ASCEND.IO

2023 DataAware™ Pulse Survey

Introduction

Ascend has operated one of the leading data industry benchmark surveys for four years running.

The survey surfaces the current sentiment around trends, best practices, and priorities of data teams.

It's a great opportunity for practitioners, managers, and everyone else who relies on data within the organization to understand the current state of affairs.

ASCEND.IO

Survey Demographics

Survey Respondents Job Level

Individual Contributor	Team Lead	Manager	Director	VP/Executive
8.8%	7.7%	33.3%	30.2%	20.0%

Survey Respondents Job Level

Job Focus (Non-Exec Only)

Data Analyst	Data Engineer	Enterprise Architect
25.0%	50.0%	25.0%

Job Focus (Non-Exec Only)

3
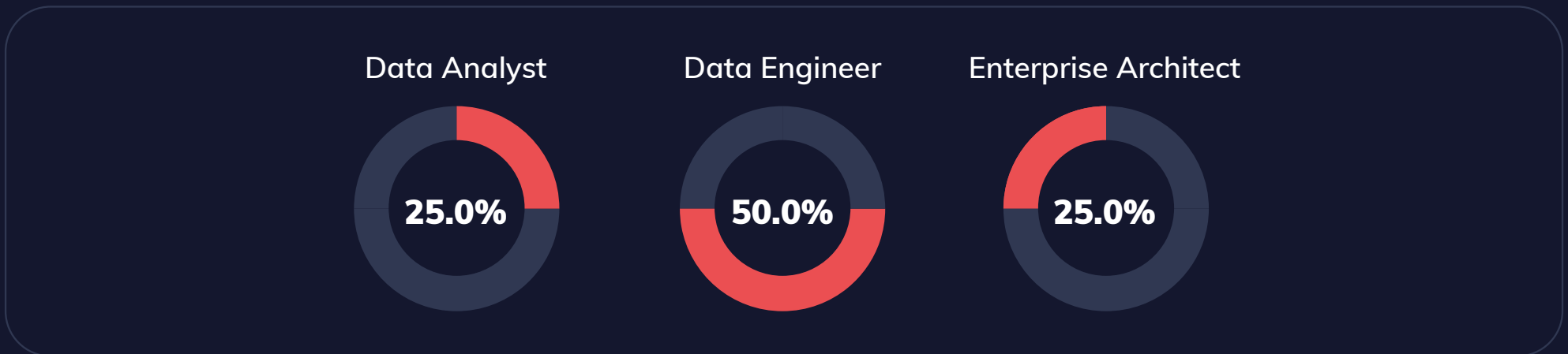
Survey Firmographics

Company Size
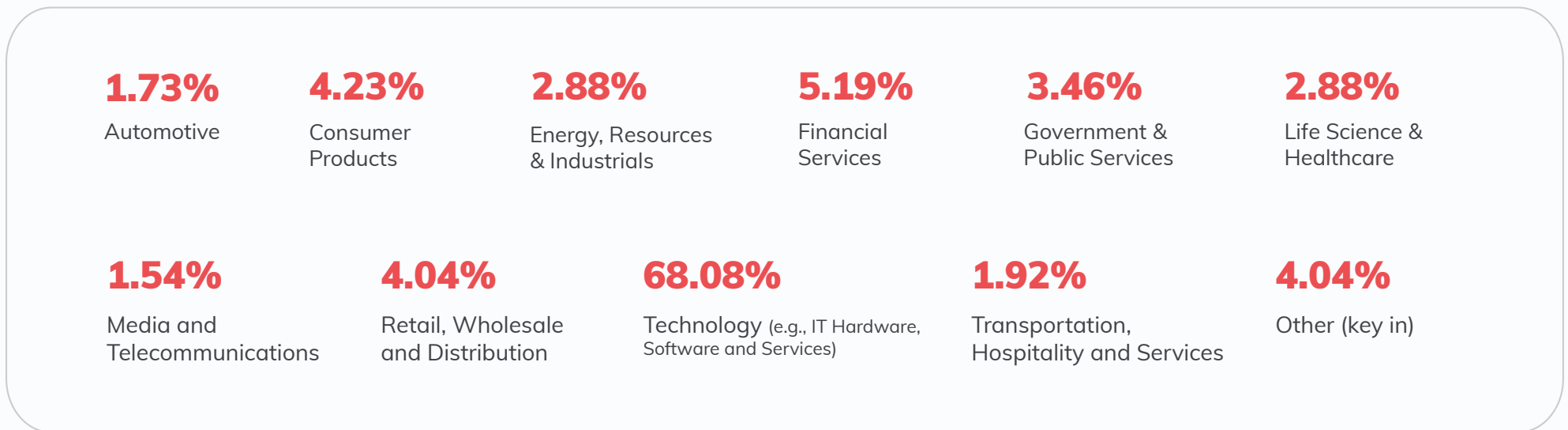
1-50	51-200	201-500	501-1000	More than 1000
8.08%	12.31%	22.88%	30.00%	26.73%

Number of employees

Industry

1.73%
Automotive

4.23%
Consumer Products

2.88%
Energy, Resources & Industrials

5.19%
Financial Services

3.46%
Government & Public Services

2.88%
Life Science & Healthcare

1.54%
Media and Telecommunications

4.04%
Retail, Wholesale and Distribution

68.08%
Technology (e.g., IT Hardware, Software and Services)

1.92%
Transportation, Hospitality and Services

4.04%
Other (key in)

ASCEND.IO

4

Result 1

Data engineering teams are at capacity, forcing them to work smarter not harder using automation technologies.

ASCEND.IO

Data teams are red-lining
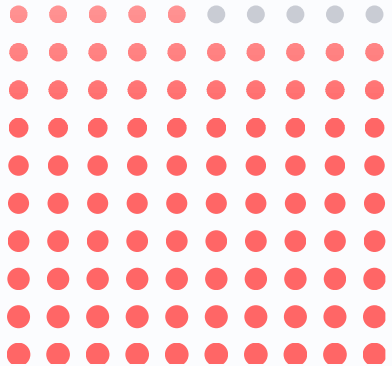
95%

of data teams report being at or above their work capacity for the fourth year in a row.
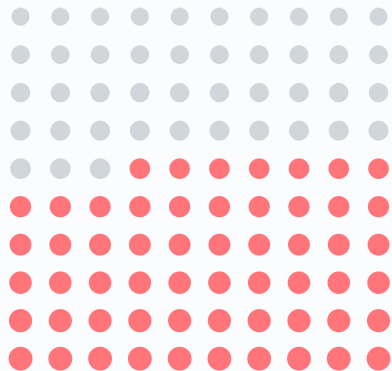
57%

feel they are somewhat above or significantly over their capacity.

ASCEND.IO

Data engineers bear the brunt of this workload

Data engineers are

● ● ● **3x**

more likely to report being significantly overworked compared with colleagues in the data architecture and data analyst teams.

Data engineer hiring is not keeping up

44% — Data Engineer

35% — Enterprise Architect

37% — Data Analyst

Data engineers are **27%** more likely to report a severe gap between team growth and company demands.
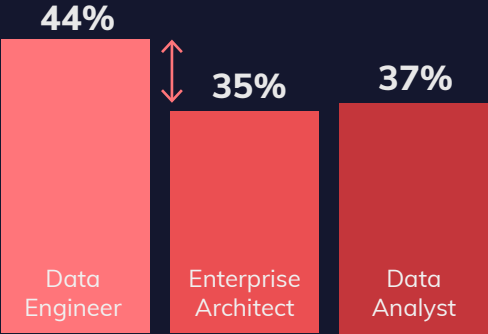
The need for data is growing significantly faster than our team size.

Busy-work is catching up

Respondents were **2x** as likely to report decreased productivity than previous years.
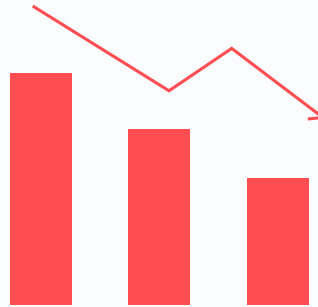
Data Engineers feel their productivity slipping the most

Data engineers were **2.2x** more likely to report decreased productivity over the past year.

ASCEND.IO

Data engineers spend too much time on maintenance

Time Spent Building New Things Vs. Maintaining Existing Ones

Legend: Data Engineer, Enterprise Architect, Data Analyst

Callouts: 44.23%, 21.15%, 1.92%

X-axis categories:
- maintain existing things
- build new things ¼ of time / maintain existing ¾ of time
- build new things ½ of time / maintain existing ½ of time
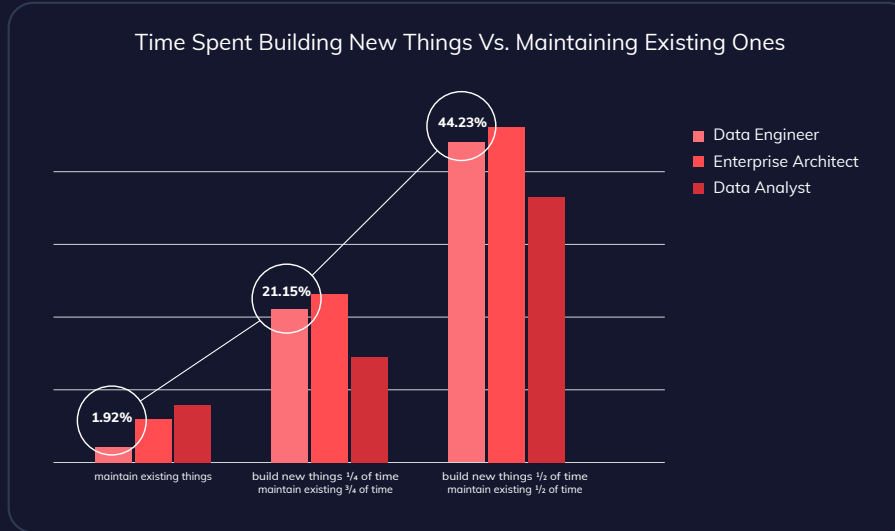
The majority of data engineers spend

50% or more
of their time just maintaining existing programs.

Our tools are letting them down

What is Limiting Your Productivity

Legend: Data Engineer, Enterprise Architect, Data Analyst

Callouts: 18.75%, 21.63%

X-axis categories:
- lack of automation
- limitations in our tools and technologies
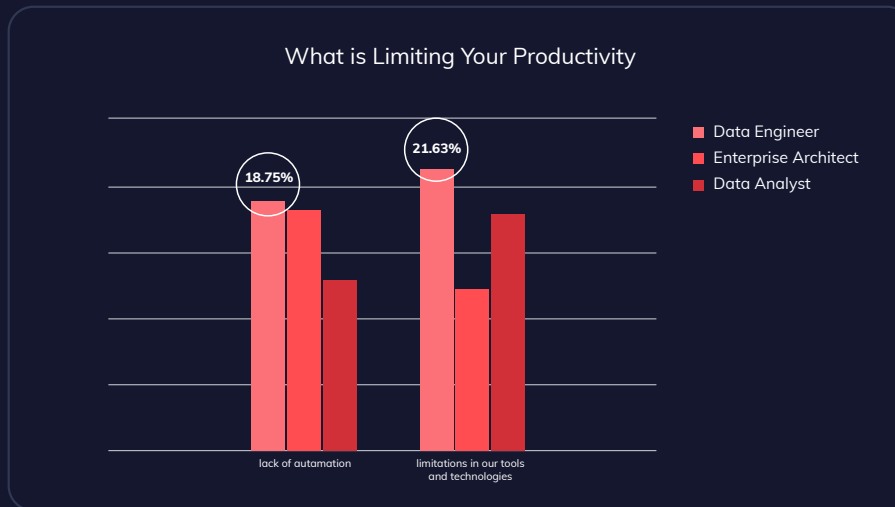
Data engineers point to limitations in tools and lack of automation as the two biggest hindrances to their productivity.

ASCEND.IO

Data teams need to get out of maintenance mode

42%

of data teams are stuck in reactive mode, with the number of errors fixed ranking as the highest impact measure for the group.

Data engineers are most impacted by this trend

46%

of data engineers report "number of errors fixed" as the most common metric used to measure their success.

ASCEND.IO

" Sean's Take

It's been a tough year for data teams... Many were planning to hire their way out of the crunch back in mid-2022 and have been thrown for a loop.

As they grappled with how to do more with less this year, we can see that initial productivity gains from last year are starting to erode as workloads shift to maintenance.

But don't give up hope yet. There are still ways for teams to take control of their workloads through increasing levels of automation!

Sean Knapp
Founder & CEO, Ascend.io

Result 2

Early pioneers in automation are seeing success, but many need help to get there.

Data teams are squarely focused on automating their work

91%
of those surveyed said they are likely, very likely or have already implemented data automation technologies.

Automation adoption accelerates, but needs platform-level support

Plans to Implement Data Automation Technology in the Next 12 Months

There is a **20%** increase in the number of respondents who indicate their team is 'very likely' to implement data automation technologies in the next 12 month.

There is a **110%** increase in the proportion of respondents who report currently using data automation technology compared with 2022 results.

+20%

+110%

2023
- Very unlikely: 1.15%
- Unlikely: 1.54%
- Not likely or unlikely: 6.73%
- Likely: 34.81%
- Very likely: 47.31%
- We already use data automation technologies: 8.46%

2022
- Very unlikely: 0.19%
- Unlikely: 2.22%
- Not likely or unlikely: 9.44%
- Likely: 44.81%
- Very likely: 39.81%
- We already use data automation technologies: 3.52%

50.00%
10.00%
0.00%

However, of the **40%** who indicated they were very likely to implement automation last year, only ~ **5%** of them managed to do so... indicating that the desire might not match the resourcing or skillsets of their team without the right supporting platforms.

ASCEND.IO

" Sean's Take

It's rare in the technology world to see something that so many people want yet so few people have today. That's where we have been with data automation for two years running.

It's great to see automation adoption growing! For those who are in the remaining 82% who want it but don't have it yet:

1. Start by defining what exactly it is you want to automate to help narrow your scope.

2. Look for solutions that give you lots of metadata, which is the fuel for automation.

3. If what you want to automate is the orchestration and maintenance of data pipelines, come check out Ascend!

Sean Knapp
Founder & CEO, Ascend.io

Automation is a top-down imperative

Very Likely to Implement

55.77%
52.87%
43.93%
34.78%
32.50%

Legend:
- VP/Executive
- Director
- Manager
- Team Lead
- Individual Contributor

Plans to Implement Data Automation Technologies in Next 12 Months
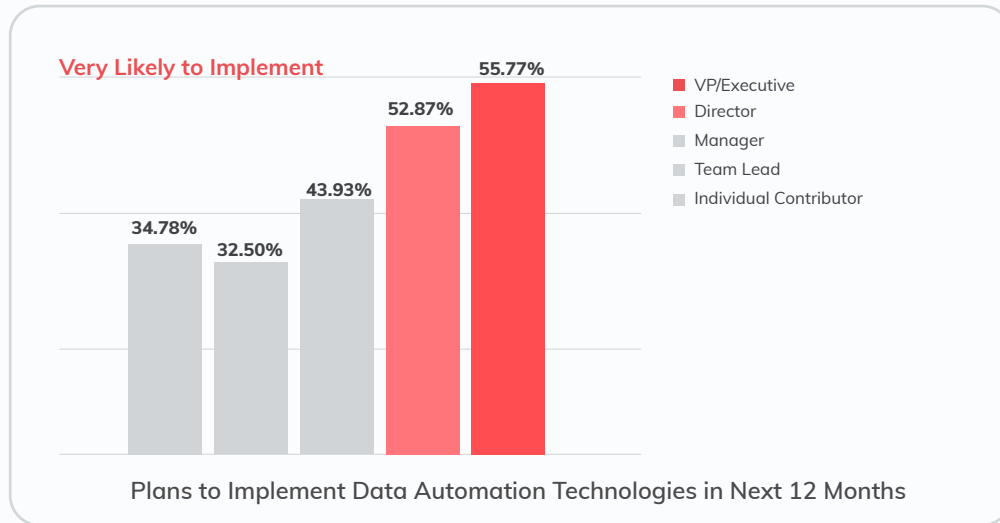
Data team executives were

74%

more likely to indicate strong plans to invest in automation.

Lack of Automation

22.30%
16.8%
16.03%
14.06%

Legend:
- 51-200 employees
- 201-500 employees
- 501-1000 employees
- more than 1000 employees

What is Limiting Your Productivity
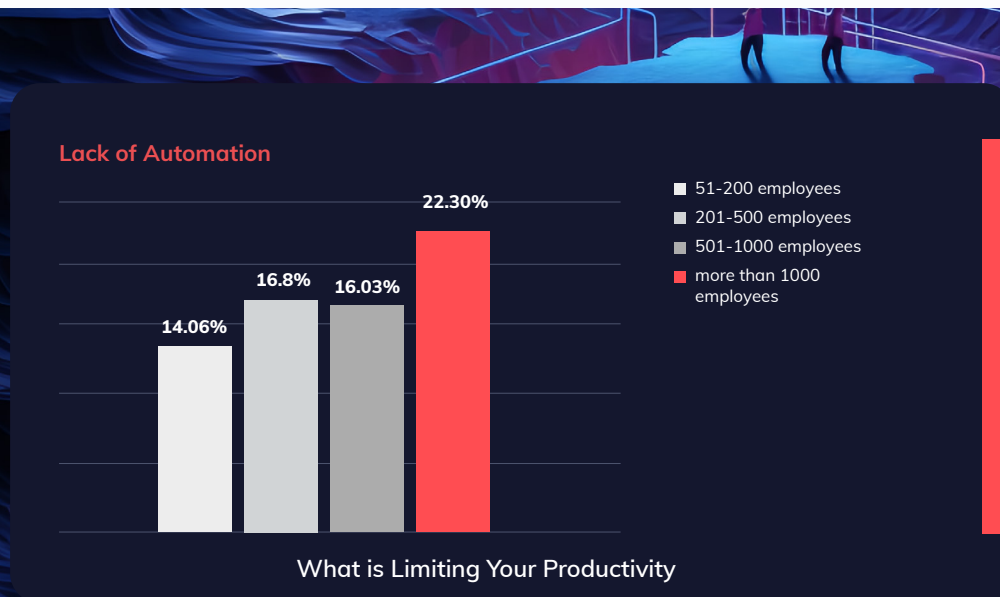
The largest companies and teams are up to

1.3x

more likely to feel the pain from lack of automation.

ASCEND.IO

Data engineers and architects see the potential for automation

33.65%

Data Engineer

30.77%

Enterprise Architect

16.35%
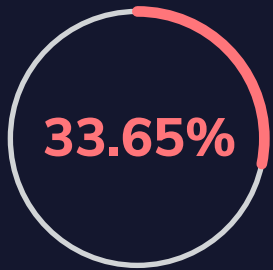
Data Analyst

Data engineers and architects are **twice** as likely to focus on automation to improve team capacity compared with data analysts.

Teams are asking for automation to meet business demands

What is your primary strategy to increase team capacity?

Legend: Individual contributor | Team lead | Manager | Director | VP/Executive

Our team has sufficient capacity to complete all of our work in the allotted time: 2.17%, 5.00%, 1.16%, 2.55%, 8.65%

We do not know how to increase bandwidth at this time: 17.39%, 7.50%, 2.89%, 1.91%, 1.92%

We plan to buy new products or tools: 10.87%, 20.00%, 26.59%, 24.20%, 25.96%

We plan to hire more staff: 26.09%, 32.50%, 29.48%, 21.66%, 34.62%

We plan to implement automation technologies: 32.61%, 15.00%, 26.01%, 33.76%, 20.19%

We plan to replatform and retire legacy technologies: 10.87%, 20.00%, 13.87%, 15.92%, 8.65%

Individual contributors and directors agree: **implementing automation technologies** is a **top priority** to increase team capacity.

ASCEND.IO

16

Result 3

Top KPIs of Data
Programs Revealed -
Does your team agree?

Overall KPIs of a data team ranked

What are the key Impact Measures of Your Role

42.77%	Number of errors fixed
40.77%	Time taken to respond to business requests
40.58%	Number of new dashboards created/supported
38.65%	Number of tickets closed
36.35%	Number of new data models delievered
35.96%	Number of new datasets delievered
35.38%	Number of presentations created/supported
33.85%	Revenue from data-driven products
29.81%	Business cost savings from data-driven optimization
25.00%	Overall business targets met by your organization

0.00%   10.00%   20.00%   30.00%   40.00%   50.00%

ASCEND.IO

KPIs breakout by specific roles

What are the key Impact Measures of Your Role

Legend:
- Data Engineer (blue)
- Enterprise Architect (red)
- Data Analyst (green)

KPI	Data Engineer	Enterprise Architect	Data Analyst
Number of errors fixed	45.67%	44.23%	34.64%
Time taken to respond to business requests	37.98%	47.12%	39.42%
Number of new dashboards created/supported	41.35%	30.77%	40.38%
Number of tickets closed	38.46%	33.65%	43.27%
Number of new data models delievered	39.90%	36.54%	28.85%
Number of new datasets delievered	32.21%	37.50%	36.54%
Number of presentations created/supported	32.21%	27.88%	37.50%
Revenue from data-driven products	37.02%	37.50%	23.08%
Business cost savings from data-driven optimization	29.33%	39.42%	20.19%
Overall business targets met by your organization	25.96%	25.96%	18.27%

ASCEND.IO

Overall KPIs of a data team ranked

Data teams in the largest organizations are up to

110%

more likely to feel their KPIs are aligned with the overall goals of the business.

What are the key Impact Measures of Your Role

Legend: 1-50 employees | 51-200 employees | 201-500 employees | 501-1000 employees | More than 1,000 employees

Category	1-50	51-200	201-500	501-1000	More than 1,000
Number of new datasets delievered	28.57%	42.19%	29.41%	39.74%	36.69%
Number of new dashboards created/supported	21.43%	26.56%	39.50%	48.72%	44.60%
Number of presentations created/supported	28.57%	29.69%	36.97%	46.79%	25.90%
Time taken to respond to business requests	40.48%	35.94%	38.66%	39.74%	46.04%
Number of tickets closed	33.33%	34.38%	36.97%	40.38%	41.73%
Number of errors fixed	33.33%	46.88%	31.09%	50.00%	43.88%
Number of new data models delievered	28.57%	31.25%	36.13%	33.33%	44.60%
Revenue from data-driven products	33.33%	40.63%	30.25%	33.81%	33.81%
Business cost savings from data-driven optimization	16.67%	21.88%	31.93%	33.97%	30.94%
Overall business targets met by your organization	21.43%	17.19%	23.53%	20.51%	35.97%

ASCEND.IO

Result 4

Executives, meet your teams. You two have some things to discuss...

ASCEND.IO

Alignment of KPIs leaves much to be desired

Misalignment Across the Board

Executives
- 49% Dashboards Created
- 47% Presentations Supported

Directors
- 45% New Data Models Delivered

Team Leads
- 50% Tickets Closed

Indv. Contr.
- 50% Errors Fixed

Executives were 3x more likely to rank presentations created, and 1.5x more likely to cite number of new dashboards created as their top impact measures compared with individual contributors, who rank errors fixed and tickets closed.

ASCEND.IO

" Sean's Take

Managers and directors are the connective tissue between the doers in the engine room and the planners in the C-Suite. This gap in alignment speaks to the need for them to do more to bridge between both worlds.

They need to help executives understand why these things are hard, and where they can invest to help it go faster.

They need to help their team understand the connection between errors & tickets and the overall objectives of the business.

Sean Knapp
Founder & CEO, Ascend.io

Executives see basic data tasks as being too slow

What Slows Down The Work (Select All That Apply)
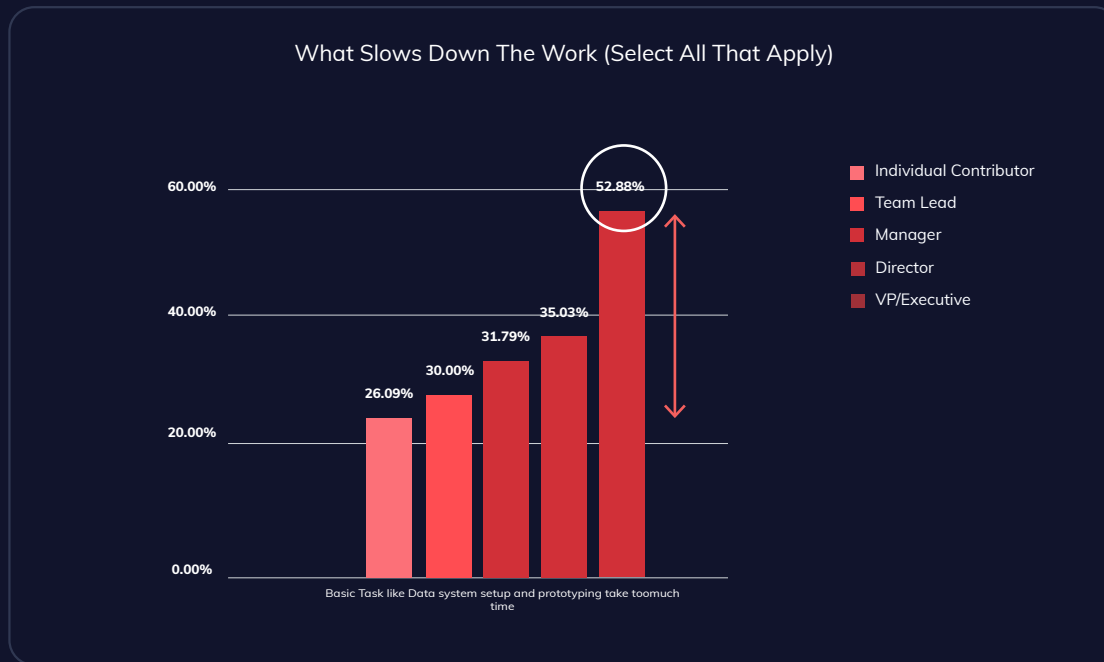
Executives are twice as likely to perceive basic tasks with data as taking too long.

Leaders spend 5 more hours/week gathering data

17.8 hrs — VPs/Executives
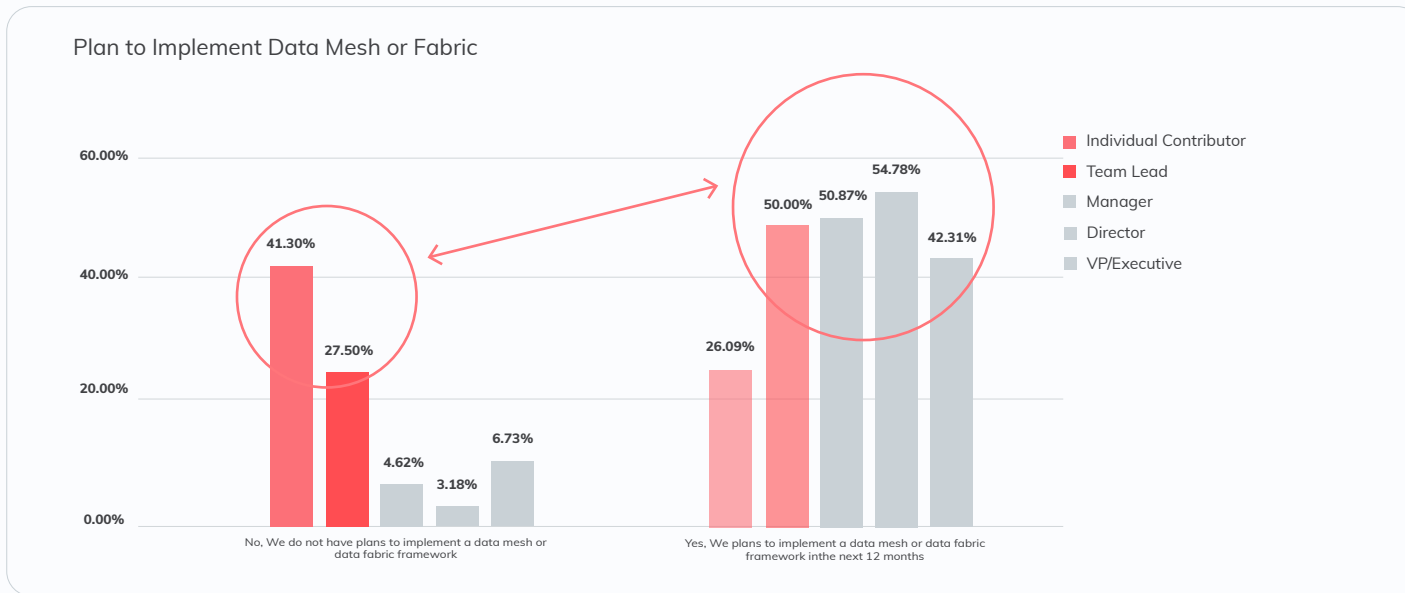
12.2 hrs — Individual Contributors

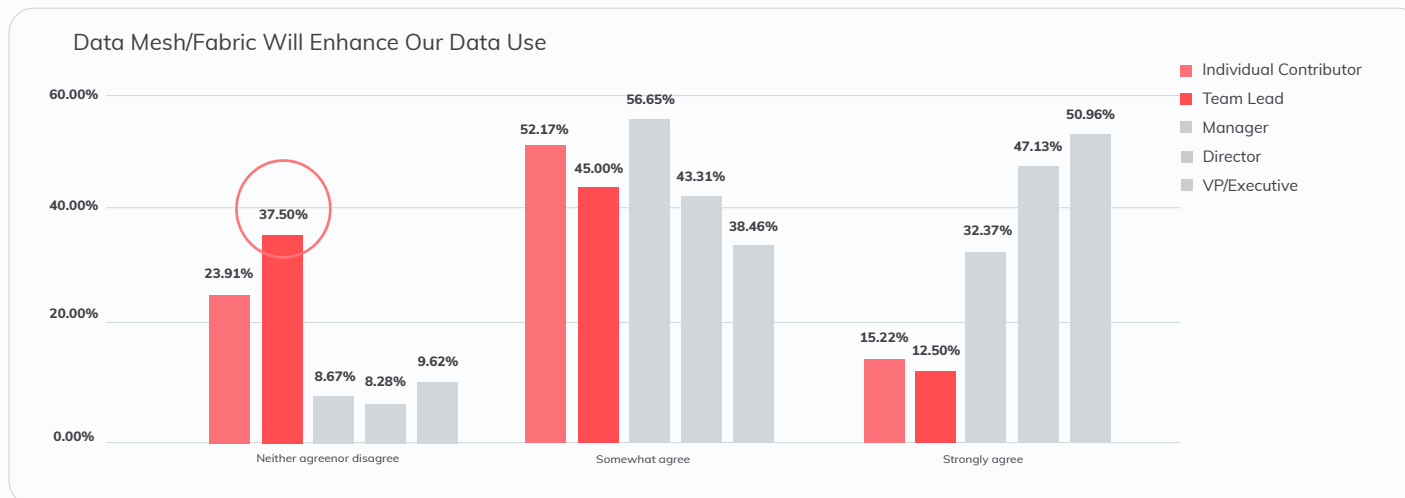Per week spent trying to get access to data to do my job (on average).

This indicates they feel the impact of inefficient, manual data pipelines the most in their organization.

Disagreement over major initiatives is present

Plan to Implement Data Mesh or Fabric

Legend:
- Individual Contributor
- Team Lead
- Manager
- Director
- VP/Executive

No, We do not have plans to implement a data mesh or data fabric framework: 41.30%, 27.50%, 4.62%, 3.18%, 6.73%

Yes, We plans to implement a data mesh or data fabric framework inthe next 12 months: 26.09%, 50.00%, 50.87%, 54.78%, 42.31%

Individual contributors are **5x** more likely to deny having any plans to implement data mesh/fabric compared with executives, who predominantly report plans to implement one within the next 12 months.
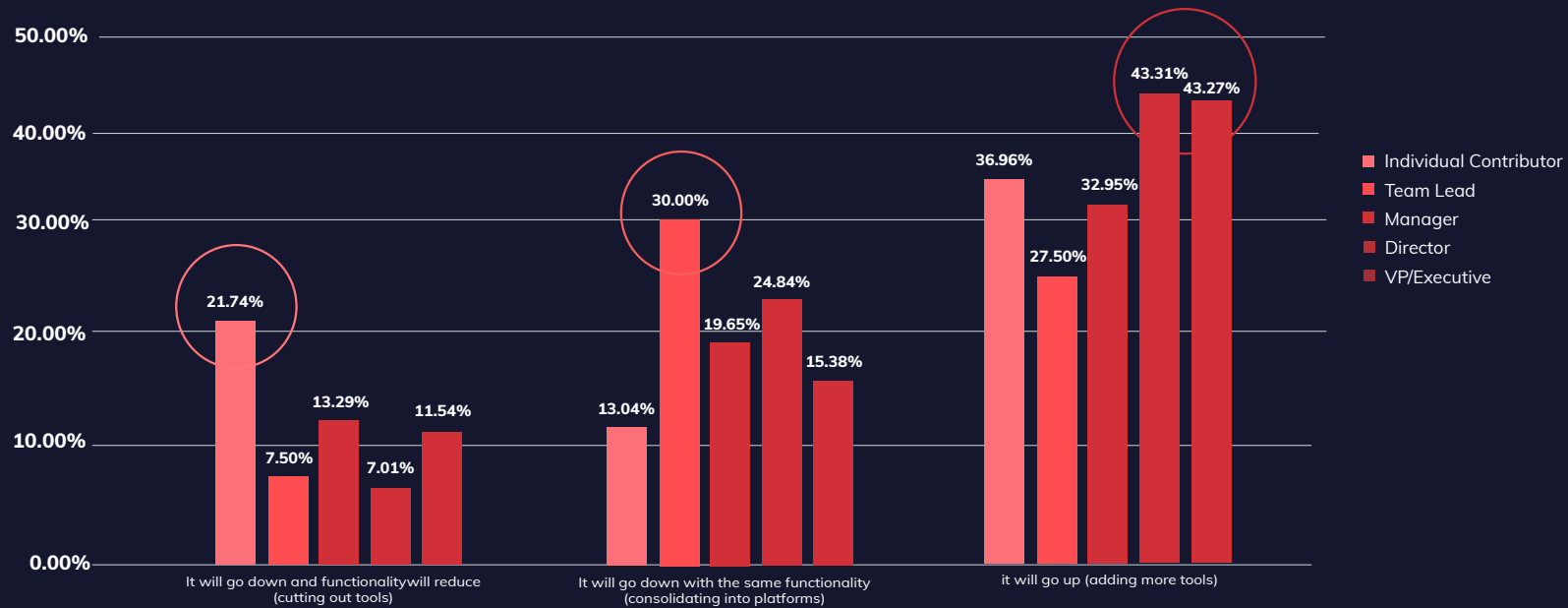
Alignment on strategy shows sign of strain as well

Data Mesh/Fabric Will Enhance Our Data Use

Legend:
- Individual Contributor
- Team Lead
- Manager
- Director
- VP/Executive

Neither agree nor disagree: 23.91%, 37.50%, 8.67%, 8.28%, 9.62%

Somewhat agree: 52.17%, 45.00%, 56.65%, 43.31%, 38.46%

Strongly agree: 15.22%, 12.50%, 32.37%, 47.13%, 50.96%

Team leads are **2.8x** more likely to report being skeptical of the benefits of data mesh than their executives.

ASCEND.IO

25

Tug of war over data stack sizes

What Will Happen To the Number of tools In Your Data Stack Next 12 Month

Legend:
- Individual Contributor
- Team Lead
- Manager
- Director
- VP/Executive

It will go down and functionality will reduce (cutting out tools)
- 21.74%
- 7.50%
- 13.29%
- 7.01%
- 11.54%

It will go down with the same functionality (consolidating into platforms)
- 13.04%
- 30.00%
- 19.65%
- 24.84%
- 15.38%

it will go up (adding more tools)
- 36.96%
- 27.50%
- 32.95%
- 43.31%
- 43.27%

Individual contributors are twice as likely as their executives to favor removing tools from their data stack.
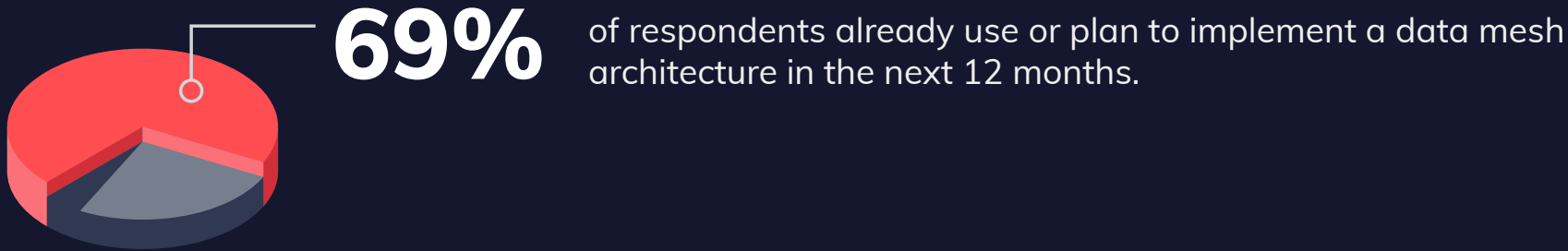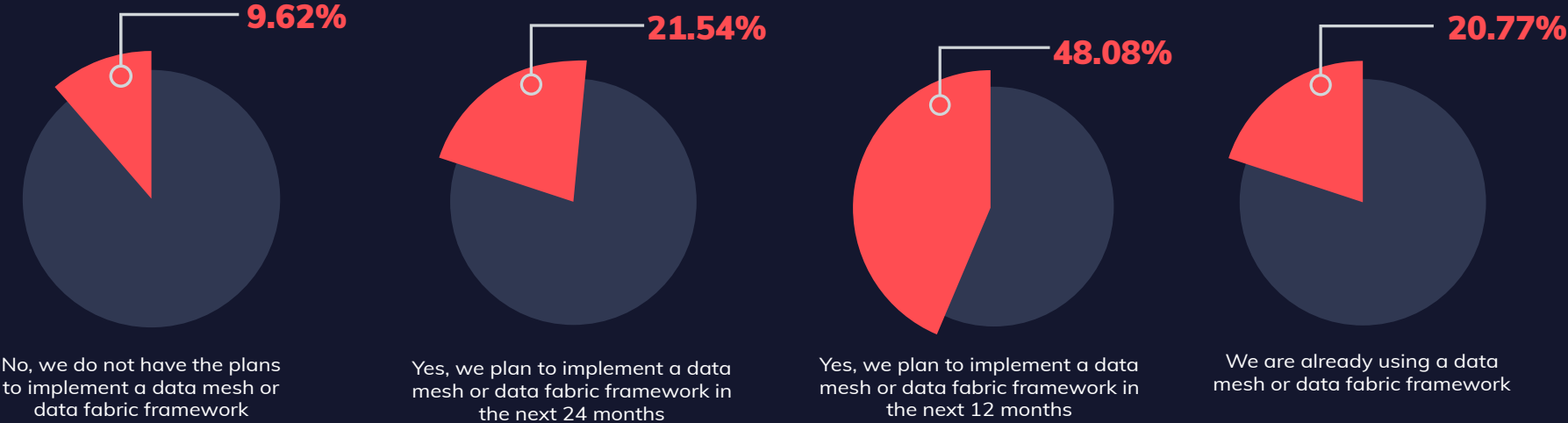
Team leads want to consolidate on platforms.

Executives favor adding more tools!
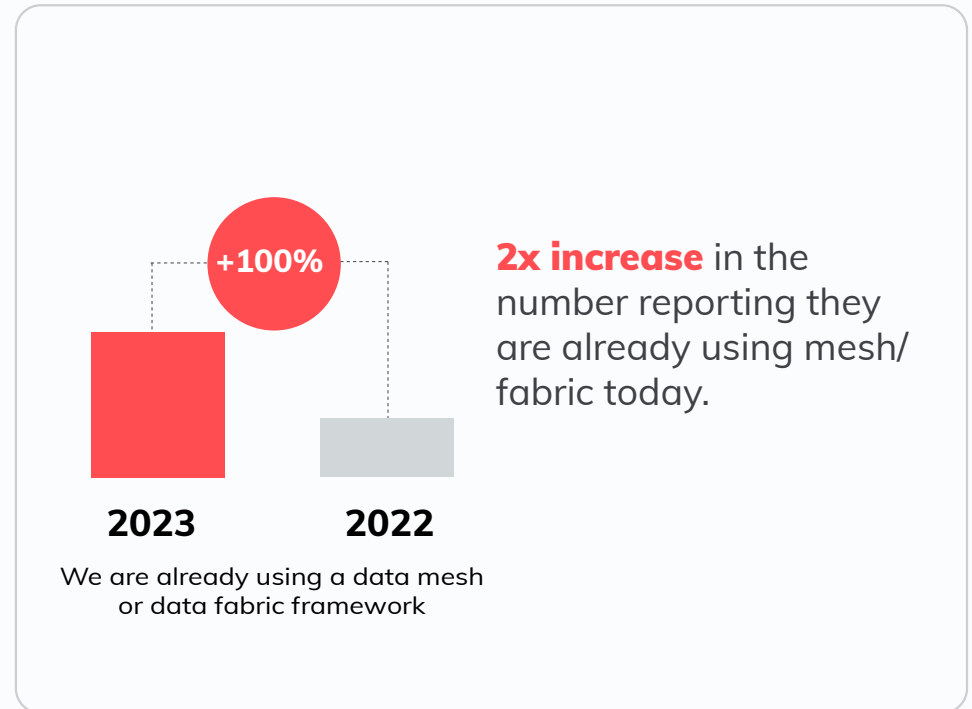
ASCEND.IO

Result 5

Data Mesh:
It's here!

Data Mesh/Fabric architectures are here to stay

Plans to Implement Data Mesh or Fabric

9.62%

21.54%

48.08%

20.77%

No, we do not have the plans to implement a data mesh or data fabric framework

Yes, we plan to implement a data mesh or data fabric framework in the next 24 months

Yes, we plan to implement a data mesh or data fabric framework in the next 12 months

We are already using a data mesh or data fabric framework

69% of respondents already use or plan to implement a data mesh architecture in the next 12 months.

Huge momentum in implementations during 2022

Plans to Implement Data Mesh or Fabric

+12%

12% increase in the number reporting plans to implement mesh/fabric in the next 12 months.

2023   2022

Yes, we plan to implement a data mesh or data fabric framework in the next 12 months

+100%

2x increase in the number reporting they are already using mesh/fabric today.

2023   2022

We are already using a data mesh or data fabric framework
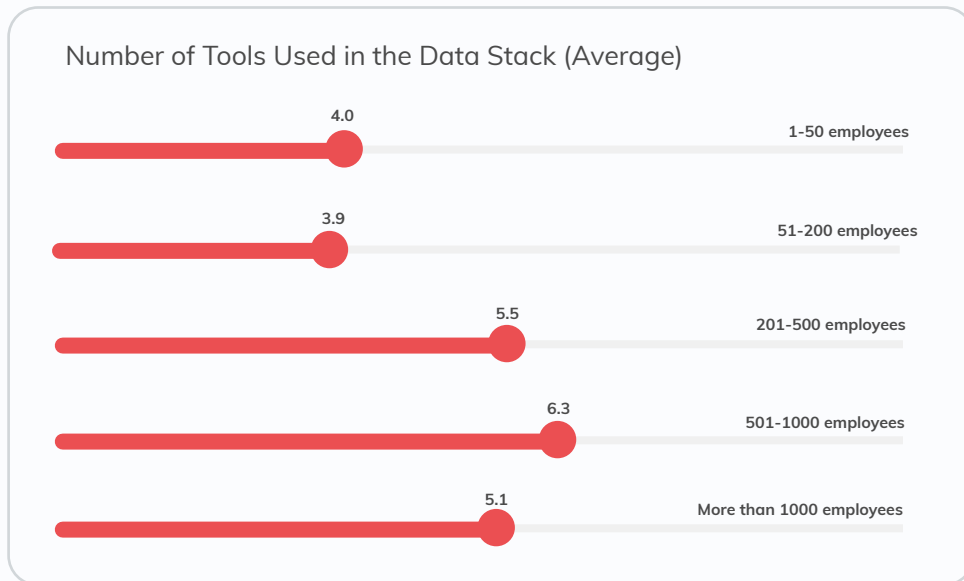
ASCEND.IO
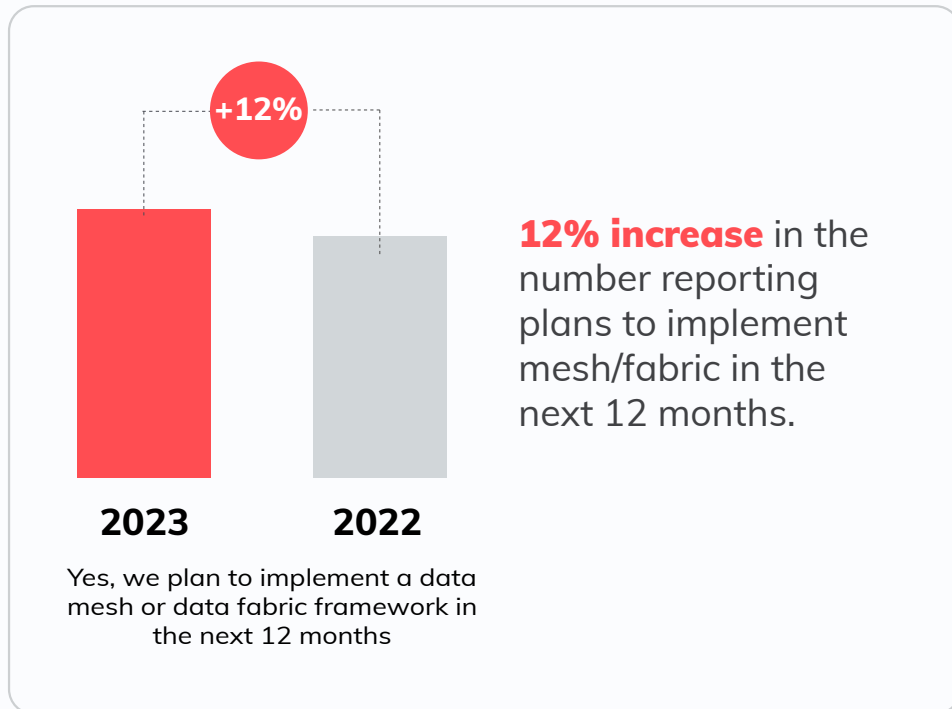
Result 6

Tool sprawl in the data stack continues to grow, hindering efforts to drive automation.

ASCEND.IO

The Modern Data Stack is big

5.1 tools

Used in the data stack
(on average)

Larger organizations are more tool heavy

Number of Tools Used in the Data Stack (Average)

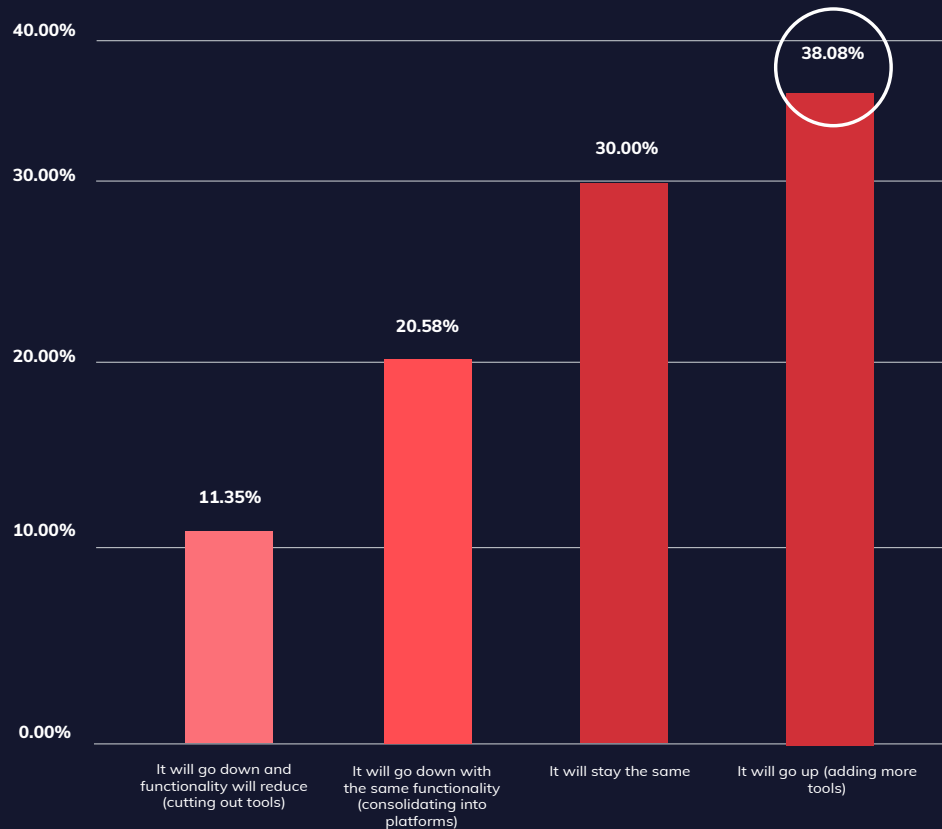|	
4.0	1-50 employees
3.9	51-200 employees
5.5	201-500 employees
6.3	501-1000 employees
5.1	More than 1000 employees

Large companies report using an average of 2 more data tools in their stack (6) than small and midsize companies (4).

ASCEND.IO

Q21: What Will Happen To The # of Tools in 12 Mos

What Will Happen To the Number of Tools In Your Data Stack Next 12 Months

40.00%	
	38.08%
30.00%	30.00%
20.00%	20.58%
10.00%	11.35%
0.00%	

It will go down and functionality will reduce (cutting out tools) — 11.35%
It will go down with the same functionality (consolidating into platforms) — 20.58%
It will stay the same — 30.00%
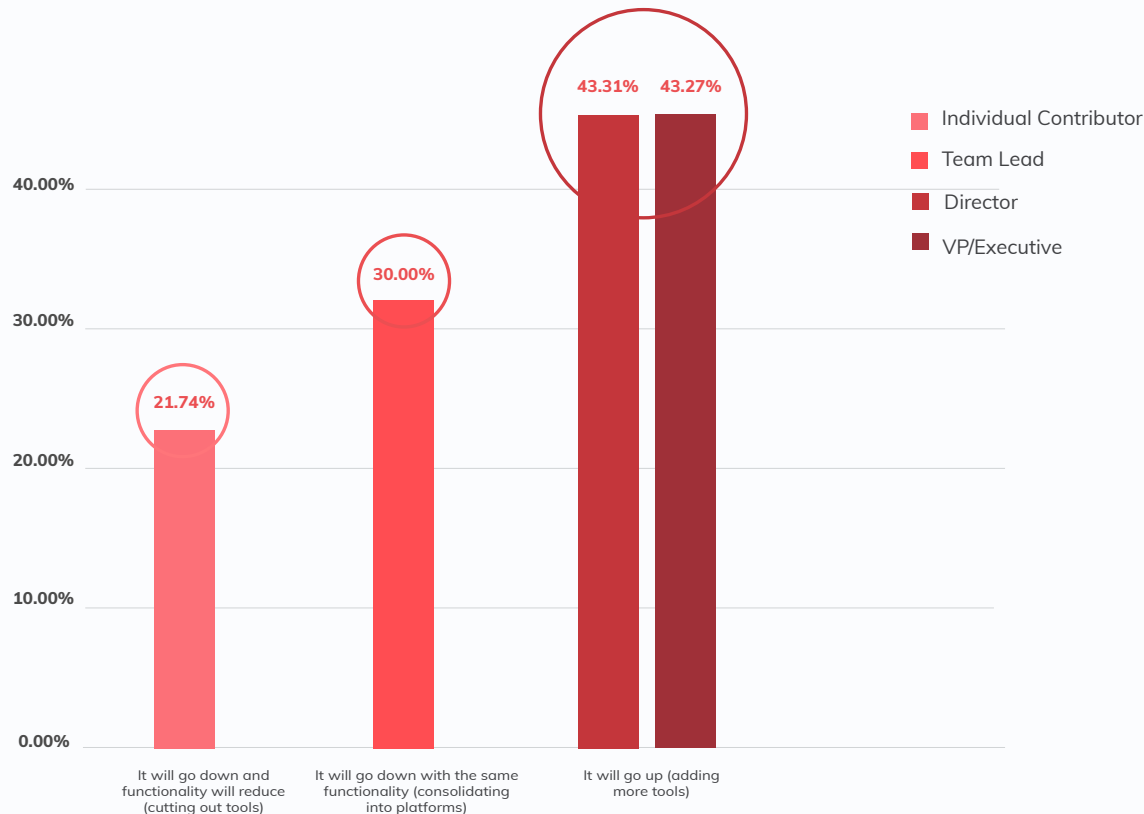It will go up (adding more tools) — 38.08%

68%
of respondents believe their data stack will stay the same size or increase in the number of components over the next 12 months.

Three groups, three different tool strategies

What Will Happen To the Number of Tools In Your Data Stack Next 12 Months

Legend:
- Individual Contributor
- Team Lead
- Director
- VP/Executive

Chart values:
- It will go down and functionality will reduce (cutting out tools): 21.74%
- It will go down with the same functionality (consolidating into platforms): 30.00%
- It will go up (adding more tools): 43.31%, 43.27%

Y-axis: 0.00%, 10.00%, 20.00%, 30.00%, 40.00%

Individual contributors are up to

2.1x

more likely to favor cutting tools out of their data stack at the expense of functionality.

Team leads favor consolidating to platforms while retaining all functions.

Executives are up to

16%

more likely to favor adding more tools to the stack.

" Sean's Take

Clearly there's a need for some conversations to be had in these teams!

It's understandable for individuals who are already overwhelmed and hindered by the shortcomings of their current tools to want to throw them overboard.
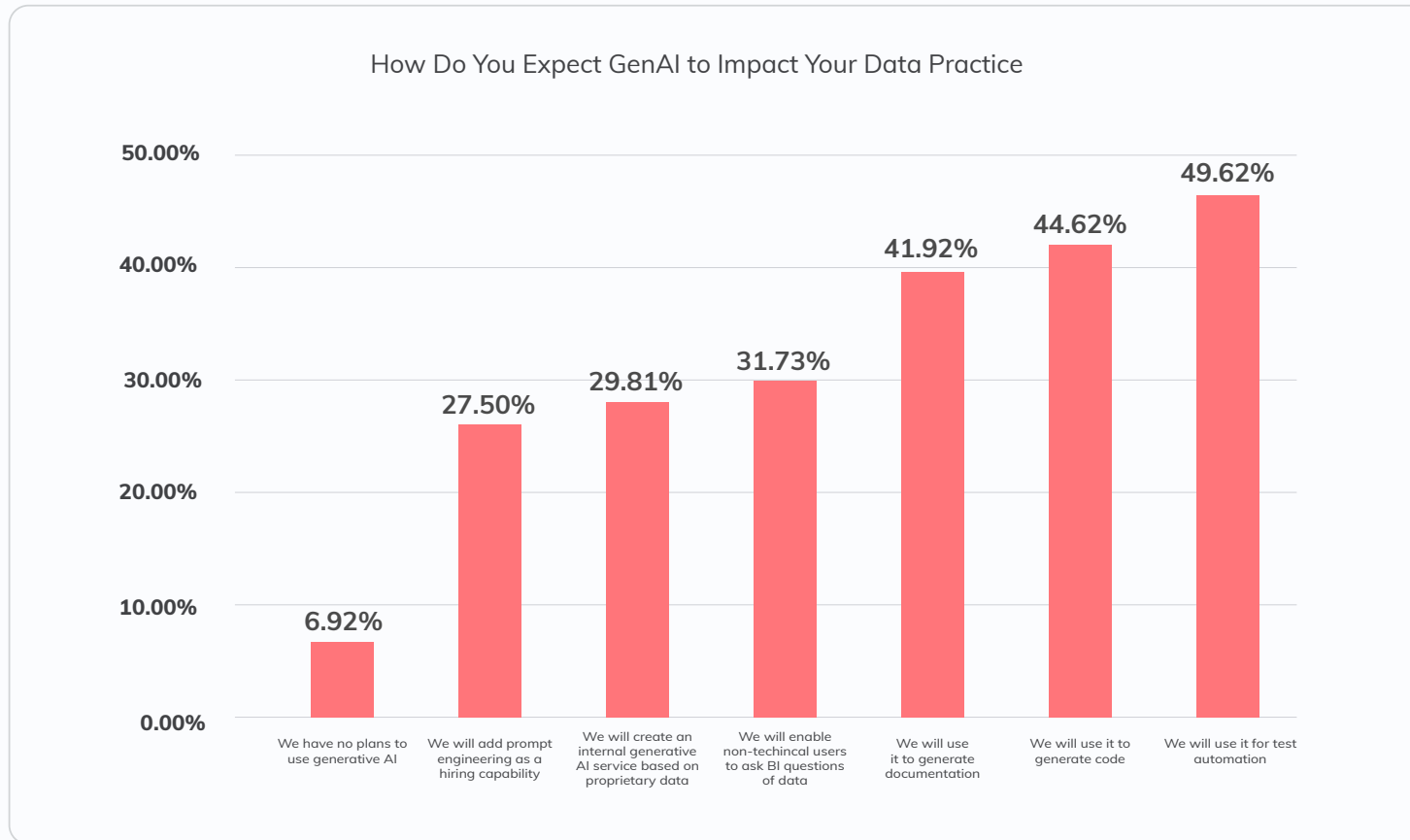
The onus will be on team leads to temper that desire by steering towards integrated platforms that can still maintain the right levels of functionality and add in a bit of automation.

It's also feasible to think that executives are looking towards the future and thinking about new challenges like GenAI and ML Ops and planning for new tool purchases in those areas.
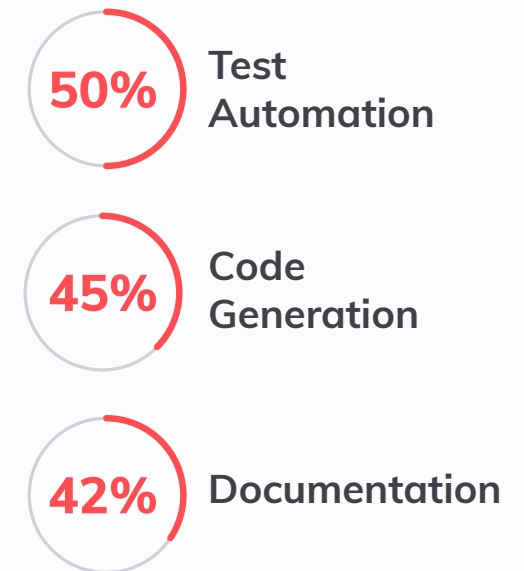
Sean Knapp
Founder & CEO, Ascend.io

Result 7

Generative AI Use in Data Teams.

The data team sees high potential for Generative AI

How Do You Expect GenAI to Impact Your Data Practice

Category	Percentage
We have no plans to use generative AI	6.92%
We will add prompt engineering as a hiring capability	27.50%
We will create an internal generative AI service based on proprietary data	29.81%
We will enable non-techincal users to ask BI questions of data	31.73%
We will use it to generate documentation	41.92%
We will use it to generate code	44.62%
We will use it for test automation	49.62%

Top use cases for Generative AI in data teams:

- **50%** Test Automation
- **45%** Code Generation
- **42%** Documentation

Highest agreement on using it to generate code

How Do You Expect GenAI to Impact Your Data Practice

Highest Agreement

50.96%

28.26%

32.61%

Legend:
- Individual Contributor
- Team Lead
- Manager
- Director
- VP/Executive

X-axis categories:
- We will use it to generate code
- We will use to generate documentation
- We will use it for test automation

Y-axis: 0.00% – 60.00%
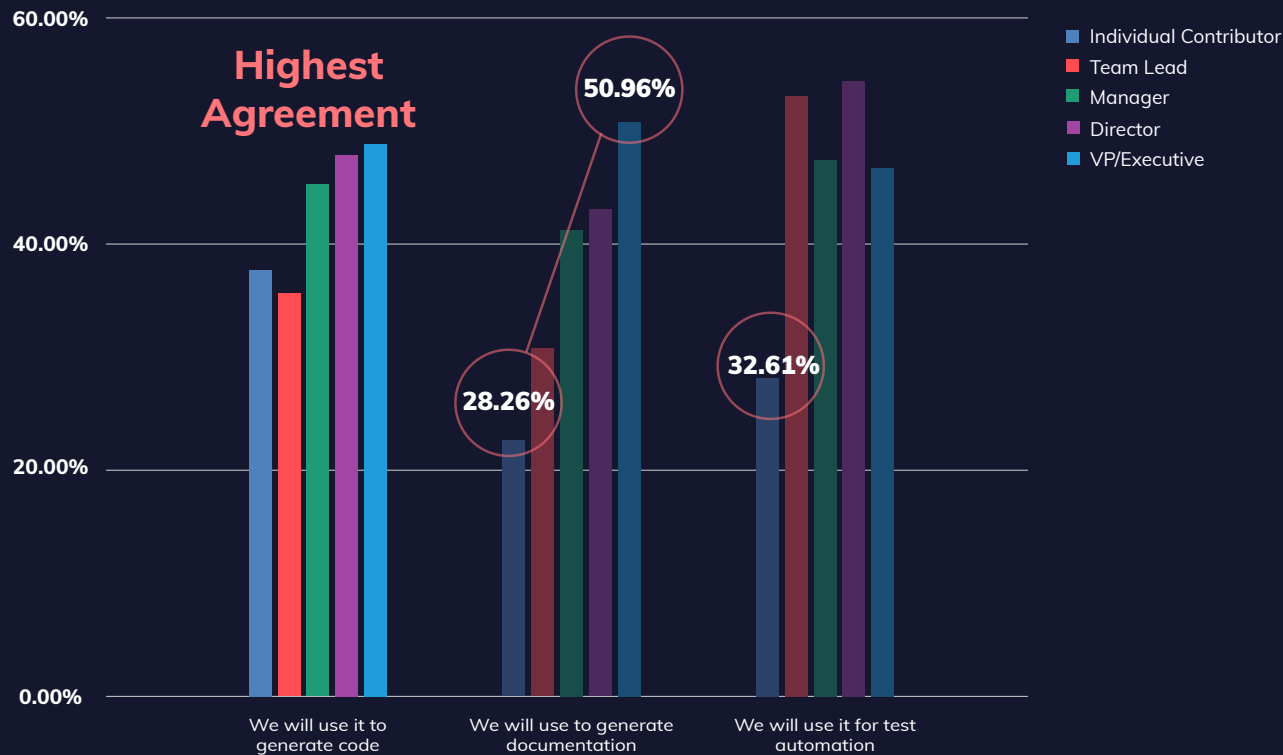
Data team executives are

80%

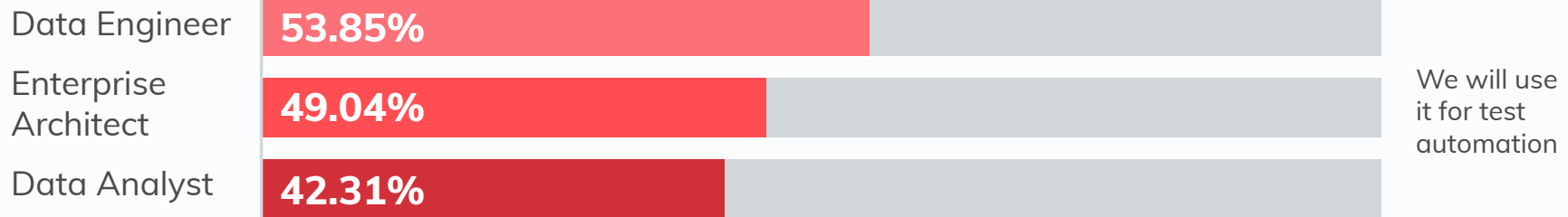more confident in using Generative AI to create documentation than individual contributors.

Individual contributors are

66%

less confident in plans to use Generative AI for data testing.

ASCEND.IO

The data team sees high potential for Generative AI to improve data pipeline reliability

How Do You Expect GenAI to Impact Your Data Practice

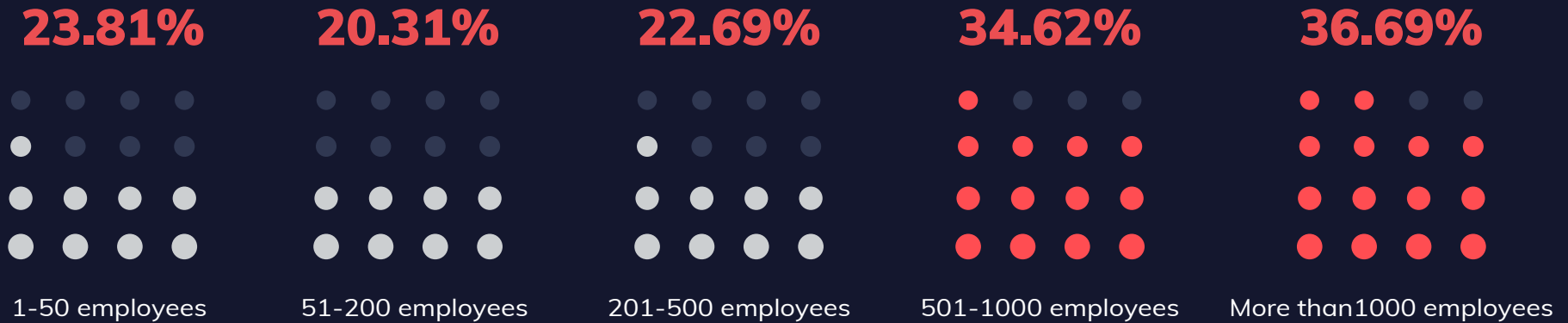|	
Data Engineer	**53.85%**
Enterprise Architect	**49.04%**
Data Analyst	**42.31%**

We will use it for test automation

54%
Data engineers are the most optimistic about the use of Generative AI for creating test automation.

The race is on to create custom LLMs

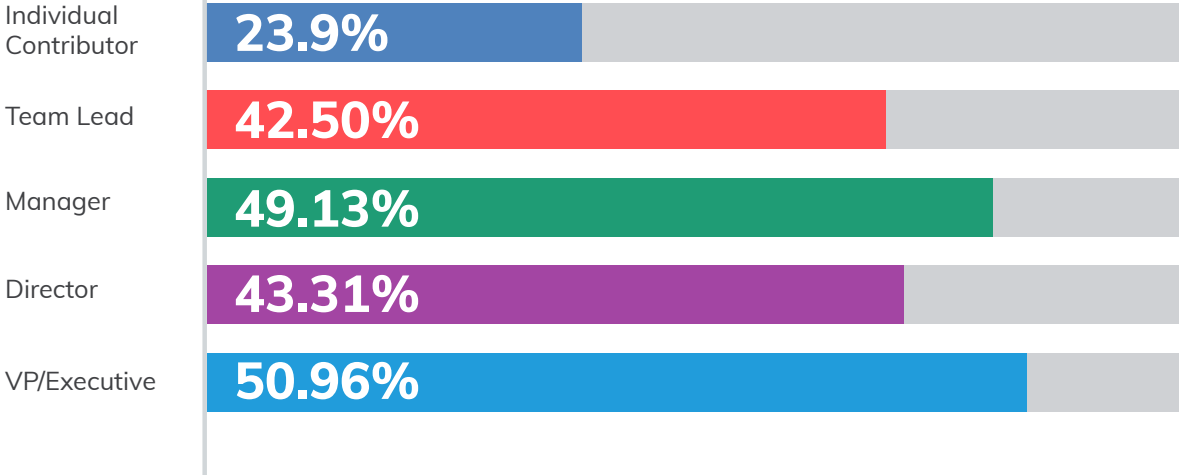We will create an internal generative AI service based on proprietary data

23.81%	20.31%	22.69%	34.62%	36.69%
1-50 employees	51-200 employees	201-500 employees	501-1000 employees	More than1000 employees

Smaller teams can leverage data pipeline automation technologies to help them catch up.

Larger companies (<500 employees) are up to

80%

more likely to have plans to create a proprietary generative AI service based on internal data.

ASCEND.IO

39

Individuals expect less business impact from generative AI

How Do You Expect GenAI to Impact Your Role in Next 12 Months

Role	Percentage
Individual Contributor	23.9%
Team Lead	42.50%
Manager	49.13%
Director	43.31%
VP/Executive	50.96%

It will increase my impact on the business

Individual contributors are **200%** less likely to report believing Generative AI can increase their impact on the business.

" Sean's Take

There's a lot to unpack here. We expect this sentiment to rapidly evolve in the coming months as people get more familiar with the technology.

On one hand, we can see how those who are the closest to these new ML models and understand the most about how they work seem to be the most skeptical of their success. Individual contributors and team leads are usually the best technical minds on the team, and we see lower scores from them almost across the board in this section.

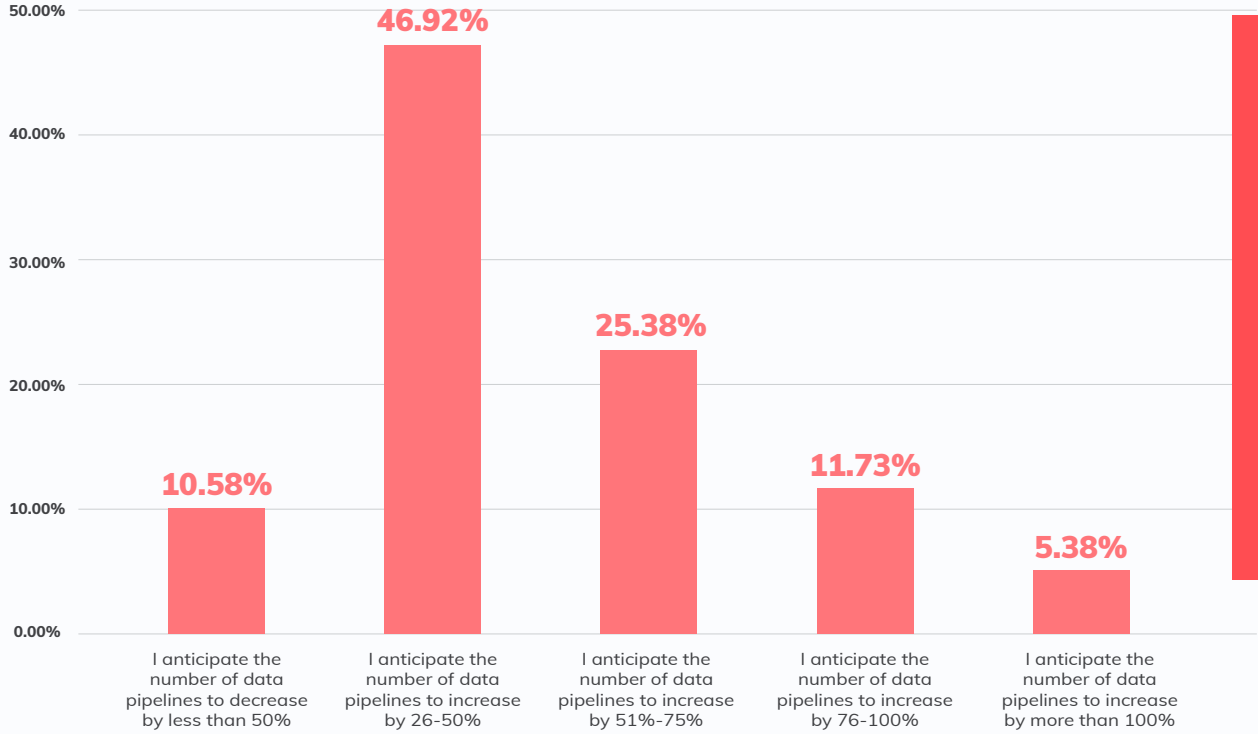It's also important for executives to do more to inspire their teams with the potential of this new technology. When deployed in a consistent manner with the right enablement, it can provide an individual with more autonomy and job satisfaction in their role. We've certainly seen this in many different roles here at Ascend.

Sean Knapp
Founder & CEO, Ascend.io

Result 8

As Data Grows in the Cloud, so do Expenses.

Data pipelines will continue to grow

Data Pipeline Growth in Next 12 Months

46.92%

25.38%

11.73%

10.58%

5.38%

50.00%

40.00%

30.00%

20.00%

10.00%

0.00%

I anticipate the number of data pipelines to decrease by less than 50%

I anticipate the number of data pipelines to increase by 26-50%

I anticipate the number of data pipelines to increase by 51%-75%

I anticipate the number of data pipelines to increase by 76-100%

I anticipate the number of data pipelines to increase by more than 100%

89%
of respondents expect the number of data pipelines to grow over the next 12 months.

ASCEND.IO

Cost Saving Measures Next 12 Months

What Measures Are You Targeting To Save Money Next 12 Months

Measure	Percentage
Reducing headcount on the team	14.81%
We have no plans for cost-sacing measures at this time	22.12%
Reducing the number of datasets supported in production	23.85%
Reducing the number of experimental projects	28.46%
Renegotiating contracts with tool vendors	28.85%
Consolidating separate tools into lower cost all-in-one platforms	32.50%
Implementing charge-backs on cloud compute consumption costs to project sponsors in other departments	32.69%
Consolidating data sources to reduce data volumes processed	38.08%
Optimizing data pipelines to reduce cloud compute consumption	48.08%

48%
Optimizing data pipelines to reduce cloud computing costs is the favorite money saving tactic.

38%
Followed by consolidating data sources to reduce data volumes.
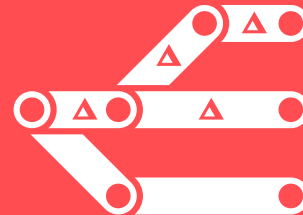
This growth is driving a variety of cost saving strategies
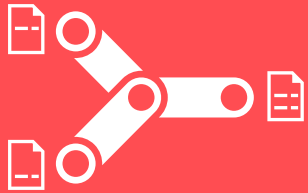
Data team executives are up to

1.9x

more likely to feel their teams can consolidate data sources.

The rest of the technical team favors pipeline optimizations to drive down costs.

" Sean's Take

Teams will need to help guide their leaders here. It's easy to see data volumes as the problem when costs appear directly correlated with size. But simply removing data sources from analytics is usually counter-productive.

Instead, teams should look to get greater cost visibility into their data pipelines to help them find low-hanging fruit for optimization. Modern data clouds will tell you how much individual queries cost, but good pipeline platforms will show you so much more. Armed with this information, you can approach your leader with a program that will drive the highest immediate impact without jeopardizing business results.

Sean Knapp
Founder & CEO, Ascend.io