation and delivery, analytical data quality is often treated as a separate activity to operational data quality. While it may leverage a common technology platform, analytical data quality typically requires a higher level of statistical and mathematical knowledge in comparison to operational data quality.

The Knowledge of Insight

Given a rich source of data, generating insight is where most organizations start. Unfortunately, it's also where they tend to finish. This activity is focused on finding answers to questions or generally looking for interesting insights. Experience plays a massive role in this; knowing what to look for or how to apply the right techniques is critical. Because of this, it's usually highly iterative and weakly defined.

Generating insight from big data requires four activities: exploratory analysis, exploratory data preparation, insight generation, and asset development.

Exploratory analysis usually starts without a defined endpoint in mind—the main objective is discovery. It can range from being as simple as browsing through data to get a feel for it, to using cross-tables and correlation plots to look for interesting relationships. Usually, the analysts doing the exploration have little idea what they're looking for upfront. All they have is some data, a lot of curiosity, and possibly some hypotheses. Unsurprisingly, this is an area where data scientists add tremendous value.

Exploratory data preparation usually involves extracting and structuring data to support model development or report creation when the used cases are ill-defined or unknown. It is a highly iterative process that is repeated until the end-state can be defined. A good example involves trying to find the right data structures to help a particular business unit make better decisions. They might not know what they need. However, they'll almost always know it when they see it.

A common pattern might involve extracting a set of data, deriving a series of measures such as the average sales over a particular time period, and then socializing the results with them and recutting the data as necessary. Another common example involves developing the input tables needed to develop a model. While the analyst may have some assumptions or beliefs as to what behavioral characteristics drive particular outcomes, it's not until they can test those assumptions with a statistical model that they can validate or disprove them. And so, they will repeatedly create and test these tables with new derived fields until they finalize their model.

On finding what they're looking for, analysts will move on to developing models or reports. The tables created during exploratory data preparation are used as inputs to generate insights and answer questions. The major difference between this and exploratory data analysis is that during this activity, the analysts have a defined objective. They may be trying to identify the major reasons behind customer churn or they may be trying to identify the levers that have the greatest impact on getting someone back to work after a major occupational injury.

Finally, these insights are ideally transformed into assets in their own right. Unsophisticated organizations miss this step entirely. Instead, the analysts give these insights to decision makers as indirect sources of information. Rather than build a recommendations process that tracks action, they'll just pick up the phone and give the answer or send through a spreadsheet. This creates two problems.

First, while the team can ensure that the information is delivered to the right decision makers, they have no way of ensuring that the information was actually used. With no tracking mechanism in place, they've no way of knowing the value of the information.

Second, the team is limited by their ability to manually communicate their findings. Every time they generate new insights, they need to spend more time making sure the right people get the right information. This heavily limits their ability to capitalize on economies of scale and reduces business analytics into an interesting, if somewhat erratic source of minor value for the organization.

Transforming insights into assets involves taking insights and turning them into objects that can be accessed by other people or systems. Most people are familiar with the idea of automated reporting. However, fewer people are aware that more advanced forms of analytics such as predictive modeling or optimization can use the same approach.

In this situation, the models themselves can be turned into a series of formulas that the organization can deploy into operational processes. However, doing so requires analysts to convert their personal skills into automated processes, often facilitated by purpose-built software. Getting to this point requires both automation and supporting systems that allow the use of analytics within operational processes.

The Need for Action

Realizing any real value requires one more step: taking action on insight to drive a specific outcome. There are two major activities that go along with this: presentation and decisioning. Presentation is primarily concerned with getting relevant and concise information in front of decision makers while decisioning is primarily focused on automating operational and microdecisions across the enterprise.

Most organizations are fairly mature when it comes to establishing and managing structured presentation (or business intelligence) technologies. However, it's also true that many organizations could benefit from better visualization practices—creating a report and creating an *effective*, *relevant*, and *concise* report are not necessarily the same thing. Regardless, presentation on its own has one fundamental flaw—it is impossible to directly link the insight to the outcome. The information contained within the report or dashboard may well have influenced the behaviors of the decision maker, leading to a better outcome. Or, it may not have—it's possible that on that particular occasion, they decided to follow the advice of their hairdresser!

While it's not always the case, presentation systems usually rely on the assumption that it's up to the decision maker to synthesize all the information made available to them and from that, make an independent decision. This decision happens independently of the presentation system, breaking the link between insight and outcome and making it extremely hard, if not impossible, to quantify the value of business analytics in a true accounting sense. Instead, the organization needs to quantify value by making broad assumptions about how information is used.

This dilemma also highlights one of the reasons why it is often so hard to objectively quantify the value of business intelligence. Few will disagree that having immediate access to better information drives better decision making. However, it's also true that most struggle to explain how much decisions have improved once the organization has access to better information.

Decisioning systems strengthen the link between insight and outcome by moving away from insight and toward recommendation. They blend a variety of rules and models to either provide specific recommendations to decision makers or automatically make decisions on behalf of the organization. These decisions drive specific outcomes such as flagging potentially fraudulent transaction, identifying whom to contact to drive down churn, or recommending the most appropriate program to an individual in need of social services. If these decisions are acted on, the organization can quantitatively determine the value of the action. If they are ignored or overridden, the organization can track the value creation or destruction of the alternative decision. By doing so, the organization directly links insight to outcomes and gains the ability to quantify the value of business analytics.

It's important to note that closing the value chain does not necessarily require measuring outcomes in a comprehensive way—it simply requires actioning insight. This seems like a small distinction, but it's an important one. Many organizations successfully operationalize their insights and drive real value without operationalizing the corresponding value measurement processes. Instead, they manually determine the benefits they've derived as a one-off activity. Establishing a comprehensive and automated value measurement framework is one of the factors that distinguish organizations at level two from those at level three of the cultural imperative.

What's The Enabler?

Big data gives organizations countless opportunities to create value. Unfortunately, there's only so many hours in a day. Without becoming more efficient, there's just not enough time to solve every problem and realize every opportunity.

Productivity enables external value. Unfortunately, productivity alone does little to the bottom line unless the organization uses that efficiency to reduce structural cost through downsizing or otherwise reducing operating costs. It does, however, enable organizations to do more with less. This is the main benefit of internal value; it's the organization's ability to scale through being efficient and responsive.

Internal value comes from:

- Improving the efficiency through automation
- Reusing analytical assets
- Understanding the need for governance

The Efficiency of Automation

Automation is possibly the single biggest enabler for productivity. This happens in two ways:

1. Automating information management assets
2. Automating analytical assets

The cornerstone to this approach is moving away from manual activities. Copying and pasting data within Excel is a common example of a manual activity. Every time new data arrives, the analyst needs to spend time massaging the data into the right shape. It can't be automated, it's highly inefficient, and it carries extremely high support costs.

By comparison, code-based approaches (such as using SQL) create an asset, albeit one that still carries a fairly high support cost. This asset can be embedded in other systems and executed without any manual interaction. They do allow automation. However, their efficiency depends on the skills of the person who created them.

Purpose-built tools take this to another level. They usually offer the best solution, albeit at the highest entry cost. They're often built to expressly facilitate automation. They are tailor-built with efficiency in mind and usually reduce support costs by providing prebuilt migration and asset management functions.

Analytical assets are no different. Exploratory data analysis tools can also usually be used to build models. These models are accurate but need to be used interactively. Some tools offer some degree of scripting or coding. These help transform the model into an asset, but they also increase support costs and, unless skills are common in the organization, link the asset to the person who created them. They're also not always compatible with existing IT assets, forcing redesign work.

More sophisticated organizations create dedicated operational analytics and decision orchestration platforms. These carry the highest upfront costs but also reduce support costs, increase efficiency, and enable systems-level integration and automation.[2]

Most organizations are well aware of the benefits of automation. Common examples include operational data management, reporting, and sometimes operational data quality. Warehousing and business intelligence are mature disciplines. Because of this, the benefits of automating data management and reporting processes are well understood.

Unfortunately, the same can't be said for many of the assets that link into decisioning systems. Efficiency comes from establishing the frameworks, processes, and architectures to support automated scoring and decisioning. In practice, this may take the form of the following:

- Scheduled scoring processes that automatically take recent behavioral information and generate customer-level probabilities that indicate propensity to churn.
- Automatically monitoring transactions in real time to identify potentially fraudulent activity based on a series of rules and propensities and, if flagged, automatically actioning the transaction with the fraud team and putting a hold on the credit card.
- On becoming an outpatient after an emergency ward presentation, dynamically assessing medical history and prescribed medications to identify whether entry to a consultative care program would reduce the odds of a future representation at the emergency ward and, if so, assigning a case worker and scheduling the first visit.

In each of these cases, a variety of analytical assets are deployed operationally to automatically make business decisions based on the most recent known and predicted behaviors. This approach not only links insight to outcomes but also drives significant economies of scale. Rather than investigating accounts based on a random sample, the organization can assess every single transaction individually.

Arguably more than anything else, it's automation that transforms business analytics from something that augments existing processes into an enabler for competitive differentiator in its own right. Without it, scale is impossible.

Reusing Assets

Once these assets have been created, they need to be managed. This usually happens in two ways:

1. Managing information assets
2. Managing analytical assets

It's important to remember that this activity spans the full breadth of analytical assets produced by the organization. Common examples include reports, models, information management processes, datamarts, and all their associated documentation. While there is no reason that these couldn't (or shouldn't) be combined into one asset repository, current skills, practices, and technologies tend to make this harder than it needs to be.

Efficiency comes from:

- Tracking these assets in a centralized, metadata-driven system
- Understanding how well they're performing
- Understanding how much value they're creating
- Understanding how frequently they're used
- Managing them through their full lifecycle

Creating any type of asset takes time and effort. It incurs real costs. Building a house involves a wide variety of specialist skills, including builders, electricians, and plumbers. It also requires significant capital investment for bricks, electrical equipment, and concrete. On top of that, it takes time.

Information assets are no different. They require specialist warehousing, modeling, and business skills. They require capital investment for the right information management, model development, and decisioning tools. And, they also take time to develop, time that carries an opportunity cost.

Similar to how we monitor and manage our return-generating real-world assets, we also need to monitor and manage our information assets. Real-world assets depreciate. Once built, the house inevitably suffers normal wear-and-tear that eventually requires maintenance costs. Some of the structural characteristics may not be appropriate for changing environmental conditions. A clay bed may contract during a period of dry weather, putting cracks into the walls and requiring a minor rebuild. Keeping on top of these and ensuring the asset is performing well is an obvious good practice.

In an ideal world, information assets carry none of this depreciation. As long as things remain constant, the asset will continue to perform exactly as designed. We don't, however, live in an ideal world. Customers change over time, rendering the assumptions that underpin the asset inaccurate. Business models change, rendering the rules that drive decisions irrelevant. And, people sometimes deliberately change their behaviors to avoid detection. For example, criminals change their patterns as soon as they know what intelligence agencies are looking for. Staying on top of this asset depreciation is a key component of achieving best practices in asset management.

While automation is a key part of being able to monitor assets efficiently, it's more than that. Efficiency and productivity comes from *not* reviewing assets. When the measures are set up correctly, the organization need look only at assets that are underperforming or underused. By grounding the philosophy in value rather than effort, the organization can reduce maintenance costs and increase time available for better value-creating activities.

The Need for Governance

Coordinating all the groups involved in business analytics can be extremely challenging. The greater the overhead needed to coordinate activities, the harder it is for organizations to leverage business analytics as a real competitive rather than just a useful source of insight.

Productivity and efficiency comes from:

- Defining an engagement model that identifies the handover points between individuals
- Establishing standardized development and deployment processes

Getting this right helps drive process efficiencies, ensure quality control, and simplify the application of competencies across new problems. It also ensures that governance is tailored appropriately. Too much, and innovation suffers; too little, and operational risks increase.

It's a key part of an effective operating model. The major focus in this area is on managing workflow and facilitating collaboration and, just like integration and asset management, this occurs at two levels:

1. Coordinating the development of analytical assets
2. Coordinating the deployment of analytical assets

As with asset management, there is no reason why both of these activities can't take advantage of a common technology platform and management approach. It's important to remember, however, that while the vast majority of their requirements overlap, the level of emphasis placed on specific requirements varies between the two.

Achieving best practice within this process requires at a minimum:

- Establishing a clear operating model that outlines roles, responsibilities, and handover points
- Documenting and following standardized processes
- Having well-defined points of ownership with the power to make decisions
- Ensuring a high degree of quality through explicit quality control activities

Unclear processes almost always create highly variable outcomes and process inefficiencies. It's hard for an organization to drive continuous improvement when everyone follows a different process. Some people will naturally do things more efficiently than others. Unfortunately, when everyone does things differently, it's almost impossible to replicate those efficiencies. Having standard processes not only increases agility through making sure everyone has clarity on how to execute but also ensures that everyone benefits when team members find new efficiencies.

This does not necessarily mean that activities need to be defined to the lowest level of detail possible—a certain degree of pragmatism and realism needs to be applied when working out an appropriate level of granularity. It's also true that too much rigidity stifles innovation; when people aren't given the freedom to experiment, improvements tend to be the first thing that suffers.

Underpinning these processes are roles and responsibilities. To be effective, everyone must be crystal clear on what they are responsible for delivering as well as when they need to get involved. This helps provide certainty as well as reduce transaction costs. By linking roles to activities, the workflow system itself can automatically notify stakeholders when their interaction is required. If a champion model has been submitted by an analyst, the next logical step would be for the information management team to deploy that model into production and validate the predictions against the original model. Ideally, the system itself handles all the necessary notifications based on a combination of asset registration or workflow and process management.

Finally, effective governance requires a high degree of quality control. When it comes to dealing with operational systems, repeatability and transparency are critical. Every process must be exhaustively tested prior to moving it into production lest it fail and cost the organization real money. Minimizing risk involves ensuring that standard tests are applied, making sure that appropriate signoffs are followed, and ensuring that outputs and predictions are the same in production as in development. While the checks will vary between information assets and analytical assets, the need for a high degree of quality control is a constant.

[2]For more detail on the use of operational analytics and decision orchestration, see Chapters 6 and 9 of Evan Stubbs, *Delivering Business Analytics: Practical Guidelines for Best Practice* (Hoboken, NJ: John Wiley & Sons, 2013).

How Does It Create Value?

The fastest way to discourage innovation is to make it hard. Everyone has a day job; if you're trying to get people to go above and beyond the call of duty, you need to make it easy. As covered in the last two sections, there are many moving parts in creating value. And, each of these parts involves multiple parties. Because of this, simply setting up a group isn't enough; it needs to be accessible, responsive, and valuable. If it isn't, it'll fail.

The best answer is one that blends flexibility with rigor, aligned against an operating model that gradually transitions from exploration into execution. Rather than building the approach around the funding model, the funding model should be dictated by the operating model.

The Wheel of Value

In moving from insight to execution to improvement, every organization needs to follow the *wheel of value* and go through six key stages, as shown in Figure 6.2.

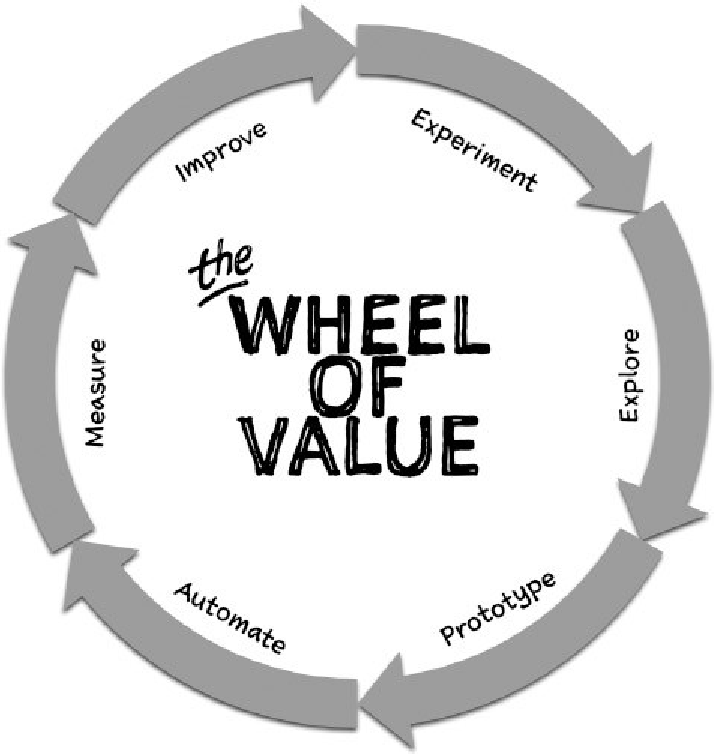

Figure 6.2: The Wheel of Value

Value only ever comes from the ability to execute. In cases where this involves coordinating multiple parties, this is only possible when there's a clear operating model. A well-defined operating model ensures that:

- Processes are aligned to support agility through prototyping, with "process hardening" happening only once solutions have been validated.
- Insight is acted on, thereby allowing the potential for better outcomes and impact.
- Measurement supports self-awareness, improving focus in the right areas and allowing for pragmatic effort and investment prioritization.
- Institutionalized learning processes enable and support growth and continuous improvement.

One of the biggest advantages of big data lies in its ability to expose "unknown unknowns." By mashing up novel combinations of information, data scientists can discover insights that the organization may have never even considered. Experimentation usually takes place in the absence of a defined business problem.

Once a business problem has been defined, the organization moves on to exploration. The business faces a challenge that requires some form of analysis. Again, this is deeply within the realm of the data scientists. Through blending qualitative and quantitative evidence, they seek to validate or disprove some hypothesis. It might be as simple as testing whether some customers prefer email over phone contact. It might be as complex as identifying the root cause for revenue leakage within a highly complex supply chain and manufacturing process.

This, along with experimentation, is the core of analytics as most think of it; it's complex, it's scientific, and it's usually highly numerical. It's also highly uncertain; data scientists rarely know the answer before they start. At best, they'll have the experience to know what will likely get them to the right answer. In practice though, it's usually a voyage of discovery, one where novel insights frequently abound.

Because of this, it's characterized by weakly defined processes. Success usually comes down to the creativity and capability of the individual. While some control measures can and should be put in place, at best they're usually guidelines and milestones. While everyone should still be working from common technologies and data sources, there's still a strong need for flexibility. The fastest way to inhibit outcomes in this stage is to mandate heavyweight and standardized processes. Kill creativity and you kill the Golden Goose.

Eventually, this creative process generates an answer. It may not always be the answer people were expecting, but it's an answer, nonetheless. What might have started out as a fraud investigation might eventually turn out to be a sales opportunity. Knowing the answer is half the battle; to make the answer worth something, it needs to be acted on. The best approach to doing so is to integrate the analytics into operational processes.

For example, unhappy customers rarely enjoy being sold to. By incorporating customer sentiment into the recommendations it makes, the organization can better decide whether to focus on sales or service by customer. Rather than sell to an unhappy customer, it might be better to tell them ways that they can better use their existing services. And, rather than tell happy customers about the benefits of the services they've subscribed to, it might be better to take the opportunity to offer other services that they'll be even happier with.

This represents a change in delivery. Insights usually come from a creative process, one involving weakly defined processes. To automate these processes, they need to be strongly defined. Without a series of steps that involve clearly defined inputs and outputs, it's impossible to turn these manual processes into automated processes.

Unfortunately, the people with the skills to create these insights are often not the people who are best placed to create these automated processes. This doesn't represent a lack of vision of understanding; it's simply the reality of an increasingly fragmented skills base created through hyper-specialization. Building the skills required by a high-performing data scientist can take a decade or more. Building the skills required by a high-performing enterprise architect can equally take a decade or more. Rather than setting the unrealistic goal of hiring someone with perfect skills, it's usually easier to split functions between prototyping and automation, thereby increasing the size of the available labor pool.

Prototyping involves building an asset that is characterized by:

- *Algorithms and logic* rather than guidelines and weakly defined processes
- A high degree of *encapsulation*, in that the asset is portable and can be handed to other people or systems for controlled use without breaking functionality or outputs
- A high degree of *abstraction*, in that the asset takes a known and finite set of inputs and delivers a known and finite set of outputs without the user needing to understand the internal complexities of the asset

Exploration and prototyping require agility and flexibility. They're focused primarily on user acceptance. Requirements are rarely known up front and delivering to business requirements is a highly iterative process. Because of this, while these prototypes reflect an accurate representation of the logic needed to deliver the outcome, they are rarely:

- Scalable
- Robust
- Ready to be integrated with operational systems

Automating these assets typically involves going through strongly defined development, test, and production processes that progressively:

- Optimize their underlying logic to achieve higher levels of performance
- Validate their results against expected results
- Integrate their logic into operational systems while ensuring business continuity and overall systems stability

Once automated, these assets provide regular recommendations to management, operations, and frontline staff through expected delivery channels. Automation is frequently the domain of IT and tends to focus more on unit and integration testing. The goal at this point is not to create something new. It's to take what's already been created and make it bulletproof.

Closing the loop involves ensuring that the impact of these recommendations is understood and that actions (or inactions) are adjusted to support continuous improvement. This involves ensuring that activities are measured and evaluated and, based on results, adjusted to support continuous improvement. These carry through the final stages: measurement and improvement. All parties have a role to play in these stages, focusing on their personal areas of expertise.

By mapping roles, responsibilities, and funding models to the six stages of the "wheel of value," organizations make it clear how insights will eventually move into production. They make it easy for other business units to get engaged. And, they provide clarity on where the handover points exist and what's expected at each point. A major point of competitive differentiation comes from reducing the time it takes to move through this entire cycle. The faster the wheel, the greater an organization's ability to out-think, out-act, and out-learn its competitors.

When this operating model is broken, the business inevitably experiences four pains. First, insight without action destroys value. Having too much insight without the ability to act on it creates confusion and introduces delays through "analysis paralysis." Typically, it eventually leads to organizations rejecting the use of their information assets. When it becomes too hard to leverage quantitative insight in any meaningful way, people will revert to gut-feel and subjective opinions.

Second, action without insight is guesswork. Insight can stem from qualitative or quantitative sources and can be intuitively or analytically based. Critically though, actions that are not based on clear linkage to supporting evidence are no better than guesswork and, more often than not, lead to suboptimal outcomes.

Third, outcomes without measurement prevent improvement. When the effectiveness of actions in driving outcomes or impact is not measured, organizations have no way of knowing what is or is not working. This actively inhibits improvement.

Finally, measurement without learning creates stagnation. Measures are worthless unless they are actively used to drive better behaviors. It may be that particular services are known to have minimal impact on getting long-term beneficiaries off the welfare system. This knowledge does little unless it is put into practice and operational staff are discouraged from offering those services.

Ensuring Sustainability

To be sustainable, every group needs to cover its costs. Unfortunately, this sometimes acts as a major barrier to the organization getting engaged. It'd be nice to offer services for free. Unfortunately, everything costs money and data scientists aren't cheap. Changes to operational systems or data warehouses are even more prohibitive. It's not unheard of for a single data feed to cost over a million dollars in internal resourcing costs.

Every group is established as either a cost center or a profit center. Cost centers are covered through a defined budget with no expectations that they'll cover their costs. As a shared service center, they support the business and are very easy to engage with. They also struggle to justify ongoing investment; to an accountant, they're still a cost. Profit centers need to be self-sustaining. Through negotiation, transfer prices, or direct cost recovery, they need to be able to demonstrate their contribution to the organization's bottom line. While they're often proactive, every integration they have with another unit "costs" money. As such, while they rarely struggle to get investment as long as they can demonstrate financial performance, their constant focus on cost recovery can act as a disincentive for other groups.

This seems to force an impossible choice. If the group tries to recover its costs from the start, it'll compound cultural resistance with financial resistance. No one likes being asked to pay for something upfront where the value is not necessarily known yet. Unfortunately, analytics is mainly about dealing with uncertainty. However, never recovering *anything* greatly limits what the team can achieve. Making changes to operational systems costs real money and without this funding, it's impossible to embed analytics in operational processes.

Clearly, neither model is more effective than the other; each offers different advantages and disadvantages. The mistake most organizations make is to assume that it needs to be one or the other. By starting with the funding model rather than the desired outcomes, they make it that much harder to succeed.

The best solution is to align the funding approach against each activity's objectives. Early on, the net should be cast as wide as possible. While the expected value from any given project is usually quite low, some will offer significant value. These need to be identified and validated if a case is to be made for change. Once validated, the expected return should be significant enough to justify the investment that will be needed to change operational processes.

Following the trajectory of expected value creates the process described in Figure 6.3.

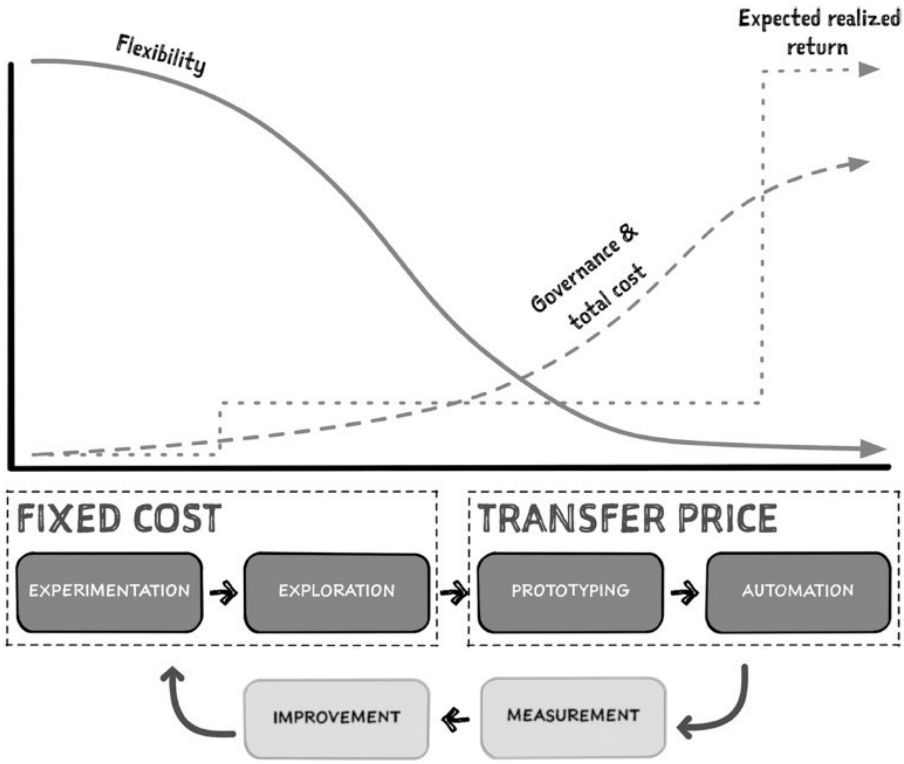

Figure 6.3: The Return Cycle

During early stages, the main objective is simply to determine whether a solution exists and, if so, whether it's feasible to develop. Rather than create barriers to adoption or innovation at this point, it's better to make it as easy as possible to engage with the group. By treating the people who support exploration as a cost center, other groups are free to engage without having to allocate budget or otherwise reprioritize their activities. Flexibility is key during this stage and as such, simply defining the engagement approach is usually enough. Activity and focus are best managed by prioritization, often supported through an oversight executive committee comprised of key senior stakeholders from the group's customers.

If there's merit in investigating whether it's worth converting the results of exploration into an operational process, that same group can assist with feasibility and prototyping. Agility and relevancy are key during this stage and, as such, methodologies such as Agile are well suited. Even better, the target result of this process should be a clear understanding of what the value of the change would be to the business. This helps build the business case and justify more direct and significant funding. Activity and focus are again best managed through prioritization, again through the oversight management committee.

Converting the prototype into a robust operational process is where the real costs start to mount. It's also where the real value of business analytics starts to emerge. Because of this, it also makes sense to treat the group responsible for automation as a profit center with their time accounted for either through cost recovery or direct recognition via a true or shadow profit-and-loss statement. Efficiency and repeatability are key during this stage and because of this, service design approaches such as the IT Infrastructure Library (ITIL), software development methodologies such as the waterfall method, and project management methodologies such as PRINCE2 are well suited. Activity and focus is managed through direct investment and program of work management.

There's good reason to separate exploration from execution. By having an investment model designed to both encourage use *as well as* support return, business analytics becomes self-funding without unintentionally establishing barriers to adoption. The goal is to make it self-sufficient and profitable, not just a cost center with no clear direction.

Notes

1. For a far more comprehensive review and explanation of these concepts along with how to quantify them, see Chapters 4 and 6 of Evan Stubbs, *Delivering Business Analytics: Practical Guidelines for Best Practice* (Hoboken, NJ: John Wiley & Sons, 2013) and Evan Stubbs, *The Value of Business Analytics: Identifying the Path to Profitability* (Hoboken, NJ: John Wiley & Sons, 2011).

2. For more detail on the use of operational analytics and decision orchestration, see Chapters 6 and 9 of Evan Stubbs,

Delivering Business Analytics: Practical Guidelines for Best Practice (Hoboken, NJ: John Wiley & Sons, 2013).