

HOW TO BUILD AN EFFECTIVE **DATA SCIENCE** TEAM THAT DELIVERS BUSINESS VALUE



Executive Summary

What skills do you need on your team to realize your data science vision? In data analytics, the first role that comes to mind is that of a data scientist. Organizations hunt for Ph.D. data scientists to heavy-lift their projects.

They look for that one magical person with expertise spanning statistics, math, machine learning, business, information design, and communication, among other skills. They assume that such a prized hire will take care of their data science journey.

This is far from reality.

There are several [accounts](#) of organizations that have burnt their fingers due to this misconception. Data science is a team sport. Every team needs five primary roles and skills to deliver business value.

This whitepaper is targeted at Executives, Chief Data Officers, Chief Analytics Officers, and Business heads. It begins by shedding light on the three phases essential to convert data to valuable business decisions.

It then covers the five roles you need to complete this journey from data to business value. Each of the five roles is detailed out with their associated responsibilities and skills required. You will get a sneak preview into three additional, emerging roles in this space.

A common question from leaders is whether you should hire from the market or upskill your existing team. The whitepaper answers these questions. It concludes by presenting five best practices to improve the collaboration in your teams.

At Gramener, we have built our data science teams by screening tens of thousands of candidates. We have advised clients on building effective teams. This whitepaper shares wisdom from a decade of our experience on the frontline.

What does this Whitepaper Cover?



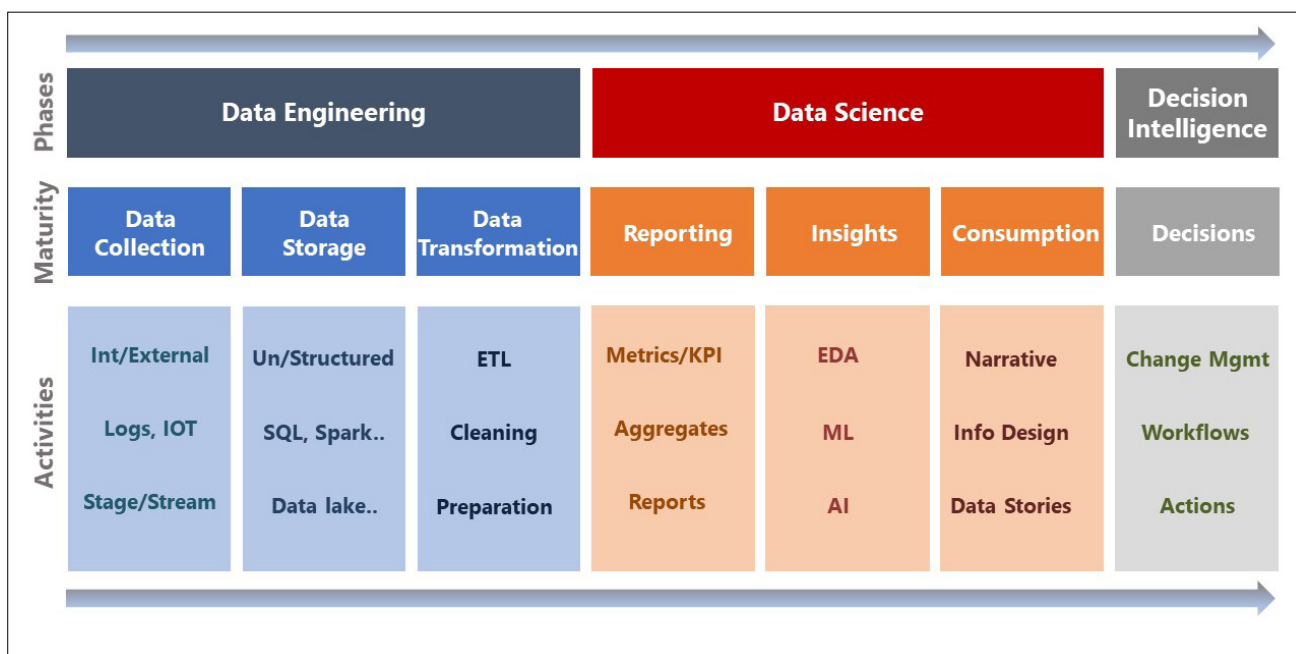
- The Journey from Data to Decisions
- The Five Roles that every Data Science Team Must Hire
- The Three Emerging Roles in Data Science
- What should you look for while Building your Data Science Team?
- Best practices to Promote Collaboration and Improve the Effectiveness of Teams?

Section 1: The Journey from Data to Decisions



Businesses deal with data in various shapes and sizes, and they come from multiple data sources. With data, the ultimate goal is to extract actionable insights and convert them into a form that can drive business decisions.

Irrespective of the industry, this journey from data to decisions follows a standard path. There are three phases in the journey.



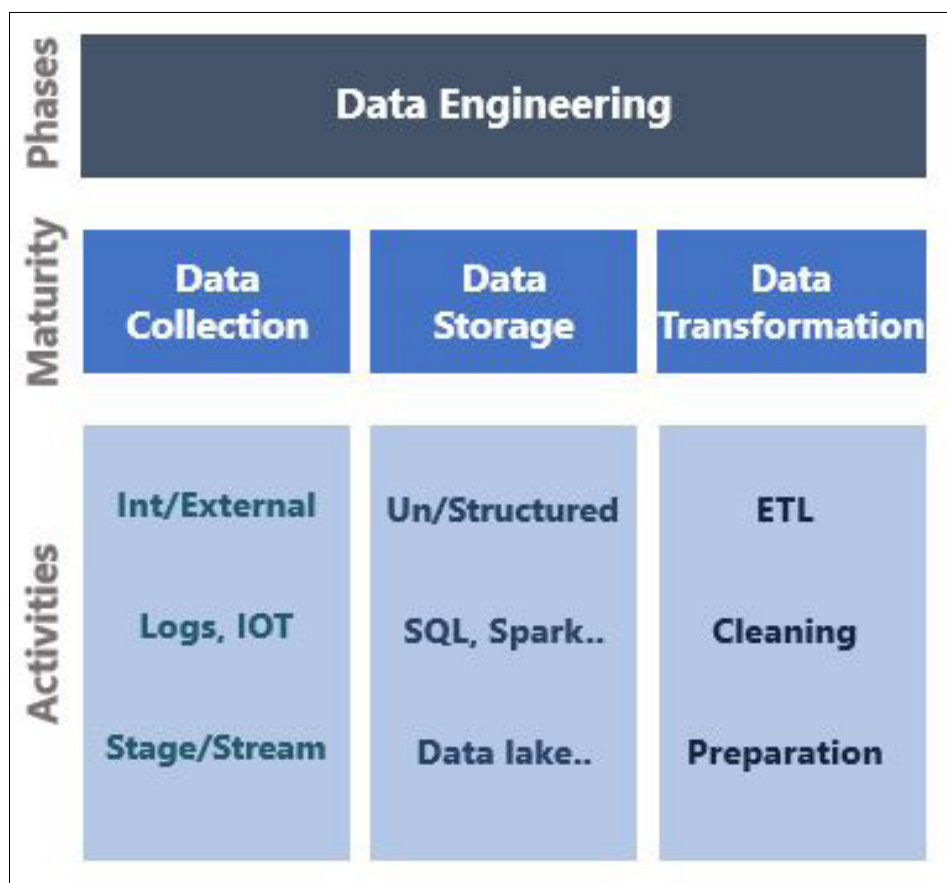
The 3 phases in the journey from data to decisions

Phase 1: Data Engineering

Every organization begins the journey by collecting data. The data could come from internal sources such as transaction system logs, sensors on the factory floor, or customer survey feedback. They could also come from external sources such as weather data or market research data.

The data ingested must be inspected and stored before it can be processed further. It could be stored in structured or unstructured data sources. Then the data is cleaned, prepared, and transformed. This processed, usable data is stored again in the same or new data storage layer.

These activities of data collection, data storage, and data transformation make up the **Data Engineering** phase. This is the first phase and an essential one that ensures that the right data is available to the enterprise.



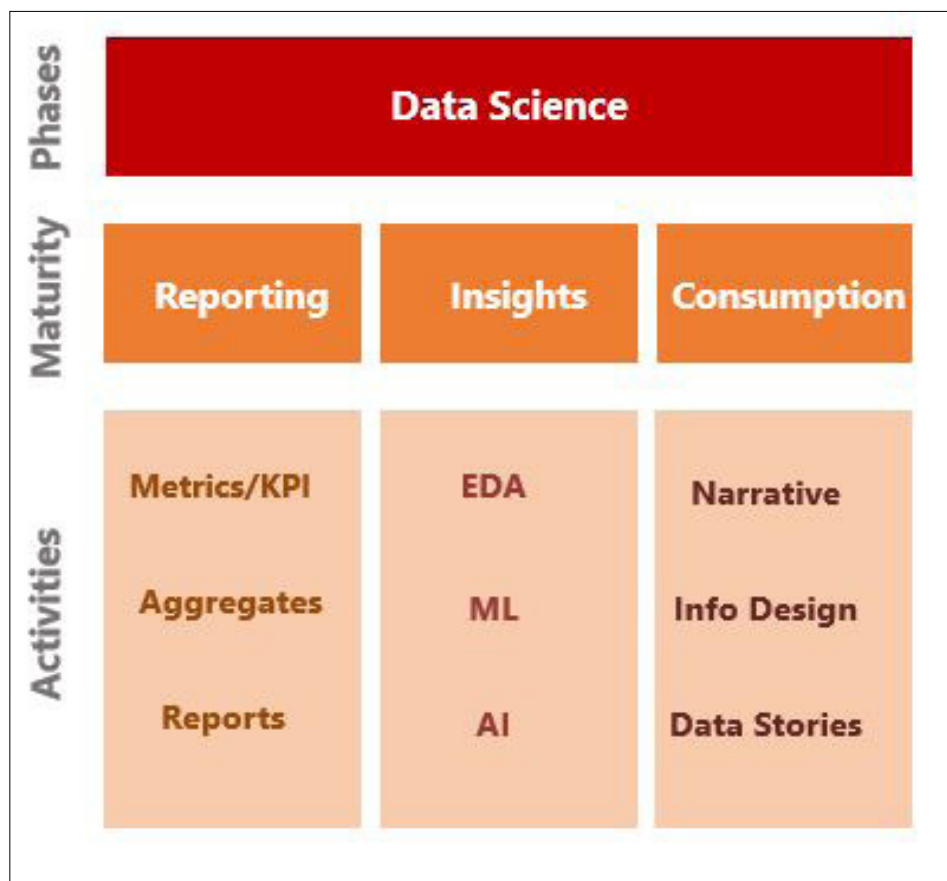
PHASE 1

Phase 2: Data Science

Data Engineering is a prerequisite for data science. The objective of data science is to extract value from the collected data. Organizations put the data to basic use by reporting metrics. This is usually in the form of simple KPI reports or tabular spreadsheets.

They aggregate the data and perform the necessary computations. The findings are simple, backward-focused, and descriptive. However, they help get an understanding of how the business is performing.

To get deeper insights from data, one must leverage exploratory data analysis (EDA), statistics, machine learning (ML), or artificial intelligence (AI) techniques. Depending on the business problem to be solved, the right technique must be chosen and applied.



PHASE 2

Analytics insights serve no purpose if you cannot consume them. Techniques such as information design, data storytelling, narratives help present the insights in the right format for the relevant users.

Phase 3: Decision Intelligence

Often most organizations stop after generating analytics insights or don't think beyond creating consumable data stories. These aren't sufficient to enable business decisions. Organizations need decision intelligence, which is the final phase.

Decision intelligence is the application of data science within the context of a business problem. It involves the influencing of stakeholders to promote the adoption of insights to make the right decisions.

To make this a reality, organizations must build capabilities in behavioral science and managerial science. It is the confluence of these two disciplines and data science that ensures actionability and business value from data.



PHASE 3

Section 2: The Five Roles that every Data Science Team Must Hire



We've seen the different disciplines and skills you need in your team. You need a multidisciplinary team. What roles do they map to? What should be the mix of responsibilities for each role? We'll talk about the five roles that every team needs.

Each of the data science team members will have a unique primary skill. However, everyone in the team must have some standard, non-negotiable skills. What are they?

- Interest and curiosity in data
- Ability to understand data and draw inferences
- Understanding of the business domain
- Ability to present information

Here are the five roles, along with their skill sets and responsibilities:

1. Data Translator

A [data translator](#) offers the best hope to a business in protecting their investment in data science. The data translator understands a user's business needs and helps identify the most relevant projects to execute. She translates the requirements into a format that the data science team can understand. Her role continues throughout the project and is crucial in creating an actionable end-product that users can adopt for decision making.



Responsibilities

- Translate across domain & data
- Own solution from inception to adoption
- Ensure end-user adoption



Skills

- Domain expertise
- Business analysis & solutioning
- Interpersonal & mentoring skills



Closest role

- Business Analyst
- Domain Consultant

2. Data Scientist

A data scientist designs and creates the heart of a data science application. With responsibility for producing business-relevant, actionable insights, he harnesses the power of data analytics. He uses various statistics and machine learning techniques to [embed intelligence](#) and continuous learning ability into solutions.



Responsibilities

- Devise analytics approach
- Analyze data & identify insights
- Build ML models



Skills

- Statistics and machine learning
- Identify & interpret insights
- Scripting skills (R, Python)






Closest role

- Statistician
- Data Analyst

3. Information Designer

An information designer makes the data science solution functional and pleasing to use. Starting with the information architecture, she develops mockups and detailed design prototypes. She makes the data insights consumable by identifying the right [kind of charts](#), interactivity, and visual design to use. She is a master storyteller with data.






 Responsibilities	<ul style="list-style-type: none">• Ensure consumption of insights• Design information architecture• Understand user, drive adoption
 Skills	<ul style="list-style-type: none">• Information design• User centered design• Aspects of visual design
 Closest role	<ul style="list-style-type: none">• UX Designer• Interaction designer

4. Machine Learning Engineer

A Machine Learning (ML) Engineer builds out the working application. He connects to the data sources at the backend, packages the machine learning modules, and integrates with all other systems. He brings the front-end to life through a functional and efficient user interface. He documents the code, maintains logs, and adopts software engineering standards.

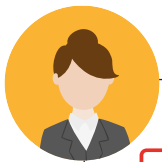


 Responsibilities	<ul style="list-style-type: none">• Package data science solution• Productionizing, DevOps• Data pipelines/integration
 Skills	<ul style="list-style-type: none">• Software engineering• Data handling• Front-end/Back-end coding
 Closest role	<ul style="list-style-type: none">• Software Engineer• Data Architect

5. Data Science Manager

A Data Science Manager is the shepherd of the data science team bringing all the roles together. She empowers the team members and ensures that the solution meets the objectives in spirit. By keeping all client commitments and maintaining communications, she ensures timely, quality deliveries. More importantly, she is responsible for change management and adoption of the solution by business users. She brings in the managerial science aspect into the solution.

This is a role that is underemphasized in the industry today. Most data science projects have a probabilistic nature, and it is challenging to guarantee definitive deadlines and outcomes. Data science managers can understand this and influence other stakeholders to achieve the needed results.



Responsibilities

- Identify roadmap & scale maturity
- Ensure business value from data science
- Drive a culture of data



Skills

- Project management
- Business analysis, solutioning
- Team handling



Closest role

- Project Manager
- Business Analyst

Section 3: The Three Emerging Roles in Data Science

Data science is an emerging field, and we see the need for new skills from varied disciplines. Here are three upcoming roles that will get critical in the coming years. Interestingly, the people best suited to play these roles come from arts backgrounds like humanities, journalism, and law.



1. Data storyteller

Data storytellers are not the same as visualization specialists. Gartner [says](#) that data stories need to have three elements: data visualization, narratives, and context. Stories are emotional, memorable, and actionable. Storytellers craft captivating narratives from the data insights and make them digestible for users who may not have any technical background. For this role, look for people with a background in data journalism or creative arts, such as theatre or creative writing.

2. Behavioral psychologist

Today, most data science applications aim to make sense of human behavior. You need experts who understand why people behave the way they do. Behavioral psychologists can help understand purchase decisions or interpret employee attrition insights. For example, in initiatives to predict employee attrition, they can go beyond obvious factors like ratings or tenure. They can tap into insights around softer aspects such as demographics, employee career stage, or past emotional responses.

3. Data ethicist

Ethics and fairness in data science is a topic that needs urgent attention. When data science teams focus on prescriptive recommendations from past data, bias creeps in. The human world is anything but perfect, and machines learn and mimic them. When implemented at scale, they magnify the bias many times over.

Rather than just focusing on making models explainable, data ethicists bring in the human element. They work with other stakeholders to make ethics and fairness the responsibility of everyone in the organization. They usually have a humanity or law background.

Section 4: What should you look for while Building your Data Science Team?

There is a demand-supply mismatch for data science talent. Buzzwords such as 'Big data' and 'AI' have created a lot of hype over the past decade. While this has increased the awareness and attractiveness of a data science career, candidates with the right skill are hard to come by.



Hiring teams often get hung up on specific tools or programming languages. They screen for knowledge in named machine learning techniques. They blindly look for a certain number of years of experience. Or, they may look for all of these in a single candidate. This can set organizations on a wild goose chase, leading to a lot of effort but minimal results.

Use the skill sets and responsibilities provided above as a guideline to screen for the role. More importantly, you must hire people who have a genuine interest in data. Look for a problem-solving mindset and check learnability instead of depth in narrow skill sets.

For example, check whether aspirants have a public code repository or an information design portfolio. Do they attend data science meetups, or do they contribute to blogs? In the interviews, check the candidates' ability to come up with an approach to solve business problems. Provide sample scenarios and evaluate their thought process. To assess learning ability, look at their career progression. Have they picked up new skills or roles? Check if they take up regular training programs and certifications.

Balance the hiring by retraining your team



Should you just hire for these roles or retrain your existing team. You must do both.

As you bring in the five specialized data science roles, you must balance hiring by looking internally. A good guideline is to look for the same three attributes we discussed earlier: a genuine interest in data, a problem-solving mindset, and healthy learning ability.

Find who on your team demonstrates a passion for data in internal initiatives. Look for problem solvers. Find people who have shown a drive for out-of-the-box thinking in the past. Spot the quick learners, folks who are passionate about new technologies and those who demonstrate a zeal for learning.

This process of spotting internal talent is a continuous process. Think of the scouts in professional sports who travel around the country looking for promising talent. Entrust your data science leaders or managers to take on this additional role of a scout.

When you have a mix of external hires and internal employees in your data science team, your team understands the organizational context and business realities. They can also bring in fresh perspectives using the latest skills available in the market.

Section 5: Best practices to Promote Collaboration and Improve the Effectiveness of Teams?



When you have a team of experts from disparate disciplines, how do you ensure collaboration? Communication breakdown and the forming of silos is a significant risk in such teams. Here are five best practices to improve collaboration and effectiveness of outcomes:

1. 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 closer to the business organizational unit to keep them closer to the market and end-users. The business must own this team under a business specialist's leadership to ensure user actionability and adoption.

2. 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 improves the effectiveness of teams.

3. Rotate team members across initiatives and business units

To ensure your team members continually get adequate exposure to new business problems and problem-solving scenarios, rotate them. Move them across projects within a business unit, or transfer them to entirely new divisions. This will help them work with new team members and acquire new perspectives.

4. Cross-train the teams across skillsets

You must regularly conduct training across skills. For example, put engineers through design boot camps, and have designers take an introductory crash course on analytics. If teams can better understand their peer's perspectives, it will help bridge the gap amongst teams. Cross-functional training helps achieve that objective.

5. Setup governance and process frameworks for repeatable execution

Use governance and processes as tools to make your data science initiatives effective. Processes are crucial for the creation of data science solutions. They are equally essential to ensure end-users follow set guidelines and adopt the recommendations shared. Robust processes make it easy to ensure the administration of responsibilities in the teams and onboard new talent.

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

- Ten red flags signaling your analytics program will fail, Mckinsey: <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/ten-red-flags-signaling-your-analytics-program-will-fail>
- Analytics translator: The new must-have role, McKinsey: <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/analytics-translator>
- Use Data and Analytics to Tell a Story, Gartner: <https://www.gartner.com/smarterwithgartner/use-data-and-analytics-to-tell-a-story>
- The missing link in many data science projects: Decision intelligence, The Enterprisers Project: <https://enterpriseproject.com/article/2020/11/decision-intelligence-key-better-business-decisions>
- Go Beyond Artificial Intelligence: Why Your Business Needs Augmented Intelligence, Forbes: <https://www.forbes.com/sites/ganeskesari/2020/10/19/go-beyond-artificial-intelligence-why-your-business-needs-augmented-intelligence>
- Gramener Advisory [whiteboard video series](#) (5-minute videos) targeted at executives on getting value from data.

Want to know more?

- Check out our [article on Forbes](#): The 5 Roles That Every Data Science Team Must Hire
- [Watch our webinar](#) on structuring data science teams for the best outcomes.
- Watch out 5-minute [whiteboard video](#) on “5 Critical Roles in Data Science”
- Check out our [article on TechCrunch](#): When and how to build out your data science team

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

Copyright © Gramener Inc. All Rights Reserved.