Gramener
Insights as Stories

HOW TO STRUCTURE YOUR DATA SCIENCE TEAMS FOR THE BEST BUSINESS OUTCOMES

By Ganes Kesari

Co-founder and Chief Decision Scientist
Gramener

Executive Summary

For organizations pursuing data science, the top question executives grapple with is, “How to achieve the best business ROI from our data investments?” A critical factor that influences the effectiveness of outcomes is how you organize your data science teams.

This whitepaper demystifies the structuring of data science teams for Technology leadership roles such as CIO (Chief Information Officer), CDO (Chief Data Officer), and CAO (Chief Analytics Officer).

We’ll discuss the two most common ways of organizing data science teams - centralized and decentralized. Covering the rationale behind these structures, we will examine the pros and cons. We will see how to address the typical challenges with an efficient hybrid structure, the hub-and-spoke model.

Every organization has unique needs, and one size doesn’t fit all. Some teams make a conscious choice to go with one of the two traditional structures. Many others have successfully adopted the hub-and-spoke structure. We will trace data science teams’ evolution to show how this organizational structure helps meet the requirements at each stage of growth.

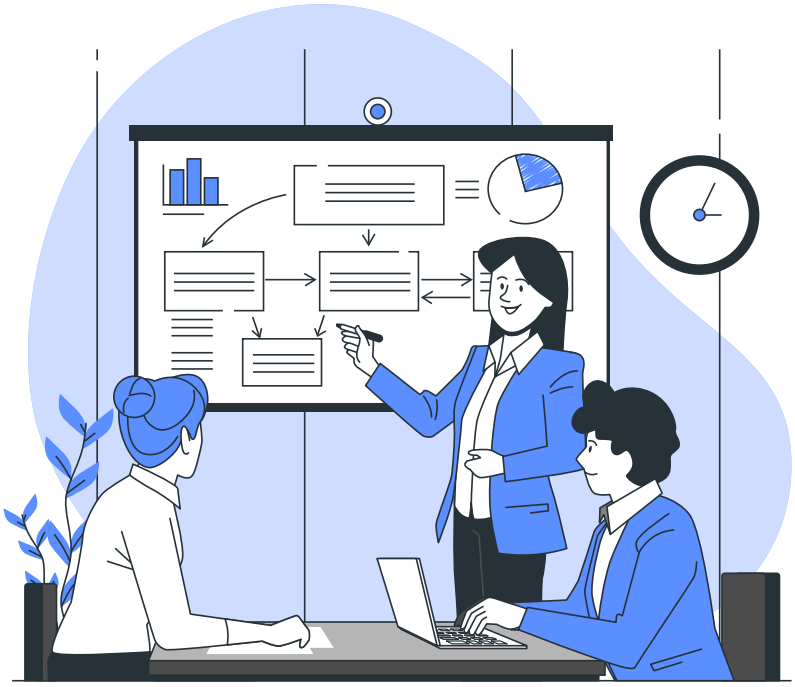
Finally, we’ll present the five best practices to promote collaboration within teams. It shows how to form project teams, what processes they should follow, how to balance accountability with authority, and how organizations can ensure business alignment with user-centricity.

> **“ While most companies understand the importance of analytics, fewer than 20 percent have maximized their potential and achieved analytics at scale ”**
>
> *- McKinsey*

What does this Whitepaper Cover?

- The Data Science Journey of an American Manufacturer
- How do Data Science Teams Mature?
- What are the most common Organizational Structures for Data Science Teams?
- Industry view: Evolution of the Hub-and-Spoke Organizational Structure
- 5 Best Practices to Promote Collaboration in Data Science Team

The Data Science Journey of an American Manufacturer



A large U.S. manufacturer decided to enable data-driven decision-making across the organization. With an executive mandate and highly aspirational goals, they set up an analytics Center of Excellence (CoE). They hired a team of data science professionals and invested in enterprise-wide licenses to implement machine learning and visualization.

This team's first transformational initiative was implementing a customer 360 command center for the sales and marketing teams. The aim was to provide superior, contextual customer intelligence to improve managers and field force decisions.

After spending three months building a robust analytics solution with predictive capabilities, they rolled it out to the users with great fanfare. The sales team's response was lukewarm since they felt that the solution only partially met their needs.

But, the marketing team never used the solution.

They were comfortable working with an internal data science team set up within their function. This team was local, agile, and more flexible in meeting business needs than the central unit. Though the analytics CoE had better funding, robust analytics capabilities, and better processes, the marketing team viewed it as bureaucratic and less aligned to their needs.

The organization was staring at a failed analytics initiative and duplicated data investments. There was an internal political struggle brewing as to which team could prevail in this situation.

Such dilemmas play out in most organizations that pursue data science. Here are some of the common questions executives face:

- *What organizational structure works best for my data science team - should it be centralized or closer to the business?*
- *How do we standardize data science processes and share the best practices while avoiding bureaucracy?*
- *When an organization must rapidly innovate with data, where should you house your data teams?*

We'll answer these questions and more. Let's discuss the industry best practices on how to get value from data science by organizing your teams into the right structure.

> **“ The war is not won with bayonets, but with effective organization ”**
>
> *- Anonymous*

How do Data Science Teams Mature?

Data science teams mature through various stages as they grow and deliver analytics and visualization solutions. Through each of these stages, the organizational structure of teams continues to evolve. Based on the priorities set in each of these stages, the teams acquire distinct characteristics and face specific challenges.

Gartner divides this journey into five stages, also called the Gartner Data & Analytics Maturity Model. Here's a quick preview of the stages on what they tell us about data science teams:

Level-1
Basic

Data & analytics capabilities are ad hoc and unplanned. The efforts are primarily siloed across the organization, often leading to multiple versions of the truth. There are pockets of transactional efforts and experimentations to extract business value from data.

Level-2
Opportunistic

After tasting initial success with data science, individual business units pursue their initiatives. There are some attempts to formalize parts of the process, but the efforts are not standardized and are still restricted to siloes.

> **"Around 87% of organizations are in low levels of data science maturity (levels 1 & 2)"**
>
> *- Gartner*

Level-3
Systematic

At this stage, a clear vision for data science emerges, with strong backing from business executives. Standardization starts setting in with a combination of centralized and shared services offered across the organization.

Level-4
Differentiating

Data science gets a boost with the addition of data leadership roles such as the CDO (Chief Data Officer) or CAO (Chief Analytics Officer). Teams are performance-oriented and operate with a clear business-innovation framework.

Level-5
Transformational

This is the stage where data science is embedded in the business strategy and seamlessly translates into tactical and operational decisions. Data science teams are integrated with the business functions, and data-driven decision making thrives as a culture.

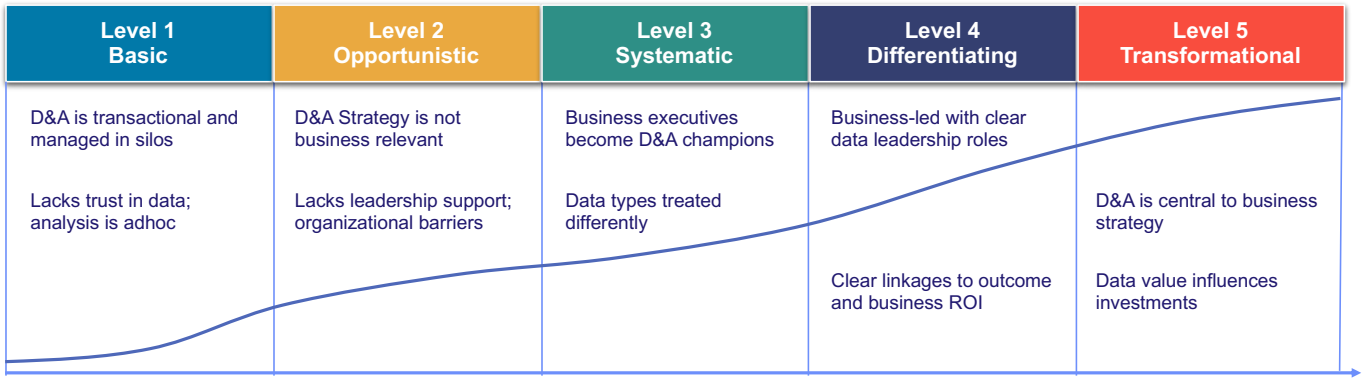


Illustration based on Gartner Maturity Model for Data and Analytics (D&A)

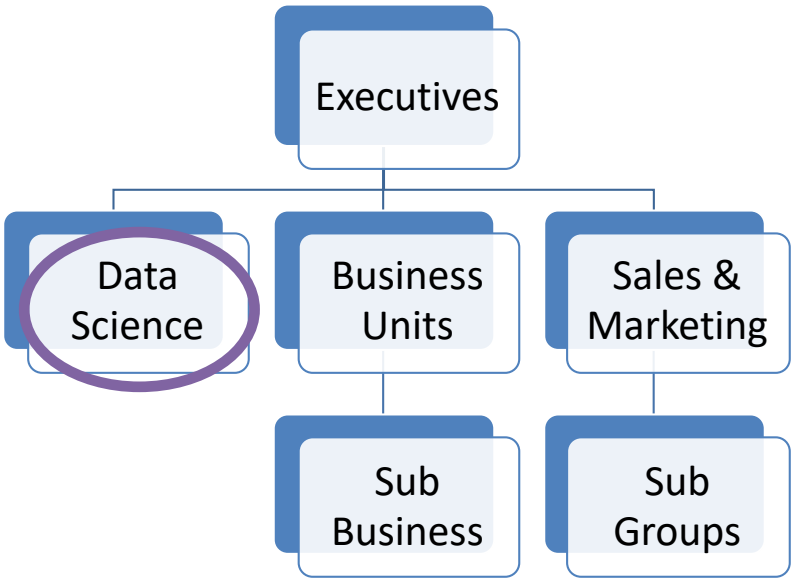
We've traced the evolution of data science teams and have seen how they grow from siloed efforts into those with an organization-wide footprint. As teams scale, they usually settle into one a few popular organizational structures. Let's discuss them and examine their pros and cons.

What are the most common Organizational Structures for Data Science Teams?

You will find two common organizational structures in the industry, and they make up two opposing ends of the spectrum.

Centralized Structure

Many organizations start data science top-down, driven by an executive mandate to leverage data. Such teams often adopt a centralized structure, with strong executive sponsorship and visibility. The team goals are set centrally, and they operate with a mandate to provide standardized data science services across the organization.



Benefits

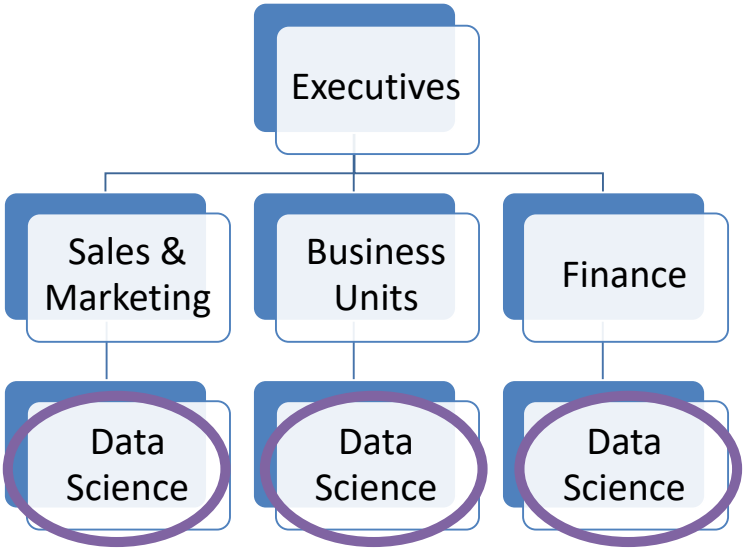
- Centralized teams have strong alignment with corporate goals and priorities.
- The focus is on the standardization and consistency of shared services for business functions. This leads to efficiency and low redundancy of efforts across teams.
- Best practices are shared more readily with excellent retention of talent and knowledge levels.

Challenges

- There's a perception of bureaucracy and tardy response to business requests.
- With teams not tightly integrated with the business, domain alignment is suboptimal.
- Business teams have lower flexibility for experimentation and innovation.

Decentralized Structure

When business executives set up data science initiatives, they move towards a decentralized approach. They are funded and managed directly by business units whose key priority is to have better control and enable rapid response. Decentralized teams are often seen in mid-to-large organizations where multiple data science teams are set up locally within each business unit.



Benefits

- Strong alignment to business needs leads to more effective project outcomes. Data science teams pick up domain knowledge faster
- Data science efforts percolate within the business unit quicker and reach more end-users
- Better control over the data science team enables rapid experimentation and demonstration of quick wins

Challenges

- Localization of data science teams within each business unit leads to duplication of efforts and redundancies.
- Scaling challenges arise when initiatives grow beyond a few projects due to suboptimal alignment with the central leadership.
- Data science professionals' retention is often a challenge due to the mundane nature of tasks and less scope for rotation.

While we have seen the two extreme organizational structures, many enterprises fall somewhere in between. They adopt a hybrid structure that balances the attributes, but they tend to lean towards one of the two models in reality. Hence, they face some of the same challenges we've discussed.

Can hybrid organization structures play the balancing act well? Can they enable standardization and sharing of best practices without compromising on the flexibility and business alignment? Here's a structure that does just that and adapts well to different organizational sizes.

> **“You can manage a large force the same way you manage a small one. It is a matter of communication and formations”**
>
> *- Sun Tzu*

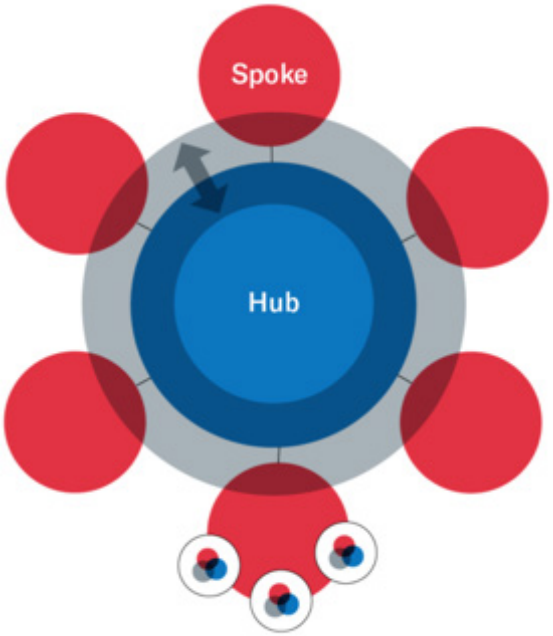
Hub-and-Spoke Structure: The best of both worlds

McKinsey recommends a hybrid organizational structure to build and scale a data science organization. There are three critical aspects to a hub-and-spoke structure:

The Hub

A central team owns the strategic data science priorities, led by a data leader, such as a CDO or CAO. This team defines the vision for data science based on organizational goals and business priorities.

This team sets the tone for all data science activities across the organization, such as hiring, training, defining processes, standardized methodologies, data governance, and working with third-party data software and services providers.



Picture: “Building the AI-Powered Organization”, HBR, Aug 2019

> **“Companies that have implemented AI at scale are three times more likely than their competitors to have a hub”**
>
> *- McKinsey.*

The Spoke

Spokes are business teams, functions, or geographical units that use data science solutions. They are the closest to the end-users, hence manage the data science project and own its outcomes. They are responsible for activities such as the definition of goals, oversight of execution, workflow redesign, end-user training, adoption, and measuring impact.

While the hubs enable standardization and sharing of best practices, the spokes drive business alignment and adoption of initiatives on the ground. The specific split of responsibilities between these teams depends on the organization and their stage of maturity.

The Gray Area

In data science initiatives, there is always an overlap of responsibilities across teams and roles. As teams grow, they must continuously adapt and redefine themselves. This is where the gray area comes in handy.

Some activities could be owned by either the hub or spoke, or could be shared by both. Some examples include translating a business problem into a data science project, insights discovery through analytics, design storytelling, or packaging solutions through DevOps.

McKinsey recommends decisions based on three factors:

1. **Maturity of data science teams:** Companies early in their journey could start with a heavy central hub and gradually transition responsibilities to the spokes while retaining few strategic activities at the hub.
2. **Business model complexity:** When the number of business lines and functions is high, the preference is to house core data science expertise in a hub. When the complexity is not high, this could be distributed to the spokes.
3. **The extent of innovation required:** When rapid experimentation is a critical need, companies build specific capabilities in the hub to test out solutions and quickly deploy them in the market.

The hub-and-spoke structure addresses the challenges of centralized and decentralized structures by bringing in some much-needed flexibility. To leverage its strengths, you must periodically adapt your organization by evaluating and moving responsibilities between the hubs and spokes.

Let's look at how the industry uses this organizational structure.

Industry View: How does an Organization's Hub-and-Spoke structure evolve?

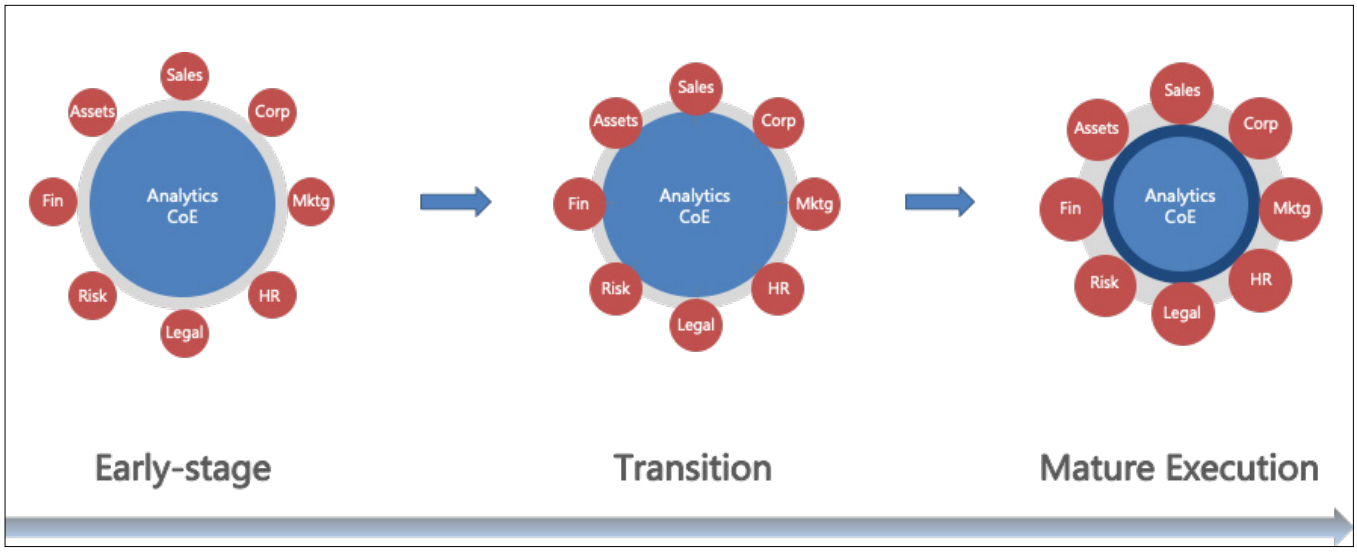
What can we learn from the evolution of data science structures across organizations? How can teams adopt a hub-and-spoke structure right through their stages of growth?

Based on our work with hundreds of clients over the past ten years, we have seen that data science team structures evolve in a certain way. From market studies and our work with clients, there are three distinct stages that the organizational structure passes through.

Early-stage

When organizations get started with a data science team, they have a large central hub with most capabilities bundled at the center. The hub has a mix of data science professionals with skills in machine learning, design, and development.

With most strategic and execution capabilities at the center, the spokes or market-facing units are lean. However, the spokes bring in the domain leader who owns the data science initiative and influences its outcomes. The gray area or zone of overlapping responsibilities is minimal in this stage.



Picture: Evolution of the Hub-and-Spoke Organizational Structure

Transition

Over time and by gaining project execution experience, some data science capabilities are transitioned to the spokes. The hub continues to specialize and deepen the data science skills to act as a Center of Excellence.

Project teams continue to be dynamically formed with leadership at the spokes. The areas of overlapping responsibilities or the grey area start to grow. The specific split of responsibilities is decided based on the three factors we've seen - maturity of spokes, the complexity of projects, and the level of innovation required.

Mature execution

Eventually, the spokes mature in capabilities and take over many of the data science roles and responsibilities. There are dedicated teams of data scientists, design specialists, and technical leaders who permanently reside at the spokes. The dynamic project teams now have majority participation from the personnel at the spokes.

The central hub becomes lean but continues to set the tone for data science within the organization. It owns critical functions such as technical hiring, training, and definition of standard governance and process frameworks. In this stage, the zone of overlapping responsibilities is large. Role clarity and project responsibilities are set firmly by a leadership council drawn from experts in the hubs and spokes.

As we've seen so far, the tenets of a Hub-and-Spoke structure can help us balance control and flexibility. We'll conclude this whitepaper by looking at some time-tested best practices to promote collaboration in data science. Irrespective of the kind of organizational structure you choose, these tips will help you get the most from your team.

5 Best Practices to Promote Collaboration in Data Science Teams

1 Form a governing coalition of Business, IT and Data Science leaders

The role of executives is crucial in ensuring integration, adoption, and business value from data science. Create a joint task force with leadership from the three functions: business, IT, and data science. This task force regularly convenes to review the direction and progress of initiatives throughout the journey towards maturity.

2 Setup governance and process frameworks for repeatable execution

Use governance and processes as tools to make your organizational structure work. Processes are crucial to initiate and manage change across teams. They are equally essential to ensure people follow set guidelines and adopt best practices shared. Robust processes make it easy to birth competencies in the hub and transition them successfully to the spokes.

3 Form dynamic multi-functional project teams, led by a business specialist

Data science is a team sport. Several competencies must come together for a well-rounded solution. Staff every project with a dynamic team at the spoke to keep them closer to the market and end-users. The spokes must own this team under the leadership of a business specialist to ensure user actionability and adoption.

4 Ensure accountability of the team along with full decision-making rights

Organizations often don't place decision-making rights with the data science teams. This can be counterproductive and go against the very purpose of data-driven decision making. Define the realm of ownership for teams. Empower them to action the insights they discover and measure the value of those actions. This combination of autonomy and accountability builds the strength of teams at the spokes.

5 Keep end-users in the loop to build trust and ensure adoption

Data science teams often operate on behalf of the users without actively talking to them. Define processes to involve end-users throughout the data science journey, not just at the time of rollout and user training. Encourage the market-facing teams at the spokes to collaborate with the users, educate them on data literacy, and build their confidence in data-driven decisions.

Whitepaper authored by

Ganes Kesari

Co-founder and Chief Decision Scientist | Gramener

About Gramener

Gramener is a design-led data science company. We have helped over 200 clients across industries across the globe unlock business value from their data by generating actionable insights that are Big, Useful, and Significant. We transform insights into visual data narratives that enable faster evidence-based decision-making.

Gramener has a strong advisory and consulting practice to help executives get business value from data. We help organizations define their data science strategy, and we assist them in picking the most impactful projects. With our strong execution experience, we handhold teams in executing the projects. We work with the executives in tracking ROI from their data science projects and help them make the right interventions to build an organizational culture of data.

References

- Gartner press release on Organizational maturity levels:

https://www.gartner.com/en/newsroom/press-releases/2018-12-06-gartner-data-shows-87-percent-of-organizations-have-low-bi-and-analytics-maturity

- Gartner IT Score for Data & Analytics:

https://www.gartner.com/document/3909048

- Building the AI-Powered Organization, HBR, Aug 2019:

https://hbr.org/2019/07/building-the-ai-powered-organization

- Create a Hybrid Centralized and Decentralized Data and Analytics Organizational Model, Gartner:

https://www.gartner.com/document/3980295

- 3 Approaches to Build a Distributed Hybrid Organizational Model for Data and Analytics, Gartner:

https://www.gartner.com/document/3987665

- Building an effective analytics organization, McKinsey:

https://www.mckinsey.com/industries/financial-services/our-insights/building-an-effective-analytics-organization

Want to know more?

- Watch our webinar on structuring data science teams for best outcomes
- Check out our blog post: How to structure a data science team
- Watch our Advisory whiteboard video series with 5-minute episodes that educate business leaders on getting value from data.

Are you making the most of your data?
Book a free discovery session to find out how data
can power your business decisions

Reach Us

Gramener
Insights as Stories