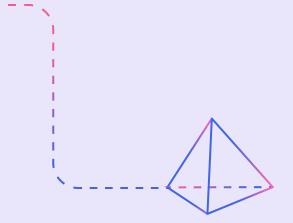


## From 1 to IPO: **Growing the Data Team and** Data Culture at GitLab/

Taylor A Murphy, PhD

Data Council - 2023-03-28

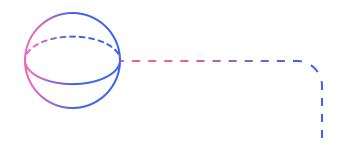




## Taylor A Murphy, PhD

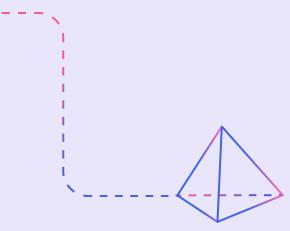
Current: Head of Product & Data / Founding Team @ Meltano Prev: Staff Data Engineer, Manager - Data & Analytics @ GitLab

Twitter: @tayloramurphy Email: taylor@meltano.com

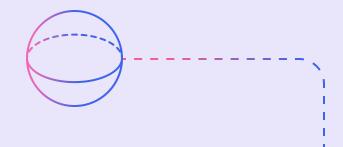




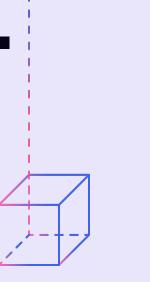


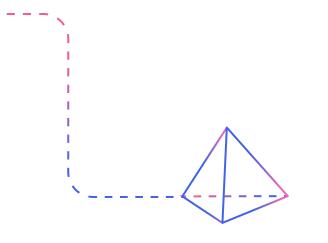


# My lessons learned from GitLab that are valuable regardless of the culture.

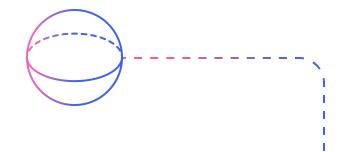




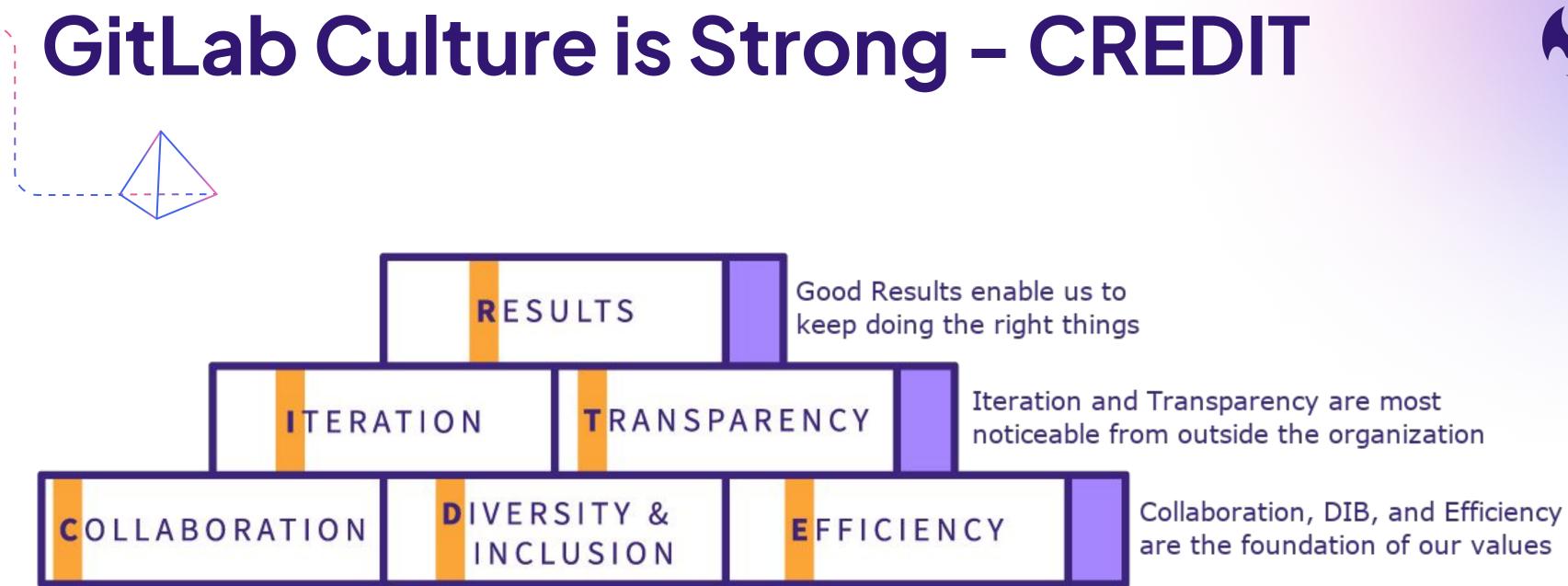


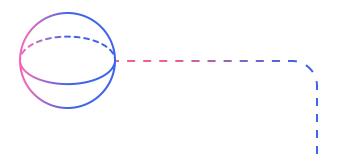
















### Public handbook

Public merge requests

Public livestreams

# GitLab



### **GitLab Unfiltered**

@GitLabUnfiltered 19.7K subscribers

Data Team

You are here: Business Technology > Data Team

Search through the handbook...

### Welcome to the Data Team Handbook

- Our Mission is to Deliver Results That Matter With Trusted and Scalable Data Solutions.
- Read our Direction page to learn what we are doing to improve data at GitLab.
- Our Principles inform how we accomplish our mission.
- Watch our <u>Data Recruiting Video</u> to learn about the growing Data Program.

Would you like to contribute? Become a Data Champion, recommend an improvement, visit Slack #data, watch a Data Team video. We want to hear from you!

### Maintained by



Contribute to this page <u>View source</u> • <u>Open in Web IDE</u>

### On this page

- Welcome to the Data Team Handbook
- Navigating The Data Team Handbook
- How Data Works at GitLab
   Data Program Teams
  - Data Program Collaboration Hub

  - How Data Teams Work Together
  - Analytics Engineering Team

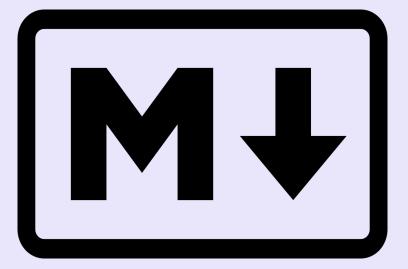


## GitLab Culture

Every employee knows how to open an issue, make a merge request, and edit YAML and markdown

# GitLab







## GitLab Culture

## Largest all remote company pre-COVID

Async first

Reinforced by CEO and execs

# GitLab



### GitLab Unfiltered

@GitLabUnfiltered 19.7K subscribers

Data Team

You are here: Business Technology > Data Team

Search through the handbook...

### Welcome to the Data Team Handbook

- Our Mission is to Deliver Results That Matter With Trusted and Scalable Data Solutions.
- Read our Direction page to learn what we are doing to improve data at GitLab.
- Our Principles inform how we accomplish our mission.
- Watch our <u>Data Recruiting Video</u> to learn about the growing Data Program.

Would you like to contribute? Become a Data Champion, recommend an improvement, visit Slack #data, watch a Data Team video. We want to hear from you!

### Maintained by



Contribute to this page <u>View source</u> • <u>Open in Web IDE</u>

### On this page

- Welcome to the Data Team Handbook
- Navigating The Data Team Handbook
- <u>How Data Works at GitLab</u>
   Data Program Teams
  - Data Program Collaboration Hub

  - How Data Teams Work Together
  - Analytics Engineering Team

## Lessons

01.

Leadership & Strategy Matter 02.

Transparency is Powerful





## 03.

## People and tools matter

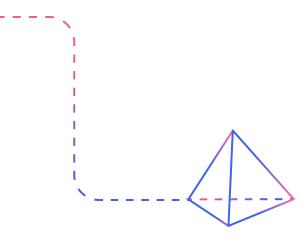
# Simple lessons that are easy to say but hard to practice.

## Each of these are essential, but they are not sufficient.



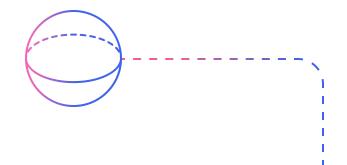






## Let's set the stage /It's 2018/

I'm sad to report that 2018 was, in fact, half a decade ago 💀







- DataEngConf, not Data Council
- dbt was in version 0.9.1
- Snowflake valued at \$1.5B
  - Bob Muglia still CEO
- Data Startup Landscape was very Different



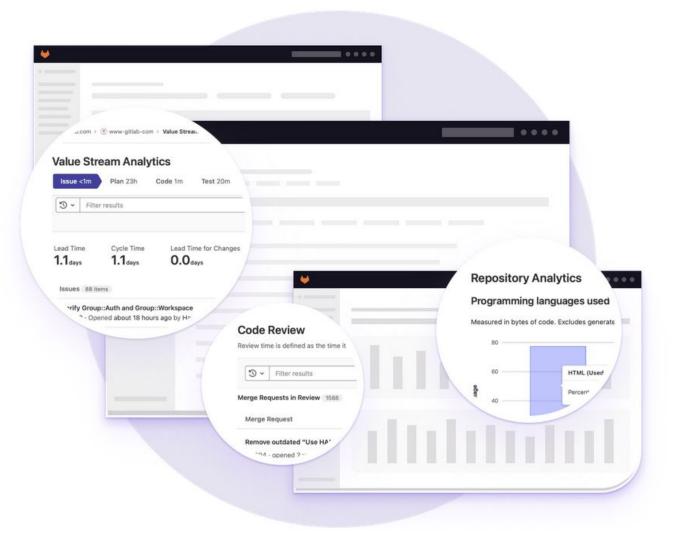
## snowflake



- Just raised Series C
- ~200 people
- All Remote

## GitLab

### **DevSecOps platform**

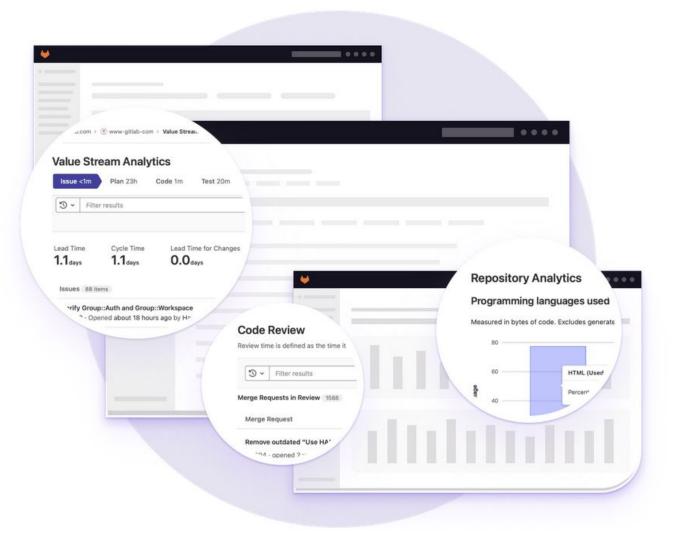


## **پ** 2021 Facts

- 6x headcount
- Data Team < 10
- IPO in late 2021

## GitLab

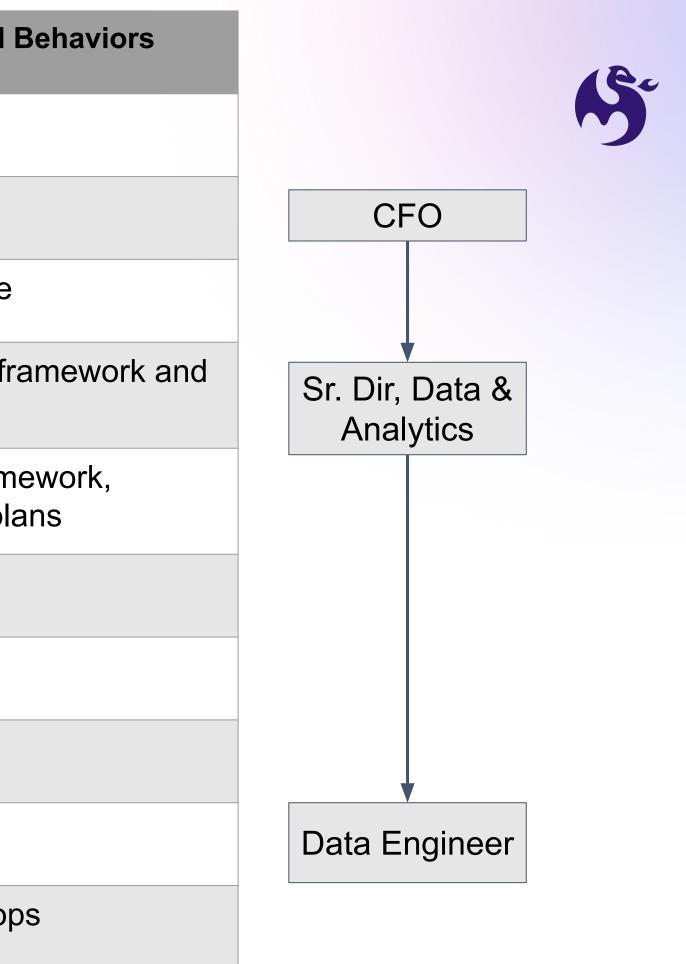
### **DevSecOps platform**



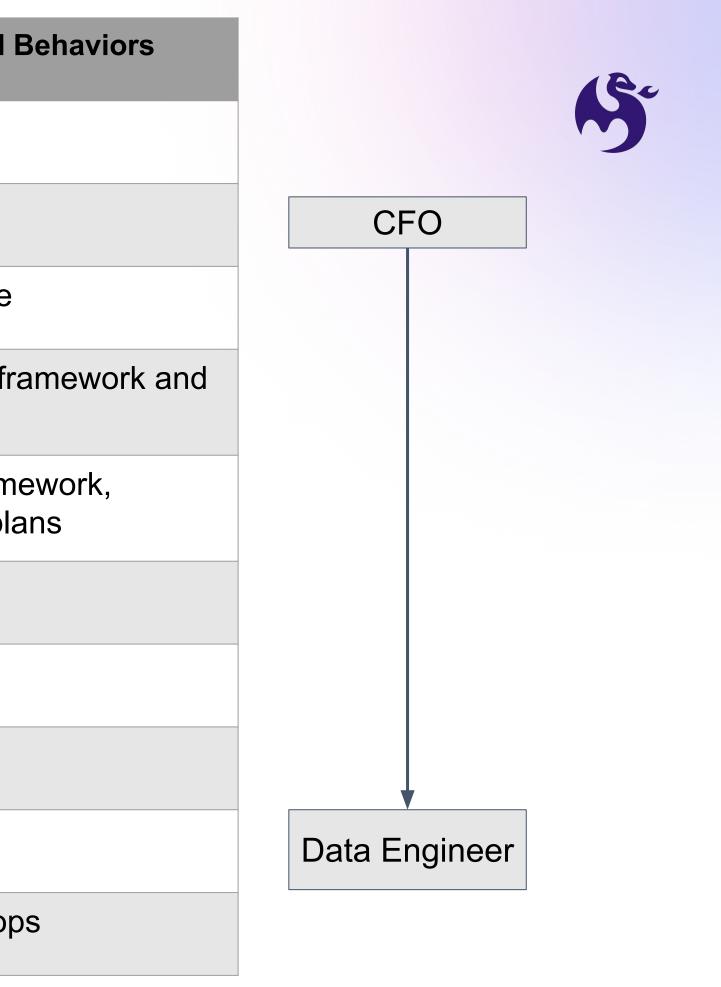
Level (People Manager/IC)	Example(s)	Scope of impact	Expected Behaviors
CEO	Chief Executive Officer	GitLab Global Organization	Champions
Executive	<u>Chief People Officer</u> and <u>Chief</u> <u>Technology Officer</u>	Division	Champions
VP/Fellow	VP of Global Channels	Department(s)	Drives Change
Senior Director	Senior Director, Engineering	Sub-department(s)	Develops the framework and strategy
Director/ Distinguished	Director of Customer Success Operations	Sub-department/multiple teams	Drives the framework, strategy and plans
Senior Manager/Principal	Principal Engineer	Across Sub-departments	Fosters
Manager/Staff	Engineering Manager	Across Teams	Implements
Senior	<u>Senior People Connect</u> <u>Specialist</u>	Cross functional work	Models
Intermediate	Intermediate Backend Engineer	Work within team	Grows/Acts
Associate	Business Development Associate	Own work	Learns/Develops

B

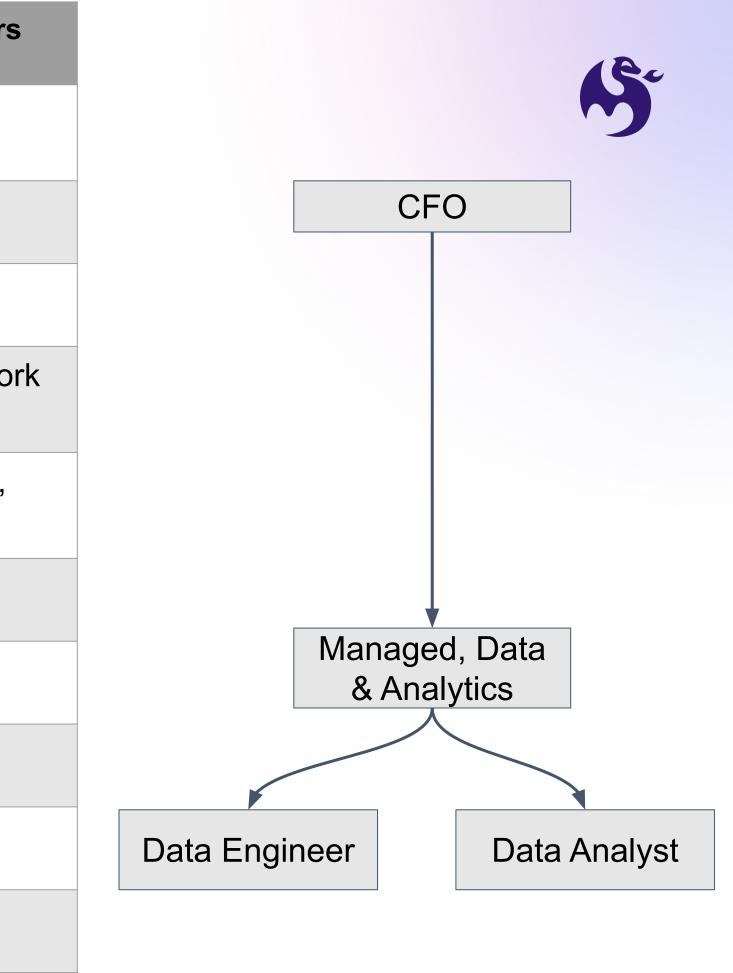
Level (People Manager/IC)	Scope of impact	Expected I
CEO	GitLab Global Organization	Champions
Executive	Division	Champions
VP/Fellow	Department(s)	Drives Change
Senior Director	Sub-department(s)	Develops the fra strategy
Director/ Distinguished	Sub-department/multiple teams	Drives the fram strategy and pla
Senior Manager/Principal	Across Sub-departments	Fosters
Manager/Staff	Across Teams	Implements
Senior	Cross functional work	Models
Intermediate	Work within team	Grows/Acts
Associate	Own work	Learns/Develop

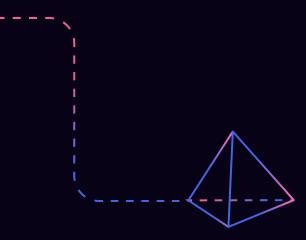


Level (People Manager/IC)	Scope of impact	Expected I
CEO	GitLab Global Organization	Champions
Executive	Division	Champions
VP/Fellow	Department(s)	Drives Change
Senior Director	Sub-department(s)	Develops the fra strategy
Director/ Distinguished	Sub-department/multiple teams	Drives the fram strategy and pla
Senior Manager/Principal	Across Sub-departments	Fosters
Manager/Staff	Across Teams	Implements
Senior	Cross functional work	Models
Intermediate	Work within team	Grows/Acts
Associate	Own work	Learns/Develop

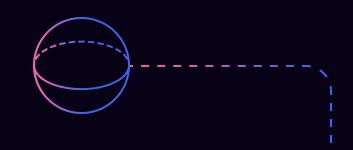


Level (People Manager/IC)	Scope of impact	Expected Behaviors
CEO	GitLab Global Organization	Champions
Executive	Division	Champions
VP/Fellow	Department(s)	Drives Change
Senior Director	Sub-department(s)	Develops the framewor and strategy
Director/ Distinguished	Sub-department/multiple teams	Drives the framework, strategy and plans
Senior Manager/Principal	Across Sub-departments	Fosters
Manager/Staff	Across Teams	Implements
Senior	Cross functional work	Models
Intermediate	Work within team	Grows/Acts
Associate	Own work	Learns/Develops

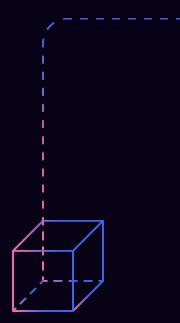


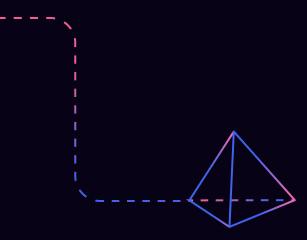


## The Director Gap was a /problem/

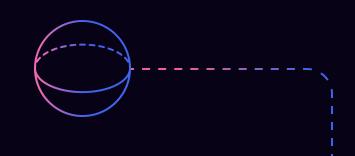








# So was reporting into / Finance/

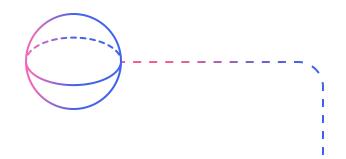




**Combination lead to suboptimal outcomes** 

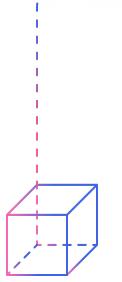
- 1. Lack of leadership (lack of power)
- 2. Lack of situational awareness
- **3. Under-serving Product, Marketing,** Sales, CS, Engineering... Everyone but

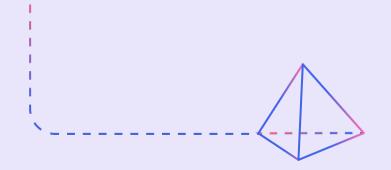
Finance







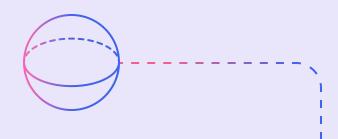




## Leadership & Strategy Stories

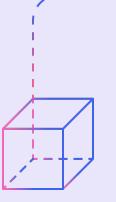


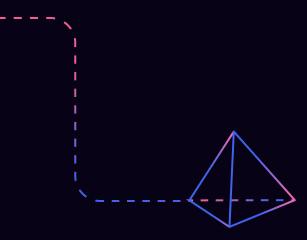
### Strategy without leadership



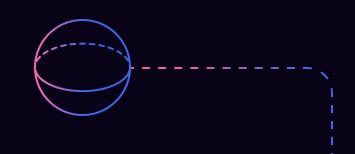


### Leadership without Strategy

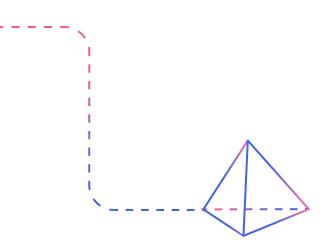




## Some Strategy No/Leadership/



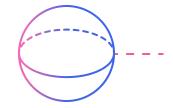




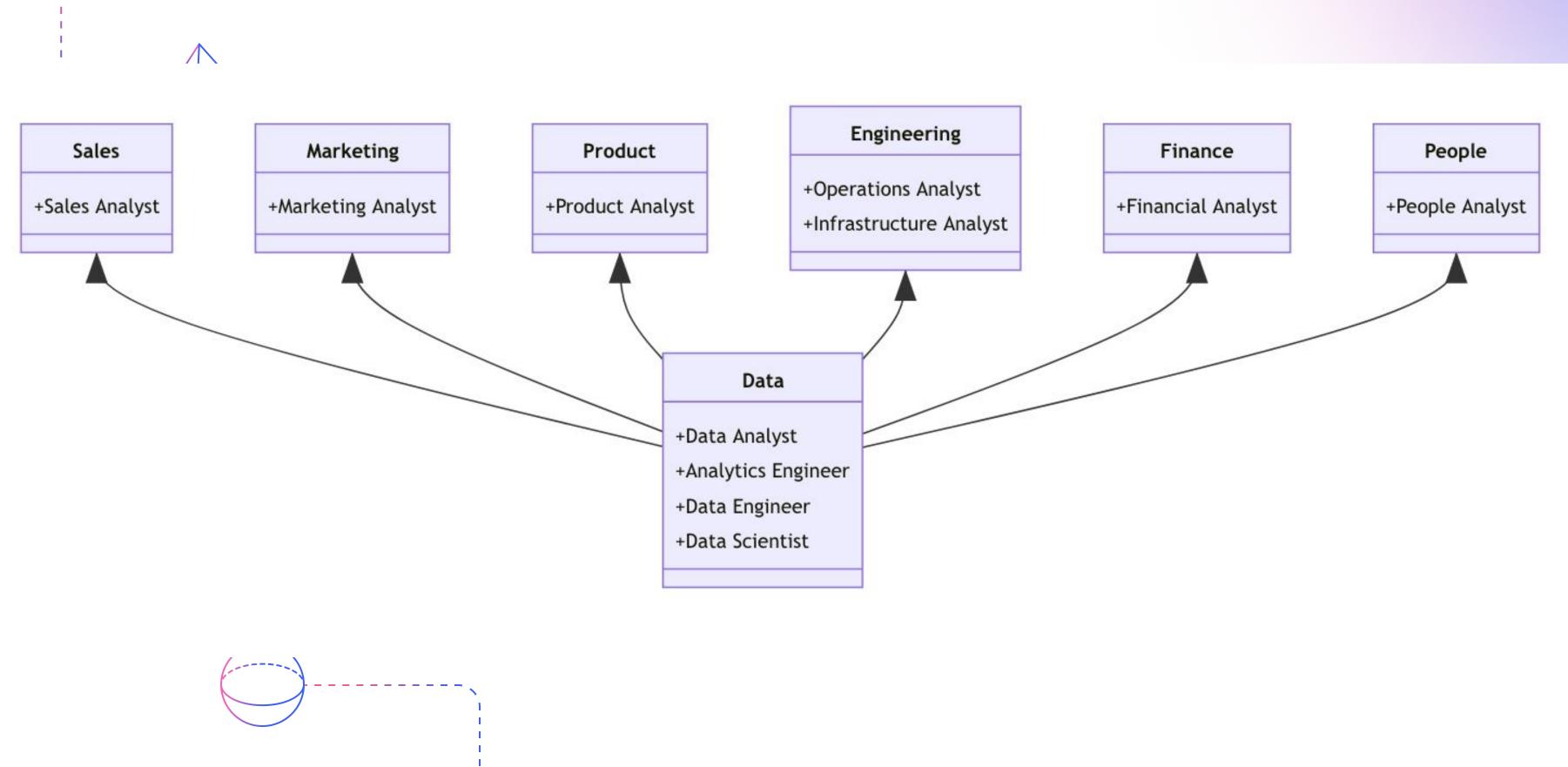
### **Data Program Teams**

The teams which compose the GitLab Data Program include:

- <u>Customer Success Operational Data Team</u>
- Data Team
- Engineering Analytics
- Finance Analytics & Insights
- <u>Marketing Strategy and Performance</u>
- Marketing Web Analytics
- People Analytics Team
- Product Analysis Group
- Product Intelligence Group
- Sales Analytics

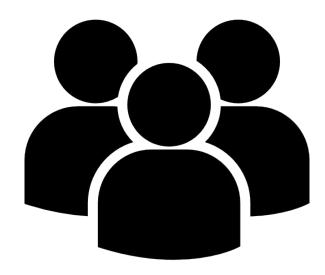












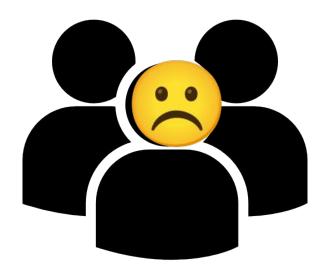
Marketing

- Solved their problems
- Marketing leader used it before
- Contract ready to go





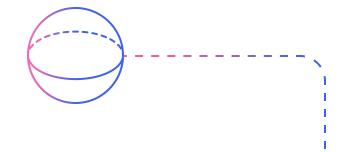




Data

Not open source

- Unclear data access and control
- Data team not consulted or informed



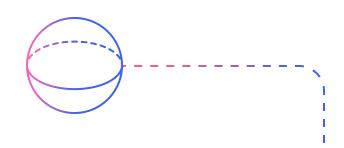






VS











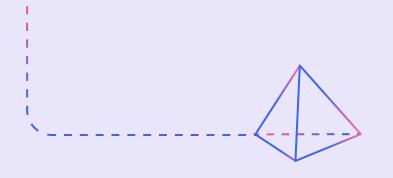


- Likely over-promised as a small data team
- Would better org visibility led to a different

outcome?

• Not sure we made a ton of friends



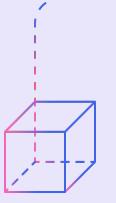


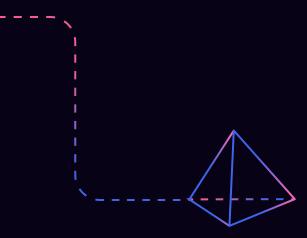
## 1. Hunger for data doesn't stop at your capabilities

## 2. Have representation at Director level and above

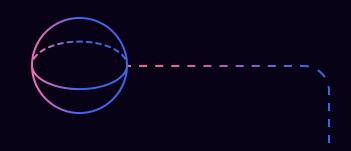
3. Don't report into Finance



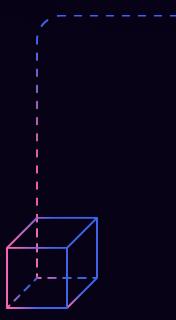


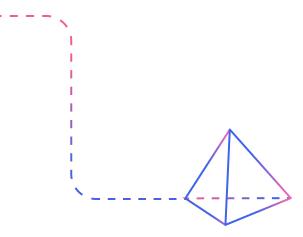


## Strong Leadership Minimal / Strategy /





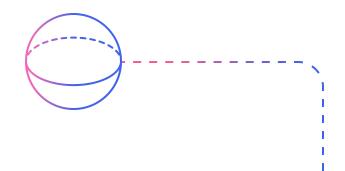




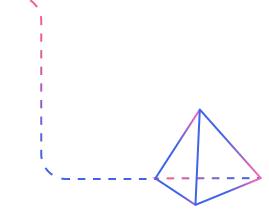
inside GitLab

## Hey, data teams - We're working on a tool just for you

Jacob Schatz and Taylor A. Murphy, PhD · Aug 1, 2018 · 5 min read · Leave a comment



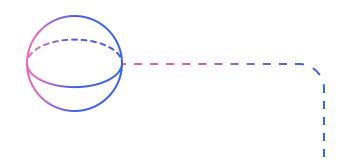




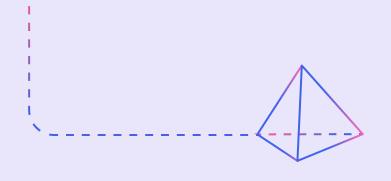
### Driven by CEO, relatively large investment

### Too focused inward on our own tooling

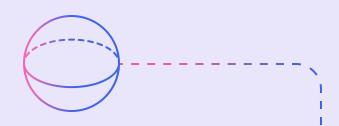
### Costing us time, therefore money



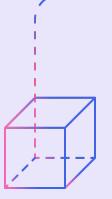


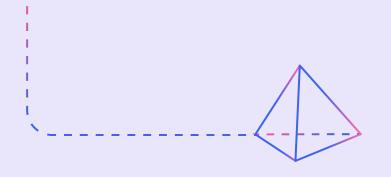


# A lacking broad data strategy was holding us back.





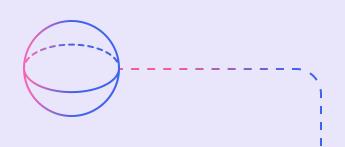




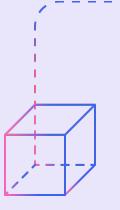
## 1. You have to know what you're trying to do

### 2. You have to focus

3. You have to say no to distractions

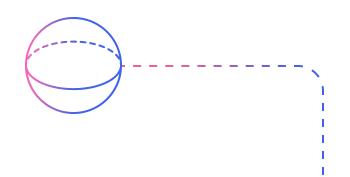








- Helps you have more context and influence
- Enables you to know what to focus on
- Should help you be more efficient





## Lessons

01.

Leadership & Strategy Matter 02.

Transparency is Powerful

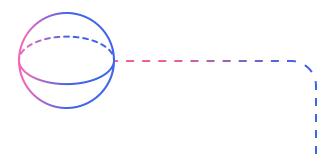




## 03.

## People and tools matter







Iteration and Transparency are most noticeable from outside the organization





## Public handbook

Public merge requests

Public livestreams

# GitLab



#### **GitLab Unfiltered**

@GitLabUnfiltered 19.7K subscribers

Data Team

You are here: Business Technology > Data Team

Search through the handbook...

#### Welcome to the Data Team Handbook

- Our Mission is to Deliver Results That Matter With Trusted and Scalable Data Solutions.
- Read our Direction page to learn what we are doing to improve data at GitLab.
- Our Principles inform how we accomplish our mission.
- Watch our <u>Data Recruiting Video</u> to learn about the growing Data Program.

Would you like to contribute? Become a Data Champion, recommend an improvement, visit Slack #data, watch a Data Team video. We want to hear from you!

#### Maintained by



Contribute to this page <u>View source</u> • <u>Open in Web IDE</u>

#### On this page

- Welcome to the Data Team Handbook
- Navigating The Data Team Handbook
- How Data Works at GitLab
   Data Program Teams
  - Data Program Collaboration Hub

  - How Data Teams Work Together
  - Analytics Engineering Team



## Transparency

## Solve Problems

## **Distributes Power**

## Is Efficient



You are here: Business Technology > Data Team

Search through the handbook...

#### Welcome to the Data Team Handbook

- Read our Direction page to learn what we are doing to improve data at GitLab.
- Our Principles inform how we accomplish our mission.
- Watch our <u>Data Recruiting Video</u> to learn about the growing Data Program.

Would you like to contribute? Become a Data Champion, recommend an improvement, visit Slack #data, watch a Data Team video. We want to hear from you!

#### Data Team

#### • Our Mission is to Deliver Results That Matter With Trusted and Scalable Data Solutions.

#### Maintained by



Contribute to this page View source • Open in Web IDE

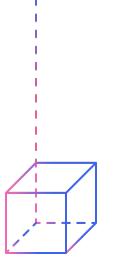
#### On this page

- Welcome to the Data Team Handbook
- <u>Navigating The Data Team Handbook</u>
- How Data Works at GitLab
  - Data Program Teams
  - Data Program Collaboration Hub
  - How Data Teams Work Together
  - Analytics Engineering Team

## **Transparency solves problems**

- 1. What was the definition of that metric?
  - a. It's in the handbook
- 2. What happened in that meeting?
  - a. There was an agenda and it was recorded
- 3. What's our policy on X?
  - a. It's in the handbook





#### Add clarifying language to the gross and net retention

Merged Taylor A Murphy requested to merge tayloramurphy-mast	ter-patch…	រៀ into m	aster 3 y	/ears	ago
erview 9 Commits 3 Pipelines 3 Changes 1					
2 👎 0 😨					
Merge request pipeline #83800807 passed for 1bebf8ef 3 year	ars ago		$\bigcirc$		•
<ul> <li>Approval is optional</li> </ul>					
View eligible approvers					
Merged by 鶲 Paul Machle 3 years ago					
<ul> <li>Merge details</li> <li>Changes merged into master with <u>leb6f9c2</u>.</li> <li>Deleted the source branch.</li> </ul>					
Pipeline #83859900 failed for 1eb6f9c2 on master 3 years ago	D		$\odot$		»
tivity			Sort	or filte	er ~
Taylor A Murphy @tayloramurphy · 3 years ago         FYI @broyer1 @wwright @iweeks	Author Co	ntributor	©	Ø	•
Paul Machle @pmachle · 3 years ago	Co	ntributor	<b>v</b>	$\bigtriangledown$	:
Resolved 3 years ago by Paul Machle					
<b>@tayloramurphy</b> i dont believe this is an accurate calculation gross retention should be 86% using this formula ((50+100)/(5		ention. in t	this exam	nple	

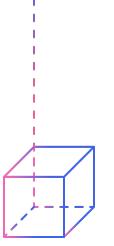
⊌ GitLab.com >	www-gitlab-com > Merge req					
Add cla	rifying language					
so Merged	aylor A Murphy requested to n					
Overview 9	Commits 3 Pipelines					
Le Compare	master ~ and latest version					
<ul> <li>source/handbook/customer-success/</li> </ul>						
<u>↑</u>	00 -38,11 +38,40 00 We m					
38 38	* Retention, Gross & Ne					
39 39	We measure Net and Gros					
	/#customers).					
40 40						
41	- Gross Retention (%)					
42	- Net Retention (%) =					
41	+ For an individual c					
43 42						
44	- <b>**A**.</b> MRR 12 month					
45	- <b>**B**</b> . Current MRR					
43	+ Gross Retention (%)					
44	+ Net Retention (%) =					
45						
46	+ <b>**A**</b> = MRR from 12					
47	+ <b>**B**</b> = Current MRF					
48	+ <b>**C**</b> = Gross retai					
49						
50	+ Individual customer					
	retention across all cu					
51	+ Because customers h					
52						
53	+ For all customers,					
	calculate as follows:					
54	+					
55	+ Gross Retention (%)					
56	+ Net Retention (%) =					
57	+					
58	+ **A** = MRR from 12					
59	+ <b>**B**</b> = Current MRF					
60	+					
61						
62	+ + Example:					
63	+ Example:					
64	+ There are two custo					
65 66	+ 12 months ago, X ha					
60	+ Gross retention for					
68	+ Net retention for )					
69						
09						

```
uests > !30499
e to the gross and net retention
                                                                                      :
                                                                   Edit
                                                                          Code 🗸
nerge tayloramurphy-master-patch... [<sup>o</sup>] into master 3 years ago
3 Changes 1
                                                   All threads resolved!
                                                                             Add a to do
                                                                                  Ô
                                                                1 file +33 -4
 V
/vision/index.html.md
                                                              +33 -4 🗍 Viewed
et (Dollar Weighted)
ss Retention aggregated by month, for the [three levels of customers](/handbook/sales
 = (min(B, A) / A) 🗴 100%<b
 (B / A) 🗶 100%
customer:
hs ago from currently active customers<t
from the same set of customers as A.
) = C / A * 100%<br>
= B / A \star 100%
 months ago from active customer<br>
R from the same customer in A<br>
ined dollars calculated as min(B, A)
r retention calculations cannot be averaged together directly to determine the
ustomers.
have different values, the retention percentages don't represent the same magnitude.
first calculate gross retained dollars for each individual customer and then
) = sum(C) / sum(A) * 100%<br>
= sum(B) / sum(A) \star 100%
 months ago from all active customers<br>
R from the same set of customers in A<br>
ined dollars for each customer (see individual example)
omers (X and Y) who each have $100 in MRR in the current month (B).<br/>
ad $50 in MRR and Y had $125 in MRR (A).
r X is (min(100, 50) / 50) * 100% = 100%<br>
K is (100 / 50) <mark>*</mark> 100% = 200%
```

## **Transparency distributes power**

- 1. Extreme visibility on policy and definition changes
  - a. Visible merge request
- 2. Visible roadmaps and backlogs
  - a. Made priority conversations easier (not easy)
- 3. If you could see the code, you could make proposals





### Update parental leave policy

Emilie Schario requested to merge update-parental-leave 1 into master 3 years ago So Merged

**Overview** 25 Commits 5 Pipelines 3 Changes 1

#### This MR

This MR:

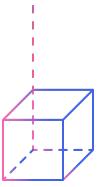
- updates the parental leave policy to be 16 weeks
- encourages parents to take "the time they need"
- does not affect the additional 4 weeks unpaid offer

#### The Business Case for 16 week Parental Leave

- 1. By increasing paid parental leave at GitLab to 16 weeks, the company sets itself up to be in the top 15% of employers offering this benefit, a huge competitive advantage in this hiring landscape. The result: improved employer branding.
- 2. By increasing paid parental leave at GitLab to 16 weeks, parents, especially women, taking the leave are more likely to be retained long-term. The result: improved employer retention, leading to decreased hiring costs and increased diversity.
  - Women who take paid maternity leave are more likely to be in the workforce a year later (Source).
  - Men take paternity leave at twice the rate and for longer periods of time when the leave is paid (Source).

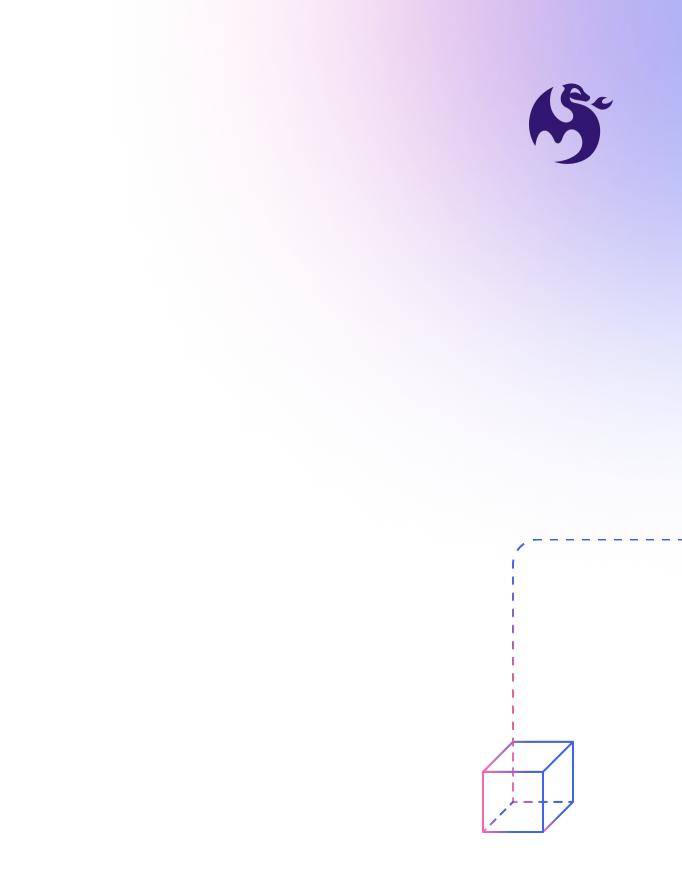


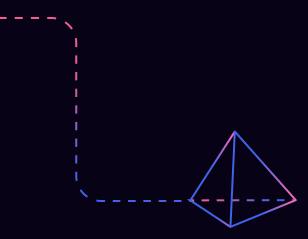
1 unres



## Transparency is efficient

- 1. Invites people into the process
  - a. Increases collaboration
- 2. Makes changes easier
  - a. Diffs are lovely
- 3. Enables async communication
  - a. Not blocked by others

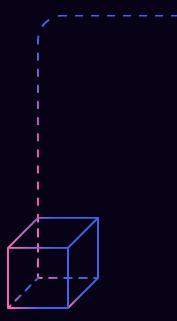


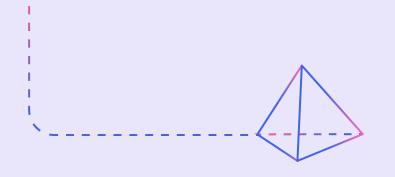


# How did this manifest in the /Data Team/?





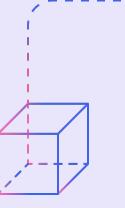


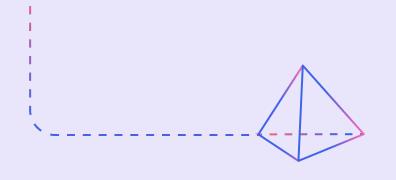


## I want to version control all the things.



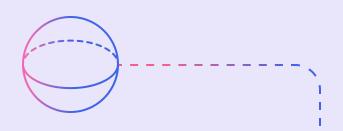




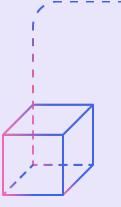


## I want to version control all the things.

# I want ever change to happen in via a pull request.

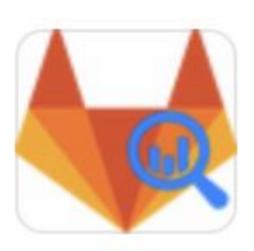






## **Change Control ALL the things**

GitLab Data > 🔰 GitLab Data Team



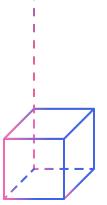
## GitLab Data Team

Project ID: 4409640 Request Access

This is the primary project for the GitLab Data team.







## Data Team Handbook



## Data Team

You are here: Business Technology > Data Team

Search through the handbook...

#### Welcome to the Data Team Handbook

- Our Mission is to Deliver Results That Matter With Trusted and Scalable Data Solutions.
- Read our <u>Direction</u> page to learn *what* we are doing to improve data at GitLab.
- Our <u>Principles</u> inform how we accomplish our mission.
- Watch our <u>Data Recruiting Video</u> to learn about the growing Data Program.

Would you like to contribute? Become a Data Champion, recommend an improvement, visit Slack #data, watch a Data Team video. We want to hear from you!











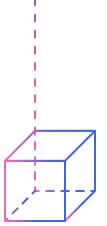
#### Maintained by



- Contribute to this page
- View source Open in Web IDE

#### On this page

- Welcome to the Data Team Handbook
- Navigating The Data Team Handbook
- How Data Works at GitLab
  - Data Program Teams
  - Data Program Collaboration Hub
  - How Data Teams Work Together
  - Analytics Engineering Team



# **Change Control ALL the things Meltano**

# Löcker dbt

😢 GitLab Data > Permifrost

postgres\_pipline (pgp)





# =====
# CI Stages
# =====
stages:
- 🗱 Snowflake
- 🊈 Extract 🛛 # extract,
- 🕸 dbt Run 🛛 # transfo
- ゲ dbt Misc 🛛 # transfo
- 📚 dbt Docs
- 🛑🖢 Python Critical
- 迄 Python
- 🛑 Snowflake Stop
- triage
- triage run

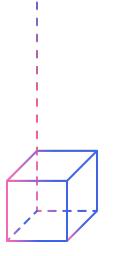




## **Story Time**







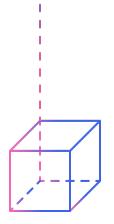
## **Story Time**





# PART OF Google Cloud



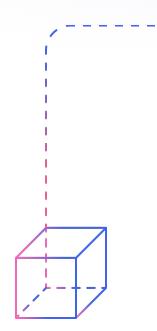


## **Story Time**



# **Periscope** Data







## **Transparency is Powerful**

## Solve Problems

## **Distributes Power**

## Is Efficient



You are here: Business Technology > Data Team

Search through the handbook...

#### Welcome to the Data Team Handbook

- Read our <u>Direction</u> page to learn what we are doing to improve data at GitLab.
- Our Principles inform how we accomplish our mission.
- Watch our <u>Data Recruiting Video</u> to learn about the growing Data Program.

Would you like to contribute? Become a Data Champion, recommend an improvement, visit Slack #data, watch a Data Team video. We want to hear from you!

#### Data Team

#### • Our Mission is to Deliver Results That Matter With Trusted and Scalable Data Solutions.

#### Maintained by



Contribute to this page View source • Open in Web IDE

#### On this page

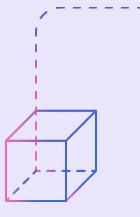
- Welcome to the Data Team Handbook
- Navigating The Data Team Handbook
- How Data Works at GitLab
  - Data Program Teams
  - Data Program Collaboration Hub
  - How Data Teams Work Together
  - Analytics Engineering Team

1. Commitment to transparency requires discipline, but it is so valuable

2. It forces conversations and shares power

3. Power structures still exist (strategy and leadership still matter)





## Lessons

01.

Leadership & Strategy Matter 02.

Transparency is Powerful

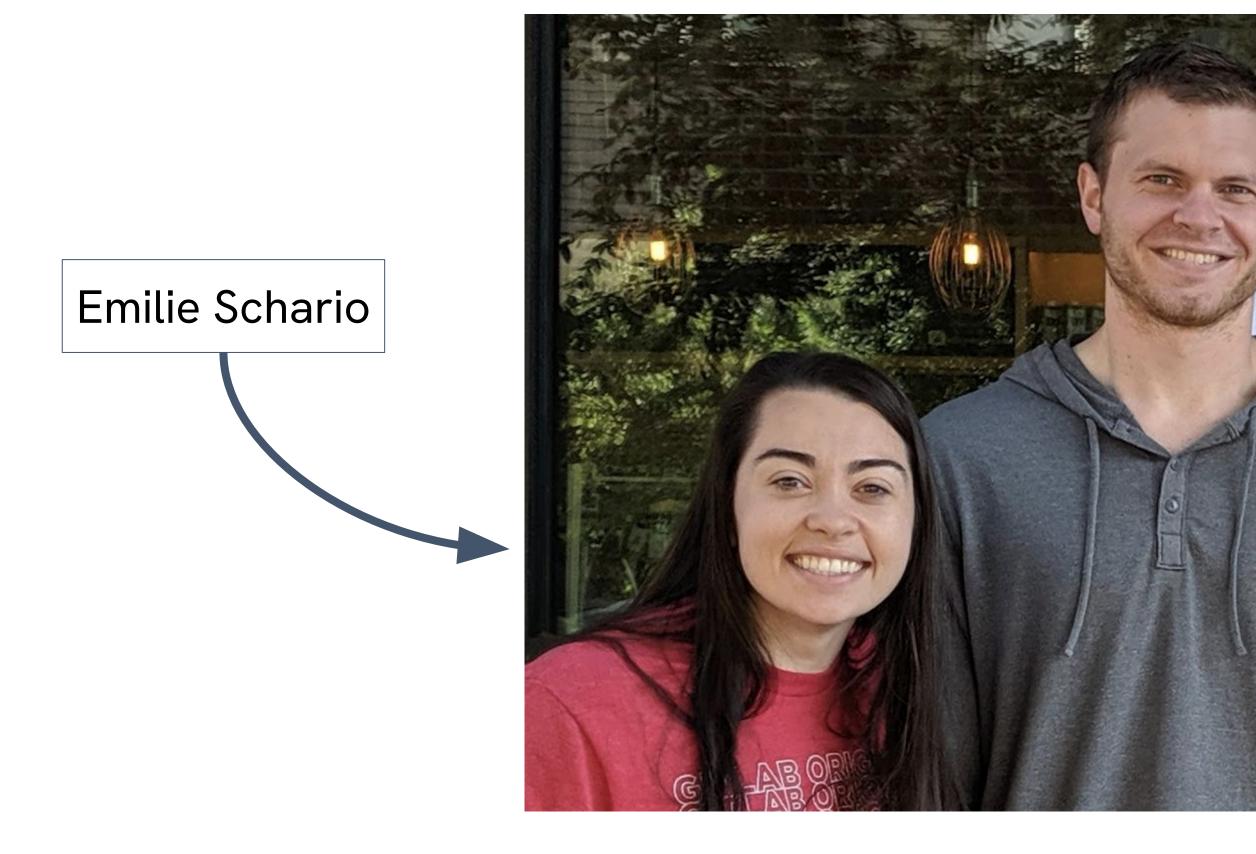




## 03.

## People and tools matter

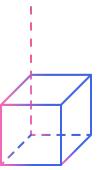
## **Best Thing | Did as a Manager**







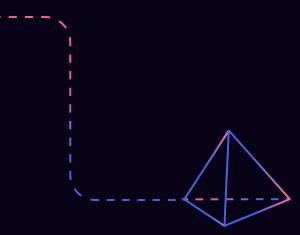
### Thomas La Piana



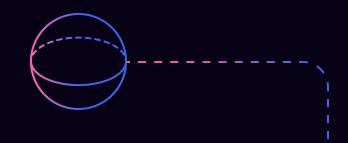
## **Hire Great People**

- Candothejob
- Can talk to other people
- Match the stage of company
- Eager and want to learn/grow



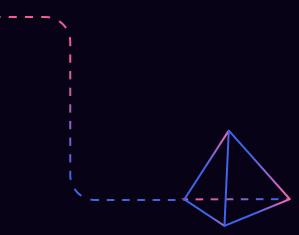


## /People/reinforce the /culture/ in every moment

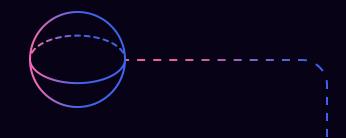








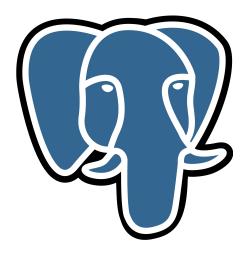
## /Tools/ reinforce your /people/ every time they're used

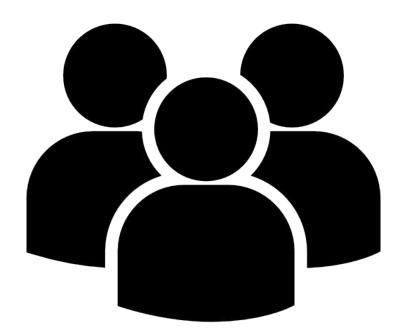




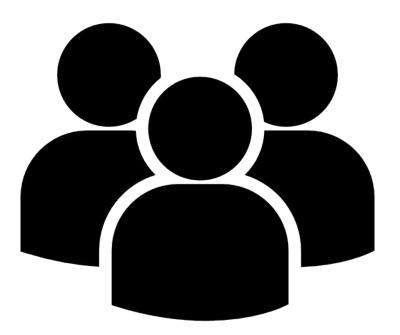


## **Story Time - MRR**







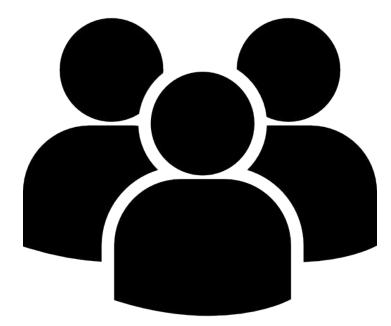


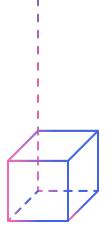
Engineering

Sales



# ZUOro





#### Finance

# Story Time – MRR

Customers DB tied out actual license key

Defined Account Hierarchy and how everything rolls up

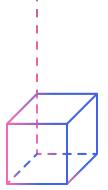
Engineering

Sales

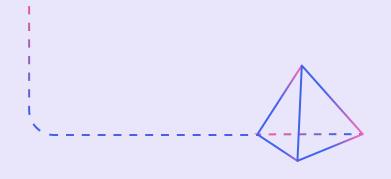


# ZUOro

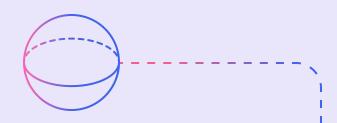
### Subscription and SKU data over time



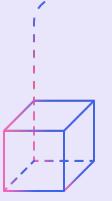
Finance

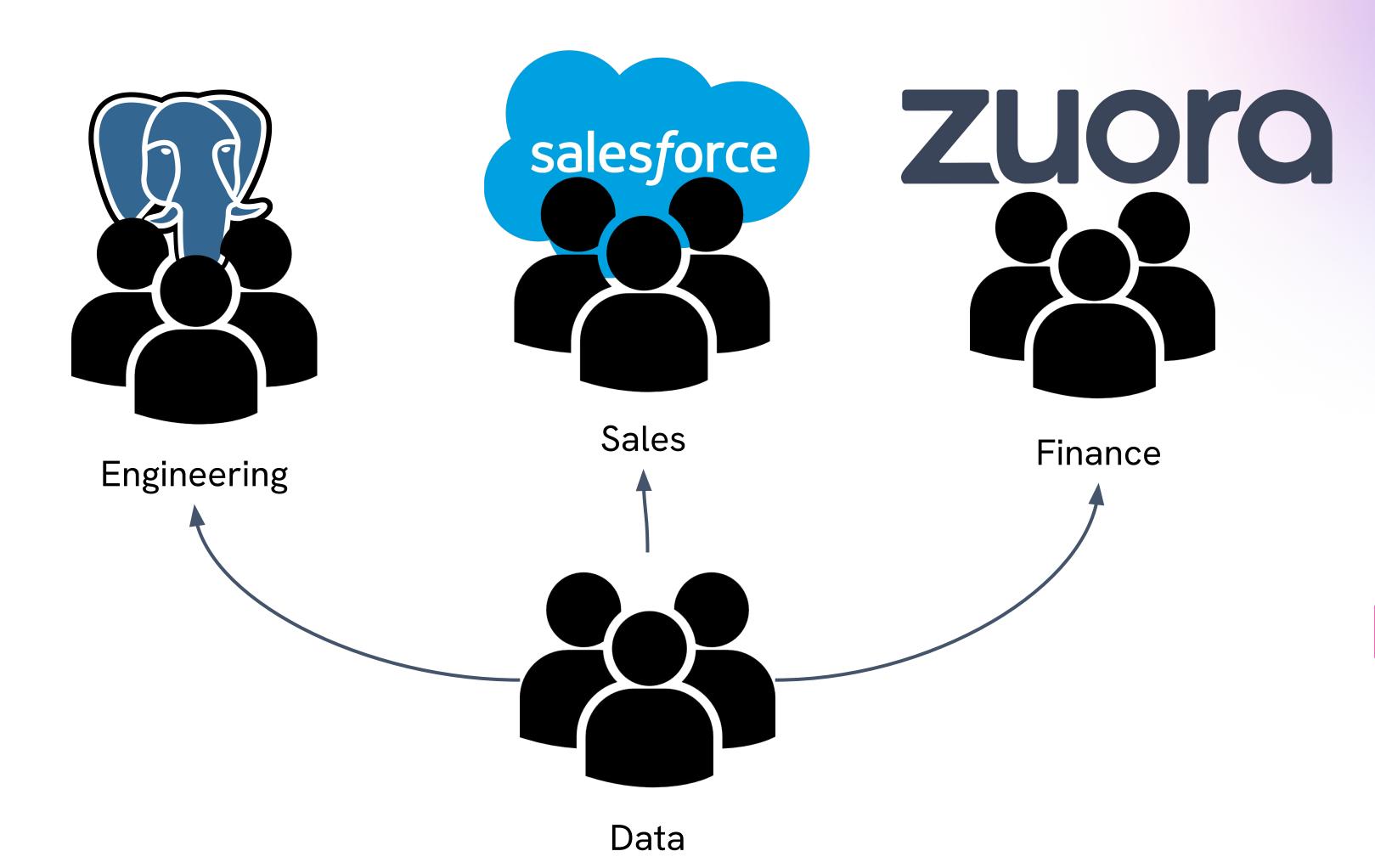


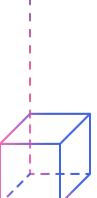
## This is where a data team shines.







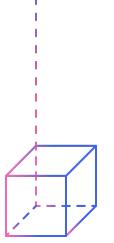


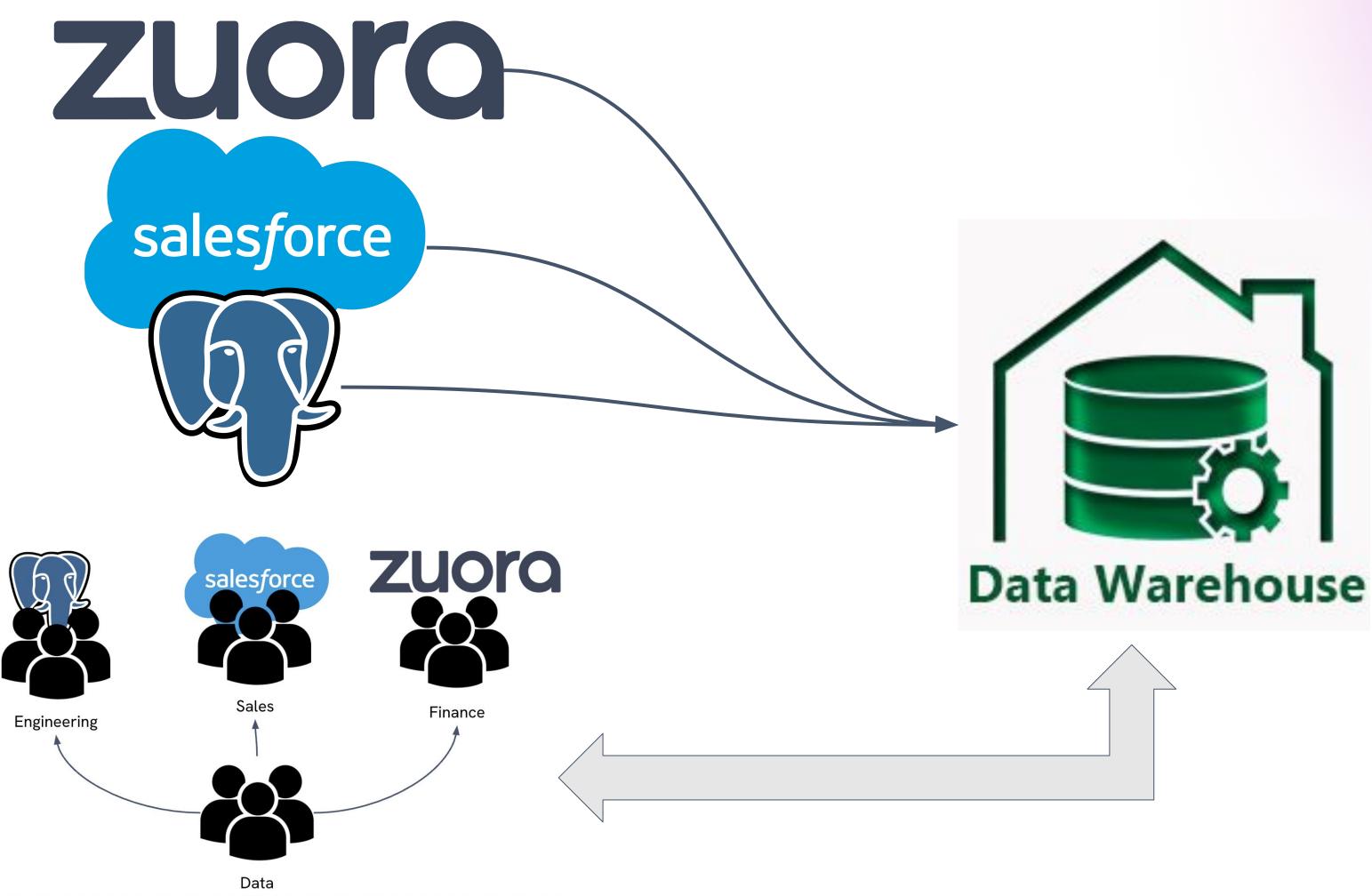




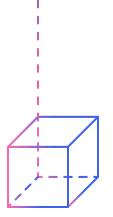




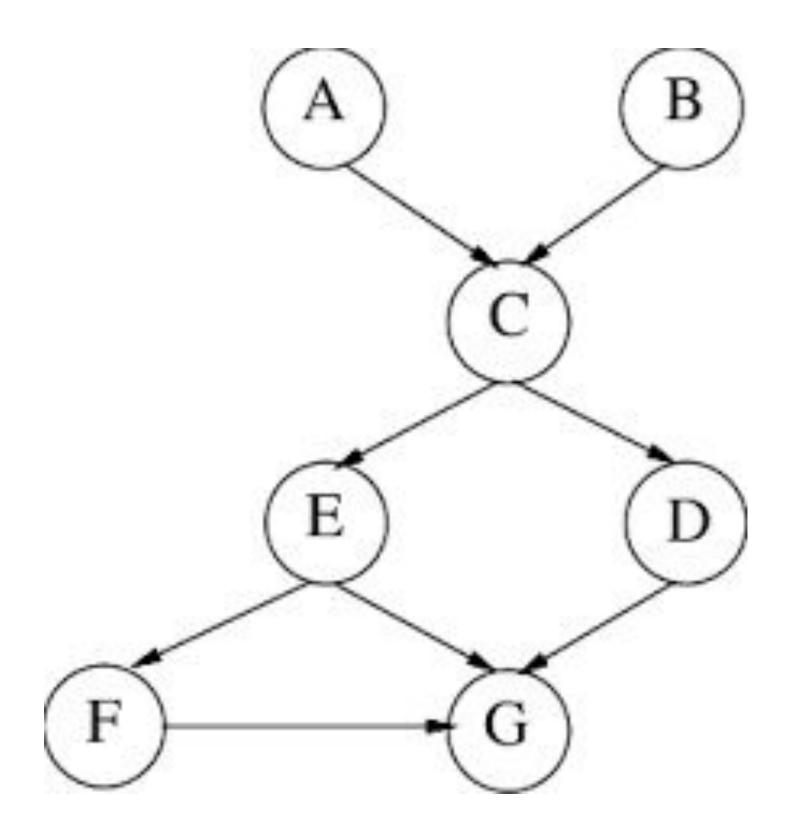


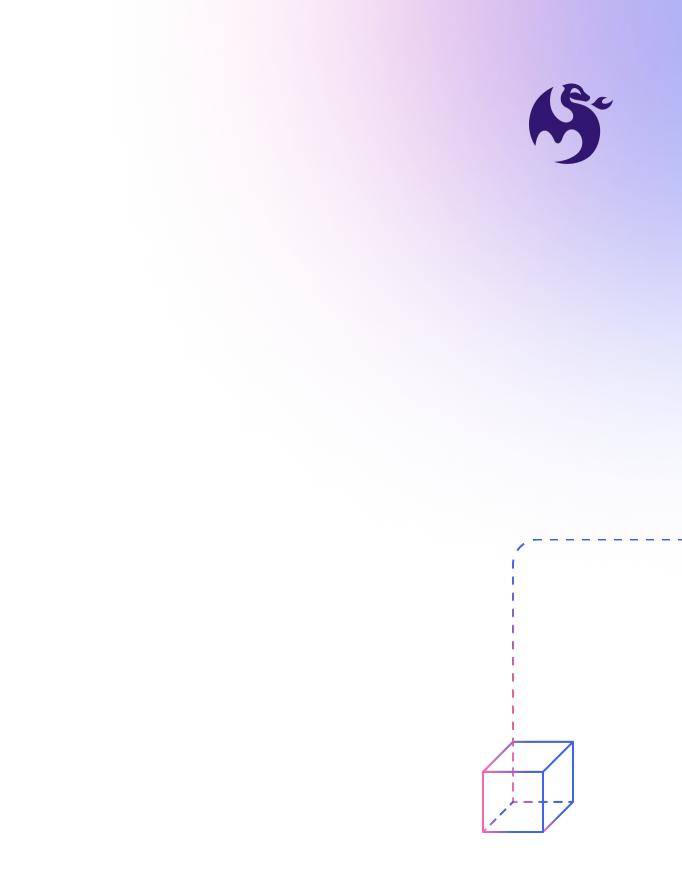






## Subscriptions were a DAG



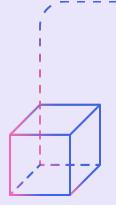


## Improvements from this Project

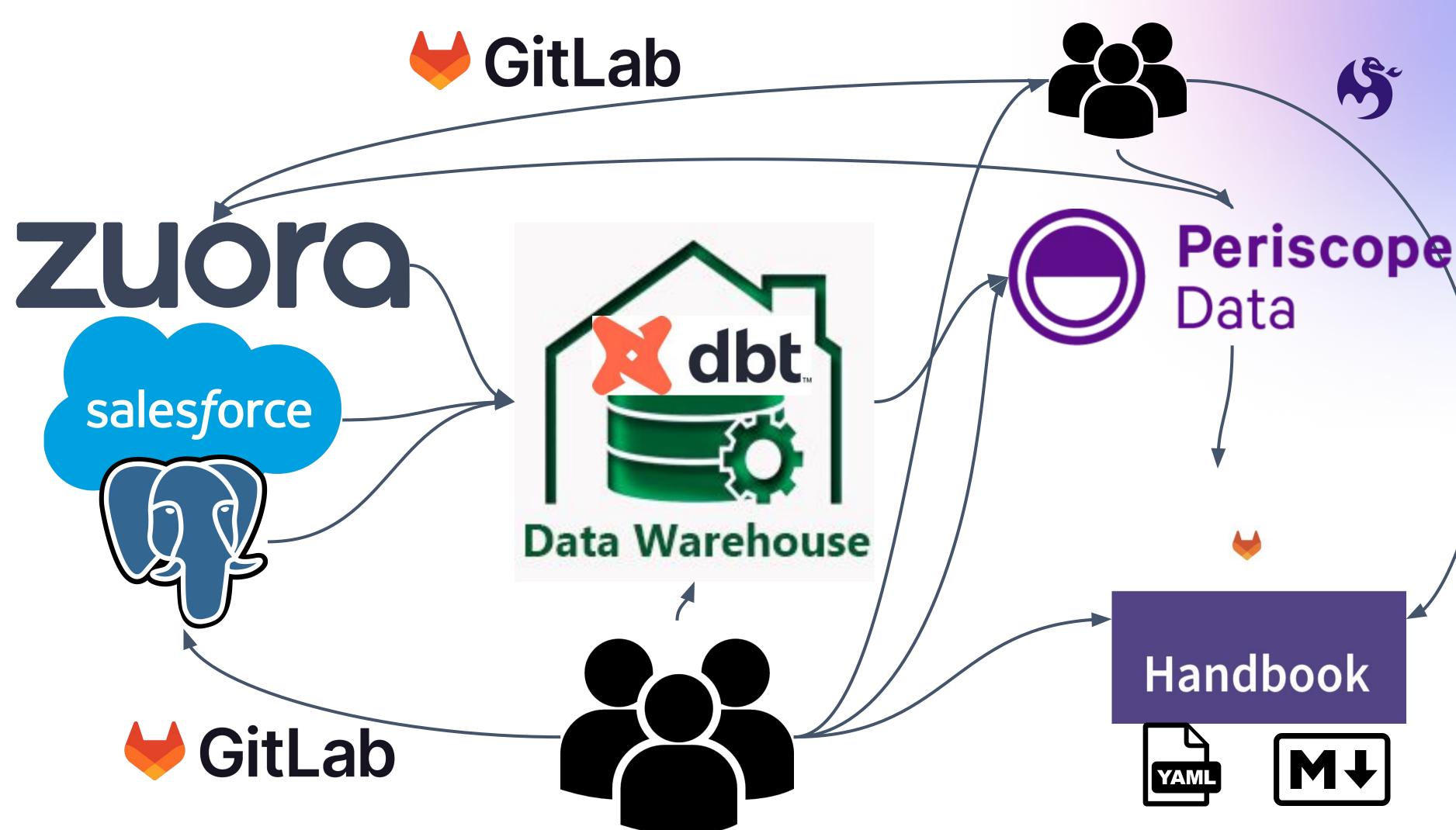
- 1. Helping to standardize SFDC account hierarchy
- 2. Standardizing Zuora subscription linking
- 3. Identifying bad database practices in **Customers DB**
- 4. Multiple metric updates and definitions
- 5. Cleaning a ton of bad data

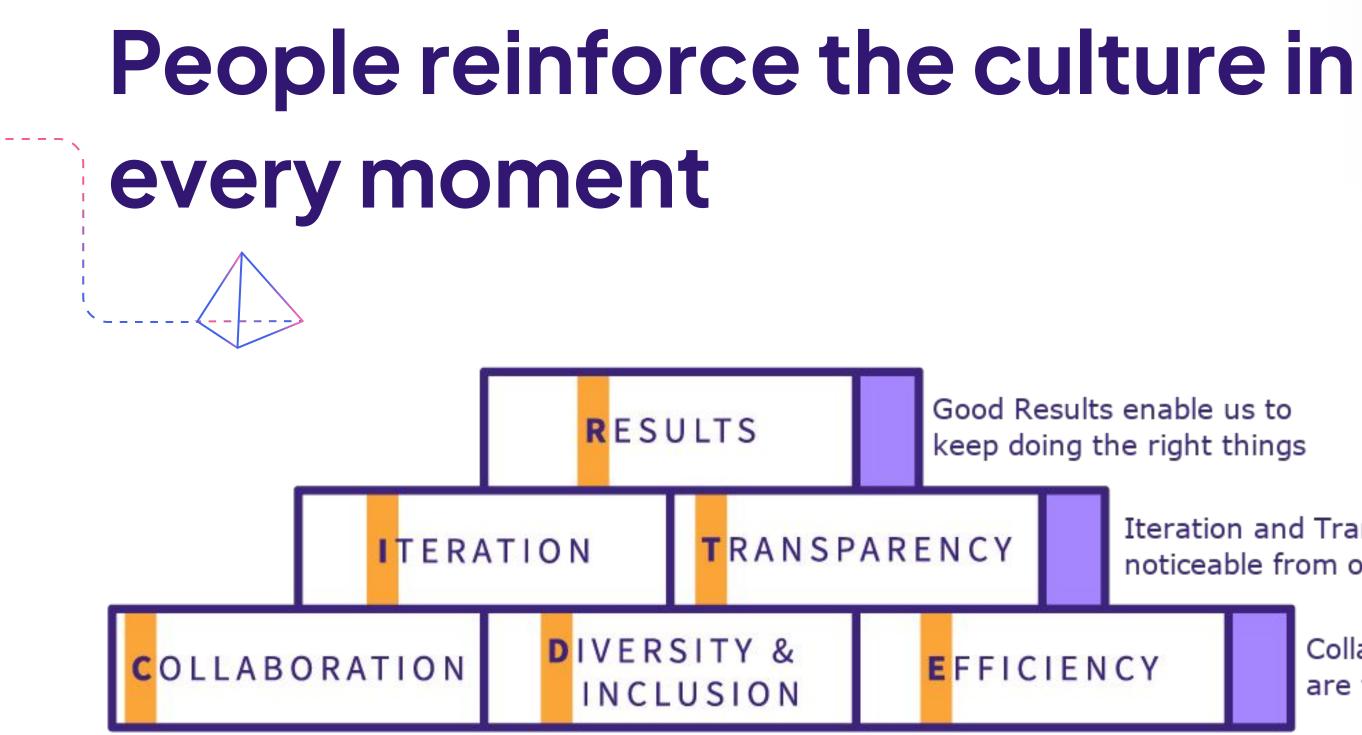


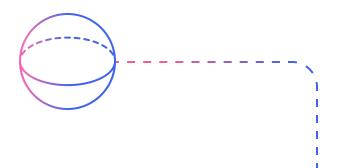








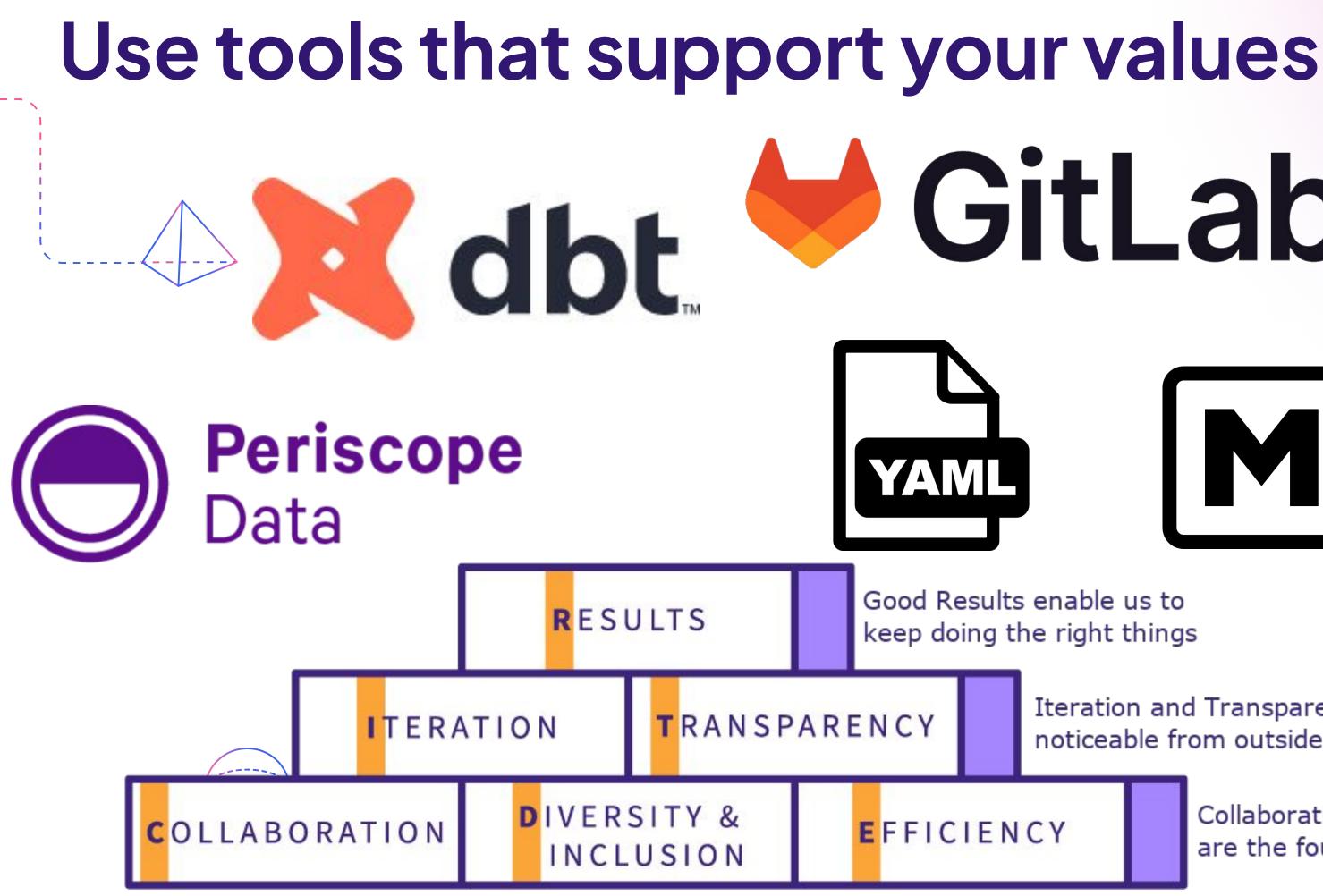






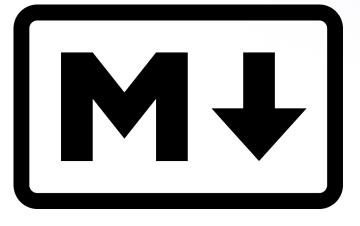
Iteration and Transparency are most noticeable from outside the organization





# GitLab





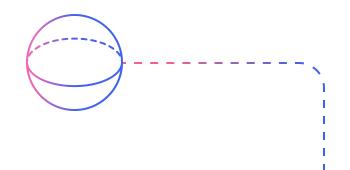
Good Results enable us to keep doing the right things

> Iteration and Transparency are most noticeable from outside the organization

> > Collaboration, DIB, and Efficiency are the foundation of our values



- People matter they reinforce culture
- Tools matter they support your people
- Great people + great tools solves real problems





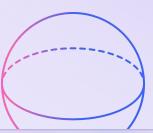
## Lessons

01.

Leadership & Strategy Matter 02.

Transparency is Powerful





## 03.

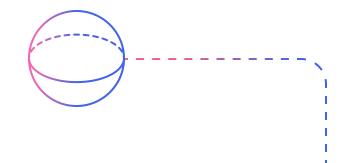
## People and tools matter



## Consider how they apply for your org and

## team culture

Keep them in mind when facing problems





# Simple lessons that are easy to say but hard to practice.

## Each of these are essential, but they are not sufficient.







# Thank you!

Twitter: @tayloramurphy Email: taylor@meltano.com



Links: https://meltano.com/data-council-2023

