

Developing Your Data Strategy: A practical guide

Gregory S. Nelson

ThotWave Technologies, Chapel Hill, NC

Abstract

The ever-growing volume of data challenges us to keep pace in ensuring that we use it to its full advantage. Unfortunately, often our response to new data sources, data types and applications is somewhat reactionary. There exists a misperception that organizations have precious little time to consider a purposeful strategy without disrupting business continuity.

Strategy is a phrase that is often misused and ill-defined. However, it is nothing more than a set of integrated choices that help position an initiative for future success. With that in mind, this presentation will cover the key elements defining data strategy. Key topics include:

- What data should we keep or toss?
- How should we structure data? (warehouse vs. data lake vs. real-time/event stream)
- How do we store data? (cloud, virtualization, federation, cloud, Hadoop), the approach we use to integrate and cleanse (ETL vs. cognitive/ automated profiling)
- How do we protect and share data?

Addressing these topics ensures that the organization gets the most value from data and has a plan to prioritize data feeds and adapt the strategy to meet unanticipated needs in the future. As with any strategy, there must be a roadmap for execution, so we will specifically address the tools, technologies, methods and processes useful in designing a data strategy that is both relevant and actionable.

INTRODUCTION	2
OUR DATA STRATEGY STATEMENT.....	2
STRATEGY VERSUS IMPLEMENTATION	3
STRATEGY DEVELOPMENT PROCESS.....	4
DEVELOPING YOUR DATA STRATEGY ROADMAP	7
WHAT SHOULD BE INCLUDED IN YOUR DATA STRATEGY	7
AN AGILE APPROACH TO DEVELOPING YOUR DATA STRATEGY	13
SUMMARY.....	13
ACKNOWLEDGEMENTS	14
BIOGRAPHY	14
CONTACT INFORMATION.....	15
REFERENCES	15

Introduction

Strategy is a term that we see used in every day conversation. Everyone seems to understand the word as “something important” or “not tactical”. Despite its commonplace usage in our vocabulary, I would argue that strategy is perhaps one of the most misunderstood words used in business. The implications of which can be seen in various ways:

- Inability to respond to emergent questions
- Poor customer satisfaction around analytic products
- Massive backlogs of data requests
- Analytic staff turnover

In business strategy, we talk about the components of an organizational statement (including Mission, Values, Vision, and Strategy) as a way to help ensure alignment between who we are, the values by which we operate and the strategies that we use to accomplish our goals. These organizational statements outline the fundamental positions that you hold as a company which help guide your everyday work and keep you focused on your goal.

We recommend that each and every analytic team have a strategy statement that describes not only why they exist and who they want to be, but also a strategy statement that clearly outlines how they intend to achieve those objectives. By articulating this, the team is setting down a flag to say not just who they are, but who *they are not* which can be an incredibly helpful compass when rationalizing potential project opportunities. Furthermore, analytic teams should ensure that they have a plan for how they intend to measure success and whether their strategy is working. When this doesn't happen, which is very common, the entire strategic planning process is undermined and staff understandably start to question its value.

We believe that there is a strong relationship between the lessons that we learn in business strategy and data strategy. In business (or organizational) strategy, we define elements that articulate our purpose (mission) and paint a picture of the future (vision) as our foundational goals. The strategy is the set of choices that will be made to get there. Similarly, data strategy is the integrated set of choices that we use to position our firm for analytic success. This is because data strategy is the guide that we use to support decisions on what data we choose to pursue, manage, use, and govern.

Our Data Strategy Statement

The key to the development of a data strategy statement is to provide clarity around what you intend to do. In general terms, a Strategy Statement details what our specific game plan will be and has three component parts: objective, scope, and advantage.

- The strategy statement begins with a definition of the ends that the strategy is designed to achieve – that is, the **Strategic Objective**. This is a single precise objective that will drive the organization over the next 5 years or so and has enough clarity that there is no room for misinterpretation.

- Next, we have the **Scope** which outlines the boundaries of operation for the organization (or what data we intend to actively pursue.)
- Finally, **Advantage** is the unique differences of the organization that make it distinctively unique and capable of delivering value. There are two parts to the advantage statement – the first is the customer value proposition and the second is the set of unique activities that position the organization to deliver on that value proposition.

In sum, the strategic statement is a clear statement outlining the organization’s objective, where it will operate, what value it creates for the customer, and what set of activities it will perform to achieve differentiation. The diagram below (adapted from Collis and Rukstad (Collis, 2008)) outlines the relationship between the Mission, Values, Vision, Strategy Statement, and the Execution and Measurement approaches for an organization’s business strategy.

The Fundamental’s of Organizational Alignment



Figure 1: The elements of business strategy that can be used to drive data strategy

So with the general strategy statement in mind, the data strategy parallel would include a strong understanding of what we want to accomplish then defines the boundaries of what we intend to include in our data strategy and what lies outside and concludes with our advantage. In the following sections, we outline a framework that can be used to develop your data strategy that borrows liberally from business strategy development.

Strategy versus implementation

One of my favorite quotes regarding the difference between strategy and execution is from [Strategy and the Fat Smoker](#) (Maister, 2008)

“

Just because something is obvious doesn't make it easy. Real strategy lies not in figuring out what to do, but in devising ways to ensure that, compared to others, we actually do more of what everybody knows they should do.

The data strategy development process is designed to ensure that we have a plan for the use of corporate data as a vital asset for strategic and operational decision-making. Through our consulting experience we know that investing in a formal data strategy frames the organization's intentions for the inevitable data issues that will crop up in any organization. These include issues around data quality, metadata management, access and data sharing, performance, ownership, provenance, maintainability, usability, security, and privacy.

The data strategy involves aligning the aspirations of your analytics organization with the overall organization's strategy. That linkage is critical so that there is clear line of sight between the "data and technical" efforts that you undertake and the business context in which you will operate.

“*Not having a data strategy is analogous to allowing each person within each department of your organization to develop their own chart of accounts and use their own numbering scheme.*

Sid Adelman, Data Warehousing Expert (Adelman, Moss, & Abai, 2005)

In practical terms, strategy is our game plan. Note, strategy is not something we do at the beginning of a planning cycle only to be forgotten once the realities of everyday operation consume us. As in business, we need to ensure that there is a clear linkage between strategy and execution. A 2006 survey conducted by Palladium (founded by Harvard Professors Kaplan and Norton) shared a striking statistic where only 25% of companies said they performed as well as, or better than, the average of their industry peer group when it came to linking strategy with execution. This means that 75% or organizations were less than or equal to being average! In their January 2008 Harvard Business Review article "Mastering the Management System," Kaplan and Norton tackle this issue directly, describing an integrated process for linking strategy and operations." (Kaplan, 2008)

Strategy Development Process

Just as in business strategy development, there are a number of approaches that can be used to develop your data strategy. What we present here is based on years of consulting experience with customers and we have found this to work well. Note, however, we utilize an agile approach so these stages are iterative and not a series of sequential steps. This is based heavily on the philosophy of design thinking.

Design thinking is a user-centered process that begins with understanding the user. In our case, the users are beneficiaries of a well-executed data strategy. The process leverages the collective expertise of the key stakeholders and establishes buy-in amongst the participants in the data strategy process – not just the IT folks responsible for managing the data platforms (SAS®, Hadoop, Oracle, Teradata, Netezza, etc.) We have found that design thinking encourages innovation by exploring multiple avenues for the same problem and getting to a feasibility much more rapidly.

We relate these stages to data strategy as one way to potentially think about how an organization might create the same alignment between data strategy development and the execution of a data roadmap. The table below highlights what we see as the critical steps in linking data strategy to implementation.

Process Area	Deliverables/ Milestones	Description of Activities
Develop the strategy	<ul style="list-style-type: none"> • Approved Data Strategy Statement • Analytics Organization Mission and Vision (who we will serve and for what value) • Design thinking workshops (empathize, define, ideate, prototype, test) 	<p>Your analytics organization must be able to state exactly what business you're in, identify the key issues you face, and determine how best to compete to support the organization in achieving the best possible outcomes. Developing the data strategy uses goes beyond traditional business strategy activities such as mission, values, and vision statements; external competitive, economic, and environmental analyses; and methodologies. We utilize the business strategy to map our data strategy to the vision, mission, and strategic goals of the organization. In addition, we develop measures of success that align to the organization's KPIs. Using these as inputs, the data strategy statement outlines the purpose, scope and advantage the organization should provide clarity around what you intend to do as an analytics organization.</p>
Create the Roadmap	<ul style="list-style-type: none"> • Approved and funded Data Strategy Roadmap (plan for execution) 	<p>Once the data strategy plan is outlined, then we must translate the mission, vision and strategy into an execution and measurement plan. This is where we articulate the strategic objectives, measures, targets, initiatives, and budgets that will ultimately guide action and resource allocation. It is critical that we can describe the data strategy, measure the plan, identify plans of action, evaluate potential risks and mitigation strategies. Once detailed, that serves as input to funding and staffing and leadership decisions. Traditional business strategy uses tools such as OKRs, Balanced Scorecards and other measurement approaches.</p>
Align the Organization	<ul style="list-style-type: none"> • Change Management Plan • Stakeholder Management and Communications Plan 	<p>As you begin the execution of your data strategy, it is imperative that we link it to the strategies of individual business units engaged as stakeholders in the process. If you are developing an enterprise data strategy, that would mean the individual business units or divisions. If the data strategy is for a service line or unit, then the stakeholders would be those affected (either positively or negatively.)</p>

Process Area	Deliverables/ Milestones	Description of Activities
		To support this, we utilize strong change management methods to ensure alignment, motivation, engagement and action so that we can optimize strategy execution. Utilizing both formal and informal methods, we ensure alignment through a formal communications process, and by linking employees' personal objectives and incentives to strategic objectives impacted by the data strategy.
Execute the Plan	<ul style="list-style-type: none"> • Analytics and Data Lifecycle workflow • Data governance plan • Quality Processes 	Many projects fail when they omit the operational planning for major initiatives such as a data strategy effort. We must link long-term strategy with day-to-day operations, aligning strategy with operating plans and budgets while focusing on those process improvements that are most critical to the strategy. Practically this means that we need to ensure that the business operations account for and build in operational impacts into their workflows and workload. Work streams should account for impacts to the analytics lifecycle, data governance and quality processes.
Monitor and Learn	<ul style="list-style-type: none"> • Measurement Plan and Results 	As with most things, sustained execution requires a commitment to monitoring performance results once a strategy has been developed, planned and implemented, enabling you to determine if the strategy is being properly executed. It requires monitoring and learning about problems, barriers, and challenges. This process integrates information about operations and strategy into a carefully designed structure of management review meetings.
Test and Adapt	<ul style="list-style-type: none"> • Lessons learned • Knowledge management plan • Best practices evolution 	Learning organizations become very good at testing fundamental strategic assumptions to determine if you, indeed, do have the right strategy. This involves testing and adapting the strategy, using internal operational data and new external environmental and competitive data—thus launching a new cycle of integrated strategy planning and operational execution.

Table 1: ThotWave's Six Stages for Data Strategy Execution

In the following page, we will explore this methodology by first discussing what should be included in the data strategy and then walk through a practical example.

Developing Your Data Strategy Roadmap

What should be included in your data strategy

While there is no perfect answer here, we have found the following elements common to our consulting work. Note, we intentionally did not specify a template or format for capturing these. Our recommendation for the format that you use should follow these guidelines:

- Use whatever form you can create quickly and easily
- Ensure it has collaboration features that allow for commenting on drafts and multiple collaborators
- It should serve as a living document with changes noted and approved so that you have a history of your strategy's evolution
- It should be accessible in a form that people can digest (e.g. executives, business users, and employees) so everyone knows where it is, what is current, and what is ahead on the roadmap

There are lots of opinions as to the best place to start when formulating a data strategy. We tend to take a pragmatic approach and ask, "What data is essential for us to achieve breakthrough performance in our organization?" Since we build our data strategy to support organizational objectives, this should be prioritized based on the ways in which these data can help move the needle on our strategic initiatives. Unfortunately, one of the challenges that people face is that they often feel like "they don't know what they don't know". The folks supporting the IT side of the data often frame their questions around "what are your requirements?" The business folks address this from "how can data best help us and what data is available? What questions should I be asking?" Thus begins a pattern of 'who's on first' riddling that cycles endlessly due to lack of a common language.

That's why an agile approach based on design thinking can spur establishing a dialog based on empathy and problem definition. We often look to the business processes that support strategic objectives and ideate around what the processes would look like if access to data wasn't a problem. Blue sky thinking can often break the chains of traditional "requirements analysis." A recognition that data supports the analysis of important decisions that in turn support the organization's movement toward achieving breakthrough performance is often lost on those that push data into their data lakes and warehouses.

“Recognize that data is the lifeblood of your organization and treat it accordingly.

Michael A. Schiff, Data Management Consultant

A typical data strategy includes the following perspectives along with key questions that should be addressed.

SCOPE AND PURPOSE

- *What data will we purposefully manage? What data is out of scope?*
- *What data do you need to run your organization versus data that is needed to grow your business or meet your strategic objectives?*
- *What are you trying to achieve?*
- *How do we measure success for our data strategy?*
- *How much is your data worth?*

The data strategy must compliment the business strategy. We encourage analytics organizations to develop their own mission statement that articulates their identity and reason for existence. This not only serves to define what is in scope, but also helps to clarify what is out of scope.

In addition, defining clear expectations for how an organization will measure progress against its data strategy is key. Metrics can include data service level metrics (for data systems and data quality), agility in supporting strategic decisions or operational metrics, or milestone achievement against the plan.

Early on, we recommend that you begin to catalog key challenges or risks as well as define a communication plan to support the anticipated design decisions. Map the customer journey to understand and prioritize where processes are broken.

When looking at priorities, ask yourself, “what is the value of a key field such as an address? What incremental revenue can you expect? Putting a value on data helps to make data management a continuous improvement process and supports the value that you are creating through these efforts.

Trap: *Do not fall into the trap of accepting “all data” in answer to the question of what data is needed. Successfully managing data consumes resources and there is a limited supply of resources for any organization.*

DATA COLLECTION, STANDARDIZATION AND CLEANSING

- *How frequently does data need to be updated?*
- *What data archiving and summarization requirements are there?*
- *What data do you need where and when? Today and in the future?*
- *Where is your database of record for different data and do you require a single customer view?*
- *How do you integrate data silos? What data transfers are necessary and how frequently?*

This stage is often where things go wrong. The need to join the dots means data *connections* are becoming as important as data *collections*. This is your opportunity to create a clear definition of the types of data or information needs that are required and translate those into a prioritized list. It is no longer sufficient to

develop data warehouses and data marts that are domain specific. Architects must consider the complete customer, patient, or product journey.

We believe that you cannot have data standardization without also considering data integration. This means that not only data redundancy has to be addressed, but also master data management. If working at the departmental level, a master data management strategy may possibly be out of reach. So you have to advocate to your leadership why this is an important enterprise tool and minimally operate at the level that you can have an effect.

DATA ARCHITECTURE, VIRTUALIZATION AND INTEGRATION

- *Where do you need real-time data and what specific data should be real-time?*
- *How should data be shared between the various data silos?*
- *How do you access, share and manage data?*
- *Should data services be a SaaS or in-house solution?*
- *How should we model the data?*

While there are a tremendous number of data integration tactics, tools and technology, in your data strategy you should be asking yourself more technically agnostic questions here such as what data should reside in what format and how often should it be loaded? What should be the data's source, and how much history must be retained? It is rarely practical or necessary to have all data available in real-time in one place. Similarly, cloud based storage solutions may look attractive but without a sound data strategy, these solutions can turn into a solution without a problem.

For example, for nearly 2 decades we have seen folks espousing the concept of a "virtual data warehouse." This means that data could be accessed without first moving it into a data warehouse. Although some of the benefits include access to real-time data and reduced storage costs, one overwhelming risk is that data from multiple sources is often inconsistent, inaccurate and can have deleterious effects on operational systems.

Similarly, some data may need to be available in real-time whereas other data may not be as time critical. How and where we store the data should be based on sound requirements. Some data needs to be available on-demand while other data may not be as time critical. A data strategy should consider what needs to be archived and where to archive it.

Your data strategy must also answer the question whether to store data on premise or in the cloud? As organizations move data to the cloud and/or use SaaS applications, it is important to consider a "worst case scenario" should there suddenly be cut off from your data, for example, a massive hacking, ransom-ware or a system wide outage for your cloud storage provider (for example [the recent Amazon Storage Services outage.](#)) Furthermore, make certain you retain all rights to your data, and make sure you have timely backups in your possession.

In recent years, people have been talking about the merits of data lakes (and data swamps) as an alternative to modeling data using traditional technical (e.g., logical data model, entity relationship modes, physical database models, etc.) Regardless of what database design schema you ultimately choose for storing the data, you should still create business data models for understanding the semantics and the business rules of the data first or else it will be impossible to standardize the data.

DATA INSIGHT AND ANALYSIS

- *How should you facilitate data exploration?*

- *How will you develop, deploy, and manage analytic results?*
- *Will you be providing self-service access to data or develop a provisioning team?*
- *How will analytic leadership learn about what's working in the data strategy?*
- *How should you foster collaboration on data throughout the organization?*

While it should not have to be said, the goal of the “data” exercise is not to end up with the most data. Rather, the purpose is to put data to work for your organization to drive decisions, improve performance and drive innovation. If data is the life-blood of an organization, then analytics is the heart pumping insights to every corner of the system. Asking yourself questions like, “so what? What will we do with the results if we had them? How will this change our processes or workflows?” is critical. Recently, I had discussions with an enterprise architect who was designing a new solution that included analytic results stored back in the warehouse. This way the organization could learn what is working and not rely on external processes or tribal knowledge.

DATA GOVERNANCE AND DATA QUALITY

- *Who owns the data?*
- *Who will provide strategic leadership around the data strategy?*
- *Who is responsible for data stewardship? for data quality? for educating users? for managing security?*
- *Who ensures that the data architecture will be updated as evolving data types and needs arise?*
- *Who will decide who gets what level of access? Who can change, copy, move, or delete data?*
- *Who gets to bring into new data?*
- *What level of data quality is practical?*
- *What is the cost of cleaning data?*
- *How will we measure data quality?*

Data governance is just one part of an overall data strategy. In data governance, we answer the tough questions about roles, responsibilities, approvals, and workflows to ensure that the data is well managed. This cascades from the overarching rules of engagement down to who owns the data elements, value lists, and our verification strategies (e.g., should addresses be checked against postal databases during data entry). We address questions here of how often a value must be updated when it changes, what are the legal requirements, and so on. Although many organizations have devoted considerable effort to governance initiatives concerning customer master and financial data, data governance should apply to all data assets.

A data strategy is a strategic plan for enterprise-wide data governance. It is our belief that instituting enterprise-wide data governance is a business responsibility supported by IT. This responsibility starts with data ownership and extends to data stewardship. The data owners are usually the originators or primary users of the data. Data stewards are responsible for enforcing standards, data governance policies, and managing business rules. While they do not have authority to set policies or to create business rules, data stewards are accountable for performing data audits and resolving data disputes.

A critical part of any data governance program is a related concept around data quality. In the data strategy, we address our definition of what we mean by “high” and “low” quality data in an objective manner. How we communicate, raise awareness.

and garner support for data quality issues across the organization is part of data governance. Tactically, the data quality governance process will address questions such as who is responsible for data quality? Who will identify, evaluate, and diagnose data quality problems? Who is responsible for outlining the processes by which we validate our data integration and transformation processes?

While quality is an important concept, it should be noted that it is important to evaluate the cost of cleansing data. Not all data may be worth its cleaning costs. Determining the cost of poor data quality for your organization allows prioritization of cleaning efforts.

METADATA MANAGEMENT

- *What metadata should be managed? Who should define business metadata and technical metadata?*
- *Who is responsible for raising awareness and managing metadata?*
- *Who is responsible for capturing and maintaining metadata?*
- *How can we exploit metadata once we have it?*
- *What tools and systems generate metadata?*
- *What tools will expose metadata for business use?*

Metadata is one vehicle for achieving data standardization and integration. Metadata is contextual information about IT assets such as data, processes, and programs. Metadata components for data assets include business definitions, domains (valid values), data formats (type and length), business rules for creating the data, transformation and aggregation rules, security requirements, ownership, sources (operational files and databases), timeliness, and applicability. While many companies talk about metadata and can demonstrate the data they collect, not many companies capture these components, and those that do, don't make effective use of it. Metadata should not be a dirty word and left as a documentation exercise. Nor should all decisions surrounding metadata strategy be punted to a vendor of a tool. But rather, metadata should be the map that people use to navigate the data across the enterprise.

ACCESS, DISTRIBUTION, PRIVACY AND SECURITY

- *Who can access data and what level of detail can they see? Does that vary by country or location?*
- *How should data be protected and should it be encrypted?*
- *How will we determine responsibility for security and privacy?*
- *Who will own (and what process will be used) to establish security and privacy procedures?*
- *What is the process and how often will we audit our security policy?*
- *How will we address regulatory issues that may be highly domain relevant?*

Perhaps one of the more contentious areas in any data strategy is "who should get what access to which data?" It is in the data strategy that we outline how we decide and what methods we will use to manage, monitor, and improve the access, distribution, and security aspects of data. In today's business environment, it is no longer

appropriate to lock people out of data. Rather we need to adopt a tone of data democracy in which data can be used by the collective organization. Obviously, there are limits to that where personal, private or confidential data are concerned.

The data strategy should outline the process for how we determine users' need for data as well as define standards for what data to distribute when, where, and to whom. As with most things, we should also monitor, measure, and report the cost to distribute data and assign responsibility for administration activities related to distribution so that we understand the full cost of our efforts and can support prioritization.

There are also several global issues that we may face as it relates to where and how data can be accessed across borders. For example, some countries place restrictions on data that can be collected or transmitted as if there is a virtual 'customs' gate for data. Make sure your organization, or your cloud vendor, is not inadvertently violating governmental restrictions as data passes through international lines.

Furthermore, we should account for what data can be displayed. For example, the U.S. government and various state regulations regulate the use, storage, and display of Social Security numbers.

DATA RETENTION

- *How long should we keep the data?*
- *What data will we keep?*
- *Who gets to determine our policies for the maintenance and archival of historic data?*

We have seen that most organizations place little effort on the strategy to determine what analyses and data sources may no longer be needed for on-demand access. Managing and protecting data comes with a cost and just because it is easy to capture and store massive amounts of transactional data, should you? Could those resources be put to better use?

One of the common tactical issues to address the level of detail that should be retained. That is, what level of granularity should we keep, in what format, and what should the service level agreement entail? We may design our strategy around keeping detailed data for some weeks or months and then retain lightly summarized data for several months or years. Legal and regulatory implications should be reflected in any retention plan.

Nonetheless, make sure that you retain data at the level is required to meet current and anticipated needs. We would argue that there should be different strategies for data that has a clear business purpose versus data that merely has a potential to be useful, someday.

PERFORMANCE/ SERVICE LEVEL AGREEMENTS

- *Who is responsible for creating the requirements for performance?*
- *Who is responsible for determining the requirements for availability?*

Finally, your data strategy should outline how you intend to manage our service level agreements – that is, what performance is expected (implicitly or explicitly). The data strategy should address how we will perform

capacity planning of data, network, and security infrastructure. In addition, we should address how we intend to monitor, measure, and report on how our “data systems” are performing and what process will be used to manage service level agreements.

An Agile Approach to Developing Your Data Strategy

While this list may seem overwhelming, you should use these questions to guide the development of your data strategy and address the components that are most pressing for your organization. As we said above, make this a living document and fill out the details over time. To that end, we recommend that you have a plan for what items you intend to address over what time period so that key stakeholders understand what the destination looks like and what stops you will make along the way.

While we don’t consider these critical to the data strategy itself, we also note that you will likely need to address lots of other questions as you go from strategy to execution. Additional topics likely would include tools and technologies, data standards, the analytics lifecycle model to be used, and testing of data theories. How will we address new data sources in the future? How will we manage the software selection process and what are our criteria for software selection? Will we sell our data, and if so, why and how? If you have a data strategy to develop in your future, start thinking about these questions, now.

One of the challenges with the model above it doesn’t specify whether an organization is designing their data strategy for the first time or refining. As a practical matter, we believe that agile approaches for working through these stages (for an overall analytics strategy or the supporting data strategy) should utilize design thinking best practices and NOT fall into the trap of traditional software development life cycle practices where we capture the artifacts in a manner analogous to planning a trip to the moon. (e.g., capture requirements; get sign-off; develop a design specification; get sign-off; design the software; get sign-off; create testing plans; get sign-off; execute testing; get sign-off; go through change management procedures; get sign-off; ,etc.)

Summary

The focus of this paper was to outline the components of a data strategy. We believe that this is an essential activity that helps to ensure alignment between who we are, the values by which we operate, and the strategies that we use to accomplish our analytic goals. Think of your data strategy as the fundamental positions that you hold as a data-driven organization which should help guide your everyday work and keep you focused on your analytical goals. A data strategy is an essential and fundamental building block for all IT initiatives, regardless whether they are operational or decision-support-based in nature.

A well-defined data strategy can help you achieve the following:

- Align all areas of an organization toward a single-minded “data” purpose that everyone understands
- Serve as a focal point for individuals to identify with the organization’s purpose and direction
- Guide the daily activities by providing a consistent, clear message regarding what is important
- Determine priorities and support decision making

- Motivate those participating in the organization as they know why they are there and why they are doing the things they are doing to support the data driven efforts
- Establish the guiding principles by which they will operate
- Clarify the purpose of the data efforts to external stakeholders
- Articulate where people will and will not spend time
- Provide a basis, or standard, for allocating organizational resources
- Establish a general tone or organizational climate
- Facilitate the translation of strategic objectives into organization structure, capabilities, team organization, team composition, and work processes around data

We believe that a data strategy should be aligned with and support the organization's strategic drivers including: Increasing revenue and value; reducing cost and complexity; and ensuring survival by mitigating risk and enforcing constraints. These recommendations hold no matter the industry. Every initiative that we spend resources on, including data-related initiatives, must serve these fundamental drivers.

Acknowledgements

I would like to thank Monica Horvath for her "thot-ful" review of this manuscript. Her combined support, insights and support of the healthcare analytics work that we do continues to inspire me.

Biography

Greg Nelson, President and CEO, ThotWave Technologies, LLC.

Greg is a global healthcare and Business Intelligence (B.I.) executive with over two decades of experience and leadership in the field. Greg is a prolific writer and speaker interested in healthcare analytics and the strategic use of information technology.

He received his BA in Psychology from the University of California at Santa Cruz and advanced his studies toward a PhD in Social Psychology and Quantitative Methods at the University of Georgia. Recently, Greg completed his Masters degree from Duke University in Clinical Informatics from the Fuqua School of Business. His academic and professional interests include helping organizations mature their analytic capabilities. Founder, President, and CEO of ThotWave Technologies, a niche consultancy specializing in healthcare analytics, Greg is particularly interested in how lessons from across other industries can be applied to help solve the challenges in healthcare.

With certifications in Healthcare IT, Project Management, Six Sigma and Balanced Scorecard, Greg is also a prolific writer and has presented over 200 professional and academic papers in the United States and Europe. He won the Best Paper Award in 2013 at the Pharmaceutical SAS Users Group Conference and sits on the board of the SAS Global Users Group. In 2011, Greg was selected by SAS into their loyalty partner group. "This

program acknowledges and supports individuals who are recognized experts in their fields and have a long-term relationship with SAS.”

Married to wife Susan and living on a small “farmlet” in rural North Carolina, Greg is an avid woodworker, enjoys photography, rides a Harley-Davidson Motorcycle, and strives to be a lifelong learner.

Contact information

Your comments and questions are valued and encouraged. Contact the authors at:

Greg Nelson greg@thotwave.com

ThotWave Technologies, LLC

1289 Fordham Boulevard #241

Chapel Hill, NC 27514 (800) 584 2819

<http://www.thotwave.com>

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

thinking data® is registered trademark of ThotWave Technologies, LLC.

Other brand and product names are trademarks of their respective companies.

References

Adelman, S., Moss, L. T., & Abai, M. (2005). *Data strategy*. Indianapolis, IN: Addison Wesley.

Collis, D. J., and Michael G. Rukstad. (50th Anniversary McKinsey Award Winner for Best Article in Harvard Business Review, 2009.). (2008). Can You Say What Your Strategy Is? *Harvard Business Review*, 86(4).

Kaplan, R. S. a. N., David P. (2008). Mastering the management system. *Harvard Business Review*, 86(1), 62–77.

Maister, D. H. (2008). *Strategy and the fat smoker : doing what's obvious but not easy* (1st American hardcover ed.). Boston, Mass.: Spangle Press.