Data Team Organization

GitLab Data Team Organization

Data Team Organization

The Data Team Organization model is guided by three primary business needs:

1. The need for **bespoke data solutions** unique to the GitLab business.
2. The need for **high-performance and reliable data storage and compute** platform to support distributed analyst teams.
3. The need for centers of excellence for **data technologies** and **advanced analytics**.
4. The need for flexible data solutions driven by varying **urgency and quality** requirements.

Based on these needs, the Data Team is organized in the following way:

1. **Data Pods:** Pods are assembled to provide concentrated focus on delivering & maintaining **data products** for **strategic company initiatives**. Pods are staffed with multiple data personas including Data Analyst, Data Scientist, Analytics Engineer, and supported by Data Engineer as stable counterpart.
2. **Analytics Engineering:** Transform raw data into clean, structured, and usable formats for data decision-making. The Lead Analytics Engineer serves as a stable counterpart for business departments and functional analytics teams.
3. **Data Platform & Engineering Team: Center of Excellence** for data technologies, including owning and operating the Data Stack
4. **Data Science Team: Center of Excellence** for advanced analytics, including delivery of data science projects to the business

Data Pod Assignments

POD	Data Product Manager	Analytics Engineer	Data Analyst	Data Scient
Enterprise Metrics	@nmcavinue	@lisvinueza @chrissharp	@annie-analyst	
Customer Intelligence	@nmcavinue	@snalamaru	@jonglee1218	
Customer Product Adoption	@mdrussell	@michellecooper @utkarsh060		

Analytics Engineering - Business Stable Counterpart Assignments

Department	Functional Analytics Team	Analytics Engineer
Sales	Revenue Strategy and Analytics	@snalamaru
Marketing	Marketing Strategy and Analytics	@snalamaru
Finance	FP&A Analytics	@chrissharp
Customer Success	CS Strategy and Analytics	@mdrussell
Product	Product Data Insights	@michellecooper
Engineering	Engineering Analytics	@michellecooper
Security	Engineering Analytics	@michellecooper
Support	N/A	@michellecooper
People	People Analytics	@rakhireddy (ramping)

Data Platform Team Stable Counterpart Assignments

POD	Data Engineer
Enterprise Metrics	@juwong

POD	Data Engineer
Customer Intelligence	@rigerta
Customer Product Adoption	@rbacovic

Manager, Data

In support of the Data Pod, the Manager, Data fulfills the below responsibilities from the Senior Manager, Data Job Responsibilites:

1. Works with the Director, Data to envision and draft Quarterly Objectives, driven by requirements gathered from multiple business partners.
2. Monitor, measure, and improve key aspects of the Data Pods.
3. Regularly meet with business partners to it/data-team/stand and solve for data needs.
4. Serve as a primary or back-up Maintainer on the Data Team Project. Provide final review, feedback, and approval of Merge Requests submitted by the Data Pod and stable counterparts.

Lead Analytics Engineer (Stable Counterparts for the Business)

In support of the Data Pod and Stable Counterpart relationships, the Lead Analytics Engineer fulfills the below responsibilities from the Senior Analytics Engineer Job Responsibilites:

1. Own one or more stakeholder relationship in Go To Market, Research & Development, General & Administrative, Financial Analytics, or Engineering Analytics business functions.
2. Co-DRI of Key Results along with the Manager, Data.
3. Lead work breakdown sessions for OKRs.
4. Work with functional stakeholders to prioritze `P3-Other` issues.
5. Serve as a primary or back-up Maintainer on the Data Team Project. Provide final review, feedback, and approval of Merge Requests submitted by the Data Pod and stable counterparts.
6. Review the weekly stand-up and provide support as needed to unblock team members and answer questions.

Data Platform Team Stable Counterpart

Following the GitLab Stable Counterpart principles, every **Data Pod** have a **Data Platform Team** Stable Counterpart assigned. The Data Platform Stable Counterpart divides their time, work and priorities between the Data Platform Team and Data Pod (general an

average of 50% each, P2-OKR scheduled ahead of the quarter in collaboration with the respective Pod). The Stable Counterpart is aware of the direction and priorities of the Data Pod and when needed brought into discussion with the Data Platform Team. I.e. when there is a bigger demand than the Stable Counterpart can handle in the assigned availability or architectural direction needs to change. The Stable Counterpart recognize, flags and address this with the applicable stakeholders (in general the Lead/DRI of the Data Platform Team and the Data Pod).

The stable counterpart is expected to participate in the following meetings asynchronously or synchronously. When in doubt, please reach out to the Data Pod Manager to learn which meetings on the calendar you should participate in. In general, the meetings in scope are as follows:

1. Data Pod Iteration Planning Meetings.
2. Data Pod Team Meetings.

Data Program Recruiting

Recruiting great people is critical to our success and we've invested much effort into making the process efficient. Here are some reference materials we use:

- a GitLab Data Recruiting video to say "Hi" and give you some insight into how we work and what we work on. Enjoy!
- Data Roles and Career Development to help existing team members and prospects it/data-team/stand growth opportunities
- a Take Home Test that we ask each candidate to complete; this test is good for the candidate and for us because it represents the type of work we perform regularly and if the candidate is not interested in this work it helps them make a more informed decision about their application

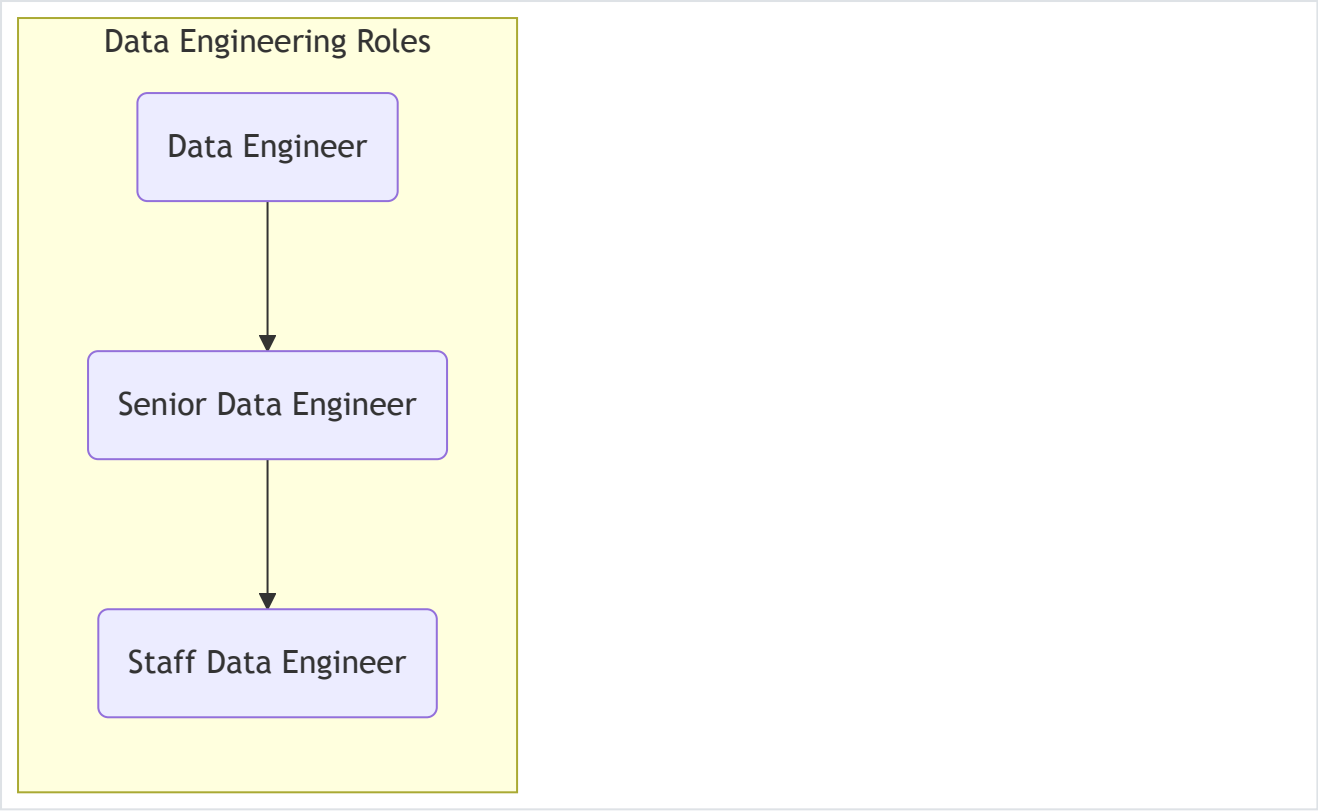
Data Roles and Career Development

Data Internships

See Data Team Internships.

Data Platform

- Data Engineering Job Family

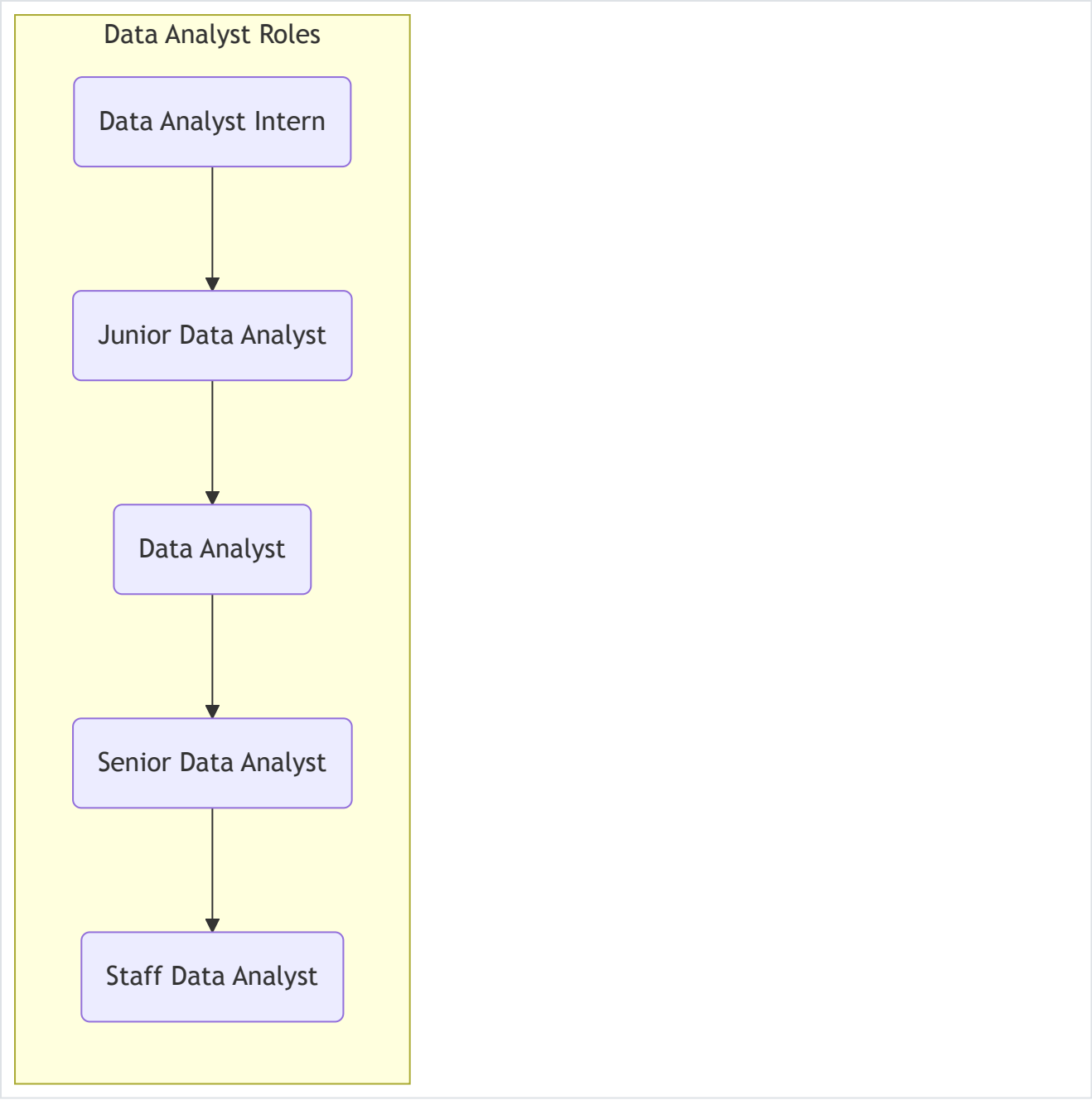


Intermediate and Senior Data Engineer Onboarding Timeline

By Day 30	By Day 60	By Day 90	By Day 120
Complete People and Data Onboarding	Perform triage activities	Extract new data sources	Own a specific area of the data platform
Create a MR to contribute to handbook or templates	Investigate incidents and issues	Work on OKR assignments	Propose new ideas and come up with Data Platform improvement initiatives
it/data-team/stand the current setup of the data platform	Make small/corrective changes to the platform infrastructure or data pipelines	Contribute on work breakdown	

Data Analyst

- Data Analyst Job Family



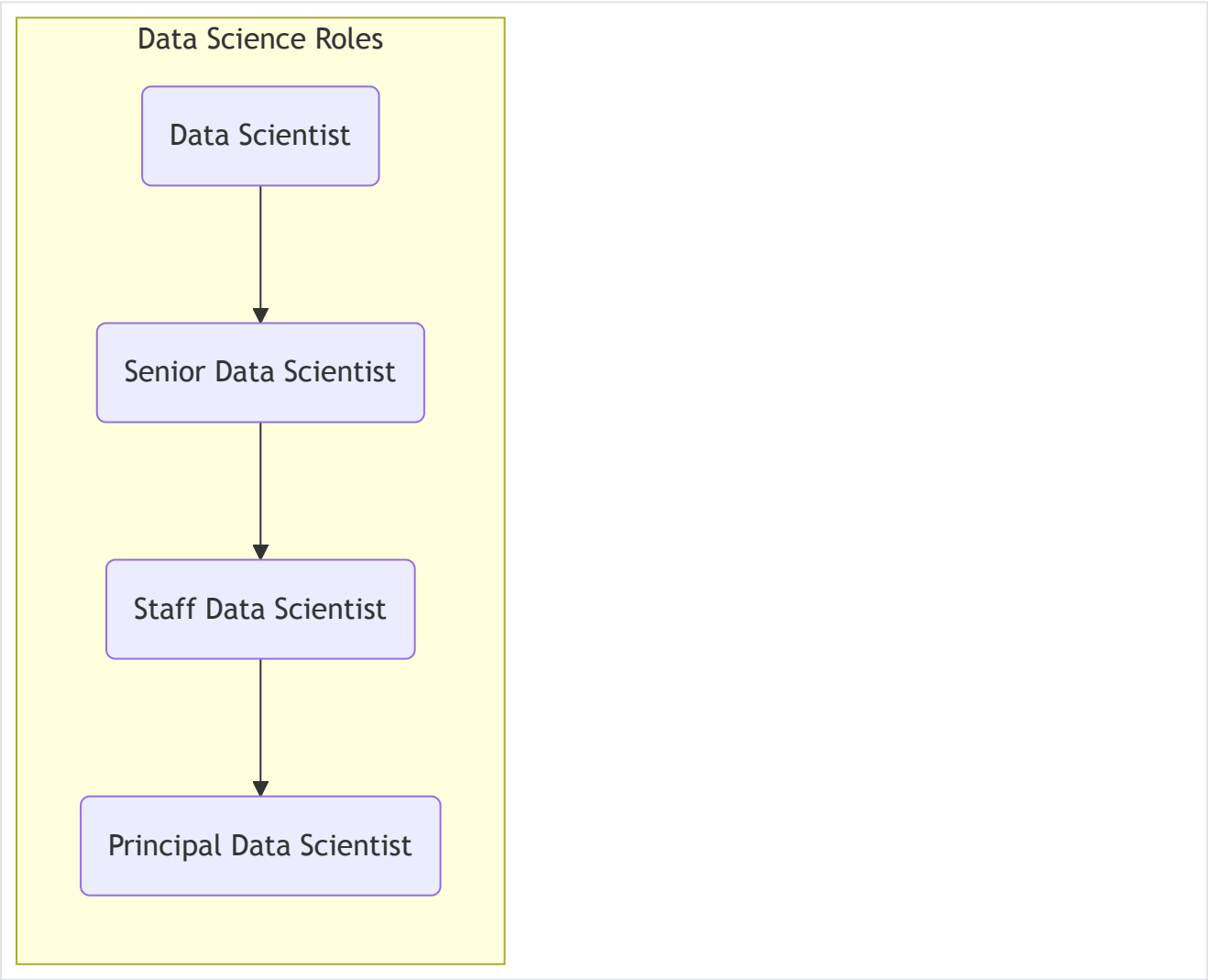
Intermediate and Senior Data Analyst Onboarding Timeline

By Day 30	By Day 60	By Day 90	By Day 120
Complete People and Data Onboarding	Extend an existing Tableau dashboard or complete the triage phase for a dbt issue	Run a project end-to-end as DRI with support from a Data Fusion Team	Create ERDs/Data Artifacts (e.g. dashboards) or complete a product evaluation

By Day 30	By Day 60	By Day 90	By Day 120
Start attending Data Fusion Team and Business Team synchronous meetings	Perform triage activities		
Complete First Issue: S to M T-Shirt Size			

Data Science

- Data Science Job Family



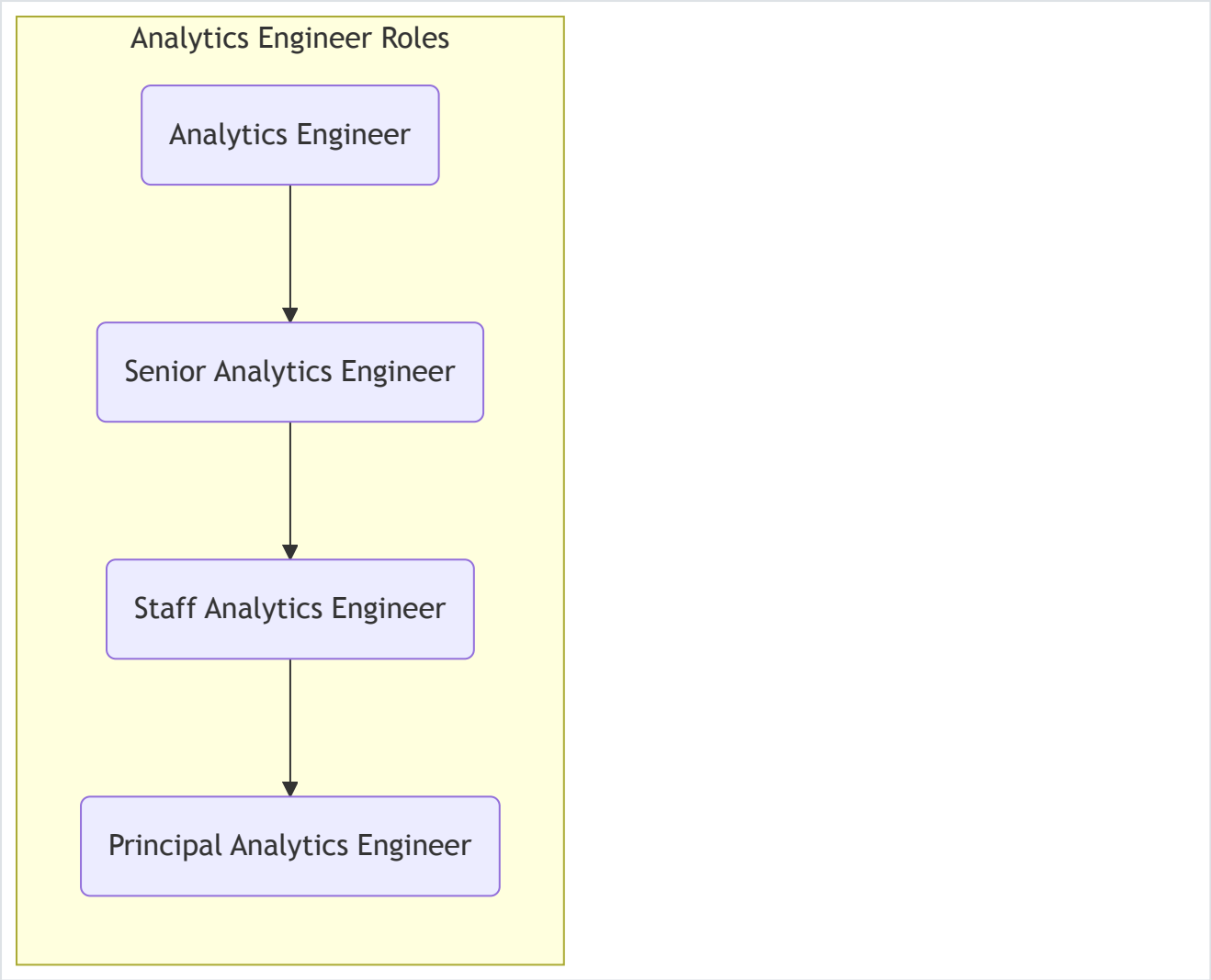
Intermediate and Senior Data Scientist Onboarding Timeline

By Day 30	By Day 60	By Day 90	By Day 120
Complete People and Data Onboarding	Meet stakeholders across the organization	Re-train or enhance an existing data science model	Make a contribution to improve the Data Science handbook, packages, or processes
Start attending Data Science Team meetings	Refine/improve one data science dashboard	Work on OKR assignments	Take ownership of at least one quarterly OKR
it/data-team/stand the current data science systems and processes			

Analytics Engineering

Analytics Engineering Job Family

Analytics Engineering Job Family

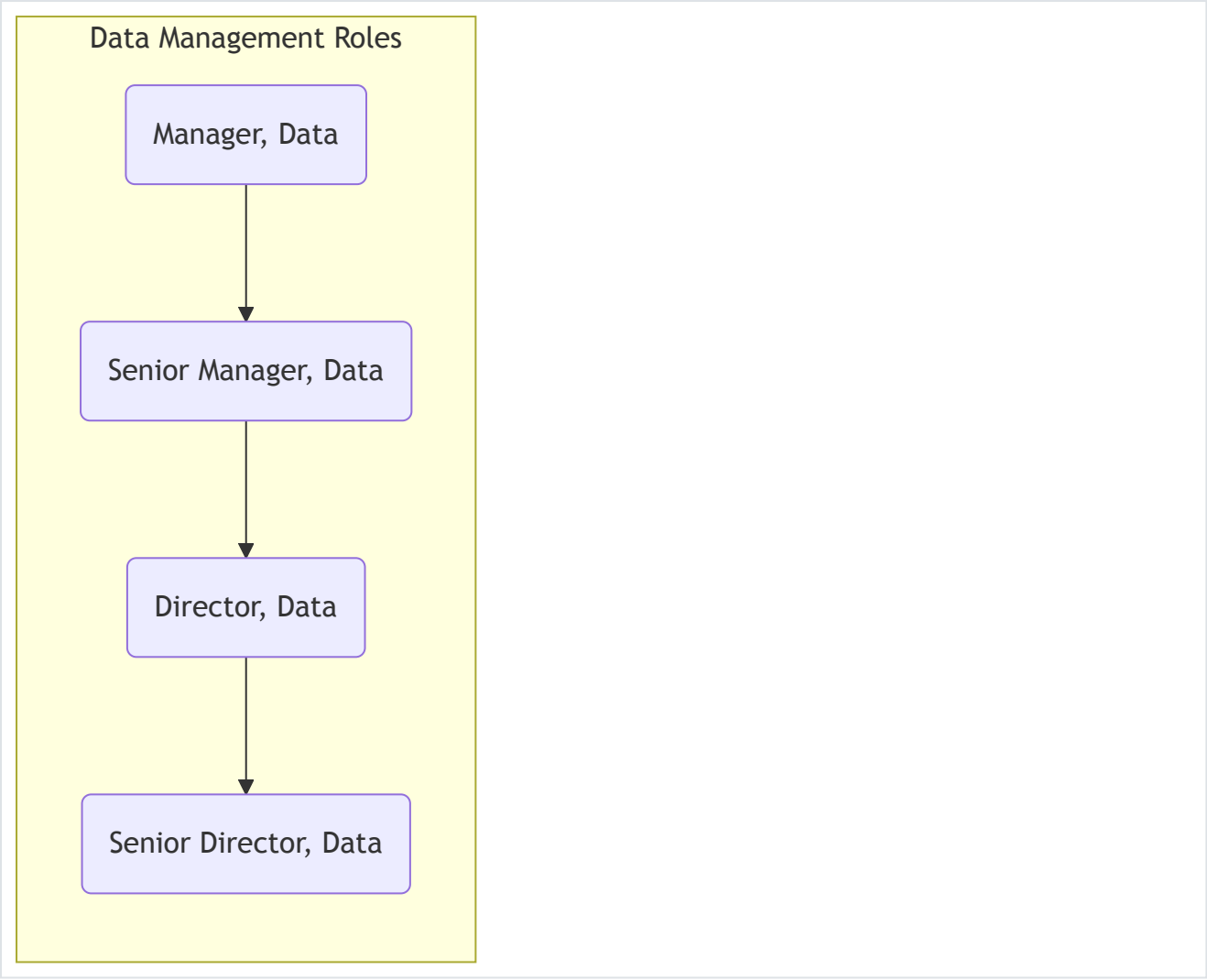


Intermediate and Senior Analytics Engineer Onboarding Timeline

By Day 30	By Day 60	By Day 90	By Day 120
Complete People and Data Onboarding	Extend an existing dbt Trusted Data Models	Run a project end-to-end as DRI with support from a Data Fusion Team	Create ERDs/Data Artifacts
Start attending Data Fusion Team and Business Team synchronous meetings	Perform triage activities		
Complete First Issue: S to M T-Shirt Size			

Data Management

- Data Management Job Family



Data Manager Onboarding Timeline

By Day 30	By Day 60	By Day 90	By Day 120
Complete People, Data, and Manager Onboarding	Meet everyone on the team and business data champions	Complete a Team Assessment	Draft a people development Roadmap
it/data-team/stand the current setup of the data platform	Work on OKR assignments and map them to the data platform	Lead discussions with Users/Stakeholders on initiatives and OKRs	Draft a program development Roadmap (Process Improvements /Future State)

By Day 30	By Day 60	By Day 90	By Day 120
Add a new page to the handbook	Make regular contributions to the handbook spanning your area of management	Become DRI for major portions of the Data Handbook	System/Application Change Control Management of one or more modules

Tool Technology Tandem

Tool Technology Tandems (TTT) are supporting to get the maximum value out of business opportunities we have in the Data Program. TTT are experts in a specific (software) tool or technology to support business opportunities or challenges we have by leveraging the tool or technology to the maximum. Although this is not the goal, we want to get the maximum value out of our technology stack. At the moment we see that we are not leveraging our technology stack to the maximum, where there are useful features or opportunities in our technology that could support in fulfilling business opportunities.

The reason is that from the technology side we don't know the business and from the business side we don't know the technology. The TTT will bridge this gap by it/data-team/standing the needs and bring this together in a technological way. We expect from TTT to do consulting, guiding and educating.

Note: TTT will **not** search for business opportunities to use any tool feature. TTT has to it/data-team/stand business opportunities and translate this into what software could bring to the table.

A single TTT consists of minimum 2 and maximum 3 GitLab Team Members with different roles. There are no requirements in which team a Team Member is part of(so this could be outside of the central Data Team as well) as long as the TTT meets the expectations described below.

Tool / Technology	Tandem
Snowflake	t.b.d.
Monte Carlo	t.b.d.
dbt	t.b.d.
Tableau	t.b.d.

What do we expect from TTT

- We expect TTT to get in touch with our business partners and all functions that contribute to the data program or work with our Data Platform, to it/data-team/stand their challenges.
- We expect TTT to get up to date with the latest in their area. They it/data-team/stand the full capabilities of the tool / technology, have regular touchpoints with the respective vendor and have a good it/data-team/standing of the latest released features.
- TTT will guide and educate our business partners.
- TTT will initiate design-spikes for quarterly planning.

Data Analytics at GitLab

GitLab Data Analytics Team Handbook

Data Platform at GitLab

GitLab Data Platform Team Handbook

Data Science Handbook

GitLab Data Science Team Handbook

Data Team Internships

GitLab Data Team Internships

Last modified August 21, 2024: Move the Data team handbook into /enterprise-data top level (`f156d4f3`)

View page source - Edit this page - please contribute.