



Data & AI Trends & Predictions

2024

Contents

•	TREND #1	
	Generative AI Goes Mainstream	2
•	TREND #2	
	The Battle Between Closed and Open Source AI Heats Up	3
•	TREND #3	
	“Boring AI” Use Cases Generate the Majority of the Value	4
•	TREND #4	
	Organizations Capitalize on Smaller, Specialized Large Language Models	5
•	TREND #5	
	The Risks of Generative AI Start Bubbling to the Surface	6
•	TREND #6	
	Generative AI Fundamentally Lowers the Barrier to a Data Culture	7
•	TREND #7	
	The AI Boom Intensifies Data Governance Efforts	8
•	TREND #8	
	Data & AI Literacy Become Table Stakes to Succeed with Generative AI	9
•	TREND #9	
	Prompt Engineering Will Become a Valuable Skillset Across Many Roles, but Not a Career Path	10
•	TREND #10	
	Generative AI Will Focus Data Professions On Value Creation	11

By almost every metric, 2023 has been a [watershed moment](#) for data and AI. The rapid pace of change has been primarily driven by the advent of generative AI, a breakthrough that has brought AI into public consciousness like never before.

The impact was palpable; according to [Gartner](#), 55% of organizations are now experimenting with or fully implementing generative AI technologies. This shift is not only reshaping how businesses operate but also how professionals view their future within the industry. Anaconda highlights that [nearly half of data specialists](#) are concerned about job security in the face of these advancements.

As we turn the page to 2024, we anticipate generative AI becoming even more pervasive, almost invisible, seamlessly integrating into our digital experiences at work and in life. In many ways, 2024 will mark the “end of the beginning” of the AI paradigm shift.

This year will be a defining moment for organizations and individuals as they navigate this new landscape. What trends and developments can we expect in 2024? How will the skills agenda evolve for organizations and individuals? How will the data industry change? Our data and AI predictions for 2024 aim to explore these questions and more.

Generative AI Goes Mainstream

TOOLS & TECHNOLOGY

SOCIETY & CULTURE

Considering the title of this trend, it's hard to dispute that generative AI is not already in the mainstream, especially with its prevalence in popular culture throughout 2023. Despite the buzz around Generative AI, it's important to note that its mainstream adoption within organizations is still nascent.

The integration and utilization of generative AI in enterprises' internal operations and customer-facing products and services have room for significant growth. MIT Technology Review's research reveals that [only 9% of organizations have successfully productionized generative AI](#). More importantly, at least [50% of organizations have struggled to deploy generative AI](#) due to high costs, lack of skills, and more.

Regarding internal usage of AI tools, around [16% of professionals regularly use ChatGPT](#) at work. Moreover, organizations are still in the process of defining their internal AI usage policies. According to [Gartner](#), 35% of companies said they hadn't finalized their AI guidance yet, while 14% had issued a blanket ban on using tools like ChatGPT.

The landscape, however, is rapidly changing in 2024. Forrester predicts a substantial shift, with over [60% of employees expected](#) to use Generative AI tools in their work by 2024. Moreover, [Goldman Sachs reports](#) that numerous industries will transition from experimenting to deploying generative AI products and services in 2024. This, coupled with the release of generative AI features across most enterprise software organizations (e.g., [Microsoft Copilot](#)), will lead to a mainstreaming of generative AI in 2024.



Sixty percent of workers will use AI to perform their job and tasks in 2024

Forrester's 2024 AI trends & Predictions

[Read More](#)

DIVE DEEPER INTO THE MAINSTREAMING OF GENERATIVE AI



[Why AI is Eating the World](#) — Hear from Daniel Jeffries, Managing Director at AI Infrastructure Alliance, on how AI will become pervasive in most software we use today.

[Why AI will Change Everything](#) — Hear from Bob Muglia, Former CEO at Snowflake, on how AI will transform the future of work.

[The Past, Present & Future of Generative AI](#) — Hear from Joanne Chen, General Partner at Foundation Capital, on how the AI ecosystem will evolve.

The Battle Between Closed and Open Source AI Heats Up

TOOLS & TECHNOLOGY

In 2023, the AI landscape was transformed by the emergence of sophisticated foundation models from proprietary giants like OpenAI, Google, Anthropic, Cohere, and others. Interestingly, we also saw the rise of notable open-source alternatives from Meta, Salesforce, Databricks, and more.

The clash between the closed-source and open-source AI ecosystems spurred a vigorous debate between advocates of open and closed-source AI.

On the one hand, prominent figures such as Yann LeCun and Andrew Ng have argued in [favor of open-source AI](#), suggesting that the AI community's collective problem-solving skills are a better route toward AI safety and value creation. Moreover, there are practical risks with using closed-source models. On a [DataCamp Webinar](#), Hagay Luplesko, VP of Engineering at MosaicML, warned about closed-source LLMs' cost and data privacy implications.

Conversely, critics of open-source AI have raised concerns about the potential negative impacts of widely democratizing AI technology. We've seen calls to [halt or pause](#) the development of stronger AI systems. We've also seen [regulators](#) potentially curtailing the development of open-source AI.

As we move into 2024, this debate is expected to intensify, and organizations are expected to diversify. Forrester predicts that [eighty-five percent of enterprises will increasingly adopt open-source AI models](#), driven by the explosive growth and experimentation with generative AI in 2023. This shift will be propelled by the adoption of open-source models like [LlaMA 2](#) (and its eventual successors) and the growing popularity of marketplaces like [Hugging Face](#). While this doesn't spell the end for proprietary models, it certainly paves the way for more diverse generative AI deployment strategies in the years to come.

These developments promise exciting possibilities within the open-sourced and closed-source AI ecosystems. As Daniel Jefferies, Managing Director at AI Infrastructure Alliance, put it on the [DataFramed](#) podcast, “I think in a long enough timeline, open source wins. However, we may also see another timeline where open-source and closed-source exist peacefully.”

“I think in a long enough timeline, open source wins. However, we may also see another timeline where open-source and closed-source exist peacefully.”

**Daniel Jefferies, Managing Director
at AI Infrastructure Alliance**



DIVE DEEPER INTO THE FUTURE OF OPEN-SOURCE AI



[Buy or Train? Using Large Language Models in the Enterprise](#) — Learn from Hagay Luplesko, VP of Engineering at MosaicML, the tradeoffs between open-source and closed-source AI.

[Demystifying AI: Unpacking the Generative AI Landscape](#) — Join Grace Isford, Partner at Lux Capital, Ollie Forysth, Global Community Manager at Antler, and Lauren Xandra, VP of Marketing & Platform at Two Sigma Ventures, on the evolving AI landscape.

[8 Top Open-Source LLMs for 2024 and Their Uses](#) — Read about the top open-source LLMs in this accessible article.

3

“Boring AI” Use Cases Generate the Majority of the Value

TOOLS & TECHNOLOGY

While the public discourse often marvels at the more sensational applications of generative AI, such as [image and video generators](#) or [automatic dubbing](#), the most substantial value for enterprises is likely to stem from less glamorous yet highly practical applications. These “[boring AI](#)” use cases, which significantly augment workflow processes, are poised to be the primary drivers of value in the initial wave of generative AI adoption within businesses.

Many of these applications often fall under the Robotic Process Automation (RPA) umbrella. They include, but are not limited to, document extraction and summarization, data entry, and other routine processes. Generative AI's potential in these areas is immense. A [Boston Consulting Group \(BCG\)](#) study revealed that knowledge workers utilizing GPT-4 experienced significant enhancements in efficiency and quality, particularly in domains like operations, customer service, legal and compliance, and technology. These are fields where routine automation presents significant opportunities for efficiency gains.

As emphasized by [Jason Feifer, Editor in Chief at Entrepreneur Magazine](#), the value of generative AI in business is not just in its ability to perform flashy tasks but in its capacity to transform and streamline mundane yet essential business operations: “The most exciting generative AI use-cases are use-cases that save them time, enabling people to liberate their time and focus on tasks that can help create long-term value.”

“The most exciting generative AI use-cases are use-cases that save them time, enabling people to liberate their time and focus on tasks that can help create long-term value.”

Jason Feifer, Editor in Chief of Entrepreneur Magazine



In 2024 and beyond, we can expect these “boring” use cases to take center stage in the first wave of applications of generative AI, generating the majority of the value for enterprises by enhancing productivity and efficiency in fundamental business processes.

An Example of “Boring AI” — AI-Assisted Customer Support Workflow

	Customer support ticket is received	Customer support addresses complaint	Ticket resolution
 AI Augmentation	Generative AI system summarizes the ticket, classifies it by urgency, and provides customer support specialist potential list of solutions.	Depending on the complaint, Generative AI system performs steps to assist the customer support specialist in resolving the issue (e.g., surfacing the most relevant documentation, updating internal systems, and more.)	Generative AI system drafts a ticket resolution response and addresses customer complaint.
 Human in the loop	Customer support specialist vets the output from the AI system and acts on next steps.	Customer support specialist supervises, edits, and updates any of the proposed actions by the generative AI system.	Customer support specialist oversees the response and provides final sign-off to

4

Organizations Capitalize on Smaller, Specialized Large Language Models

TOOLS & TECHNOLOGY

The past few years have marked the rise of massive, pre-trained language models capable of generalizing across various domains. Yet, as the open source ecosystem for generative AI matures, a new trend is emerging: the advent of smaller, domain-specific large language models. These specialized models, such as BloombergGPT, Replit's Code v1-3B, and Med-PaLM 2, offer a tailored approach to AI, focusing on specific sectors or tasks.

These domain-specific models come with a host of advantages. Firstly, they provide higher accuracy for tasks within their specialized areas. For example, Med-PaLM 2 is the first large language model to perform at an [expert test-taker level performance](#) on medical benchmarks. Secondly, these models are less costly to train, as dataset sizes are much smaller than more “traditional” LLMs. Additionally, these models grant organizations more control over their usage, allowing for a more customized application of AI technology.

As Tom Davenport, Distinguished Professor at Babson College, discussed on [DataFramed](#),

“I think in the next 12 months, there will probably be millions of customized, fine-tuned models. Just think about the legal profession. There are differences between US law, European law, and UK law. So they will be different models for different countries, regions, domains, and more.”

“I think in the next 12 months, there will probably be millions of customized, fine-tuned models. Just think about the legal profession. There are differences between US law, European law, and UK law. So they will be different models for different countries, regions, domains, and more.”

Thomas Davenport, Distinguished Professor at Babson College



In 2024, this trend is expected to gain even more momentum. Organizations will likely invest more in developing and utilizing domain-specific models, recognizing their value in precision, cost-effectiveness, and control.

The rise of small domain-specific LLMs



Law

ChatLaw, an open source LLM fine-tuned on legal data.

[Read More](#)



Medicine

Med-PaLM 2, an LLM fine-tuned on medical data developed by Google.

[Read More](#)



Finance

BloombergGPT, an LLM developed by Bloomberg specifically trained on financial data.

[Read More](#)



Coding

Replit Code v1-3B, an LLM developed by Replit specifically trained on coding data.

[Read More](#)

LEARN MORE ABOUT SPECIALIZED LLMs

[Is Data Science Still the Sexiest Job of the 21st Century?](#) — Learn from Tom Davenport, Distinguished Professor at Babson College, on the future of LLMs within organizations.

[Buy or Train? Using Large Language Models in the Enterprise](#) — Learn how to approach building specialized LLMs from Hagay Lupesko, VP of Engineering at MosaicML.

[Fine-Tuning LLaMA 2: A Step-by-Step Guide to Customizing the Large Language Model](#) — Learn how to fine-tune existing open-source AI models in this accessible tutorial.



The Risks of Generative AI Start Bubbling to the Surface

SOCIETY & CULTURE

Although 2023 was rich in discussions about AI safety and existential risk, these conversations largely remained in the realm of theoretical debate. A few high-profile incidents highlighted the risks of generative AI, such as [data leakage](#), [prompt injections](#), and [hallucinations](#). Luckily, these occurrences were relatively isolated and minor.

However, as generative AI systems become more robust and capable, so does their ability to generate harm. These harms can range from the amplification of cyber threats to the proliferation of misinformation and the amplification of bias. For example, Dr. Joy Buolamwini, acclaimed author and President of The Algorithmic Justice League, mentioned on [DataFramed](#): “We’re building AI systems interacting with humans and trained on human language. This language contains so many of our biases. Being aware of that allows us to build better technical systems because technical systems are never divorced from us.”

“We’re building AI systems interacting with humans and trained on human language. This language contains so many of our biases. Being aware of that allows us to build better technical systems because technical systems are never divorced from us.”

Dr. Joy Buolamwini, Artist-in-Chief and President of The Algorithmic Justice League

 Listen Here

Experts are raising alarms that 2024 could be a critical year for [AI-generated misinformation](#), particularly with a contentious US Presidential election on the horizon, as noted in [Air Street Capital’s State of AI Report](#). This looming threat underscores the growing importance of cultivating good digital citizenship and bolstering our defenses against AI-generated misinformation through enhanced data and AI literacy. As we advance into 2024, the focus must shift to celebrating the advancements of generative AI and proactively addressing and mitigating its potential risks.

LEARN MORE ABOUT AI RISKS



[Fighting for Algorithmic Justice](#) — Learn from Dr. Joy Buolamwini, Artist-in-Chief and President of The Algorithmic Justice League, on the ethical considerations of generative AI.

[Is AI an Existential Risk?](#) — Learn from Trond Arne Undheim, Research Scholar in Global Systemic Risk at Stanford University, on how AI can pose an existential risk to humanity.

[Building Trustworthy AI](#) — Learn from Beena Ammanath, Global Head of the Deloitte AI Institute, on how organizations can approach building trustworthy AI experiences

6

Generative AI Fundamentally Lowers the Barrier to a Data Culture

DATA CULTURE

Generative AI is proving to be exceptionally adept at data-related applications, especially in the realms of exploratory data analysis and coding tasks. Tools like ChatGPT for Data Analysis, [DataCamp Workspace](#), and GitHub Copilot showcase generative AI's ability to reduce complex data queries and coding tasks into natural language queries. These advancements are not just about automating tasks; they radically reduce the barrier to interacting with data. Generative AI's ability to understand and execute natural language queries has made data analysis more intuitive and accessible, even for those without extensive technical expertise.

This evolution could be foundational in achieving the “Holy Grail” of data transformation: building a [data culture](#). Integrating generative AI interfaces into coding-based and low-code platforms like DataCamp Workspace, Excel, Power BI, Tableau, ThoughtSpot, and more makes these tools more accessible to wider audiences within any organization. This is the essence of what [Valerie Logan, CEO of the Data Lodge, refers to as “Data Literacy by Design,”](#) where data tools are designed with enabling data literacy in mind.

“More and more data tools are allowing for these augmented capabilities, things like not allowing time series data into pie charts or providing tool tips and hovers that explain data insights. This is what our community defines as data literacy by design”

Valerie Logan, CEO of the Data Lodge

 Watch Here

As we progress into 2024, the widespread adoption of these AI-enabled tools is expected to significantly lower the barriers to engaging with data, enabling more people within organizations to leverage data for insights and decision-making. However, with great power comes great responsibility. The responsibility of critiquing outputs, methodologies, and potential hallucinations. This is why data literacy and conceptual understanding of data science remain crucial, ensuring that data insights are not unquestioningly believed but intentionally curated and understood.

LEARN MORE ABOUT DATA TOOLING INNOVATION



[Fostering Confidence with Data Across Your Organization](#) — Learn from Valerie Logan, CEO of the Data Lodge, on the intersection of tooling and data culture.

[Analyzing Top Runner Performance from A to Z with AI using Workspace](#) — Analyze a dataset fully using the DataCamp Workspace AI Assistant.

[Using ChatGPT's Advanced Data Analysis](#) — Learn how to use ChatGPT's Advanced Data Analysis Capabilities in this easy to follow code-along.

The AI Boom Intensifies Data Governance Efforts

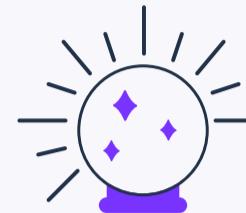
TOOLS & TECHNOLOGY

DATA CULTURE

As we venture into 2024, generative AI is set to become increasingly ubiquitous across organizations. As we've explored throughout our predictions, this widespread adoption will occur not only through the use of AI tools like ChatGPT but also via the training and deployment of large language models that interface directly with organizational data.

Moreover, given the opportunity of leveraging LLMs to analyze unstructured data, [Forrester](#) predicts that the volume of unstructured data managed by enterprises is expected to double in 2024. This surge in generative AI applications underscores the critical importance of high-quality data. As [Deloitte](#) succinctly points out:

"Generative AI is only as good as your data quality."



"Generative AI is only as good as your data quality."

Deloitte's Data governance perspectives on Generative AI

[Read More](#)

As organizations increasingly rely on unstructured data, there's a growing need to build a culture emphasizing the governed use of data. Evolving data governance needs will make organizations reassess and reinforce their data governance frameworks. In 2024, we anticipate a significant uptick in investments in data governance and quality initiatives. These efforts are not merely supportive but foundational to the success of generative AI projects.

[LEARN MORE ABOUT DATA GOVERNANCE](#)



[**Laying the Foundations: Data Quality in the Age of AI**](#) — Learn from Susan Walsh, the Classification Guru, and Scott Taylor, the Data Whisperer, on the importance of data quality in the age of generative AI.

[**Data Governance Cheat Sheet**](#) — Get the download on data governance in this accessible cheat sheet

[**Introduction to Data Quality**](#) — Learn about the importance of data quality in this 2-hour conceptual course.

Data & AI Literacy Become Table Stakes to Succeed with Generative AI

DATA & AI SKILLS

As we have previously discussed, generative AI holds immense potential for democratizing data-driven decision-making. However, one of the critical challenges with generative AI in any domain is the risk of hallucinated insights. This risk highlights the importance of data and AI literacy. Understanding how to effectively use these systems and accurately interpret their outputs, especially in the context of data work, becomes foundational.

Despite the increased focus on data and AI literacy, there is still a significant gap in education and training. According to [Randstad](#), while 50% of employees believe AI will fundamentally change their workflows, only 13% have received AI training. Moreover, according to [New Vantage Partners](#), only 1.6% of CDOs cited data literacy as their top investment priority.

In 2024, building robust data and AI literacy capabilities will be essential for organizations. It will serve as an insurance policy against the risks and inaccuracies associated with generative AI and foster better and broader adoption of these technologies across various sectors.

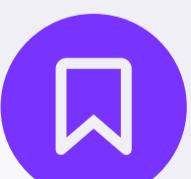
As [Cindi Howson, Chief Data Strategy Officer at ThoughtSpot](#) notes on DataFramed “AI literacy is understanding what the training data is, and then how is that being used to generate the output of an AI-generated insight. So even if you think we know generated AI has a problem with hallucinations, understanding what a hallucination is and why that may happen is foundational for AI literacy.”

“AI literacy is understanding what the training data is, and then how is that being used to generate the output of an AI-generated insight. So even if you think we know generated AI has a problem with hallucinations, understanding what a hallucination is and why that may happen is foundational for AI literacy.”

Cindi Howson, Chief Data Strategy Officer at ThoughtSpot



LEARN MORE ABOUT DATA & AI LITERACY



[Empowering Government with Data & AI Literacy](#) — Learn from Richard Davis, Chief Data Officer at Ofcom, on the importance of driving data & AI literacy within government.

[From Data Literacy to AI Literacy](#) — Learn from Cindi Howson, Chief Data Strategy Officer at ThoughtSpot, on how organizations can go from data literacy to AI literacy.

[What is Data Literacy? A Comprehensive Guide for Organizations](#) — Read about data literacy in this accessible guide.

Prompt Engineering Will Become a Valuable Skillset Across Many Roles, but Not a Career Path

DATA & AI SKILLS

Throughout 2023, there was considerable debate about the future of prompt engineering as a distinct career path. While it's still uncertain whether it will emerge as a standalone profession, it's becoming increasingly clear that prompt engineering skills are in high demand across various roles. According to [Forrester](#), an estimated 60% of employees are expected to undergo training in prompt engineering, indicating its rising importance in the workforce.

As we move into 2024, prompt engineering is set to become a highly sought-after skill set. However, it will likely be sought after within existing job roles rather than alone as a separate role. For instance, engineers working with large language models (LLMs) must be proficient in prompt engineering to deploy and manage these AI systems effectively. Similarly, marketers will find prompt engineering invaluable for creating more efficient and impactful marketing workflows. The skillset will be essential for various professionals interacting with AI daily, making it a versatile and crucial component of various job functions.

AN You

You are a data science expert; follow my instructions carefully! Provide step-by-step reasoning of why you're making your decisions! Help me explore this data and provide insights.

Message ChatGPT...

ChatGPT can make mistakes. Consider checking important information.

AN You

You are a copywriter; follow my instructions carefully! Provide step-by-step reasoning of why you're making your decisions! Create a LinkedIn post about the 2024 data & AI trends and predictions.

Message ChatGPT...

ChatGPT can make mistakes. Consider checking important information.

AN You

You are a writing expert; follow my instructions carefully! Provide step-by-step reasoning of why you're making your decisions! Create a draft outline for a blogpost on the 2024 data trends & predictions.

Message ChatGPT...

ChatGPT can make mistakes. Consider checking important information.

LEARN MORE ABOUT PROMPT ENGINEERING



[What is Prompt Engineering? A Detailed Guide](#) — Learn more about prompt engineering in this article.

[A Beginner's Guide to Prompt Engineering with ChatGPT](#) — Get the basics of prompt engineering with ChatGPT.

[ChatGPT Prompt Engineering for Developers](#) — Dive deep into the principles and best practices of prompt engineering to leverage powerful language models like ChatGPT to solve real-world problems.

Generative AI Will Focus Data Professions On Value Creation

DATA & AI SKILLS

The transformation of coding workflows by generative AI is reshaping the landscape for data professionals.

Anaconda's [2023 State of Data Science report](#) indicates a significant uptick in the use of generative AI among data science practitioners, with 63% using it as much or more in 2023 compared to the previous year. This evolution is exerting pressure on roles traditionally centered around coding, where proficiency in syntax and technical skills was paramount.

Consequently, there's a growing sense of concern among data professionals about job security, with 45% expressing apprehension about their future in the field. However, the evolving landscape also points to a necessary shift in the skillset required for success in data roles. The most successful data professionals are not merely those with strong technical skills; they are increasingly those who excel in driving value from data.

In 2024, this shift will become more pronounced. Data professionals will build on top of their technical skills and focus more and more on the importance of skills such as identifying and applying the right solutions to specific problems, project management, autonomy, and effective data storytelling. As [Cassie Kozyrkov, CEO at Data Scientific and Google's first Chief Decision Scientist](#) puts it on DataFramed "If you don't have someone who actually understands the business, understands the strategy, helping you work on the right things rather than the wrong things, chances are you won't be very impactful".

"If you don't have someone who actually understands the business, understands the strategy, helping you work on the right things rather than the wrong things, chances are you won't be very impactful"

Cassie Kozyrkov, CEO at Data Scientific and Google's first Chief Decision Scientist



Key skills to stand out as a data practitioner in a generative AI world



Problem-solving



Project management



Autonomy and collaboration



Communication and storytelling

LEARN MORE ABOUT PROMPT ENGINEERING

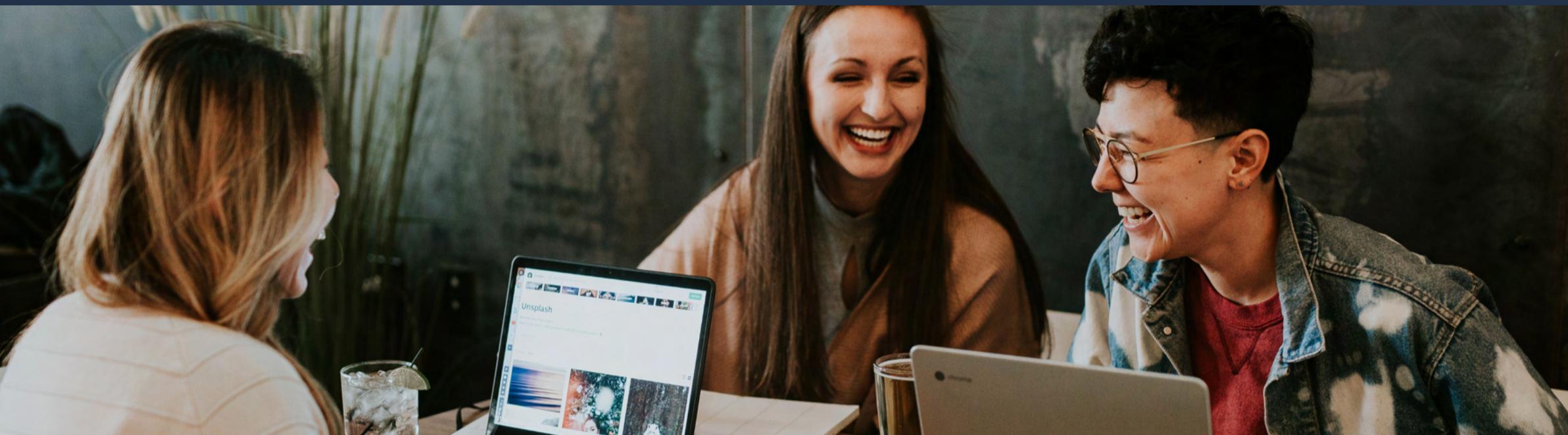


[Making Better Decisions using Data & AI](#) — Learn more from Cassie Kozyrkov, Google's First Chief Decision Scientist, on how data & AI will transform data careers.

[GPT and Generative AI for Data Teams](#) — Learn from Sarah Schlobohm, Head of Artificial Intelligence at the Citation Group, on how AI will impact data teams.

[A Guide to Using ChatGPT For Data Science Projects](#) — Learn how to apply generative AI on data science projects in this accessible tutorial.

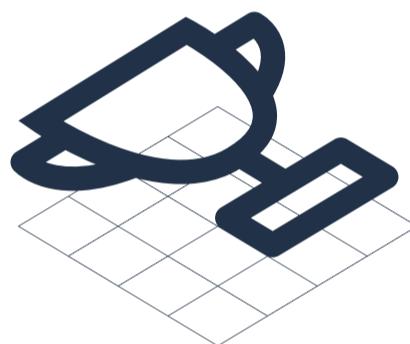
Future-proof your business with DataCamp



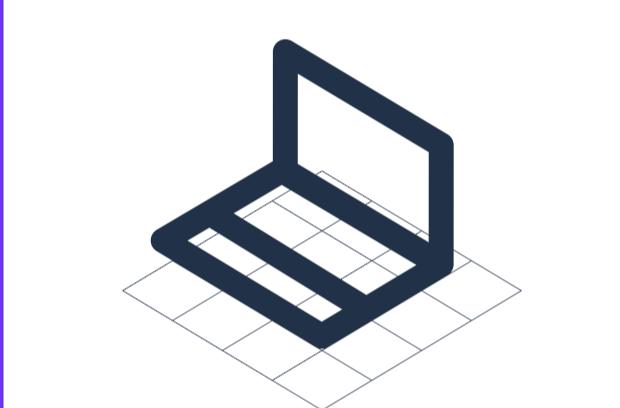
Best in class learning

Build work ready skills

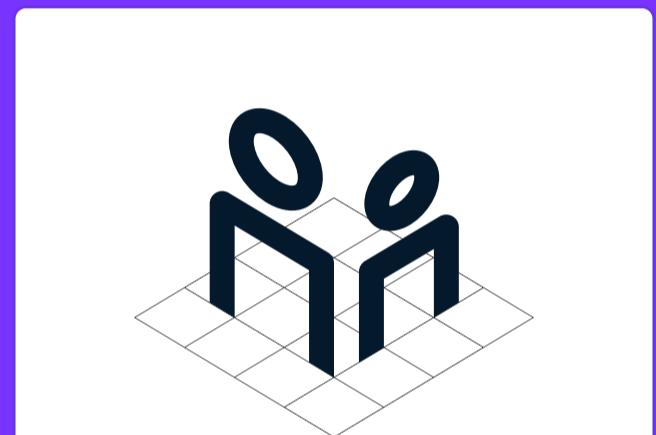
Certify your team



Give your employees access to market-leading training with DataCamp Learn



Apply your skills in a risk-free online coding environment with DataCamp Workspace



Cement your team's learning and help your teams grow with industry-leading data certifications

Trusted by more than 12 million learners and over 4,000 data-driven companies

Google

Microsoft

ebay

T-Mobile...

CREDIT SUISSE

PayPal

Uber



HSBC



Mercedes-Benz



BNP PARIBAS



Deloitte.

Thank you for reading.

Are you an organization interested in scaling your data literacy skills?
Book a meeting with our sales team today.

[Book a Meeting](#)



About DataCamp

DataCamp is a one-stop shop for building a data-driven workforce. With DataCamp for Business, you can transform how everyone in your organization uses data. With more than 450+ interactive, self-paced online courses, a cloud-based IDE that connects to your company data, and a recruitment platform for hiring data talent, DataCamp for Business will help you make the most of your organization's data.

