



Building the Dream Team:

Who Should Be Part of Your Data Science Organization?

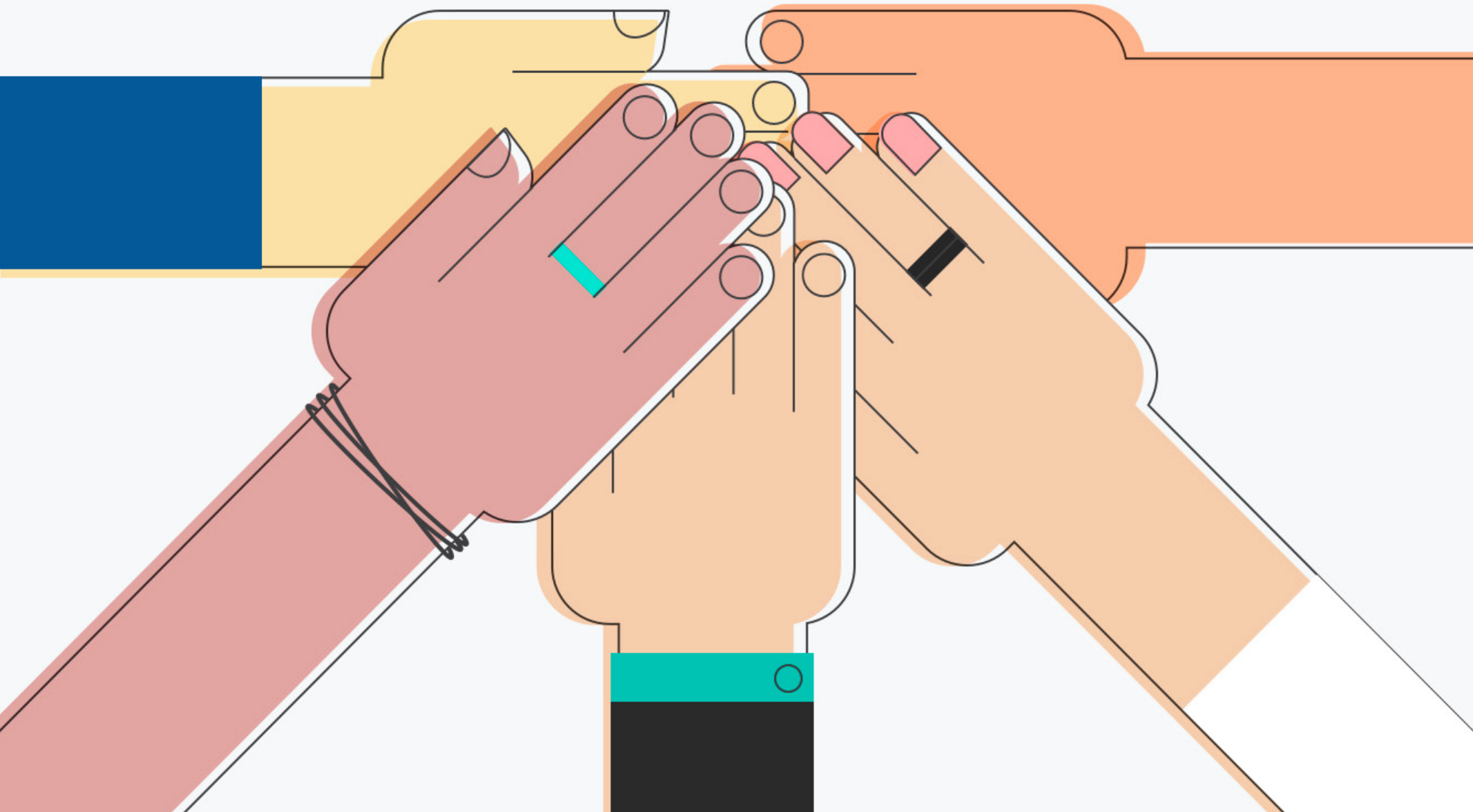


Table of contents

Introduction	3
Part one: What skills are needed by your team?	6
Part two: Who do you need on your team?	13
Final Thoughts: Easing the pressure	19

Introduction

From Jurassic Park to Independence Day to Godzilla, there's one disaster cliché that comes up time and again in sci-fi stories of all stripes: no one *listened to the scientist!*

Unfortunately, real life isn't always that different. Data science teams don't always get heard – to the detriment of everyone concerned.

Sometimes, it's because they're tucked away in a different part of the organization from the business colleagues that should benefit from their findings. Sometimes, it's because the machine learning specialists lack the skills to explain exactly why a pattern or insight should matter to a particular department head. Perhaps, data science and business priorities have drifted in different directions as a project progressed. Whatever the reason, this disconnect is often the biggest barrier to data science success.

It doesn't have to be that way. If you pick the right people for your data science team right from the start, you'll be in a great position to keep everyone invested in the project and pulling in the same direction.

That means assembling a dream team with not just technical prowess but business acumen and people skills. It means looking beyond algorithms to your company's broader vision and goals.

That sounds like a big ask, but what it really means is that you don't necessarily need to go on a hiring spree. You may find that the very best people for the job aren't AI experts but all-rounders you already have inside your organization. Business analysts with a curiosity for data science, for example, and IT champions who are interested in expanding into machine learning.

By keeping an open mind and focusing on skills rather than job titles, you can build a rock star data science team even with limited resources.

If you pick the right people for your data science team right from the start, you'll be in a great position to keep everyone invested in the project and pulling in the same direction.



Part one:

**What skills are needed
by your team?**

Chapter 1: Define your goals

Start by sitting down with your C-suite and talking about what the organization actually hopes to achieve and what budget will be available to you to build your team.

This is important for three reasons:

1. It defines your priorities so that you can figure out which combination of skills you need to bring into your team – and by extension, who you should recruit into the role.
2. It allows you to clarify the scope of your project and the resources you will have access to. This means you can set and manage realistic expectations of what your team will be able to achieve, keeping everyone on the same page.
3. It helps you to decide which tasks will need to be assigned to dedicated team members and which you can tackle with the right technology, in order to allocate resources in the most effective way possible.

Chapter 2: Getting the right technical skills

Core technical skills

To start with, you need to make sure that either your team has technical skills in the following areas – or that you have access to technology that can automate and manage some of these processes.

- IT and system architecture
- Data preparation and exploration
- Data engineering
- Machine learning operationalization
- Data analytics governance

Data management

Data management is fundamental to effective machine learning projects. It's highly likely that your internal data is spread across the organization, scattered across different databases, data warehouses, lakes, and other repositories, source applications, and files of varying types.

Your team must be able to gather and access this data, build data pipelines that are robust and scalable, optimize quality, and catalog and manage both internal and external datasets.

Data science and machine learning capabilities

Your data science team needs the quantitative mathematical skills and statistical expertise to build, prototype, and deploy machine learning models. Doing this effectively is about much more crucial than straightforward coding or programming, though. Your team needs to be adept at reasoning, problem-solving, experimenting, and optimizing models.

Chapter 3: The non-technical skills you need

For your data science team to be a success, they can't only be technical whizzes. They also need first-rate people skills, especially when it comes to communication, teamwork, collaboration, project management, and leadership. Creativity, passion, and curiosity are key ingredients, too.

Understanding business vision, objective, and success criteria

Half the challenge for this is choosing the right data science projects in the first place – and that comes down to the acumen and judgment of your team. They need to be skilled at listening carefully to challenges (that are expressed in business rather than technical terms) and translating business stakeholder priorities into the most relevant models.

In-depth domain knowledge

To derive ROI on your data science projects, your team needs to understand the industry, business sector, or function that the insights will apply to. Ideally, you will have machine learning experts on the team that also understand the broader business strategy, but you can also access domain expertise by involving your business colleagues. We'll come back to this in the next section.

Storytelling

Communicating what you do to the wider organization and getting excited about the value you offer is the key to your long-term success and growth – but it's new territory for many data scientists and machine learning specialists. Try to identify potential team members who a) understand the link between science and business objectives, and b) can present inspiring use cases from industry peers (as well as hypothetical projects) to your colleagues.

Try to identify potential team members who a) understand the link between science and business objectives, and b) can present inspiring use cases from industry peers (as well as hypothetical projects) to your colleagues.

Leadership and guidance

It will be up to your team to keep up the momentum and enthusiasm for what you do. To galvanize the organization to push ahead on projects that will deliver ROI. To educate customers on how machine learning projects will add value. To communicate the progress of projects, pilots, and prototypes. Don't underestimate the skill that goes into this.



Part two:

Who do you need on your team?

Chapter 1: Building bridges is the key

In Part one, we looked at the key skillsets you need in your data science team. However, the single most important skill is the ability to forge connections, building bridges across different departments, units, and functions in the business.

The best way to achieve this is assigning people whose skillsets cut across these barriers as key team members. For example, you could:

- Appoint a **Chief AI Officer** who combines domain acumen with solid IT skills and data science/machine learning capability.
- Source **Data Engineers** who can use their domain knowledge to inform the IT architecture. An ideal data engineer will take primary responsibility for building and the underlying IT infrastructure you need for machine learning modeling and analysis, maintaining databases and data pipelines, and overseeing production processes.
- Make sure that you have team members who can act as **Business Translators**, converting domain knowledge into machine learning

- **Process Engineers** are ideally placed to bring together IT with a focus on machine learning. They understand how the infrastructure feeds into data science use cases.

Chapter 2: Machine learning and data science specialists

Of course, you don't have a data science team at all without people who can actually do data science and machine learning tasks!

If you are in a position to hire dedicated specialists for this, two key roles to consider are:

- **Data Scientists.** These are primarily responsible for the modeling and testing process.
- **Data and Machine Learning Analysts.** As your team grows, it's a good idea to consider hiring a data or machine learning analyst to evaluate data quality, monitor processes, and oversee the performance of production models. This allows your data scientists to focus on innovating rather than maintaining machine learning models.

However, if you don't have the resources to hire data scientists, don't despair. With the right tools and platforms, it's entirely possible to conduct sophisticated machine learning tasks without a dedicated AI expert. That includes building models or applying the most suitable algorithms and accessing ready-to-go datasets.

Chapter 3: Involve domain experts

Not all the domain knowledge that drives forward your machine learning projects can come from within your core data science team. It's also important that you collaborate with colleagues from the business side of the organization on an ongoing basis.

This will help you ensure that you are still on track to deliver the insights the company actually wants and needs. It also creates plenty of opportunities for colleagues to make suggestions or raise concerns about the datasets you are using, the assumptions you've made when designing the model, or the direction of the project overall.

Talk to business managers from relevant departments, be that marketing, sales, or operations. Consult with strategists that may draw on your work to drive the organization's top-level decisions and long-term plans. Ask for the input of business analysts who may interpret data differently, suggest connections between trends and datasets, or otherwise provide useful insight that guides your work.

Chapter 4: Building links with IT

For your data science team to be a success, you also need to work with IT to ensure the underlying systems and architecture are set up in the best way for your project. That includes your approach to data storage, access, and availability. It also includes privacy issues surrounding whether to store data in the cloud and hardware limitations when tapping into huge data warehouses run on legacy systems.

Even if they aren't dedicated entirely to machine learning and AI, make sure you collaborate closely with your company's system architects, application developers, and IT managers to ensure you're all on the same page and they understand what you need. Springing huge requests on IT when they're already snowed under will mean inevitable delays to your own projects, so keep the lines of communication open.

Final thoughts: easing the pressure

Once you've assembled your crack team, you need to give them the tools and the data they need to do their jobs effectively and efficiently. It's a bit of a waste of expertise to bring together all those top strategists and data science geniuses, only to have them spend half their time cleaning and pre-processing datasets for prototyping models. Far better to look for smart ways to automate the basics using streamlined tools and platforms, leaving them with the bandwidth for the real business-boosting work of data science.

Bear in mind, too, that companies tend to think they have far more useful data than they actually do. It's the curse of dimensionality: you may have millions of data points, but if they aren't the right data points for the insights you need, your team will be stuck at the first hurdle. It's highly likely that you will need to look to external data for critical insights and truly valuable models. Again, using a platform that facilitates connections to external datasets and automates much of the heavy lifting will save your team time better spent on high-level tasks.



About Explorium

Explorium offers a first of its kind data science platform powered by augmented data discovery and feature engineering. By automatically connecting to thousands of external data sources and leveraging machine learning to distill the most impactful signals, the Explorium platform empowers data scientists and business leaders to drive decision-making by eliminating the barrier to acquire the right data and enabling superior predictive power.

**For more information,
visit www.explorium.ai**