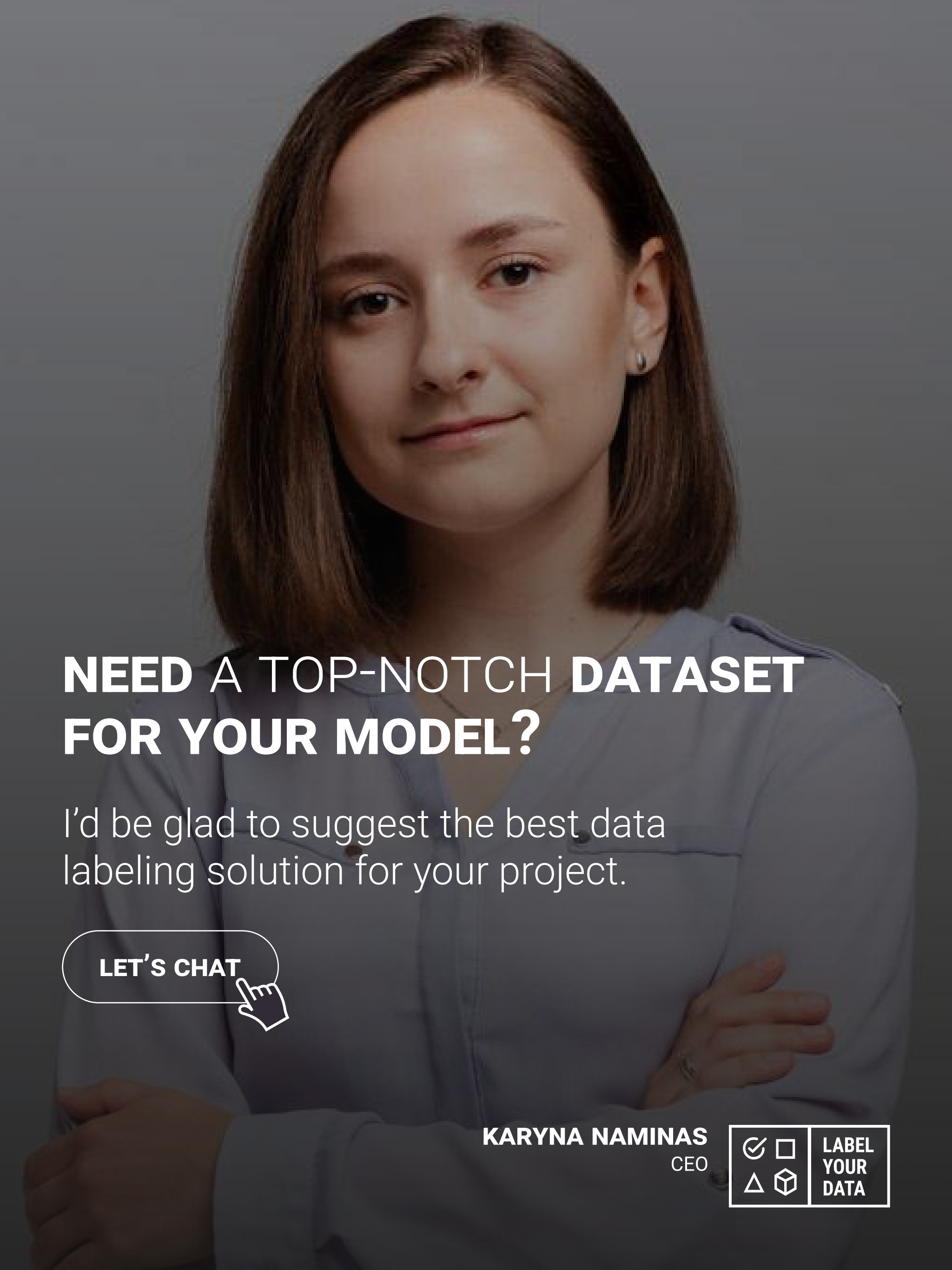
LABEL YOUR DATA

TOP 2024 AI TRENDS

IN THIS GUIDE

LABEL YOUR DATA



NEED A TOP-NOTCH **DATASET FOR YOUR MODEL?**

I’d be glad to suggest the best data labeling solution for your project.

LET’S CHAT

KARYNA NAMINAS
CEO

LABEL YOUR DATA

01

2023's ROUNDUP: THE BREAKOUT YEAR FOR GENAI

LABEL YOUR DATA

2023 has been marked as the “year of generative AI.” The technology has seized the global AI industry at a lightning fast pace.

In the beginning of 2023, we didn’t know much about this intricate technology. Today, by (almost) the end of the year, generative AI solutions have become a crucial part of our work and everyday lives.

Generative AI tools

Scribe

dCODE

GitHub Copilot

OpenAI ChatGPT 4.0

Bard

cohere

DALL·E 2

CLAUDE 2

synthesia

Sophisticated ML models powering the GenAI technology took a long time to develop. Since the debut of ChatGPT in late 2022, there have been frequent updates, with March 2023 alone witnessing six noteworthy improvements. These include **enhanced customer relationship management tools and improved support for the financial services industry.**

As you’ve probably tested yourself, current GenAI tools can produce a wide range of content. Anything from written or audio content, to images and video art. McKinsey predicts that by the end of this decade, GenAI will be about as good as the average person at these tasks. It’s expected to **compete with the top 25%** of people in various undertakings before 2040, which is way faster than experts thought. Sometimes by **up to 40 years.**

60%

of workers will use their own AI to perform their job and tasks.

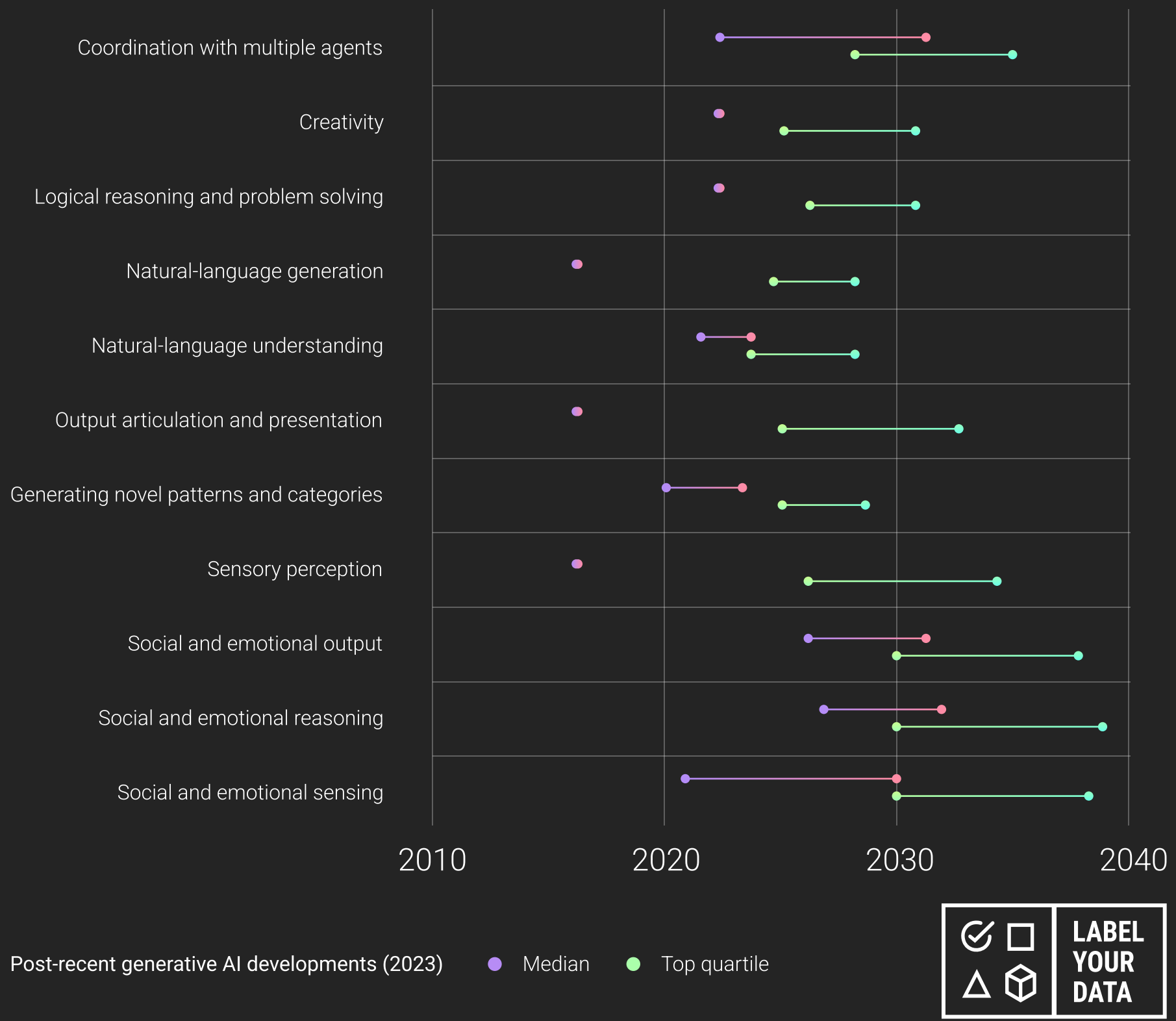
LABEL YOUR DATA

Automation has been a dream for ages. Generative AI is changing the game for knowledge-based jobs in fields like **education, law, technology, and the arts.** Some parts of these jobs might get automated sooner than we thought.

Many industries are already benefiting and will keep benefiting from using generative AI in their **marketing and sales**. More specifically, leaders in these business development niches already see a big difference in:

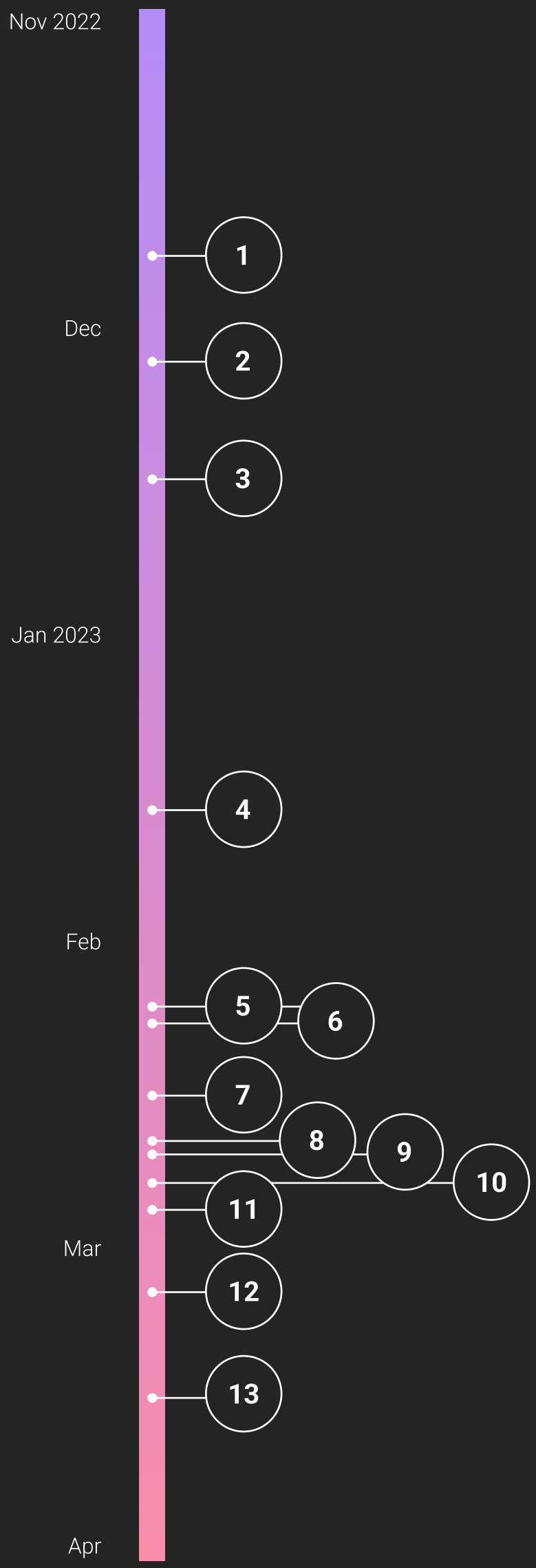
- Lead identification,
- Marketing optimization,
- Personalized outreach.

Estimated range for technology to achieve human-level performance, by technical capability



LABEL YOUR DATA

Timeline of major large language model (LLM) developments following Chat GPT's launch



1. **Nov 30, 2022:** OpenAI's ChatGTP, powered by GTP-3.5, becomes the first widely used text-generated product, gaining a record 100 million users in 2 month
2. **Dec 12:** Cohere releases the first LLM that supports more than 100 languages, making it available on its enterprise AI platform
3. **Dec 26:** LLMs such as Google's Med-PaLM are trained for specific use cases and domains, such as clinical knowledge
4. **Feb 4, 2023:** Amazon's multimodal-CoT model incorporates "chain-of-thought-prompting", in which the model explains its reasoning, and outperforms GTP-3.5 on several benchmarks
5. **Feb 24:** As a smaller model, Meta's LLaMA is more efficient to use than some other models but continues to perform well on some tasks compared with other models
6. **Feb 27:** Microsoft introduces Kosmos-1, a multimodal LLM that can respond to image and audio prompts in addition to natural language
7. **Mar 7:** Salesforce announces EinsteinGPT, the first generative AI technology for customer relationship management
8. **Mar 13:** Open AI releases GPT-4, which offers significant improvements in accuracy and hallucinations mitigation, claiming 40% improvement vs GPT-3.5
9. **Mar 14:** Anthropic introduces Claude, an AI assistant trained using a method called "constitutional AI", which aims to reduce the likelihood of harmful outputs
10. **Mar 16:** Microsoft announces the integration of GPT-4 into its Office 365 suite, potentially enabling broad productivity increases
11. **Mar 21:** Google releases Bard, an AI chatbot based on the LaMDA family of LLMs
12. **Mar 30:** Bloomberg announces a LLM trained on financial data to support natural-language tasks in the financial industry
13. **Apr 13:** Amazon announces Bedrock, the first fully managed service that makes models available via API from multiple providers in addition to Amazon's own Titan LLMs

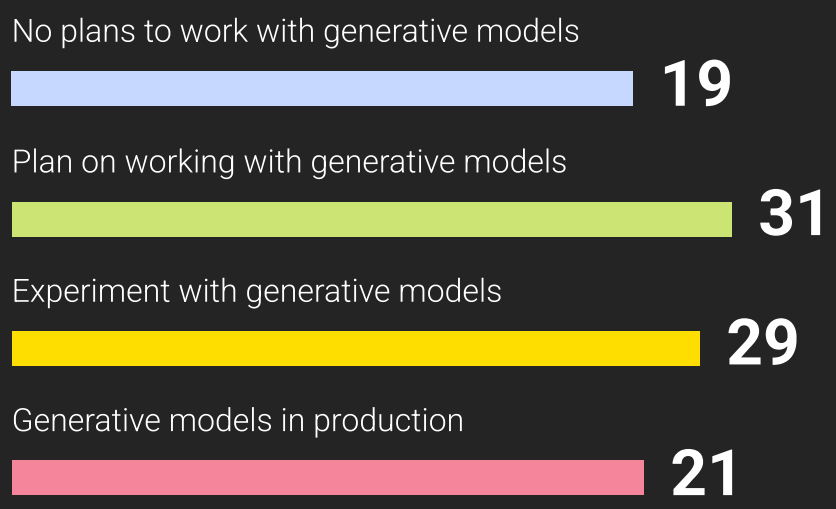
High-tech and banking, especially, are in for some big changes because GenAI makes software development faster. Case in point, 40 developers from McKinsey personally evaluated tools based on generative AI and observed significant improvements in the speed of various routine developer activities.

The potential of generative AI is exciting. However, it poses risks like biased or inaccurate content. Thus, organizations need to be cautious about reputational and legal issues. To address them, it's wise to have a "human in the loop" — ensuring a real person reviews GenAI output before it's used.

As organizations pursue advanced AI goals, they require more employees who understand this technology.

Despite the growing value of GenAI tools, there's a significant shortage of skilled workers. To address this gap, companies should prioritize effective talent management. This will create positive work experiences for attracting and retaining employees familiar with AI.

Which of the following describes how your company works with generative models?, %

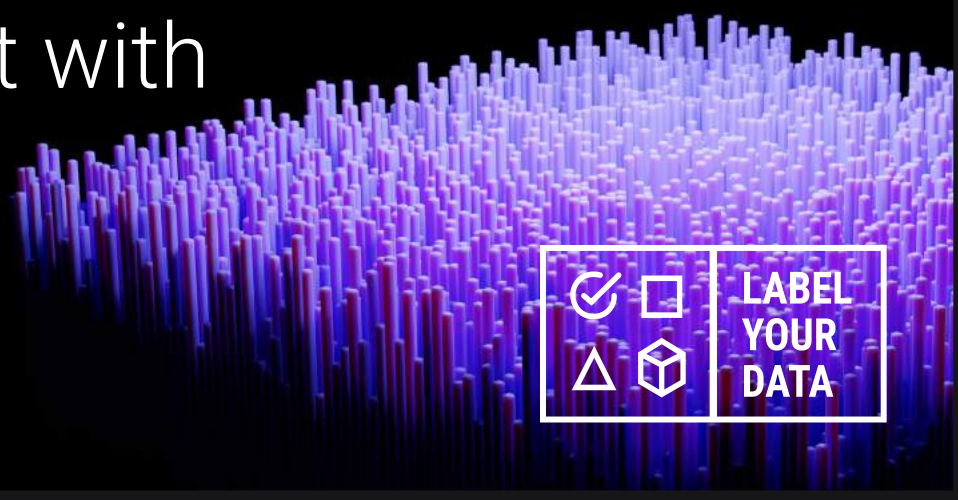


65%

either accelerated their existing strategies or created an AI strategy for the first time

Secure your ML project with Label Your Data.

RUN YOUR FREE PILOT

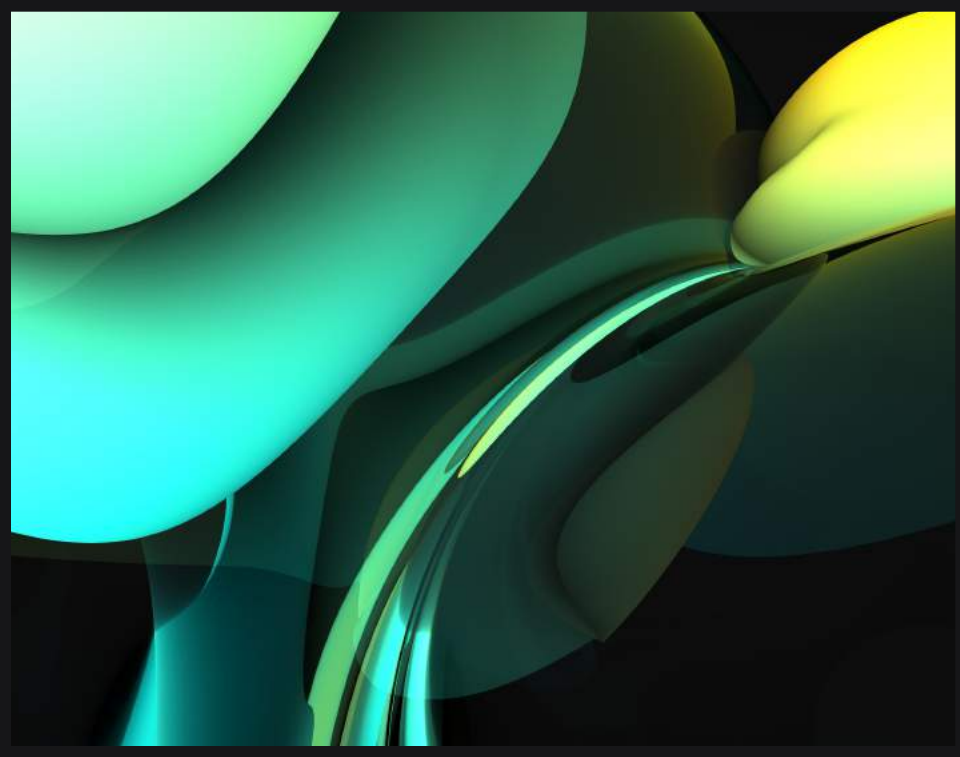


02

TOP 5 AI & ML TRENDS SHAPING 2023 AND THE YEARS AHEAD

LABEL YOUR DATA

The Label Your Data team has handpicked and analyzed the top 5 AI and ML trends that are shaping the current landscape of 2023 and influencing the industry's future! Let's explore them:



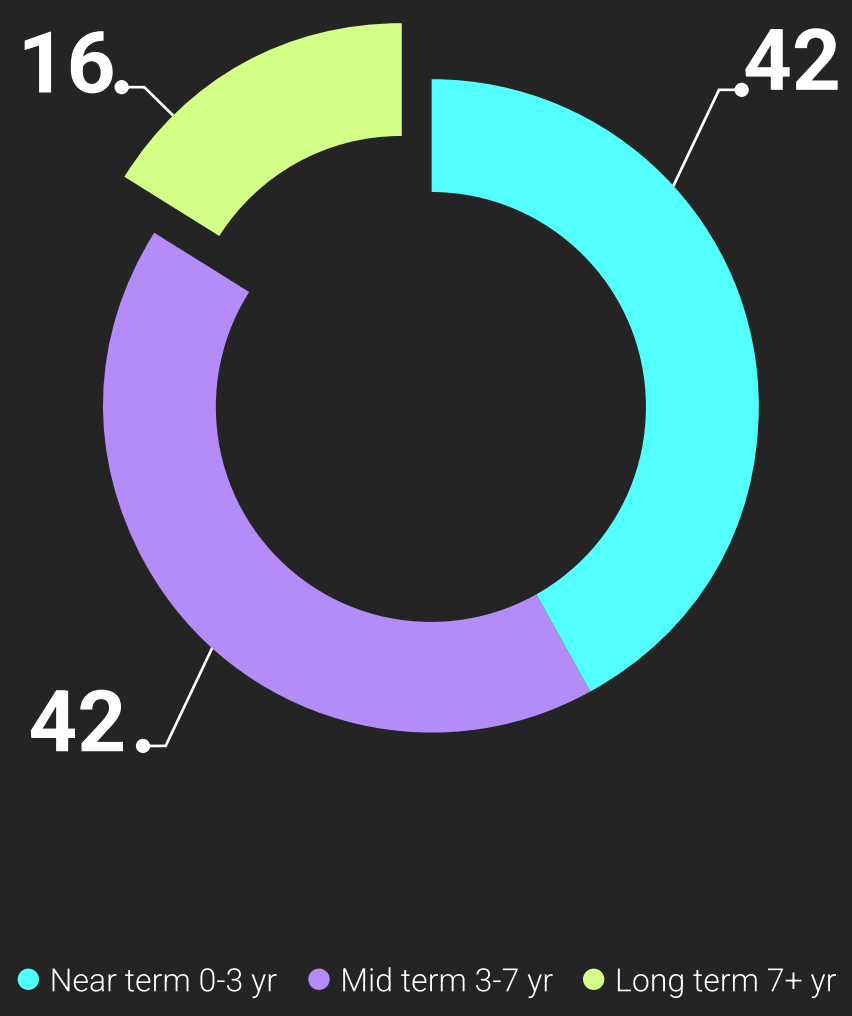
AI impact scores range from 1-5

(1 being lowest impact, 5 being highest)

Retail	
Personalisation	2.6
Utility	3.3
Data available	3.7
Time saved	2.0
Potential AI consumption Impact	3.0

Estimated use case adoption time frames

as a % of total



LABEL YOUR DATA

TREND 1: GENERATIVE AI FOR A GLOBAL IMPACT

Generative AI got a lot of attention in 2023 and is still in the process of development. This year, many new startups were using this technology to come up with innovative ideas, and we expect more of them to appear in the future.

According to Gartner, there are two facets to the GenAI movement: advancements that will be powered by generative AI and innovations that will contribute to the progress of generative AI.

Innovations that will contribute to the progress of GenAI

- AI simulation
- AI trust, risk and security management (AI TRiSM)
- Causal AI
- Data labelling and annotation (DL&A)
- First-principles AI (FPAI)
- Foundation models
- Knowledge graphs
- Multiagent systems (MAS)
- Neurosymbolic AI
- Responsible AI

Advancements that will be powered by GenAI:

- Artificial general intelligence (AGI)
- AI engineering
- Autonomic systems are self-managing physical or software systems
- Cloud AI services
- Composite AI
- Computer vision (CV)
- Data-centric AI
- Edge AI
- Intelligent applications
- Model operationalization
- Operational AI systems (OAISys)
- Prompt engineering
- Smart robots
- Synthetic data

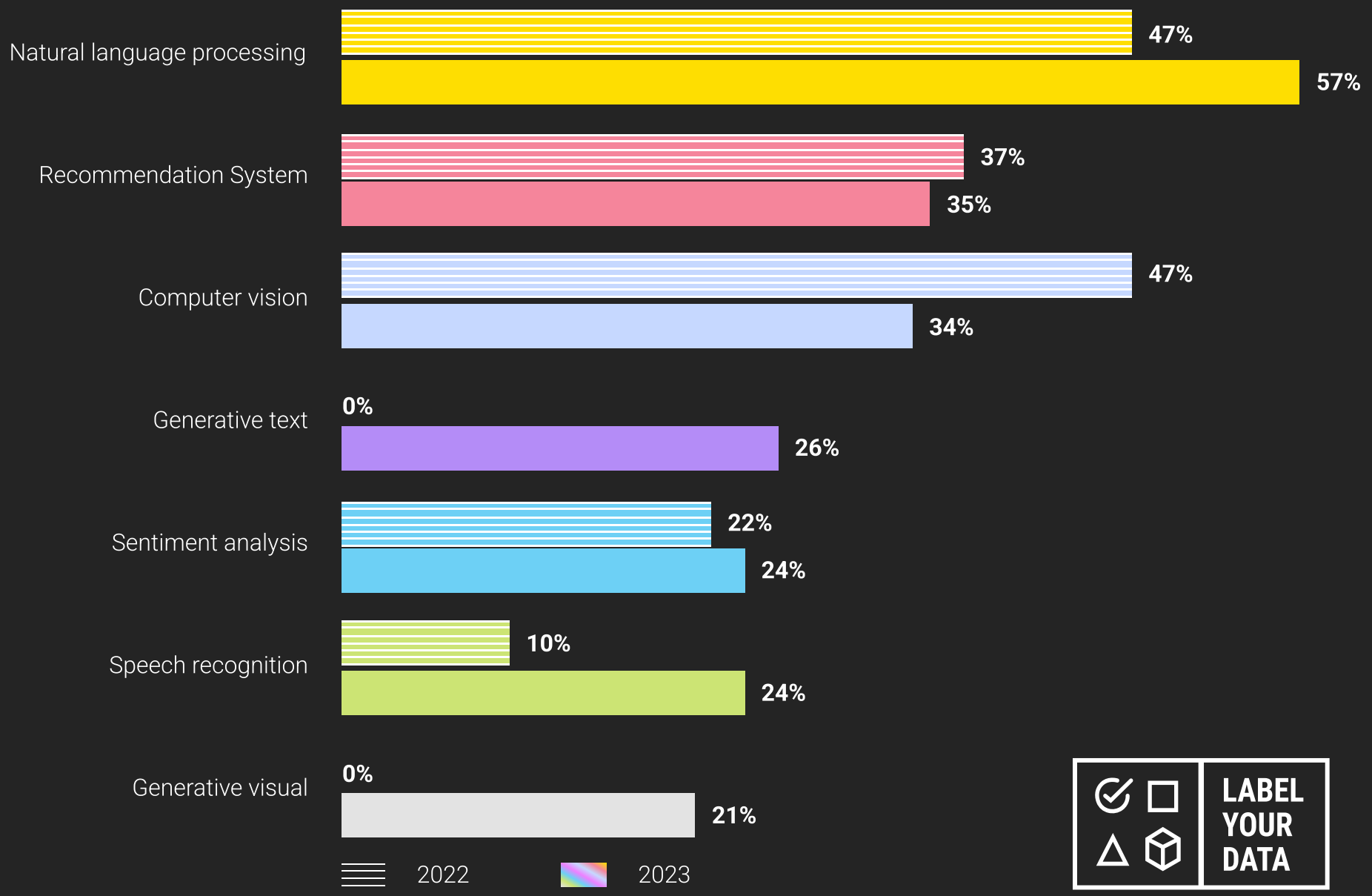
GenAI boosts global productivity, but workers need support in transitioning to new roles for sustainable, inclusive growth.

LABEL YOUR DATA

TREND 2:
THE RISE OF LARGE LANGUAGE MODELS (LLMS)

Natural Language Processing, or simply NLP, is a valuable technology that allows human-machine communication through the analysis of language patterns. Throughout 2023, NLP systems were the backbone of daily-used tech products, from search engines to voice assistants. Grand View Research projects a 40.4% CAGR for the NLP market, expecting it to reach $439.85 billion by 2030. In particular, LLMs contributed to the advancement of NLP this year by providing powerful tools for processing and generating human language. These are deep learning (DL) models that revolutionize our interaction with devices, websites, and information at large.

What type of ML system do you work on?



LABEL YOUR DATA

Since GPT-3's introduction in 2023, LLMs have seen substantial growth, ranking among the **top 14% of emerging global technologies.**

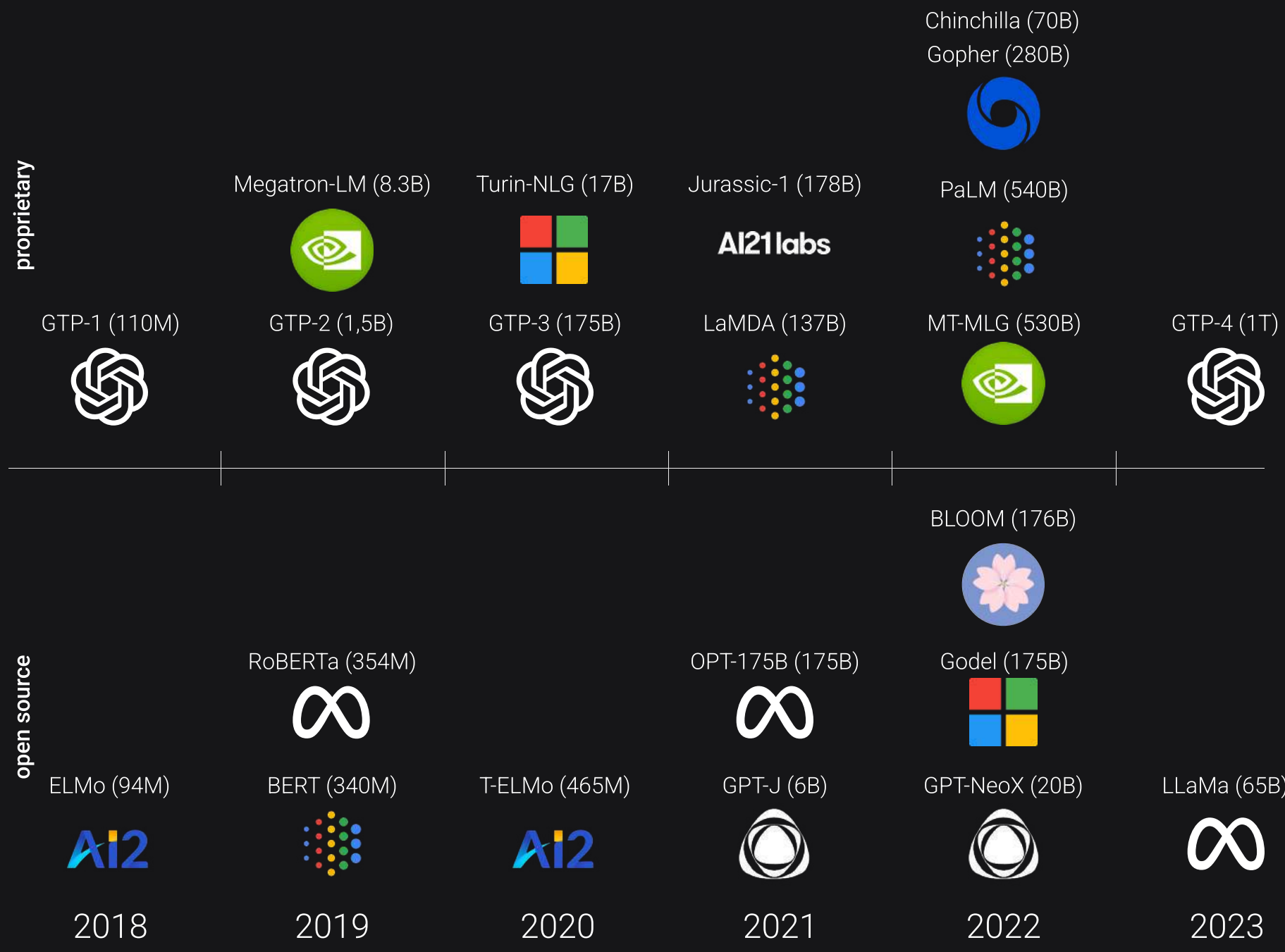
With a **trend magnitude of 85.05% and maturity of 24.5%**, these DL models hold strong potential for further impact.

LLM companies are actively involved in other **trends** like:

1. Neural Machine Translation
2. Voice Intelligence
3. Transfer Learning
4. Machine Learning as a Service
5. Automated Speech Recognition

In 2023, LLMs' **media exposure soared by 30-fold**, securing **a top 5% position** among news topics.

Key investors include **Microsoft, Tiger Global Management, and Section 32.**



LABEL YOUR DATA

TREND 3:
AUTOMATION & SHIFTING WORK DYNAMICS

Forrester's latest research forecasts the potential replacement of 2.4 million jobs in the United States alone by GenAI by 2030. There are a few crucial reasons why this might happen in the near future.

01

AI IS A SIGNIFICANT PRODUCTIVITY BOOSTER

However, this might be both a blessing and a cure for the global tech (and not) industry. A more productive economy can result from industry shifts if labor rotation is managed effectively. This involves swiftly moving workers from basic manual labor roles to more productive and value-added industries. The success of this approach depends on several factors, including the number of displaced workers.

02

GENERATIVE AI WILL CHANGE THE WAY WE WORK

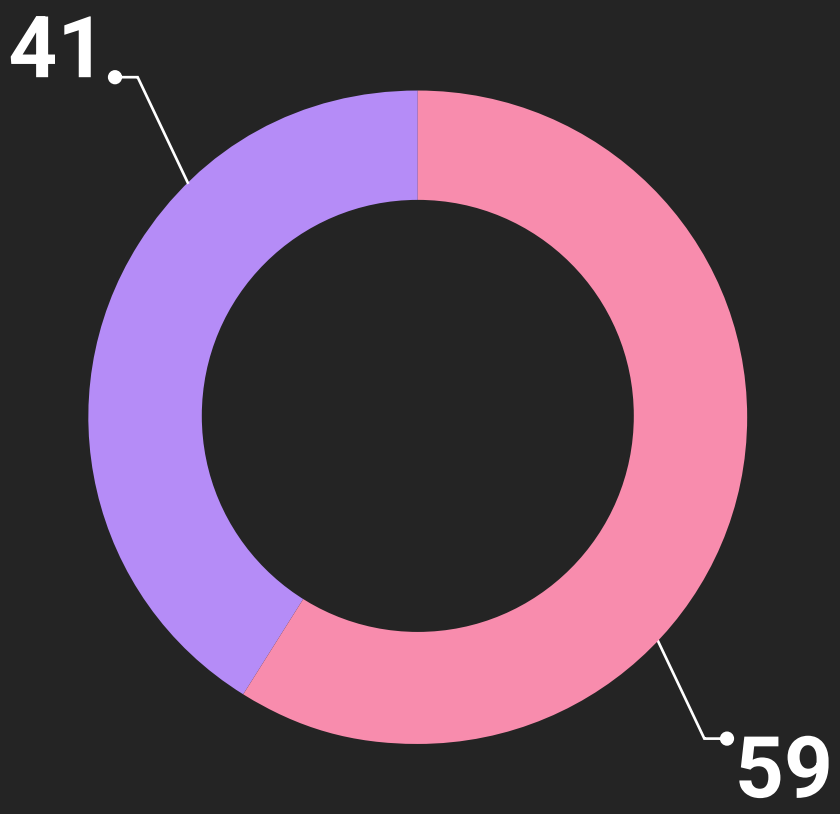
Nearly **80% of workers** expect that tools such as Microsoft Copilot, powered by generative AI, will **affect around 20 hours,** or half of their workweek. However, a majority (63%) recognize the need for **acquiring new skills** or an entirely fresh skill set **by the end of 2024** to fully leverage the benefits of GenAI.

LABEL YOUR DATA

TREND 4:
MORE STRINGENT DATA REQUIREMENTS FOR AI

AI needs data. And not any data, but high-quality, and often annotated data used for training advanced ML or DL models. Case in point, ChatGPT has been through complex processes of text data collection and annotation to serve as a valuable tool for around 180.5 million users.

Most important when developing & training ML models, %



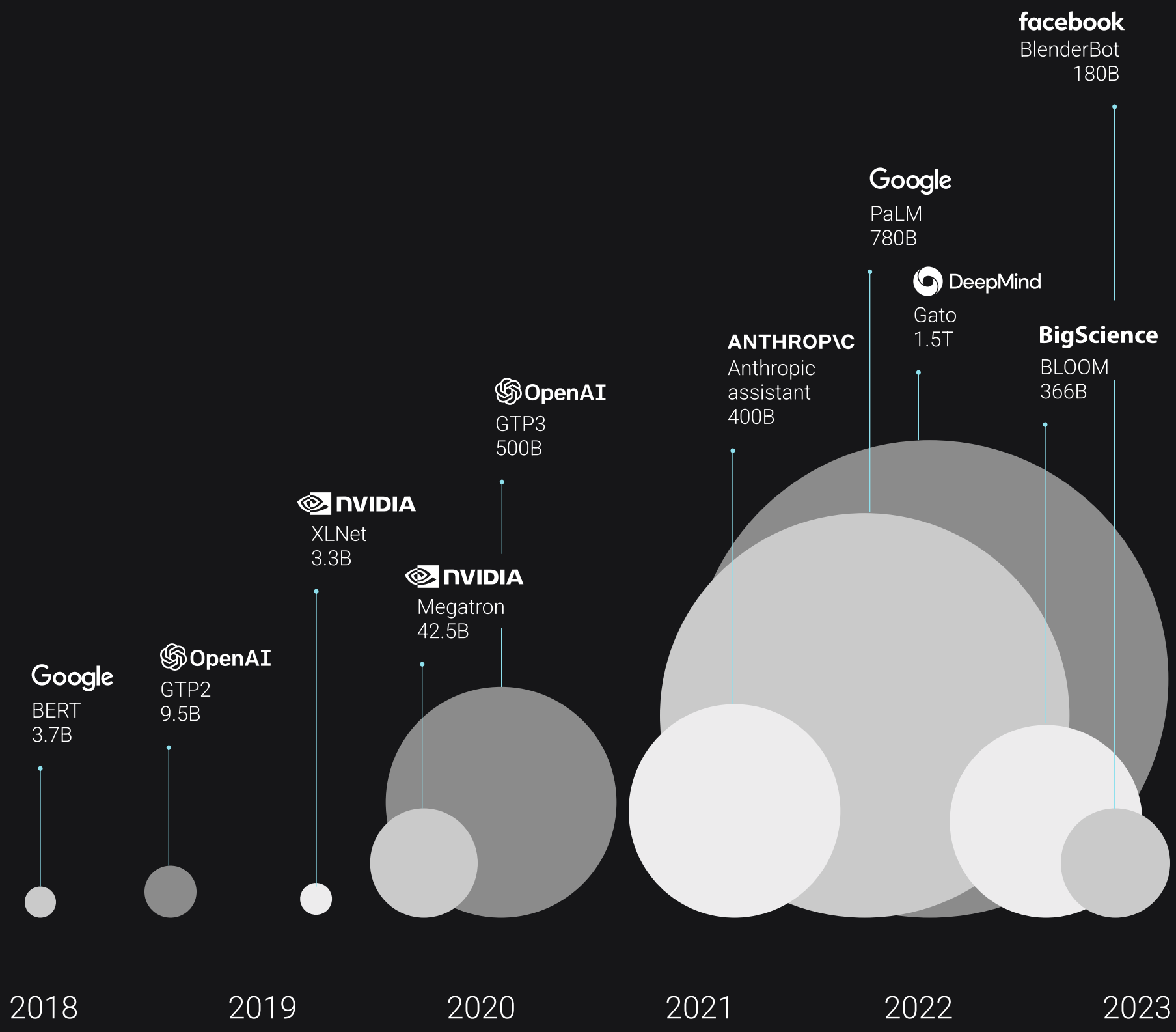
To handle innovation like GenAI, data management remains the pressing issue in 2023 and beyond for the majority of AI businesses (or those that depend on AI). Here are some aspects we considered when analyzing this trend:

- **Certain projects demand more precise data**, while others face challenges with their teams in effectively managing and organizing their existing enterprise data.
- **There's the overall lack of confidence in ML models.** Concerns typically center around the sustainability and scalability of these models. This often happens due to limited resources and a lack of thorough quality assurance.
- **Proprietary data makes generative models even more powerful.** Such models are useful by themselves, but when combined with a company's unique data, they stand out. This makes **customer experience, product development**, and **profits better.**

LABEL YOUR DATA

Models are increasing in size

model size in tokens



- BERT (2018): 3.7B tokens and 240 million parameters;
- GPT-2 (2019): 9.5B tokens and 1.5 billion parameters;
- GPT-3 (2020): 499B tokens and 175 billion parameters;
- PaLM (2022): 780B tokens and 540 billion parameters.

LABEL YOUR DATA

TREND 5:
ACCELERATED AI INVESTMENT & ADOPTION IN BUSINESS

With the increased capabilities and accessibility of GenAI, companies are swiftly integrating it into their core operations. Recognizing its pivotal role in their future, business leaders aim to adopt AI rapidly for maximum impact.

Over the next three years (including 2023), **72% of companies** plan to boost their AI investment annually.

In 2030, the potential contribution of AI to the global economy is estimated at **$15.7 trillion**. This is more than what China and India currently produce together.

Out of this, around **$6.6 trillion** is expected to come from getting more work done efficiently, and about **$9.1 trillion** from how people spend money differently because of AI.

Of companies making significant investments in AI

52% are in investing heavily in LLMs

36% are in investing heavily in generative visual models

30% are in investing heavily in computer vision models

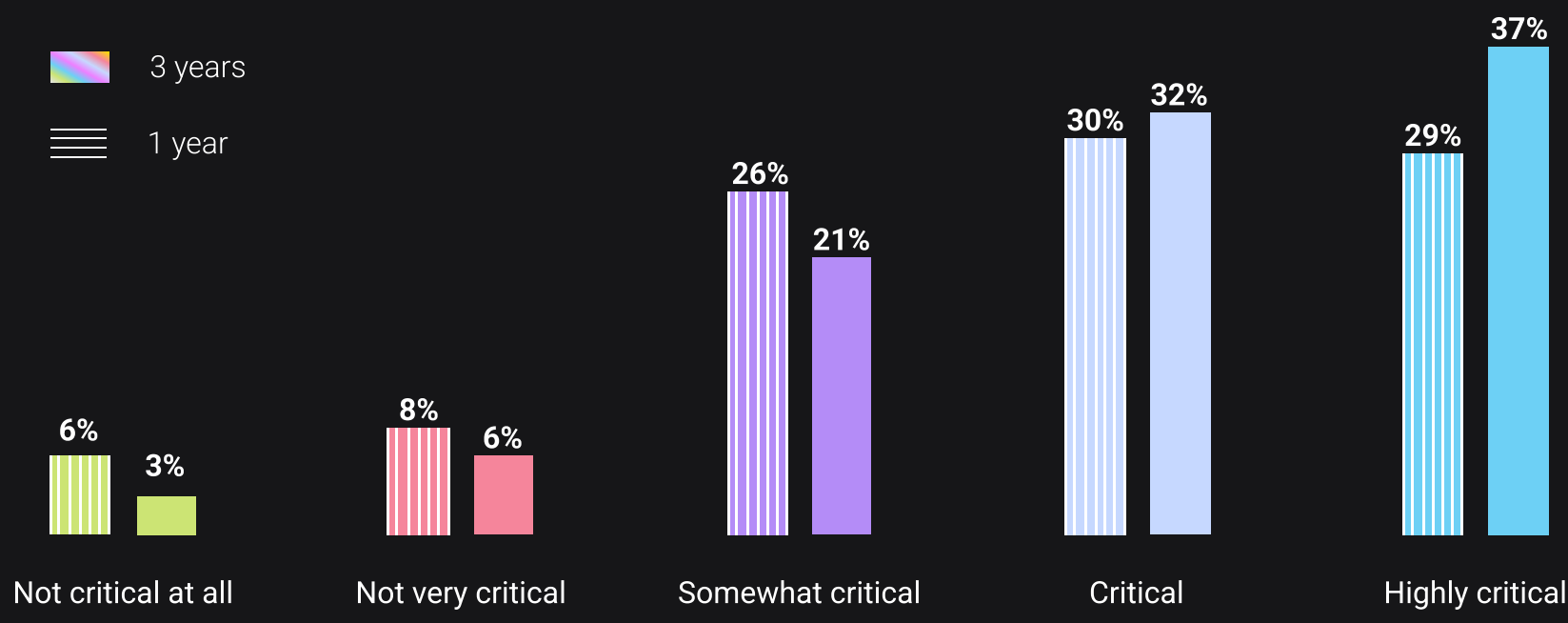
Organizations and industries are also rapidly increasing investments in AI.

Gartner predicts:

- Over $10 billion will be directed towards AI startups leveraging foundational models by the end of 2026.
- 45% of executive leaders acknowledged that the recent buzz around ChatGPT prompted an uptick in their AI investments.
- 70% of the same top management stated that their organizations are currently in the exploration phase with generative AI, while 19% have progressed to pilot or production stages.

LABEL YOUR DATA

How critical is AI to your business in the next 1-3 years?



The rise of powerful generative AI is making AI adoption faster:

- In 2023, 59% of companies deem AI crucial for their business.
- Over the next three years, this sentiment is expected to increase to 69%.

AI adoption by companies yields positive outcomes:

- Improved customer experiences,
- Refined product development,
- Enhanced collaboration across business functions.

What outcomes have you seen from AI adoption?

LABEL YOUR DATA

AI adoption by industry, top industries:

Every industry is looking to increase its AI budgets over the next 3 years.

80%	**Insurance**
79%	**Logistics & Supply chain**
77%	**Financial services**
75%	**Healthcare & Life sciences**
74%	**Retail & eCommerce**

Overcome dataset challenges with **Label Your Data** and our custom annotation solutions for ML model training!

GIVE US A CALL

LABEL YOUR DATA

03

AI & ML INDUSTRY INSIGHTS FOR 2024

LABEL YOUR DATA

A few AI predictions for 2024 in numbers:

- **60% of employees** will use their own AI for job-related tasks.
- **85% of companies** will adopt open source models to enhance their AI capabilities.
- By 2024, there will be more **AI voice assistants (8.4 billion)** than people on Earth.
- A leading insurance company is set to introduce a specialized policy covering the risks associated with **AI hallucinations** (when machines generate false or nonsensical information).
- By 2024, an estimated **60% of AI data is expected to be synthetic.**
- GenAI transformation will reshape IT operations, delivering up to **40% enhancements in software development.**
- Forward-thinking tech leaders will use resources wisely to bring out creativity in all IT roles, giving a **50% increase in time for creative problem-solving.**

Now, let's check out the major predictions for AI and ML in 2024 in more detail together with our Label Your Data team!

01 Next Gen GenAI

The next generation of generative AI goes beyond basic chatbots and memes. In 2024, the technology will be able to craft intricate narratives, orchestrate musical compositions, and potentially collaborate on bestselling novels.

A big step forward would be a **multi-modal generative AI,** which can combine text, voice, melodies, and visuals to make articles, images, music, and narrations in different languages.

This means we'll soon have AI that can create diverse content and immersive experiences, making it **harder to tell the difference between human and AI creations** as we head into 2024.

02 Conscientious AI

AI is becoming a big part of our lives, but it raises **ethical concerns.** It's not just about making AI unbiased. We also need rules to hold AI systems and their creators accountable.

As we head into 2024, there's a growing interest in teaching people about the ethics of AI. To use AI responsibly, businesses should consider factors like risk, trust, transparency, and accountability.

Gartner warns that by 2025, if 1% of AI vendors focus too much on pretrained AI models, responsible AI **could become a societal issue.** So, organizations are advised to adopt a risk-proportional approach and seek assurances from AI vendors to manage potential financial, legal, and reputational risks associated with AI implementation.

03 Next Gen GenAI

The trend for data-centric AI is a shift towards **a more focused approach on data** rather than just models and code. **AI-specific data management, synthetic data generation, and data labeling** will enhance AI systems by addressing challenges like accessibility, volume, privacy, security, complexity, and scope.

One rapidly growing area is the use of generative AI to create synthetic data, reducing the need for real-world data and making **ML model training more efficient.** By 2024, it's anticipated that 60% of data for AI will be synthetic, simulating reality and future scenarios to reduce risks.

04 Cloud Data Ecosystems

The landscape of data ecosystems is transitioning from standalone software or mixed deployments to complete cloud-native solutions. It's predicted that by 2024, **half of the new system deployments in the cloud will rely on a unified cloud** data ecosystem instead of manually integrated individual solutions.

According to Gartner, companies should check how well their data systems can handle challenges with scattered data and how easily they can connect to data sources from outside their usual setup.

05 Edge AI

Edge AI is on the rise as it facilitates the processing of data at the point of origin. This empowers businesses to obtain real-time insights, identify emerging patterns, and adhere to strict data privacy regulations.

Additionally, the technology contributes to enhancing the development, orchestration, integration, and deployment of AI for companies. According to Gartner's predictions, **over 55% of deep neural network data analysis** is expected to take place at the point of capture in an edge system by 2025. Hence, organisations should pinpoint the applications, AI training, and inferencing necessary for transitioning to edge environments located near IoT endpoints.

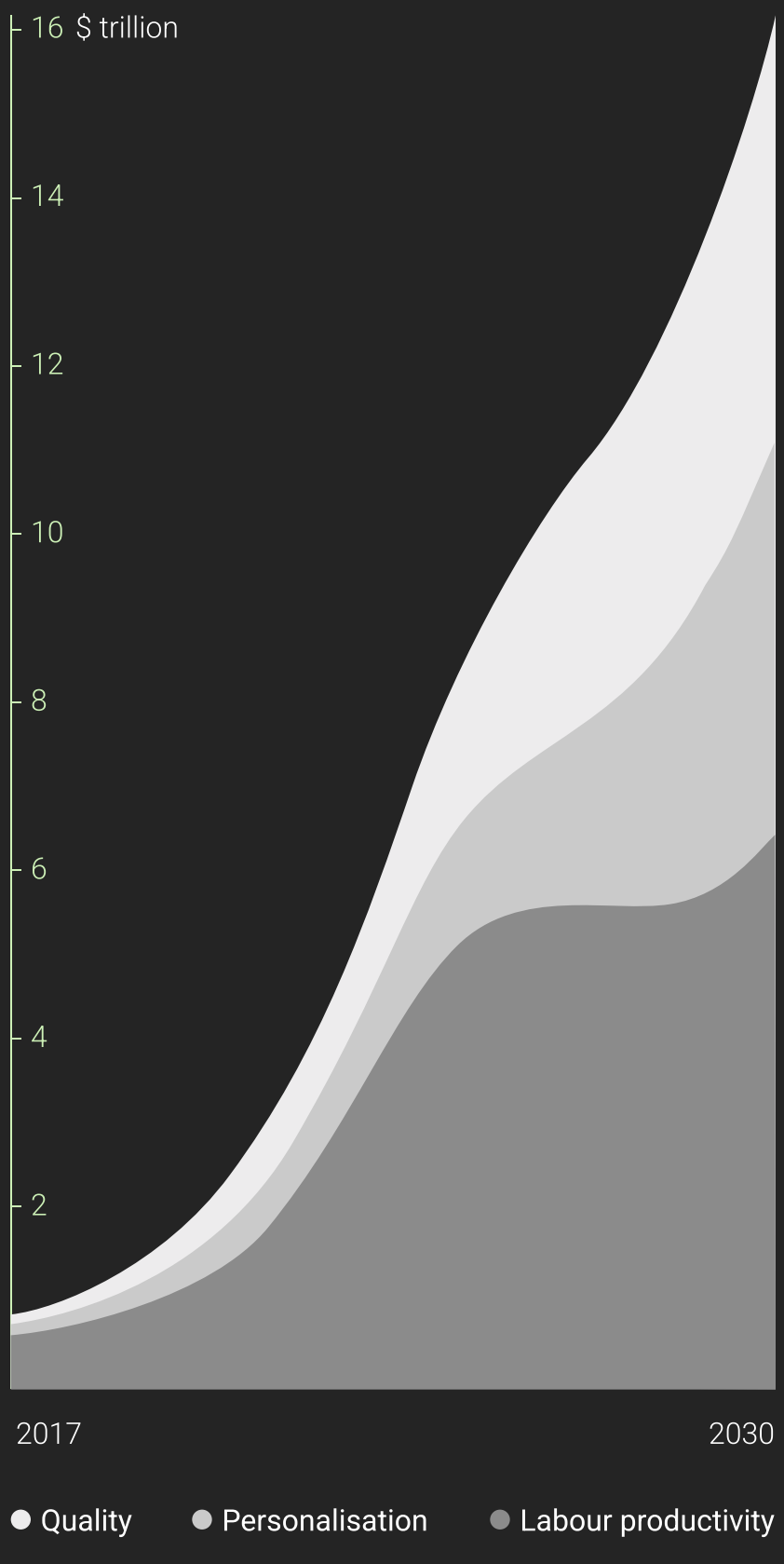
06 AI Governance

AI is advancing rapidly, and it's no longer just for fun — it needs proper governance now. In 2024, global leaders are planning to pay more attention to detailed AI policies, including countries like China, the EU, the U.S., and India.

The goal of this trend is to boost new technology, attract investments from all over the world, and make sure the society is safe from any unintended AI problems.

Tech experts are also discussing the possibility of countries working together on AI legislation rules and standards globally.

Global GDP impact by effect of AI



LABEL YOUR DATA

04

DATA ANNOTATION

TRENDS IN 2023

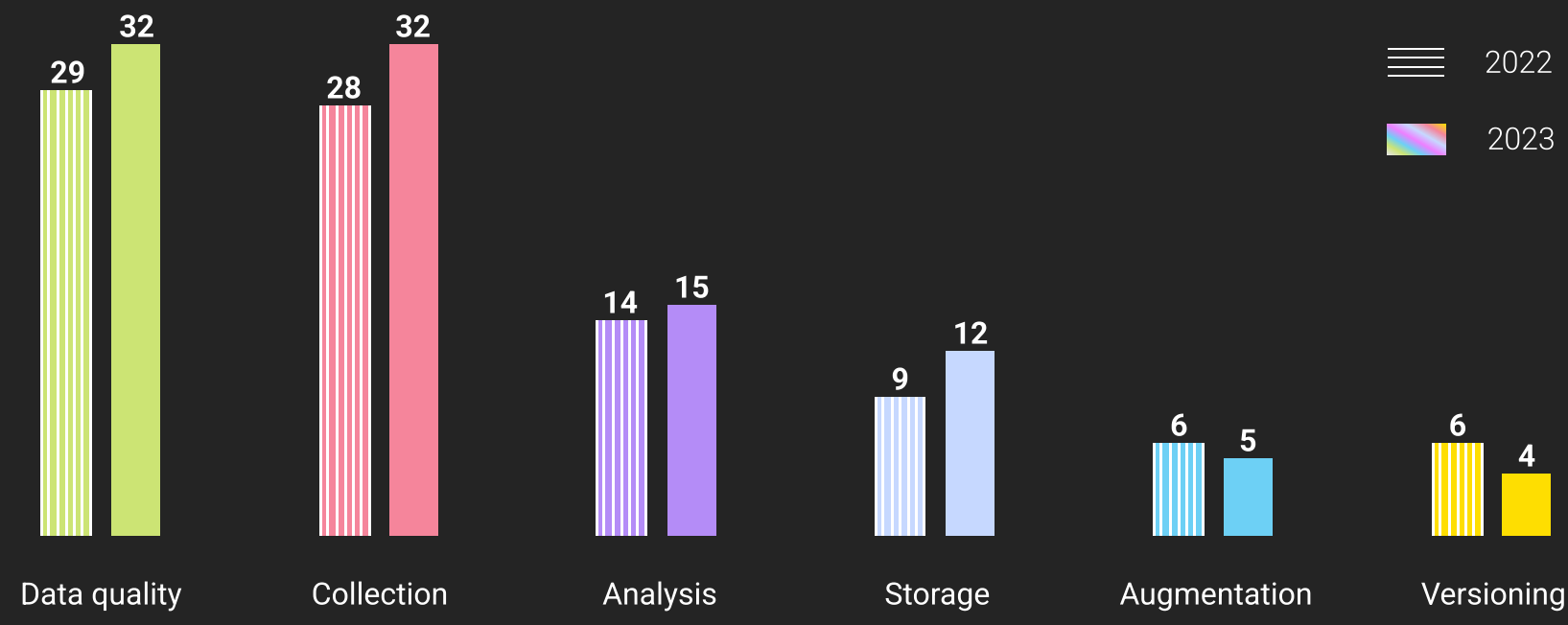
LABEL YOUR DATA

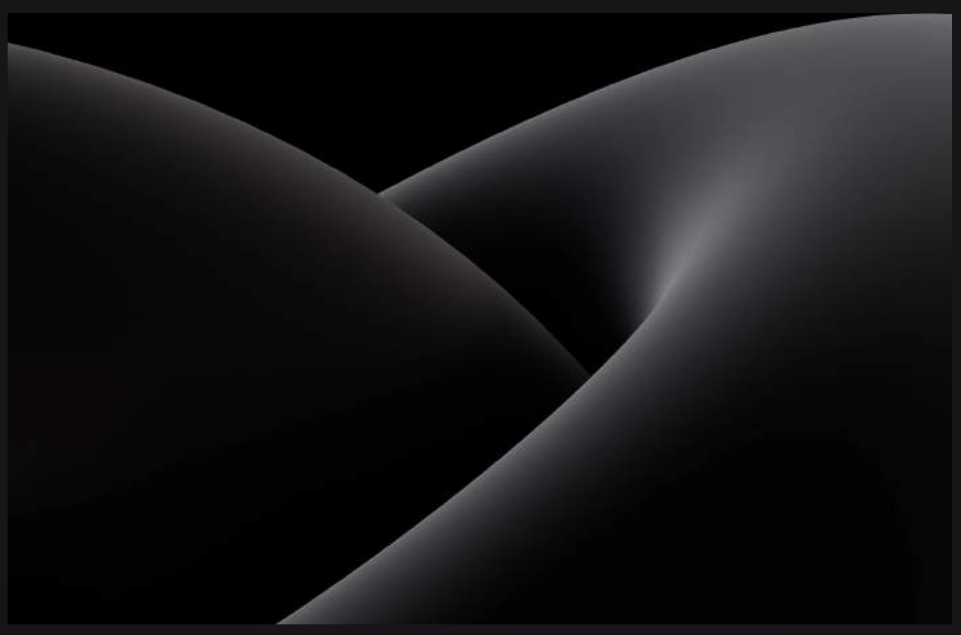
A crucial part of the entire AI ecosystem, data annotation does not lag behind the accelerated industry growth. As data labeling experts, the Label Your Data team could not miss the key trends shaping the data annotation landscape this year and the years to follow.

Let’s start with the factors that impact such trends in data annotation:

1. Massive generation of data daily and, thus, increasing reliance on it.
2. Rising popularity of facial recognition technology.
3. Increasing demand for autonomous driving solutions.

Data challenges,%





4. Current and emerging AI & ML trends shaping the course of the industry.

LABEL YOUR DATA

5. The introduction of new industries and applications for machine learning.

6. Ai is employed in the operations of nearly 40% of organisations globally.

7. By 2024, software-based facial recognition solutions will be integrated into approximately 1B devices globally.

8. Each day, over 4 billion people use the internet, generating about about 3 quintillion bytes of data.

9. 2,7M industrial robots are in operation, necessitating top-notch annotations for developing and testing CV models in robotic navigation systems

10. Text data is used by almost 70% of business, driven by the rise od chatbots and other NLP innovations.

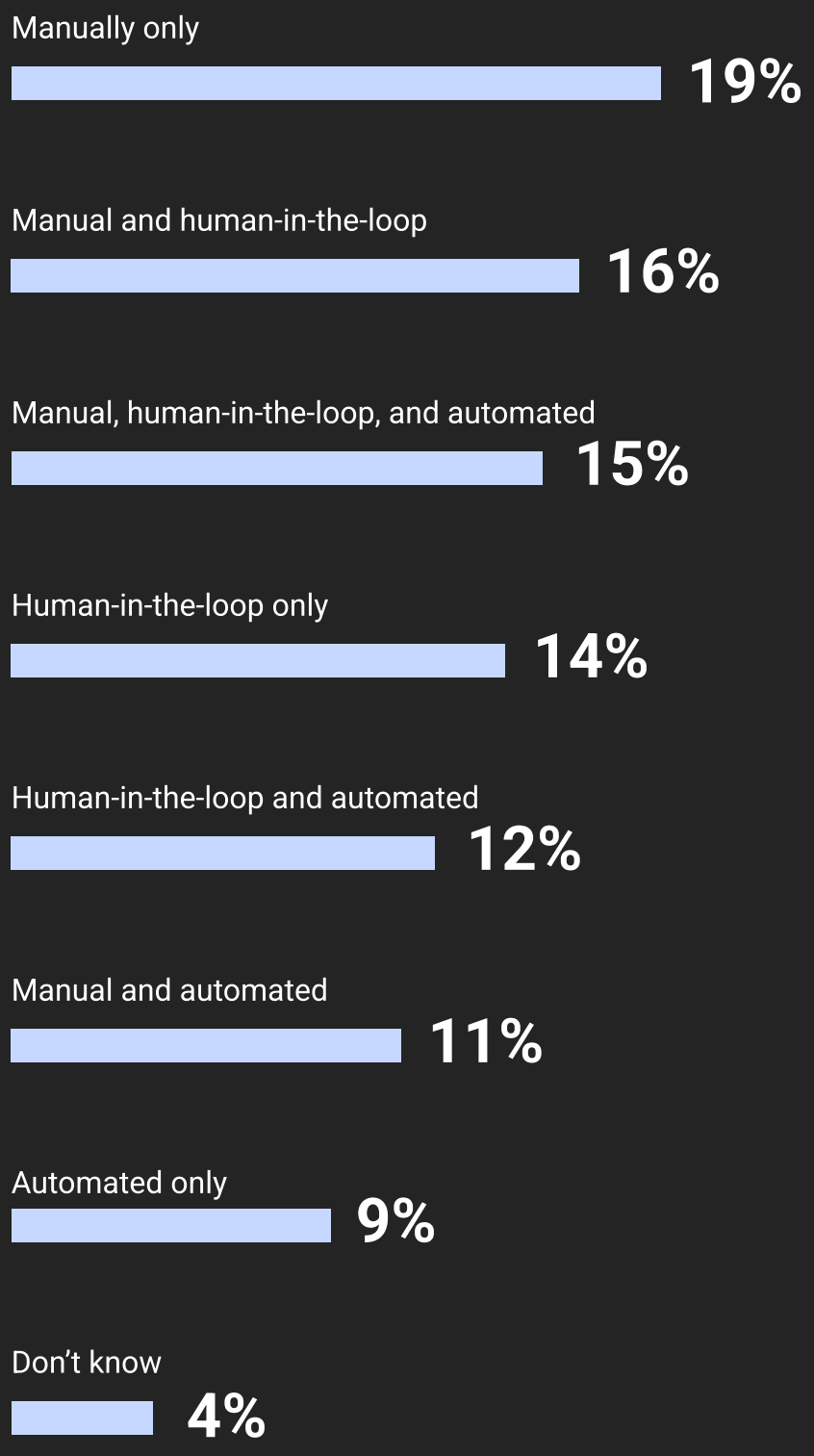
11. Increasing demand for secure and high-quality annotations.

12. The need for specialised annotation services tailored to specific domains.

13. The impact and management of remote and distributed work setups.

14. Increasing role of data privacy and security regulations.

How are you labeling the data when quality is the #1 challenge?



LABEL YOUR DATA

TREND 1:
MORE COMPLEX DATASETS

High-level machine learning requires more intricate work on data annotation to provide datasets for efficient model training. This also implies that the trend (or rather the need) for expert data labeling services is growing.

At Label Your Data, our annotators are trained for each individual project, either in Computer Vision or NLP, to grasp all the nuances of the data they are working with.

TREND 2:
MORE COMPLEX DATASETS

Data annotation has become essential during the collection phase, with a growing demand for real-time annotation. For annotators, this means operating with increased efficiency and precision. Mistakes at this stage could significantly influence the model training process results.

TREND 3:
AUTOMATED DATA LABELING

Automated annotation has been around for a while, even prior to 2023. However, the trend is growing. Using algorithms for automatic data annotation is great for many cases in machine learning, however, not without its shortcomings.

Automation is not always a reliable solution, which means that human supervision is required for such cases to ensure accuracy and precision in the annotation process. Thus, automated labeling is expected to complement, not replace, traditional human-based labeling in the future.

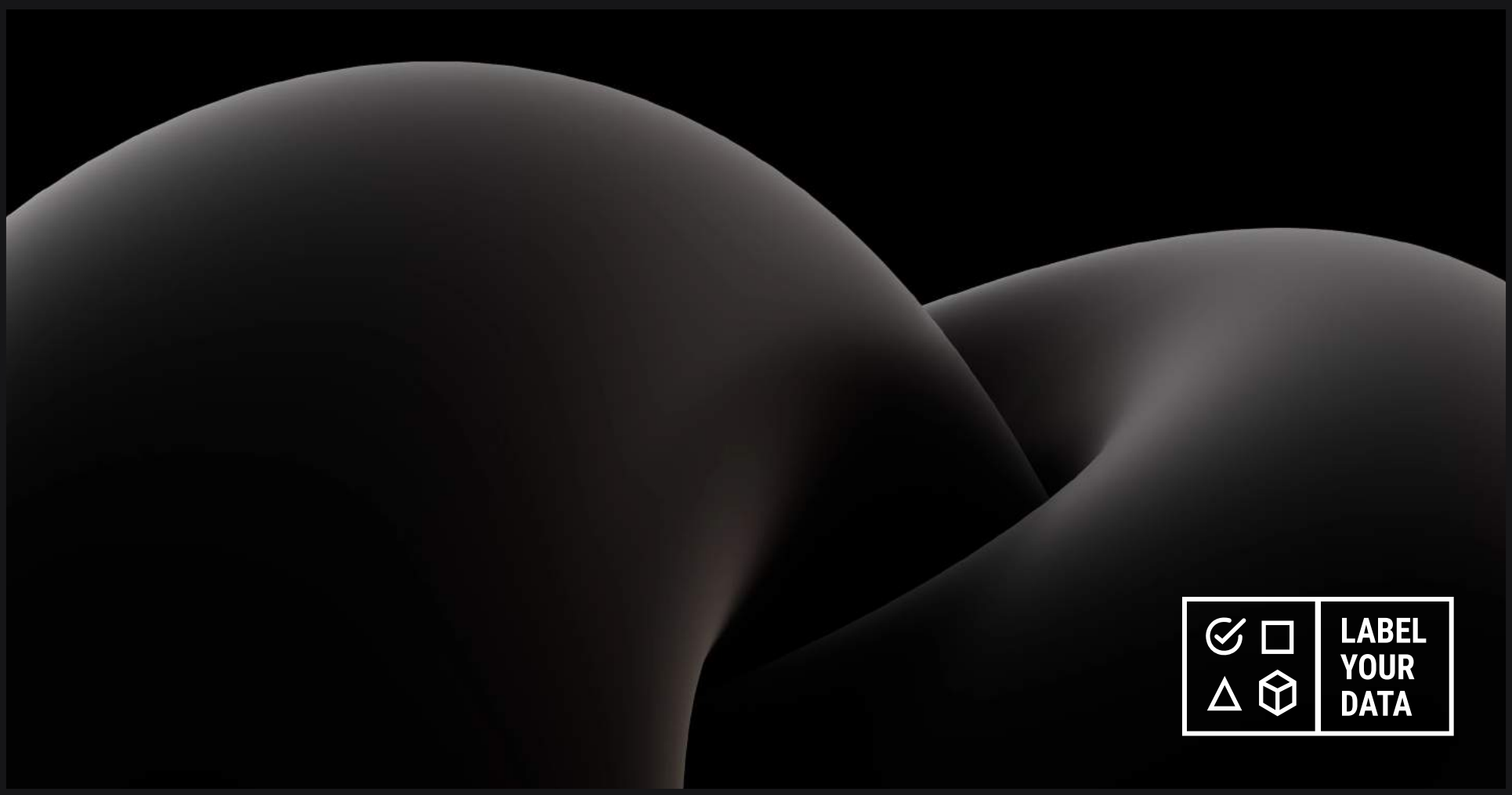
Our services at **Label Your Data** are certified with PCI DSS (level 1) and ISO:27001 and comply with GDPR, CCPA and HIPAA.

LABEL YOUR DATA

Teams that get annotated data faster tend to deploy new models to production faster and are rare able to update existing models more frequently

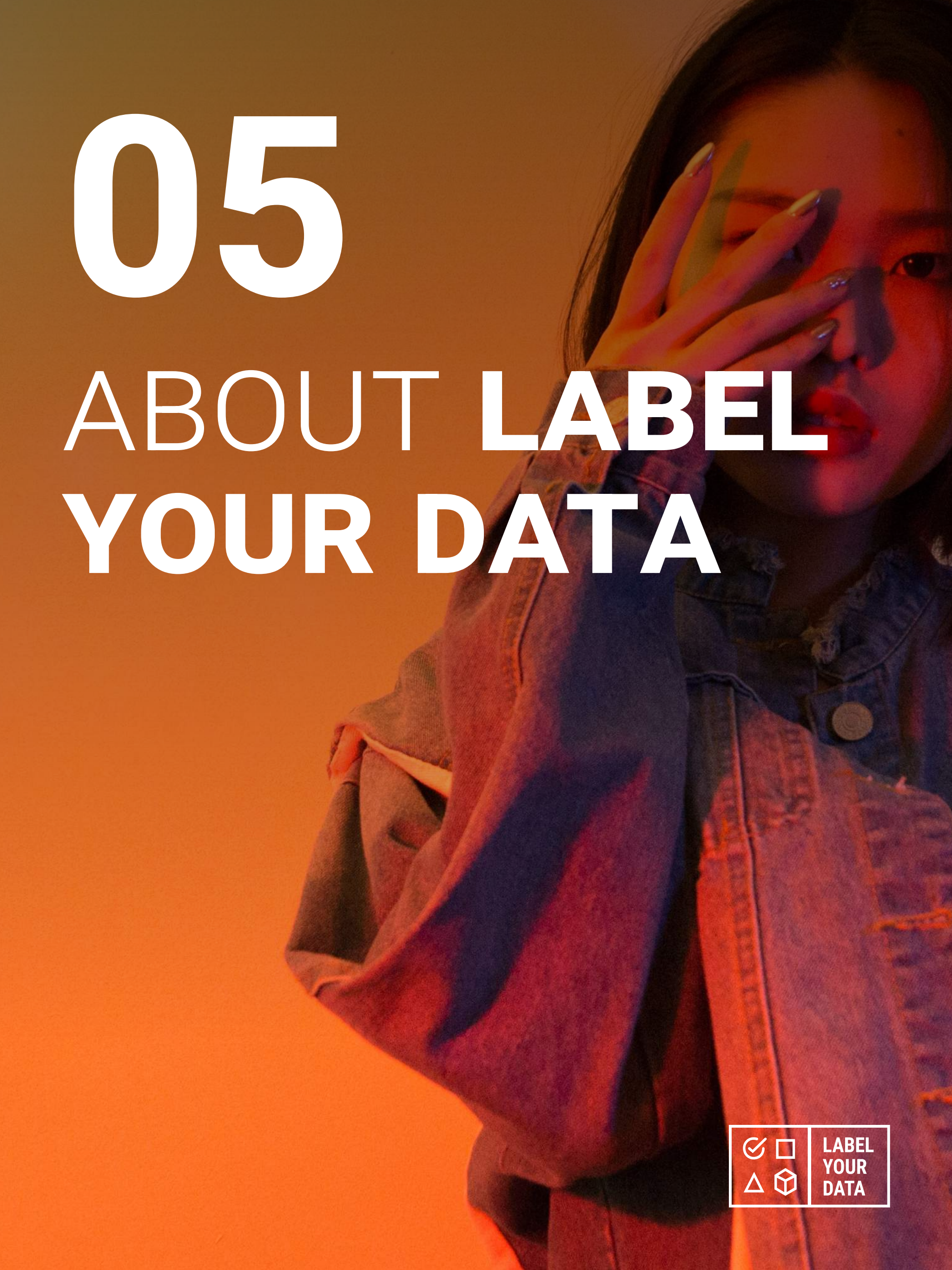
Time to get annotated data	Less then 1 month	1-3 month	3-6 month	6-9 month	9-12 month	More then 1 year
Greater then 3 month	0%	0%	3%	5%	10%	9%
1 to 3 month	2%	7%	9%	11%	2%	1%
1 week to 1 month	3%	9%	10%	4%	1%	1%
Less then 1 week	6%	5%	2%	1%	1%	0%

Time to deploy models



05

ABOUT **LABEL YOUR DATA**



For over a decade, the **Label Your Data** team has been dedicated to delivering top-notch data annotation services to help our clients scale their AI initiatives. We deliver custom labeling solutions with enterprise-class security level. Our clientele spans various industries such as Automotive, Robotics, Fintech, Healthcare, E-commerce, Manufacturing, Insurance, and more.

We prioritize the security of our clients' data. That's why our teams and facilities are certified with **PCI DSS Level 1** and **ISO/IEC 27001:2013** to ensure the safety of your datasets.

PCI Security Standards Council® Level 1 Provider

ISO 27001:2013 CERTIFIED Information Security Management

Label Your Data in numbers:

13+
years of industry expertise

500+
data annotators

55
languages

100+
clients from 25 countries

With a global workforce of over 500 specialists, we have a flexible annotation team that shares **our mission of co-building an AI-driven economy.** Our skilled annotators provide tailored solutions for both enterprise and R&D projects, covering the following services:

Computer Vision Annotation:

- Semantic Segmentation
- 2D Boxes
- Polygons
- OCR
- 3D Cuboids
- Key Points
- Video Annotation
- Image Categorization
- LiDAR/RADAR

LABEL YOUR DATA

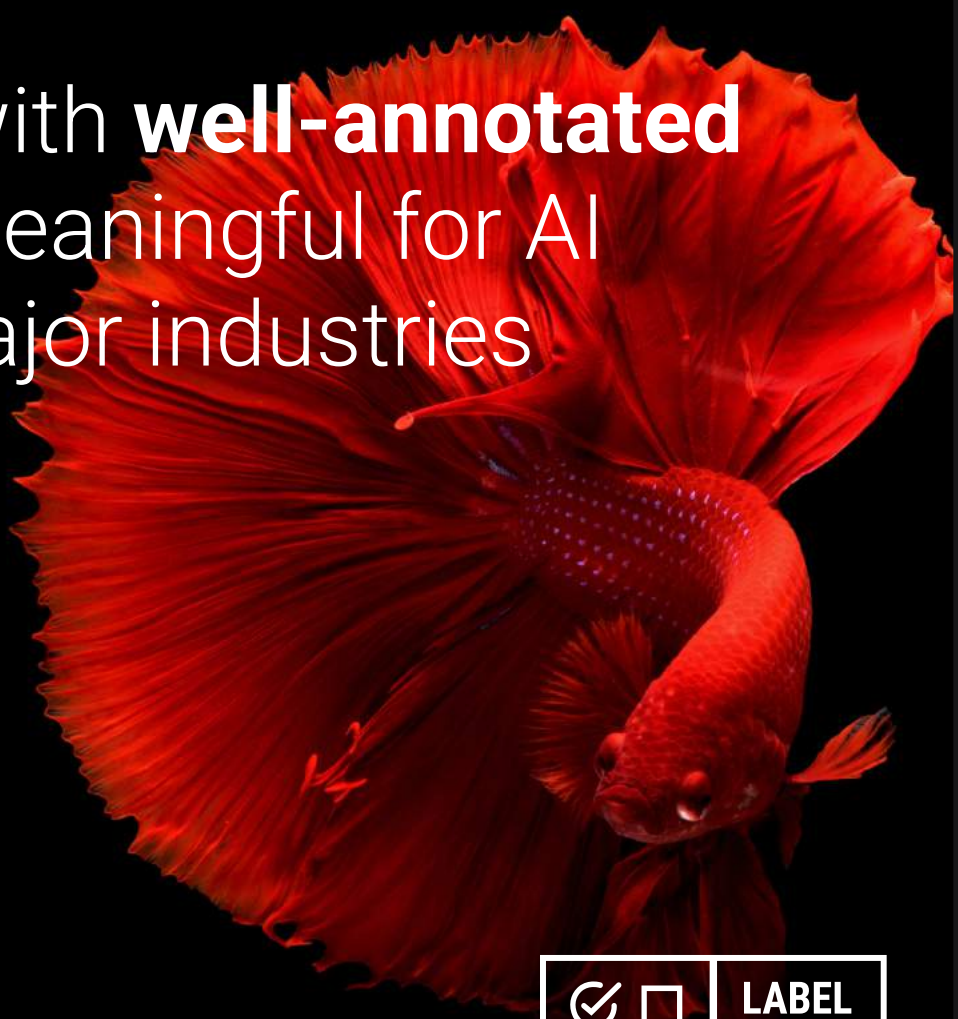
Additional Services:

- Data collection
- Model validation
- Know Your Customer (KYC)
- Data Anonymization
- Data Entry

NLP Annotation:

- Text Classification
- Named Entity Recognition (NER)
- Intent/Sentiment Analysis
- Comparison
- Audio-To-Text Transcription

Successful AI starts with **well-annotated data.** We make data meaningful for AI solutions across all major industries worldwide!



CONTACT US

LABEL YOUR DATA