draup

# Mapping the Evolving Tech Stack and Role Interactions Across Software, IT, and Data Science Functions

AUGUST 2025

Copyright © 2025 DRAUP. All Rights Reserved.

This document is solely for the use of Draup Prospect, Draup Client and Draup Personnel only. No part of it may be quoted, Circulated or reproduced for distribution outside the prospect / Client organization without prior written approval from Draup

# Executive Summary

The technology ecosystem is undergoing rapid and profound transformation driven by the convergence of generative AI, and cloud-native infrastructure. This report delivers a synthesis of how these forces are reshaping the roles, responsibilities, and skill expectations across Software Engineering, IT Operations, and Data Science functions.

## Key Trends and Findings

❖ **Generative AI as a Unifying Force**
Over 70% of enterprises have adopted generative AI in at least one function, with Software, Data, and IT leading this shift. Tools like GitHub Copilot, ServiceNow's Virtual Agent, and LangChain are becoming embedded across the stack—cutting coding time, automating infrastructure management, and accelerating AI model deployment.

❖ **Role Convergence and Tech Stack Interdependence**
Functional boundaries are blurring. Software developers now embed ML models, data scientists manage CI/CD pipelines, and IT professionals orchestrate secure, AI-enabled cloud environments. This convergence is giving rise to **hybrid roles**—such as ML-integrated developers and cloud-native data scientists—that co-own platform responsibilities.

❖ **Evolving Workloads Across Functions**
Traditional, siloed workflows are being replaced by **AI-native, integrated pipelines**. Software engineers benefit from automated testing, code generation, and cloud-first deployments. IT teams manage self-healing infrastructure, predictive observability, and zero-trust security. Data scientists leverage LLMs, AutoML platforms, and unified MLOps stacks for real-time insight generation.

❖ **Rising Demand for Hybrid Tools & Skillsets**
Job postings reflect a ~50% YoY increase in AI skill requirements across all three functions. Core languages (Python, Java, SQL) remain foundational, but emerging capabilities in prompt engineering, infrastructure-as-code, and AI observability are becoming critical differentiators.

## Implications for Team Design and Capability Development

❖ **Platform Engineering and Stack Standardization**
Shared tooling—such as MLOps platforms, cloud orchestration, and API frameworks—is central to enabling cross-functional collaboration. Investing in platform engineering will allow teams to scale innovation while reducing complexity and duplication.

❖ **Talent Strategy Shift**
Prioritize the development of hybrid talent profiles through targeted learning pathways, internal mobility programs, and mentorship structures. Cross-functional delivery pods and AI Centers of Excellence will help anchor this transition.

❖ **Future-Ready Workforce Architecture**
Aligning tech stack choices with available (or trainable) talent pools will be critical. By integrating skills intelligence into technology governance, ensure sustainable innovation, faster go-to-market capabilities, and long-term talent resilience

Copyright © 2025 DRAUP. All Rights Reserved.

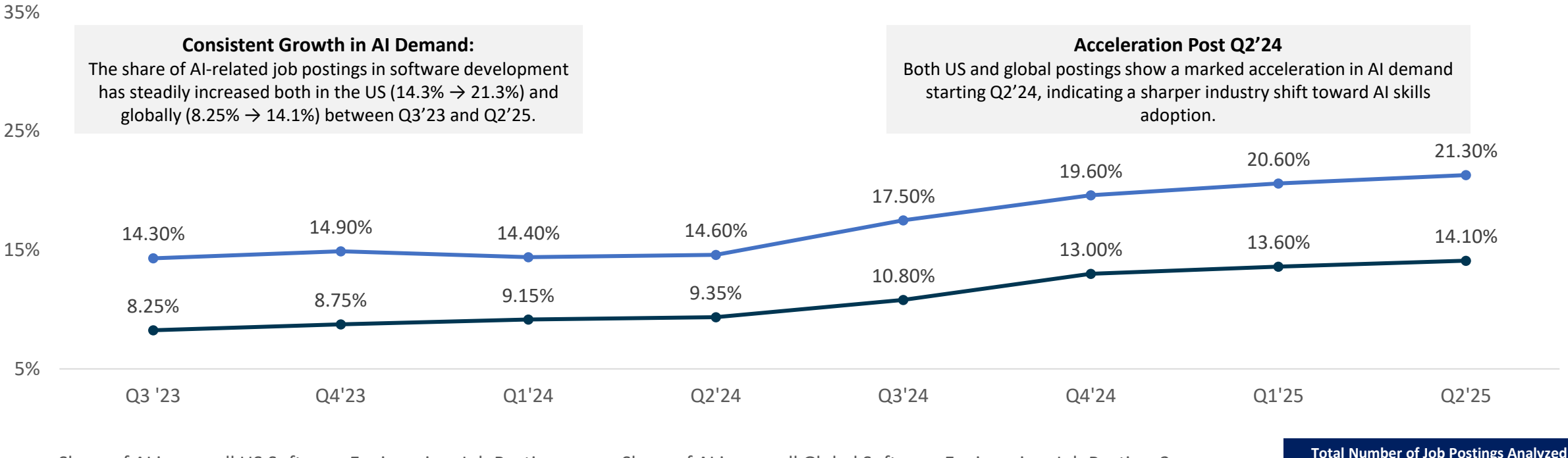
# Technology Mapping - Software

# AI adoption in software engineering roles has surged, with a ~50% growth in demand for AI-driven tools & skills in the USA , positioning AI as a core capability in modern software roles

## Share of AI related Tools / Skills in Software Development Job Postings (Jul'23 – Jun'25) – Global vs USA Comparison*

**Consistent Growth in AI Demand:**
The share of AI-related job postings in software development has steadily increased both in the US (14.3% → 21.3%) and globally (8.25% → 14.1%) between Q3'23 and Q2'25.

**Acceleration Post Q2'24**
Both US and global postings show a marked acceleration in AI demand starting Q2'24, indicating a sharper industry shift toward AI skills adoption.




| Total Number of Job Postings Analyzed | |
|---|---|
| Global | USA |
| 4.9Mn | 1.4 Mn |

## Key Insights

**Within AI jobs specifically, languages like Python is mentioned in over 50% of postings, TensorFlow (14%), PyTorch (11%) and cloud tools like Azure (15%) and AWS (10%) are also common**

**Companies, especially Top Performers, are also prioritizing specific AI-related skills in 2024. The three most highly sought after skills by U.S. engineering leaders are AI engineering (74%), integrating AI functionality into products via API (62%), leveraging AI tools while coding (56%)**

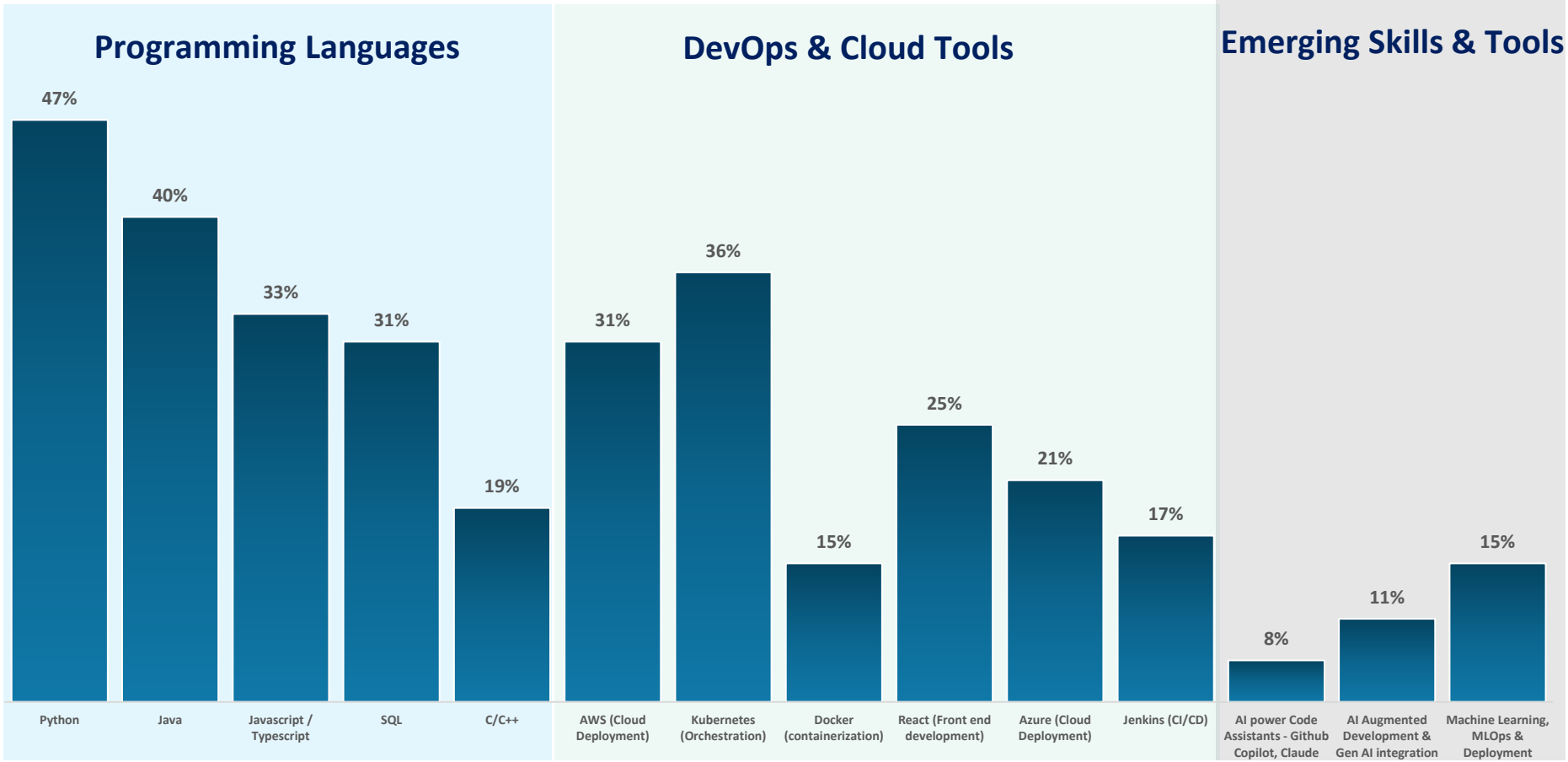
**Demand of AI based skills in Software Jobs surged by ~50% as compared from Q'3 2023 to Q2 2025, driving broader AI related hiring growth.**

**Note**: Number of Job Postings have been extracted using Draup Talent Demand modeler that tracks over 850 Mn+ JDs. We have considered only software developer workload related job roles & filtered the AI related skills to identify the relevant skills.

Copyright © 2025 DRAUP. All Rights Reserved.

# Software engineering continues to evolve rapidly, shaped by both the adoption of AI-assisted development tools and the resilience of traditional programming languages, frameworks, and architectural patterns

## Penetration of Tools Languages & Models in Software Job Postings_USA (Jul'24-Jul'25)*




% Share of Job Postings Demanding these skills USA

| Total Number of Job Postings Analyzed | 796,000 |
|---|---|

*Note: Since multiple skills can appear simultaneously within a single job description, the skills overlap across different JDs. Therefore, the total will not sum up to 100%

## Key Insights

In US, Python continues to be the most demanding programming language with (47%) share in the software related job postings. it was asked in **almost 1 out of 2 the Software related job** postings over the past one year
.

Python and Java accounts for a total of 87% of the job offers, meaning nearly 4 out of 5 job postings requires knowledge of these skills A study of GitHub commits through December 2024 shows AI **generated ~30% of Python** code in the U.S.

Despite the rise of AI, core languages and frameworks remain dominant According to Stack Overflow's 2024 Developer Survey, JavaScript (63%) and Python (49%) continue to be the most widely used programming languages, followed closely by TypeScript, Go, and Java

GitHub reports that over 55% of developers on its platform are now using Copilot regularly, with controlled experiments showing that AI-assisted coding can complete tasks 55–56% faster

Copyright © 2025 DRAUP. All Rights Reserved.

**Note**: Number of Job Postings have been extracted using Draup Talent Demand modeler that tracks over 850 Mn+ JDs. We have considered only software developer workload related job roles & filtered the required skills to identify the relevant jobs.

# Advancements in modern software architecture and automation technologies are reshaping workloads across Software Engineering—driving shifts toward API-first development, microservices, cloud-native patterns, and streamlined deployment pipelines

| Key Areas | Generative AI Tools & Libraries | Transformed Priorities With GAI | Key Insights |
|---|---|---|---|
| **Software Development** | tabnine, GitHub Copilot, Claude, React | • **Shift to API-first development with AI-generated interfaces**<br>• AI-assisted debugging & code refactoring<br>• Develop software apps using Python, TensorFlow, PyTorch, & integrating Gen AI techniques | • GitHub Copilot, has reached massive scale, platform has surpassed 1.8 million paying subscribers and has been adopted by over 50,000 organizations and enterprises worldwide. |
| **System Architecture** | Amazon SageMaker, vertex.ai, spaCy | • Design scalable, high-performance, and reliable system architecture incorporating generative AI techniques<br>• Performance & scalability simulations with AI agents<br>• **Shift to microservices & cloud-native architecture**<br>• Intelligent microservices orchestration suggestions | • Industry analysis shows a major architectural shift is underway. It's estimated that over 40% of large enterprises are now actively building or re-architecting applications to be "AI-native. |
| **Database Management** | TONIC THE FAKE DATA COMPANY, Faker | • Develop and optimize database queries and procedures tailored for generative AI data retrieval and manipulation<br>• Generate synthetic test data at scale<br>• AI-driven anomaly/fraud detection in DB activity<br>• Intelligent query optimization & indexing | • The vector database market, which is essential for generative AI features like semantic search, is experiencing explosive growth, projected to expand at a Compound Annual Growth Rate (CAGR) of over 50%. This signifies rapid, widespread adoption by thousands of companies. |
| **DevOps & Infrastructure** | snyk, circleci, COPADO, kubernetes, docker | • Maintain infrastructure-as-code (IaC) scripts and templates for automating the deployment and configuration of generative AI environments<br>• **Cloud-native deployment automation**<br>• Automated CI/CD pipeline generation with AI suggestions, Self Healing Infra | • Hugging Face has become the central infrastructure hub for the open-source AI ecosystem, essential for AI DevOps (MLOps). Over 500,000 organizations use the platform to host models, manage datasets, and deploy applications. |
| **Software Testing and QA** | Tricentis Testim, eggplant, testsigma | • Automated test plans, test cases, and test scripts to validate the functionality and performance of generative AI applications and models<br>• Intelligent bug triaging & prioritization<br>• Automated load testing with AI simulations | • CodiumAI's developer-led growth demonstrates how generative AI is automating the traditionally manual process of software testing. The platform has attracted over 500,000 registered developers who use it to automatically generate test suites. |

**Note:** Draup's proprietary models leverage 850 Mn+ JDs to identify the relevant skills associated with each of the skills clusters analyzed based on the demand for any job family. Listed Skills are not exhaustive. The list of Digital and Gen AI tools is for representational purposes and non-exhaustive

Copyright © 2025 DRAUP. All Rights Reserved.

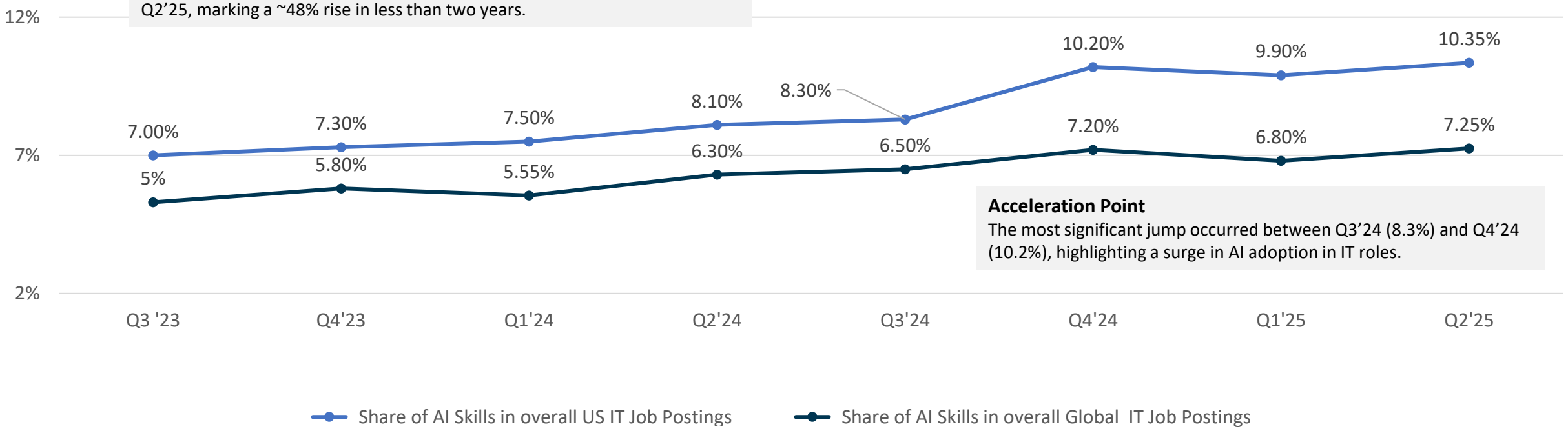
# Technology Mapping - IT

# AI is rapidly reshaping IT roles across cloud engineering, DevOps, and cybersecurity. Traditional Infrastructure-as-Code (IaC) practices using tools like Terraform, Ansible, and Kubernetes are now being supercharged with AI-driven automation

## Share of AI Skills in IT Job Postings (Jul'23 – Jun'25) – Global vs USA Comparison*

| Total Number of Job Postings Analyzed | |
|---|---|
| Global | USA |
| 7.4Mn | 2.7Mn |

**Steady Uptick**
AI-related IT job postings in the USA grew from 7.0% in Q3'23 to 10.35% in Q2'25, marking a ~48% rise in less than two years.




**Acceleration Point**
The most significant jump occurred between Q3'24 (8.3%) and Q4'24 (10.2%), highlighting a surge in AI adoption in IT roles.

## Key Insights

**In cloud engineering, AI is being deployed to optimize costs and performance in increasingly complex multi-cloud environments. IDC notes that 65% of cloud teams are already using AI/ML tools to automate workload placement and scale resources dynamically across AWS, Azure, and GCP.**

**The parallel rise of zero-trust architectures is amplifying this trend. Okta's 2024 Zero Trust report indicates that 61% of organizations have implemented zero-trust principles, with AI-driven identity and access tools**

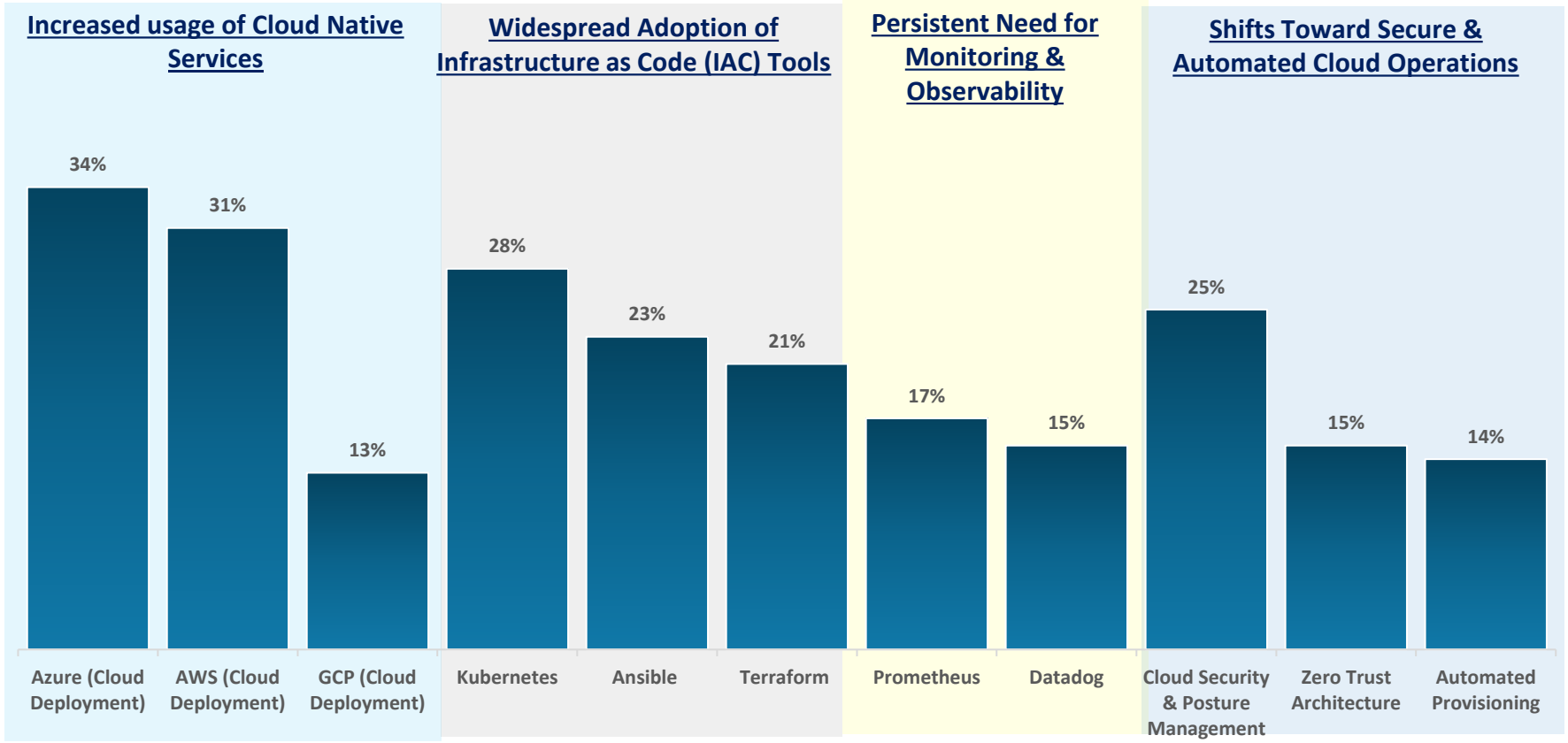
**Generative AI chatbots embedded into platforms like ServiceNow and Freshdesk can now resolve 40–60% of Tier-1 IT support tickets autonomously. A research report projects that by 2026, 80% of large enterprises will have adopted AI-powered IT operations (up from 20% in 2022).**

**Note**: Number of Job Postings have been extracted using Draup Talent Demand modeler that tracks over 850 Mn+ JDs. We have considered only IT workload related job roles & filtered the AI related skills to identify the relevant skills.

Copyright © 2025 DRAUP. All Rights Reserved.

# Cloud engineers, DevOps professionals, and security architects are increasingly leveraging tools like AWS, Azure, GCP, Datadog, and Prometheus, while automation and zero-trust architectures are rapidly expanding to complement these capabilities

## Penetration of Core IT, Tools / Libraries in IT related Job Postings USA (Jul'24-Jul'25)




**Total Number of Job Postings Analyzed**

**1.5 Mn**

## Key Insights

**Increased usage of Cloud Services**

Cloud platforms are equally integral to IT roles Flexera **2024 State of the Cloud Report**, states that **94% of enterprises use cloud services** in some form **87% of enterprises have a multi-cloud** strategy.

**Adaption of IAC Tools**

**IaC has become the backbone** of modern IT operations, with HashiCorp 2023 State of Cloud Strategy Survey **reporting that over 85% of enterprises use Terraform** for multi-cloud provisioning

**Need for Monitoring & Observability**

Alongside cloud adoption, observability has become a priority. As per external report, by 2026, **70% of enterprises will have** implemented **unified observability** platforms.

**Shifts Toward Secure Cloud Operations**

Security responsibilities are also expanding rapidly, According to Okta's 2024 survey, **61% of organizations have implemented** zero-trust principles, while external reports forecasts that by 2027, **60% of enterprises will have formalized zero-trust programs**, up from less than 20% in 2022.

Copyright © 2025 DRAUP. All Rights Reserved.

*Note: Since multiple skills can appear simultaneously within a single job description, the skills overlap across different JDs. Therefore, the total will not sum up to 100%

**Note**: Number of Job Postings have been extracted using Draup Talent Demand modeler that tracks over 850 Mn+ JDs. We have considered only IT related workload related job roles & filtered the required skills to identify the relevant jobs.

# Generative AI is driving automated provisioning, zero-trust architectures, and cloud-native security, shifting IT from reactive support to a proactive, self-healing backbone of enterprises

## Generative AI Tools Reshaping Core IT Priorities

| | Increased usage of Cloud Native Services | Widespread Adoption of Infrastructure as Code (IAC) Tools | Persistent Need for Monitoring & Observability | Shifts Toward Secure & Automated Cloud Operations |
|---|---|---|---|---|
| **Transformed Priorities** | • Growth in automated provisioning<br>• AI predicts demand spikes & scales infra automatically<br>• Automated patching & compliance checks<br>• Intelligent workload placement across hybrid/multi-cloud<br>• Self Healing Servers, AI Driven Cloud cost optimization | • Automated troubleshooting & auto remediation of Kubernetes misconfigs<br>• Auto-generation of Terraform scripts & Ansible playbooks<br>• AI-driven compliance checks in CI/CD pipelines<br>• Policy compliance checks in real-time | • AI-based anomaly detection for unknown threats, Predictive alerting<br>• Automated incident response playbooks<br>• Log summarization & incident report generation<br>• Predictive downtime alerts & Log Anomaly detection | • Increase adaption of zero-trust architectures, and cloud security postures<br>• Policy generation for zero-trust frameworks<br>• AI-driven vulnerability scanning & patching |
| **Tools Used** | Vertex AI, dynatrace, Google Cloud Platform, Amazon Bedrock, Azure OpenAI | Terraform, kubernetes, ANSIBLE, Amazon CodeWhisperer, GitHub Copilot | Prometheus, Grafana, new relic, DATADOG, elastic, AI Assistants for Observability | SentinelOne, WIZ, DARKTRACE, Microsoft Copilot for Security, PRISMA CLOUD, splunk> |
| **Key Insights** | • Amazon Web Services (AWS), Holds 31% global market share in cloud (Synergy Research, 2025).<br>• 80% of Fortune 500 run critical workloads on AWS. AWS Lambda executes 10 trillion+ requests per month (2024), showing massive reliance on serverless computing.<br>• Vertex AI adoption has doubled YoY as companies shift to AI-native development environments. | • Recent Cloud Native Computing Foundation, data shows 89% of surveyed organizations are now using cloud-native technologies, with Kubernetes adoption (in use or evaluation) at 93%<br>• Azure Kubernetes Service (AKS) is used by 60%+ of enterprise DevOps teams who run Kubernetes in the cloud. | • As per CNCF survey. 60% of organizations vetting open-source projects for active communities, and **57% using automated tools** to detect vulnerabilities.<br>• Datadog reports over 2,500 customers using at least one of their AI/ML-powered integrations & with 83% of all customers using two or more products | • As per the survey result of Darktrace Tool, it was found that **over 73% of Chief Information Security Officers (CISOs)** are actively increasing their investment in generative AI to augment their security teams |

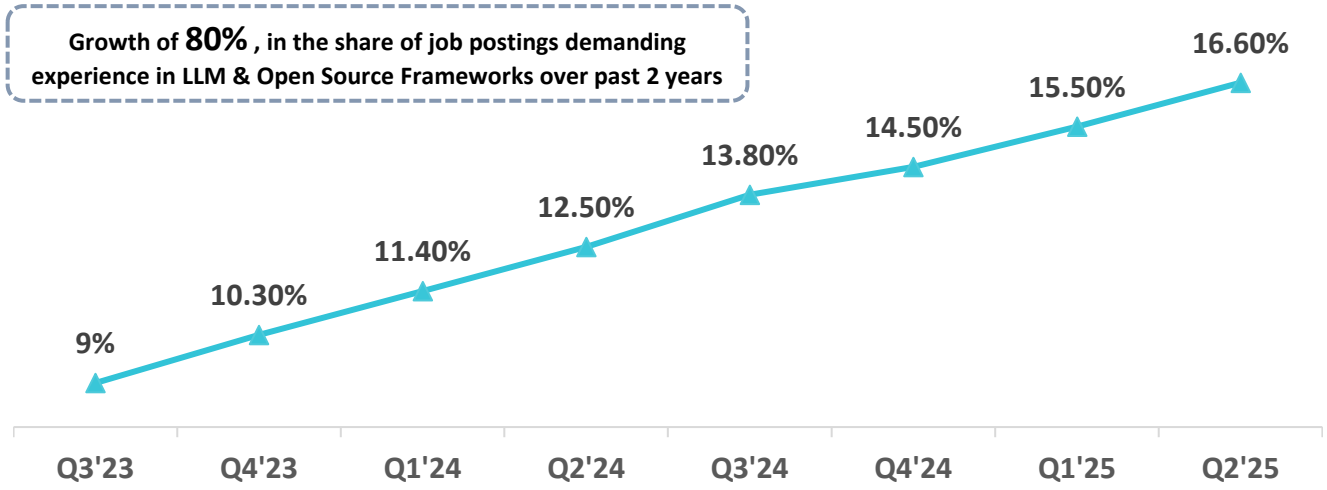
Copyright © 2025 DRAUP. All Rights Reserved.

**Note:** Draup's proprietary models leverage 850 Mn+ JDs to identify the relevant skills associated with each of the skills clusters analyzed based on the demand for any job family. Listed Skills are not exhaustive. The list of Digital and Gen AI tools is for representational purposes and non-exhaustive

# Technology Mapping – Data Science

# The data science role is rapidly evolving, driven by the power of Generative AI. Data scientists and ML engineers are now leveraging LLMs, low-code platforms and prompt engineering to redefine traditional workflows

## Increased Demand of LLMs in Data Science Job Postings past two years (2023-2025)




## Emerging Trends Reshaping the Future of Data Science

**Rise of LLMs & Open-Source Frameworks**
Open-source frameworks like Hugging Face and LangChain now power NL interfaces, with 64% of enterprises exploring or deploying GenAI/LLMs..

**AutoML & Low-Code: Democratizing AI Development**
Low-code and AutoML platforms are making ML more accessible — by 2027, 65% of new AI projects will use them, up from <20% in 2022 (External Report).

**Increased use of Prompt Engineering into Key Workflows**
LinkedIn's 2024 Global Skills Report identified prompt engineering as one of the fastest-growing new skills globally, with adoption growing by more than 1500% YoY

## Integration of prompt engineering into traditional data workflows

**Data Scientist** → **Multimodal Data Scientist Data Scientist**

- Feature engineering
- Model documentation
- Writing SQL/code for EDA

- LLM-assisted EDA and Prompt-based feature generation
- Auto-generated model insights,
- LLMs to generate synthetic data

**ML Scientist** → **Applied Scientist – GenAI/LLM**

- Designing ML models from scratch-
- Manual feature selection & tuning

- Prototyping with LLM APIs & prompt chaining
- Auto-generating model summaries

Copyright © 2025 DRAUP. All Rights Reserved.

**Note**: Number of Job Postings have been extracted using Draup Talent Demand modeler that tracks over 850 Mn+ JDs. We have considered only data science workload related job roles & filtered the required skills to identify the relevant jobs.

# Data science roles are transforming with the adoption of LLM frameworks, low-code ML platforms, and prompt engineering tools, while foundational big data technologies like Spark and Hadoop remain integral to modern AI/ML workflows

## Penetration of Data Science Tools in U.S. Job Postings (July'24 – July'25)

### LLMs & Open Source Frameworks

Tools like Hugging Face and LangChain are appearing in over **1 in 10** job postings, showing mainstream LLM adoption.

### Low Code / Auto ML Platforms

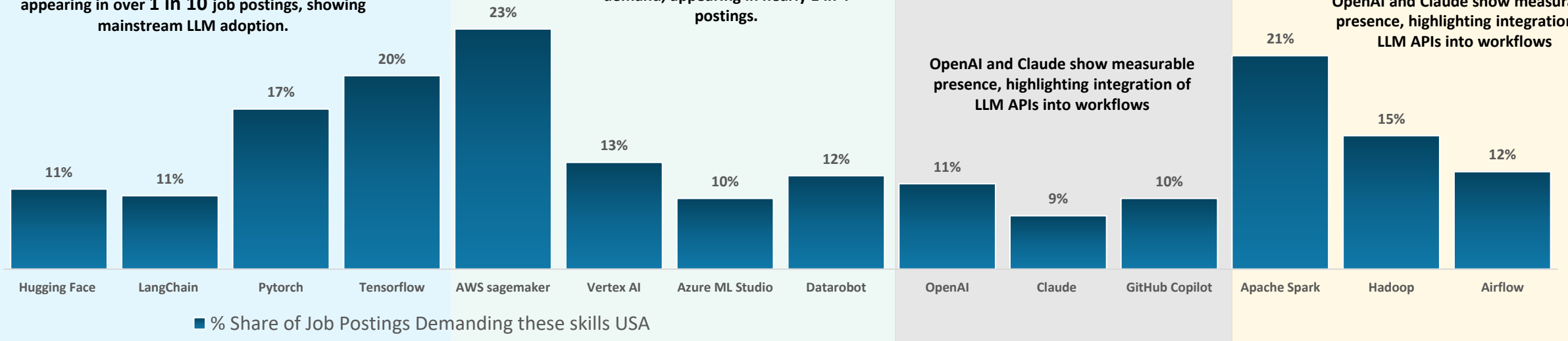
AWS Sagemaker leads low-code demand, appearing in nearly 1 in 4 postings.

### Prompt Engineering / AI Assistants

OpenAI and Claude show measurable presence, highlighting integration of LLM APIs into workflows

### Big Data Processing Tools

OpenAI and Claude show measurable presence, highlighting integration of LLM APIs into workflows




| Total Number of Job Postings Analyzed | 796,000 |
|---|---|

*Note: Since multiple skills can appear simultaneously within a single job description, the skills overlap across different JDs. Therefore, the total will not sum up to 100%

## Emerging Trends Reshaping the Future of Data Science

### Surge in Low-Code/AutoML Adoption

As per External Reports, Low-code and AutoML platforms are now mainstream — projected to drive 70% of new enterprise app development by 2025.

### Big Data Tools Remain Foundational

Even with rapid adoption of GenAI and AutoML platforms, demand for tools like Spark (21%), Hadoop (15%), and Airflow (12%) remains high. This underlines that scalable data infrastructure is still essential for modern AI/ML workflows.).

**Note**: Number of Job Postings have been extracted using Draup Talent Demand modeler that tracks over 850 Mn+ JDs. We have considered only data science workload related job roles & filtered the required skills to identify the relevant jobs.

Copyright © 2025 DRAUP. All Rights Reserved.

# The data science workflow is being reinvented, shifting from complex coding to intuitive, AI-powered tools; This transformation democratizes development and dramatically accelerates the path from raw data to actionable insights

## Generative AI Tools Reshaping Core Data Science Workloads

| | Increased usage of LLMs & ML Ops | Auto ML & Low-Code: Enabling AI Development | Integration of Prompt Engineering | Use of Copilots in Data Analysis |
|---|---|---|---|---|
| **Transformed Workloads** | • Faster model experimentation and deployment<br>• Automated model monitoring, versioning, and retraining<br>• LLMs being used for data exploration, feature generation, and code documentation | • Non-experts can now build and deploy models<br>• Frees data scientists for more complex modeling tasks<br>• Reduces development time by up to 50% | • Turns non-technical users into power users of LLMs<br>• Shift from hardcoded pipelines to flexible natural language interfaces<br>• Replaces traditional rule-based automation | • **Automates exploratory data analysis (EDA), data cleaning, and even chart creation**<br>• Embedded into IDEs, Notebooks, and BI tools<br>• Generates synthetic data samples for testing |
| **Tools Used** | LangChain, databricks, Claude, neptune.ai, Vertex AI, OpenAI GPT-4, Hugging Face | Azure No-Code AutoML, runway, DataRobot, H2O.ai, Amazon SageMaker, Google's AutoML, Data Studio, Vertex AI | LangChain, LlamaIndex, runway, OctoAI, DALL·E 2, OpenAI GPT-4, PromptLayer, Semantic Kernel | tabnine, CTGAN, Claude, Copilot in Power BI, databricks, PandasAI, GitHub Copilot, Data Studio |
| **Key Insights** | • **76% of enterprises** are exploring or adopting LLMs in production<br>• In 2024, **65% of enterprises** had deployed MLOps platforms vs. **45% of SMBs**; by 2026, SMB adoption is expected to reach **~60%, narrowing the gap**<br>• The global MLOps market reached approximately **USD 1.7–2.2 billion in 2024**, projected to grow at a **37–43% CAGR** through 2030–33 | • The AutoML market is experiencing explosive growth, with a projected **Compound Annual Growth Rate (CAGR) of over 40**%<br>• As per external reports, **45% of AI projects** now involve low/no-code platforms<br>• AutoML usage projected to **grow 3x by 2027** in small to mid-sized enterprises | • LinkedIn's 2024 Global Skills Report identified prompt engineering as one of the **fastest-growing new skills globally**, with adoption growing by **more than 1500% YoY**<br>• Over **60% of enterprise** GenAI use cases rely on prompt engineering for outcomes | • Over **60% of Fortune** 500 firms now use AI copilots specifically within their data analytics workflows<br>• As per State of Customer Experience AI report, **58% of firms use AI** for churn prediction, NPS analysis, and personalization<br>• As per CFO Pulse Survey 2024, **35% of finance leaders** adopted AI copilots for rolling forecasts |

**Note:** Draup's proprietary models leverage 850 Mn+ JDs to identify the relevant skills associated with each of the skills clusters analyzed based on the demand for any job family. Listed Skills are not exhaustive. The list of Digital and Gen AI tools is for representational purposes and non-exhaustive

Copyright © 2025 DRAUP. All Rights Reserved.

# Cross-Role Interactions and Functional Convergence

## Organizations are moving from siloed development and reactive IT to integrated, cloud-native ecosystems, where software and IT teams leverage security, automation, and AI/ML to speed delivery, optimize infrastructure, and boost resilience

### Traditional Workload

**Software**
Traditional software development was siloed, manual, and prone to delays due to isolated coding, late IT involvement, and minimal automation.

***Example**: IT managed physical data centers and resolved outages, often without visibility into upcoming software needs.*

**IT**
IT operated in static, manually managed environments, while reacting to issues after they occurred and remaining disconnected from the development process.

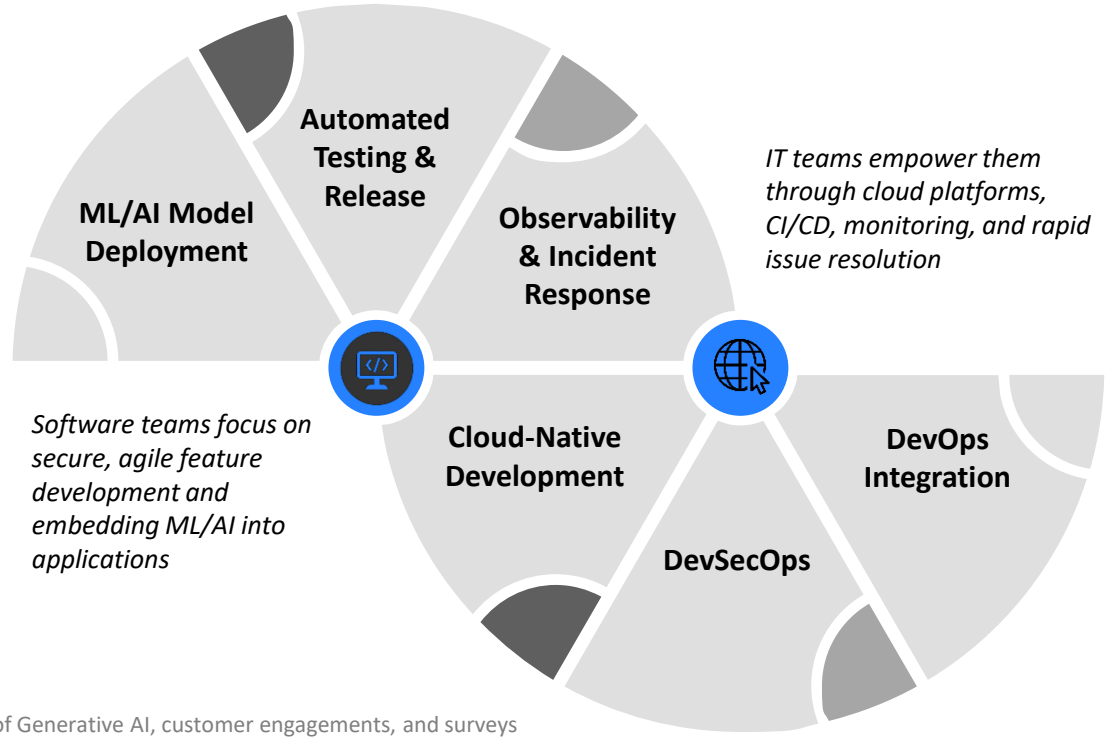
***Example:** Developers created new product features and then handed them over to IT without aligning on capacity planning, runtime dependencies, or performance implications.*

### Evolving Interdependencies - Integrated Workloads

**ML/AI Model Deployment**
Software engineers embed ML models directly into apps, relying on IT for GPU infrastructure, monitoring, and scaling

**Automated Testing & Release**
IT provides test environments, while developers automate code quality checks through shared pipelines

**Observability & Incident Response**
IT ensures monitoring and quick issue resolution; developers use feedback for rapid fixes




*Software teams focus on secure, agile feature development and embedding ML/AI into applications*

*IT teams empower them through cloud platforms, CI/CD, monitoring, and rapid issue resolution*

**DevSecOps**
IT embeds policies across the lifecycle while developers adopt secure coding practices early in development

**Cloud-Native Development**
Developers build applications as microservices using IT-managed cloud environments, which allow scaling & efficient resource utilization

**DevOps Integration**
IT teams now manage CI/CD pipelines, while software developers adopt agile workflows

**Source:** Draup's internal analysis and simulation of future capabilities of Generative AI, customer engagements, and surveys
**Note**: Draup analyses 16+ Million data attributes every day to help global HR leaders solve their challenges

Copyright © 2025 DRAUP. All Rights Reserved.

# Data Science and IT now operate in sync, with IT delivering scalable infrastructure, security, and deployment pipelines while data scientists refine and improve models in real time—enabling enterprise-wide, production-ready AI

## Traditional Workload

**Data Science**

Traditionally reporting-driven and research-heavy, focusing on insights and prototype models rather than scalable, production-ready solutions.

***Example**: A data scientist might deliver a detailed trend analysis or a high-accuracy prototype model, but it remained confined to reports or test environments without integration into business operations.*

**IT**

Traditional IT in data science played only a support role, lacking collaboration, tools, and processes to manage or scale machine learning models effectively.

***Example:** IT provisioned Hadoop clusters or on-premise servers, but these were managed separately from data science activities.*

## Current Interdependencies - Integrated Workloads

### ML/AI Model Deployment

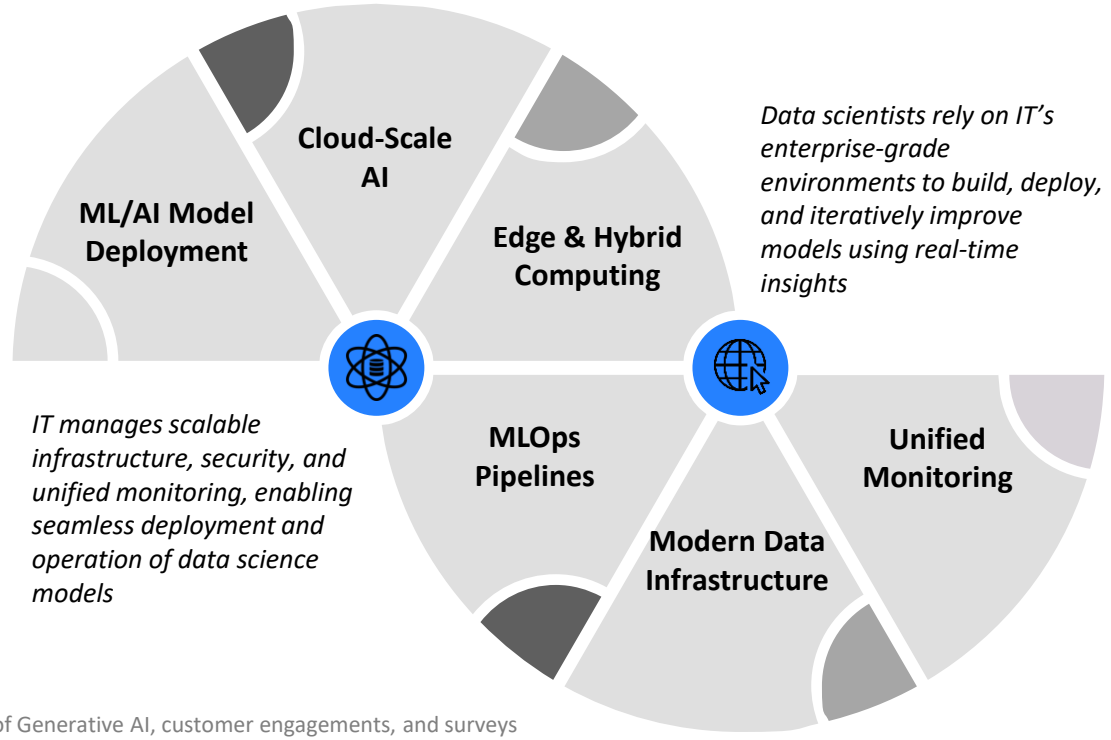
Data scientists now depend on IT to build and maintain scalable data pipelines and cloud infrastructure, shifting from small-scale, local experiments to enterprise-wide AI deployments

### Cloud-Scale AI

IT provisions high-performance computing environments, including GPUs and distributed systems, allowing data scientists to train large, complex models efficiently

### Edge & Hybrid Computing

IT enables analytics to run at the edge or in hybrid setups, allowing data scientists to deploy models where data is generated, minimizing latency




*Data scientists rely on IT's enterprise-grade environments to build, deploy, and iteratively improve models using real-time insights*

*IT manages scalable infrastructure, security, and unified monitoring, enabling seamless deployment and operation of data science models*

### MLOps Pipelines

IT now manages the deployment, monitoring, and scaling of machine learning models, while data scientists continuously update models based on real-time performance metrics

### Modern Data Infrastructure

IT teams have implemented enterprise-grade data lakes, streaming platforms, and governance frameworks, enabling data scientists to work with high-quality, real-time data at scale

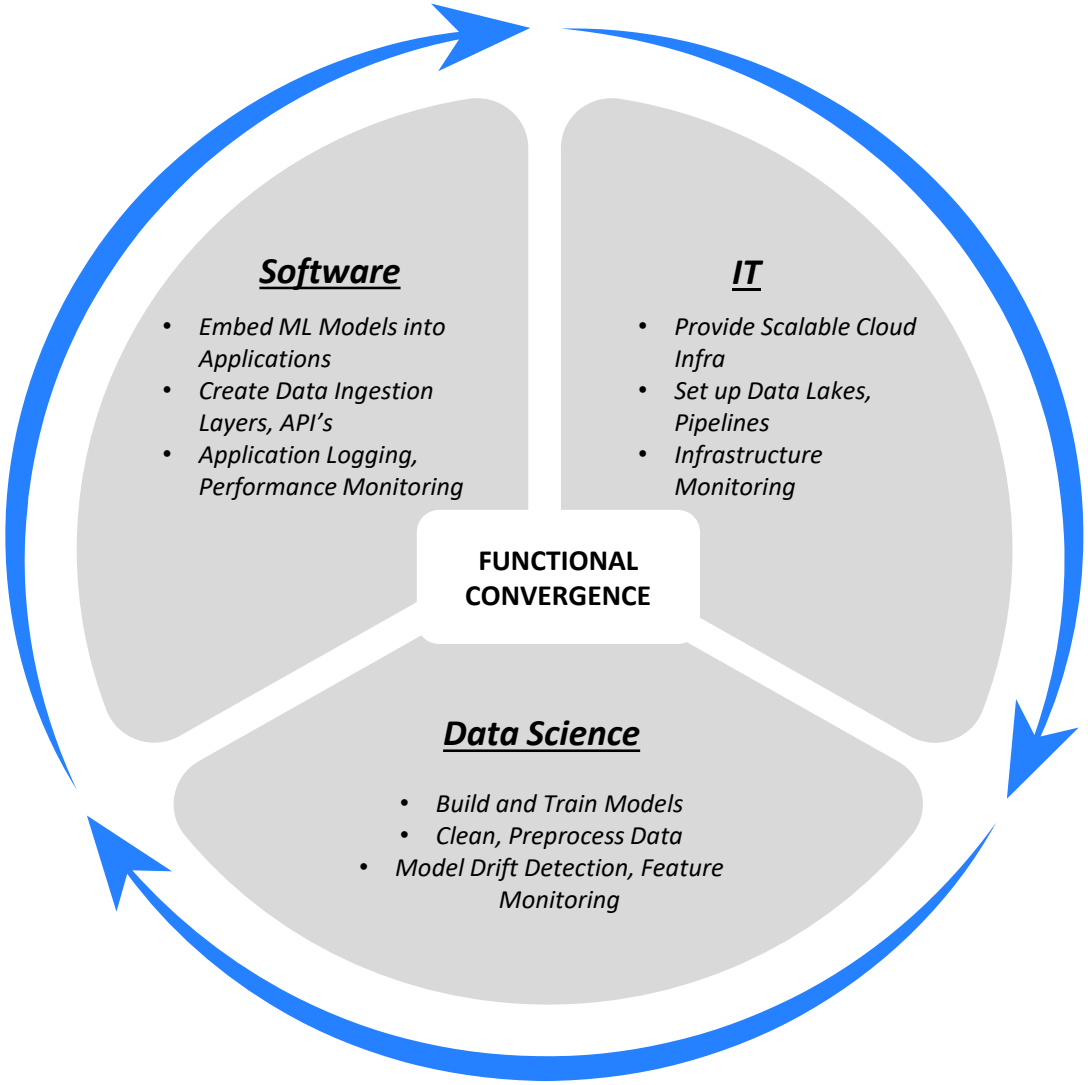
### Unified Monitoring

IT provides centralized monitoring that tracks infrastructure and model performance, helping data scientists detect drifts while IT ensures stability

**Source:** Draup's internal analysis and simulation of future capabilities of Generative AI, customer engagements, and surveys
**Note**: Draup analyses 16+ Million data attributes every day to help global HR leaders solve their challenges

Copyright © 2025 DRAUP. All Rights Reserved.

## Cross-Role Interactions and Functional Convergence: Traditionally separate roles in software, IT, and data science now integrate closely through shared responsibilities, enabling faster innovation, seamless experimentation, and stronger compliance




| Business Function | Evolved Workloads | Functional Convergence |
|---|---|---|
| Software | • AI integration has shifted software workloads toward embedding intelligent models, creating robust data ingestion mechanisms, and continuously monitoring application performance | • ***Jointly Co-owns MLOps Platforms:*** *Software engineers integrate models into applications, data scientists refine and monitor them, while IT provides the scalable infrastructure to support end-to-end MLOps*<br>• ***Develop API's, Automate Workflows:*** *Software teams design APIs, data team create automation workflows, and IT ensures these are securely deployed and maintained*<br>• ***Ensure Security, Governance Frameworks:*** *IT implements infrastructure security, software teams follow secure coding practices, and data team ensure model usage meets compliance and governance standards* |
| IT | • The adoption of AI has expanded IT responsibilities to building scalable cloud environments, orchestrating data pipelines, and proactively monitoring infrastructure to support AI-driven operations | |
| Data Science | • With AI, data science workloads have evolved to not only focus on model development but also on rigorous data preparation and continuous monitoring to detect drift and sustain model accuracy | |

Copyright © 2025 DRAUP. All Rights Reserved.

**Source:** Draup's internal analysis and simulation of future capabilities of Generative AI, customer engagements, and surveys
**Note**: Draup analyses 16+ Million data attributes every day to help global HR leaders solve their challenges

**Tech-stack Interconnect:** Technology stacks now interconnect to enable unified workflows and shared capabilities which enhances collaboration, accelerates deployment, and allows teams to leverage common tools

## Software Tech Stack

**Frontend:**

- *JavaScript frameworks (React, Angular, Vue.js), HTML5, CSS3*

**Backend:**

- *Programming languages (Java, Python, Node.js, Go, C#), frameworks (Spring, Django, Express.js)*

**Databases:**

- *Relational (SQL: PostgreSQL, MySQL), NoSQL (MongoDB, Redis)*

**DevOps/Deployment**

- *CI/CD (Jenkins, GitHub Actions), Containers (Docker, Kubernetes), Cloud (AWS, Azure, GCP)*

**APIs:** Data scientists expose their ML models via APIs (e.g., using FastAPI or Flask in Python). Software engineers integrate these APIs into the backend of applications

**MLOps Platforms**: Tools like MLflow and Kubeflow help manage the lifecycle of a model, from training (Data Science) to deployment and monitoring (Software/IT)

## Data Science Tech Stack

**Data Processing & Storage:**

- *Apache Spark, Hadoop, Data Warehouses (Snowflake, BigQuery)*

**ML & Scientific Computing:**

- *Python libraries (Pandas, NumPy, Scikit-learn), ML Frameworks (TensorFlow, PyTorch)*

**Data Visualization:**

- *Tableau, Power BI, Python libraries (Matplotlib, Seaborn)*

**Workflow & Orchestration:**

- *Apache Airflow*

**Cloud Data Services:** IT provisions and manages services like Amazon S3, Azure Blob Storage, and Google Cloud Storage, which data teams use to store raw data and model artifacts

**Managed Databases & Warehouses:** IT manages access and security for databases like Amazon RDS or data warehouses like Google BigQuery that data scientists query

## IT Tech Stack

**Infrastructure as Code (IaC):**

- *Terraform, Ansible, Puppet*

**Networking:**

- *Firewalls, Load Balancers, DNS, VPNs*

**Operating Systems:**

- *Linux (Ubuntu, CentOS), Windows Server*

**Monitoring & Logging:**

- *Prometheus, Grafana, ELK Stack (Elasticsearch, Logstash, Kibana)*

Copyright © 2025 DRAUP. All Rights Reserved.

**Interconnect across tech stack (Software + Data Science + IT)**

### Cloud Platforms – AWS, Azure, GCP

Cloud platforms provide a scalable environment for all teams—**software deploys applications**, **IT manages infrastructure**, and **data science runs advanced analyses**.

### Python

Python acts as a common language, enabling **software to build backends**, **IT to automate processes**, and **data science to analyze data and develop models**.

### Git & Version Control – GitHub, GitLab

Git tools such unify workflows by allowing **software to track code**, **IT to manage infrastructure scripts**, and **data science to version experiments for reproducibility**.

**Source:** Draup's internal analysis and simulation of future capabilities of Generative AI, customer engagements, and surveys
**Note**: Draup analyses 16+ Million data attributes every day to help global HR leaders solve their challenges

# Key Trends Shaping the Tech Ecosystem

**Generative AI Copilots are driving exponential adoption across Software, Data Science and IT, unlocking improvements in productivity, code quality, and IT efficiency, while enabling organizations to scale innovation and reduce operational friction at unprecedented speed**

## Rise of Generative AI Copilots

### Explosive Developers Usage

15 million developers were using GitHub Copilot by early 2025—a 4× increase in just 12 months

### Rising Organizations Adoption

Growth among enterprise organizations accelerated, with 90% of Fortune 100 companies now using Copilot

### Explosive User Growth

July 2025, Copilot surpassed 20 million all-time users, 5 million new trial users joined in last three months

### Increased AI Contribution

73% of open-source contributors use AI tools like Copilot, heavily supporting data science tasks in Python and notebooks

## Software

- **Accelerated Coding:** Tools like **GitHub Copilot**, **Tabnine**, and **Amazon CodeWhisperer** provide real-time code suggestions, auto-complete features, and even generate entire functions or code blocks based on natural language prompts

GitHub Copilot tabnine Amazon Code Whisperer

- **Improved Code Quality and Best Practices:** By analyzing vast amounts of code from public repositories, copilots such as **Codiga** and **Snyk Code** can suggest clean, efficient, and error-free code snippets

snyk Codiga DEEPCODE

- **Debugging and Error Resolution:** Copilots can help identify and fix errors by providing explanations for bugs and suggesting code corrections. For instance, **Cody by Sourcegraph** can search through your entire codebase and documentation to answer questions

Cody CURSOR

## Data Science

- **Streamlined Data Analysis:** Copilots such as **PandasAI** allows data scientists to use conversational language to analyze data, while tools like **Microsoft 365 Copilot** can generate formulas and analyze data in Excel with simple prompts

PandasAI Copilot

- **Enhanced Productivity and Efficiency:** Data scientists are using copilots like **GitHub Copilot** to generate code for machine learning models, saving time on coding and allowing them to focus on high-level analysis and problem-solving

GitHub Copilot

- **Error Reduction:** Copilots help reduce errors by suggesting correct syntax and function calls for libraries like Pandas, NumPy, and Matplotlib. **Julius AI** is a tool specifically designed for data analysis that can help in this regard

Julius

## IT

- **Streamlined IT Service and Support:** Copilots are revolutionizing the IT help desk. Tools like **ServiceNow's Virtual Agent** and **Moveworks** use natural language processing to automate the resolution of common issues such as password resets, account unlocks, and software installation problems

servicenow Moveworks Instant help at work.

- **Proactive Infrastructure Management:** Copilots analyze system logs and performance data to predict and prevent potential failures. **Microsoft's Copilot in Azure** helps IT professionals troubleshoot issues and optimize cloud configurations using natural language queries

Copilot

- **Automated IaC:** Copilots can generate Infrastructure as Code (IaC) snippets and command-line scripts, which accelerates infrastructure deployment and reduces errors

Copyright © 2025 DRAUP. All Rights Reserved.

**Source:** Draup's internal analysis and simulation of future capabilities of Generative AI, customer engagements, and surveys
**Note**: Draup analyses 16+ Million data attributes every day to help global HR leaders solve their challenges

**Platform engineering is becoming a critical enabler for scaling modern workloads, by building shared infrastructure, standardizing environments, and fostering seamless cross-team collaboration—ultimately driving faster deployments, reduced complexity, and greater organizational agility**

draup

## Evolving Workloads

**Complex Cloud-Native Ecosystems**
Microservices, containers, and multi-cloud setups have increased operational complexity.

**AI & ML Integration**
Software and data science teams now require scalable GPU/infra pipelines, adding new layers of dependency.

**Growing Cross-Team Interdependencies**
Development, IT, and data science teams are increasingly interconnected, requiring tighter alignment to deliver at scale.

| Limitations of Current Approach | ***DevOps at Scale Fails:*** *Tool sprawl and inconsistent processes make scaling hard* | ***Ticket-Driven IT Support:*** *Slows down developers and introduces bottlenecks* | ***Lack of Standardization:*** *Each team builds its own workflows, causing governance and cost issues* |
|---|---|---|---|

**Expansion of Platform Engineering**

**Siloed Approach**
- **Separate tools** for IT, Data Science & Software
- **Manual infrastructure setup** by IT
- **Slow deployment cycles** & misaligned workflows

**Platform Engineering Expansion**
- Builds **shared infrastructure & pipelines**
- Standardizes environments (cloud, CI/CD, MLOps)
- Enables **cross-team access & collaboration**

**Integrated Workflow**
- Faster, **automated deployments**
- **Seamless integration of ML models** into applications
- **Strong collaboration** across teams

A **unified MLOps platform** allows data scientists to **train models**, IT to **manage scalable infrastructure**, and software engineers to **integrate models into apps effortlessly**.

### Case Study

Netflix modernized its ML workflows by adopting **Metaflow**, Netflix's platform engineering approach enabled highly automated deployment pipelines and robust observability systems. As a result, the company achieved **over 4,000 daily deployments while reducing operational overhead by 70%.**

By implementing a **unified internal developer platform**, Spotify streamlined workflows and standardized tooling across teams. This transformation led to a **30% faster delivery of new features and improved developer experience**.

**Source:** Draup's internal analysis and simulation of future capabilities of Generative AI, customer engagements, and surveys
**Note**: Draup analyses 16+ Million data attributes every day to help global HR leaders solve their challenges

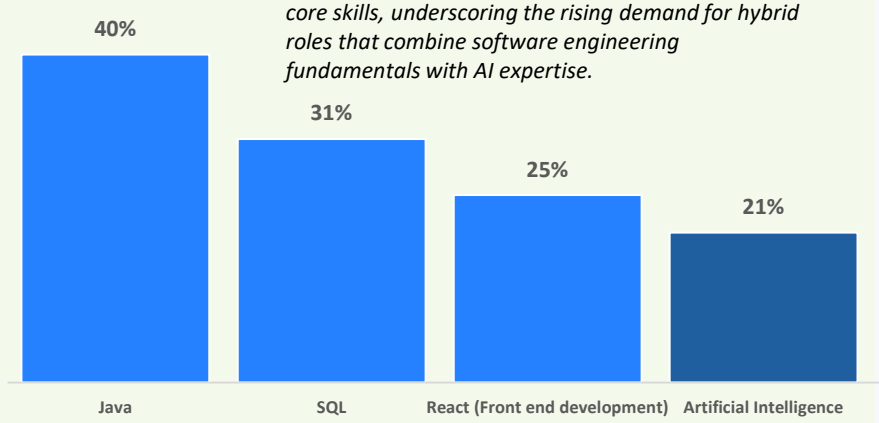
Copyright © 2025 DRAUP. All Rights Reserved.

# AI adoption is accelerating across business functions, led by IT, Data, and Software teams; This shift is intensifying demand for hybrid roles—such as AI Data Scientists and ML Specialists—where scarce cross-domain expertise is emerging as a critical constraint

## Hybrid Skillsets Redefining Tech Roles: % Share of AI and Cloud Drive Demand in U.S. Software & Data Science Job Postings

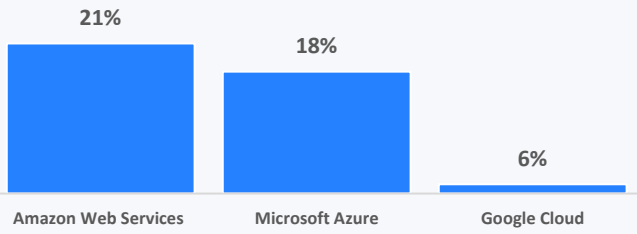
### AI Skills becoming Mainstream in Software Engineering

*The prevalence of AI skills in software engineering has surged to nearly match that of some traditional core skills, underscoring the rising demand for hybrid roles that combine software engineering fundamentals with AI expertise.*




### Cloud Expertise Becomes Core to Data Science Roles

*The growing prevalence of cloud tools like AWS and Azure in data science job postings underscores the rising demand for hybrid roles that blend data science expertise with cloud deployment and engineering skills.*




## Transformation to Hybrid Role (Software Engineer)

### Workload Evolution

Core workloads are expanding beyond application development to include **ML model deployment, real-time data processing, and AI-enabled feature engineering**

### Skills Shift

Software engineers are increasingly expected to work with machine learning frameworks **(TensorFlow, PyTorch), model integration APIs**, and data preprocessing pipelines, in addition to traditional programming skills like Python, Java, and C++

## Key Insights Highlighting the Demand for Hybrid Roles

### Explosive Growth in AI-Skilled Data Roles

- **Jobs for data scientists—a role deeply intertwined with AI/ML expertise—are projected to grow by 36% between 2023–2033**, significantly outpacing the average for other tech occupations. This highlights the criticality of AI-driven data talent

### Engineering Teams Undergoing AI Upskilling

- By 2027, an overwhelming majority of engineering teams will need to reskill and upskill to align with generative AI adoption in development workflows. This shift is already visible: **62% of developers are using AI coding tools**, with nearly half leveraging them on a daily basis

**Source:** Draup's internal analysis and simulation of future capabilities of Generative AI, customer engagements, and surveys
**Note**: Draup analyses 16+ Million data attributes every day to help global HR leaders solve their challenges

Copyright © 2025 DRAUP. All Rights Reserved.

# Section – 8: Appendix

# In Software Engineering, core languages like JavaScript, Python, along with tools such as React, Docker, & Kubernetes, remain dominant in modern software delivery. At the same time, AI-powered coding assistants like GitHub Copilot, and Claude are improving development while microservices, API-first design, and cloud-native architectures drive scalability

draup

| Core Skills Still in Demand - Software Engineering Job Family | | | | |
|---|---|---|---|---|
| **Systems Architecture** | **Software Development** | **Software Testing & QA** | **Database Management** | **DevOps & Infra** |
| System Architecture Design & Scalability Microservices) | Core Programming Languages (Java, C++, Python, Go) | CI/CD Deployment (GitLab CI, GitHub Actions) | SQL & Relational Database (MYSQL, Postgre SQL | Containerization & orchestration (Docker, Kubernetes) |
| API Design & Integration - (REST, SOAP) | Backend Development (Java Spring, Node.js) Web Development (JavaScript) | Integration Testing (Cypress) | NoSQL Databases (MongoDB, Cassandra) | Monitoring & observability (Nagios, Prometheus, Grafana), Cloud Platforms (AWS, Azure, GCP) |
| | Full-stack development with modern frameworks (React, Node.js, etc. | Automation testing frameworks (Selenium, JUnit, TestNG) | | Infrastructure as Code (Terraform, Ansible, Puppet, Chef) Configuration Management |
| | | SOFTWARE ENGINEERING EMERGING SKILLS | | |
| Serverless Architecture (AWS Lambda, Azure Functions), | AI/ML Application Development (TensorFlow, PyTorch) | AI/ML Test Automation, Testing for ML Models | Cloud-Native & Distributed Databases (CockroachDB, YugabyteDB, TiDB) | Machine learning Operations |
| Event-Driven Architecture (Kafka, RabbitMQ) | AI Assisted Coding, Quantum Programming | Chaos Engineering - (Gremlin, LitmusChaos) | DataOps & Automated Data Pipelines (dbt, Airbyte, Fivetran) | Advanced Infrastructure as Code (IaC) (Pulumi, Terraform Cloud) |
| API gateway & GraphQL Development | Cross Platform Mobile Development | Generative Adversarial Network (GAN) | | AIOps for proactive monitoring & remediation |
| Cloud-Native Architecture, Zero Trust Architecture | | Testing for ML Models (TensorFlow Extended – TFX testing frameworks) | | |

| Gen AI & Automation Tools Paving the Path for streamlining Software Engineering Capabilities (Not Exhaustive) | | | | | |
|---|---|---|---|---|---|
| **AI Code Assistance** | **Low-code/No-code Development** | **Automated Testing & QA** | **Database Automation** | **DevOps & Infra** | **Gen AI & Automation Frameworks** |
| GitHub Copilot, tabnine, codium ai | mendix, bubble, outsystems | testim, functionize | DataRobot, amazon RDS | DATADOG, spacelift, harness | Hugging Face, Polymer, LangChain, OpenAI Codex |

Copyright © 2024 DRAUP. All Rights Reserved.

**Note:** Draup's proprietary models leverage 850 Mn+ JDs to identify the relevant skills associated with each of the skills clusters analyzed based on the demand for any job family. Listed Skills are not exhaustive. The list of Digital and Gen AI tools is for representational purposes and non-exhaustive

## Software Tech Stack

| Category | Tools | Platforms | Languages |
|---|---|---|---|
| **Systems Architecture** | Enterprise Architect, Lucidchart, Draw.io, Prometheus, Grafana, ELK Stack, Terraform, Pulumi | AWS (API Gateway, Lambda), Azure (Service Fabric), GCP (Cloud Run), Kubernetes, Docker Swarm | Python, Java, Go, C#, Scala |
| **Software Development** | Visual Studio Code, IntelliJ IDEA, PyCharm, WebStorm, Git, Maven, Gradle, Webpack, Vite, Jest, Mocha, Pytest | AWS, Azure, GCP, Heroku, GitHub, GitLab, Bitbucket, OutSystems, Mendix | Python, JavaScript/TypeScript, Java, C++, Rust, Kotlin, Swift |
| **Software Testing & QA** | Selenium, Cypress, Playwright, Appium, JMeter, Gatling, Locust, Postman, TestRail, Zephyr, OWASP ZAP, Burp Suite | BrowserStack, Sauce Labs, Jenkins, GitHub Actions, GitLab CI | Python, Java, JavaScript, C# |
| **Database Management** | DBeaver, pgAdmin, SQL Server Management Studio, Apache Airflow, Talend, Informatica, ER/Studio | PostgreSQL, MySQL, Oracle, SQL Server, MongoDB, Redis, DynamoDB, Cassandra, Neo4j, AWS RDS, Azure Cosmos DB, GCP BigQuery | SQL, PL/SQL, T-SQL |
| **DevOps & Infrastructure** | Jenkins, GitHub Actions, GitLab CI, CircleCI, Docker, Podman, Kubernetes, Terraform, Ansible, Prometheus, Grafana, ELK Stack, Vault, Snyk | AWS, Azure, GCP, Oracle Cloud, Istio, NGINX, Traefik | Python, Bash, Go, PowerShell |

Copyright © 2024 DRAUP. All Rights Reserved.

**In today's IT landscape, core skills like ITSM, cloud governance, security, and end-user support remain prevalent, as the backbone of reliable operations, while emerging capabilities such as AIOps, FinOps, DevSecOps, and DEX optimization are redefining efficiency. GenAI tools now streamline workflows, shifting IT from execution to strategic orchestration.**

draup

| Still in Demand Core IT Skills - IT Job Family | | | |
|---|---|---|---|
| **IT Service Management** | **Cloud Infra & Ops** | **IT Security Governance & Compliance** | **End User Support & Monitoring** |
| ITIL Processes (Incident, Problem, Change, Release Management | Multi Cloud Platform Management (AWS, Azure, GCP), | Identity & Access Management (IAM) - Okta, CyberArk, Azure AD | Remote Troubleshooting & Helpdesk Management |
| ServiceNow (dominant ITSM platform) | Observability - Datadog, Prometheus, Grafana (cloud monitoring, Virtualization (VMware, Hyper-V) | Disaster Recovery & Incident Response, Tenable, Qualys (vulnerability management), DevSecOps | Digital Experience Monitoring (DEX) |
| DevOps & Agile-aligned ITSM | Infrastructure as Code (Terraform, Ansible, Chef, Puppet) | Zero Trust Architecture, ITIL Governance – GDPR, ISO, Cloud Security Posture Management | IT Asset Lifecycle Management, Automated Provisioning |
| ITSM Integration with Cloud & DevOps pipelines – Jira, Zendesk, Fresh service, | Containers & Orchestration (Docker, Kubernetes, OpenShift) | Tools - Splunk, QRadar, Sentinel, CrowdStrike, Palo Alto, Fortinet (endpoint & network security) | Zendesk, Freshservice (support platforms), Jira Service Management |
| **IT EMERGING SKILLS** | | | |
| AI-powered ITSM (AIOps for automated incident detection | Automated Self-Healing Systems, AI-driven Cloud Cost Optimization, GreenOps | AI-driven Threat Detection & Response, Threat Hunting & Cyber Threat Intelligence | AI Chatbots for IT Helpdesk |
| Conversational ITSM (chatbots, virtual agents for support) | Multi-Cloud & Hybrid Infrastructure Management, Predictive Infrastructure Scaling | Automated Phishing Detection, Behavior Analytics, Automated Compliance Audits | Digital Experience Monitoring & Optimization, Expérience Analytics – Systrack, Catchpoint |
| Self-Service Portals & Knowledge Bases, Intelligent Workflow Automation - Servicenow | Tools - Chaos Engineering tools (Gremlin, LitmusChaos), Anthos, Azure Arc, AWS Outposts (hybrid/multi-cloud ops) | AI for Fraud & Anomaly Detection, Cloud Security Posture Management (CSPM) | Predictive User Issue Resolution - ServiceNow Virtual Agent, |
| Tools - BMC Helix AIOPS, Moveworks, Amelia, Kore.ai (AI chatbots for IT support), Service Now Predictive Analytics | AI-driven Infrastructure Automation (self-scaling, self-healing), AI-based Resource Optimization, Edge & Distributed Cloud Management | SOAR platforms (Splunk Phantom, Palo Alto Cortex XSOAR), SentinelOne, Darktrace (AI-driven threat detection) | Nexthink Infinity, ControlUp (DEX platforms), TeamViewer AR Assist, Microsoft Dynamics Remote Assist (AR/VR IT support) |

## Gen AI & Automation Tools Paving the Path for streamlining IT Capabilities (Not Exhaustive)

| IT Service Management & Support | Software Development & DevOps | IT Security Governance & Compliance | End User Support & Monitoring |
|---|---|---|---|
| servicenow AI Platform | harness | DARKTRACE | Charlotte AI |
| Moveworks Instant help at work. | tabnine | deepinstinct | IBM Watson Assistant |
| BMC HelixGPT Generative AI services in the BMC Helix platform | codium ai | Microsoft Security Copilot | control UP |
| AISERA | DATADOG | | Nexthink Infinity Platform |
| | | | lakeside SysTrack |

Copyright © 2024 DRAUP. All Rights Reserved.

**Note:** Draup's proprietary models leverage 850 Mn+ JDs to identify the relevant skills associated with each of the skills clusters analyzed based on the demand for any job family. Listed Skills are not exhaustive. The list of Digital and Gen AI tools is for representational purposes and non-exhaustive

| Category | Tools | Platforms | Languages |
|---|---|---|---|
| **IT Service Management** | ServiceNow, BMC Remedy, Jira Service Management, PagerDuty, Opsgenie, Ivanti, Lansweeper, Power Automate | AWS Systems Manager, Azure Automation, Microsoft 365, Google Workspace | Python, JavaScript, PowerShell |
| **Cloud Infrastructure & Operations** | Terraform, Ansible, CloudFormation, Chef, Puppet, CloudWatch, Azure Monitor, Google Operations Suite, Docker, Kubernetes | AWS (EC2, S3, Lambda, EKS), Azure (VMs, AKS), GCP (Compute Engine, GKE) | Python, Bash, Go, PowerShell |
| **IT Security & Compliance** | Splunk, Elastic Security, Microsoft Sentinel, Nessus, Qualys, OpenVAS, CrowdStrike, Microsoft Defender, Metasploit, Burp Suite, Okta, Wazuh, Drata, Vanta | AWS Security Hub, Azure Security Center, GCP Security Command Center, Palo Alto Networks, Cisco Secure | Python, Bash, PowerShell |
| **End-User Support & Monitoring** | Zendesk, Freshdesk, ServiceNow, TeamViewer, AnyDesk, Microsoft Intune, Jamf, SCCM, Zabbix, SolarWinds, PRTG, Power Automate, UiPath | Windows, macOS, Linux, iOS, Android, Microsoft 365, Google Workspace | PowerShell, Python, JavaScript |

Copyright © 2024 DRAUP. All Rights Reserved.

# Core data science skills in analysis, modelling, engineering, and MLOps remain vital for building reliable data foundations. Emerging practices like DataOps, real-time streaming, explainable AI, & vector databases are redefining workflows. GenAI tools are accelerating data-to-insight pipelines, shifting data roles toward strategy and innovation.

draup

| Core Skills Still in Demand – Data Science | | | | |
|---|---|---|---|---|
| **Data Analysis** | **Data Modelling** | **Data Engineering** | **MLOps** | **Programming & Tools** |
| Database Design (SQL, NoSQL) | Data Warehousing, Augmentation | ETL Pipelines - Talend, Relational Database | Model training, deployment, and monitoring | Programming Languages – R, Python, Go |
| Statistical Analysis | Multi-Dimensional Scaling | Time Series Forecasting (ARIMA, Exponential Smoothing) | CI/CD pipelines for ML (MLflow, Kubeflow, Torch Serve) | GPU Programming |
| Business Intelligence Tools– Tableau, Power BI | Dimensional & ER Modelling – Snowflake | Big Data Frameworks – Hadoop, Spark | Anomaly Detection, Word Embedding Technique, Décision Trees | Jupyter Notebooks, RStudio for experimentation |
| Causal Inference, Data Storytelling | | Relational & NoSQL Databases (Postgres, MongoDB, Cassandra) | Support Vector Machines (SVM), KNN, Principal Component Analysis (PCA) | |
| **DATA SCIENCE EMERGING SKILLS** | | | | |
| Natural Language Querying (NLQ), | Ethical Data Handling, Metadata-driven modelling | Ethical AI Governance | Deep Reinforcement Learning, LLMOps for Deployment | Automated Code Generation, Low-code/No-code ML platforms |
| Predictive & Prescriptive Analytics, | Differential Privacy (DP) Techniques, Multimodal data modelling (text, images, audio combined) | Real-time Streaming Data (Kafka, Flink, Pulsar) | Synthetic Data Generation for Model Training, Responsible AI | Cloud Native SDKs- HuggingFace, LangChain, H2Oai |
| AI-driven Feature Engineering | Large-Scale Generative Modelling | Cloud-native Data Engineering (AWS Glue, Azure Synapse, GCP Dataflow) | AutoML orchestration at scale, Vertex AI | AI-based Query Generation (SQL-to-NL), Retrieval Augmented Generation (RAG), |
| | Bayesian Hierarchical Modeling | Prompt Engineering | Explainable AI (XAI) for model transparency | Prompt Engineering & LLM Integration |

| Gen AI & Automation Tools Paving the Path for streamlining Data Science Capabilities (Not Exhaustive) | | | | |
|---|---|---|---|---|
| **Model Deployment & Assist** | **Gen AI & Automation Frameworks** | **Data Engineering** | **Testing & Automation** | **Privacy & Responsible AI** |
| databricks, Vertex AI, OctoAI | Hugging Face, Claude, LangChain, OpenAI Codex | AWS Glue, DataRobot, Copilot in Power BI, databricks | Gemini Code Assist, EVIDENTLY AI, arize | IBM AI Fairness 360 Toolkit (AIF360), Arthur, fiddler |

Copyright © 2024 DRAUP. All Rights Reserved.

**Note:** Draup's proprietary models leverage 850 Mn+ JDs to identify the relevant skills associated with each of the skills clusters analyzed based on the demand for any job family. Listed Skills are not exhaustive. The list of Digital and Gen AI tools is for representational purposes and non-exhaustive

# Data Science Tech Stack

| Category | Tools | Platforms | Languages |
|---|---|---|---|
| **Data Analysis** | Tableau, Power BI, Looker Studio, Qlik Sense, Jupyter, RStudio, Pandas, dplyr, NumPy, SPSS, SAS | AWS QuickSight, Azure Synapse Analytics, Google BigQuery, Databricks, Google Colab | Python, R, SQL |
| **Data Modeling** | Scikit-learn, caret, Statsmodels, Prophet, SimPy, Seaborn, Plotly, Matplotlib, Great Expectations, Deequ | AWS SageMaker, Azure Machine Learning, Google Vertex AI, Snowflake, BigQuery, Redshift | Python, R, SQL |
| **Data Engineering** | Apache Airflow, Apache NiFi, Talend, dbt, Apache Kafka, Apache Flink, Apache Spark, Delta Lake, Apache Iceberg, Apache Superset, Metabase | PostgreSQL, MongoDB, Cassandra, DynamoDB, Snowflake, AWS Glue, Azure Data Factory, GCP Dataflow | Python, Scala, Java, SQL |
| **ML Ops** | TensorFlow, PyTorch, Scikit-learn, XGBoost, MLflow, Kubeflow, Seldon, BentoML, Weights & Biases, Evidently AI, DVC | AWS SageMaker, Azure ML, Google Vertex AI, Kubernetes | Python, Bash |
| **Programming & Tools** | Jupyter Notebook, VS Code, PyCharm, RStudio, Git, Apache Airflow, Prefect, Matplotlib, Seaborn, Plotly, ggplot2 | Databricks, AWS SageMaker, Azure Synapse Analytics, GCP BigQuery, GitHub, GitLab | Python, R, Julia, Scala |

Copyright © 2024 DRAUP. All Rights Reserved.

# Glossary – Software Tech Stack

| Tools | Platforms | Languages |
|---|---|---|
| Apache Airflow: Workflow orchestration tool for scheduling and managing data pipelines. | AWS (Amazon Web Services): Cloud platform offering services like EC2, Lambda, RDS, and API Gateway. | Bash: Scripting language for automating system tasks in Unix-like environments. |
| Burp Suite: Security testing tool for web application penetration testing. | Azure: Microsoft's cloud platform with services like Service Fabric, SQL Database, and Functions. | C#: Object-oriented language for building Windows and web applications. |
| Cypress: JavaScript-based end-to-end testing framework for web applications. | Bitbucket: Git-based source code hosting and CI/CD platform. | C++: High-performance language for system-level programming. |
| DBeaver: Universal SQL client for managing and querying multiple databases. | BrowserStack: Cloud platform for cross-browser and device testing. | Go: Statically typed language for scalable and concurrent systems. |
| Draw.io: Diagramming tool for creating architecture and flowchart diagrams. | Cassandra: Distributed NoSQL database for high scalability and availability. | Java: Versatile, object-oriented language for enterprise applications. |
| ELK Stack (Elasticsearch, Logstash, Kibana): Suite for log aggregation, processing, and visualization. | Docker Swarm: Container orchestration platform, simpler alternative to Kubernetes. | JavaScript: Language for web development and server-side programming |
| Enterprise Architect: Modeling tool for designing and documenting system architectures. | DynamoDB: AWS NoSQL key-value and document database. | Kotlin: Modern language for Android development and backend services. |
| ER/Studio: Data modeling tool for designing relational and NoSQL databases. | GCP (Google Cloud Platform): Cloud platform with services like Cloud Run, BigQuery, and GKE. | Python: General-purpose language for web, automation, and scripting. |
| Gatling: Performance testing tool for simulating high-load scenarios. | GitHub: Platform for Git hosting, collaboration, and CI/CD pipelines. | Rust: Systems programming language focused on safety and performance. |
| Git: Version control system for tracking code changes. | GitLab: Git-based platform for source code management and CI/CD. | Scala: Functional and object-oriented language for distributed systems. |
| Gradle: Build automation tool for Java and other languages. | Heroku: Platform as a Service (PaaS) for deploying and scaling applications. | SQL: Query language for managing relational databases. |
| Informatica: Data integration platform for ETL/ELT processes. | Istio: Service mesh platform for managing microservices communication. | Swift: Language for iOS and macOS app development. |
| IntelliJ IDEA: IDE for Java and other languages, with advanced coding features. | Kubernetes: Container orchestration platform for automating deployment and scaling. | TypeScript: Superset of JavaScript with static typing for large-scale applications. |
| Jenkins: Open-source CI/CD server for automating build, test, and deployment. | MongoDB: NoSQL document database for flexible data storage. | PL/SQL: Procedural language for Oracle Database programming. |
| Jest: JavaScript testing framework for unit and integration tests. | MySQL: Open-source relational database management system. | T-SQL: Microsoft's extension of SQL for SQL Server. |
| JMeter: Performance testing tool for load and stress testing. | Neo4j: Graph database for relationship-based data queries. | |
| Locust: Open-source load testing tool written in Python. | NGINX: High-performance web server and reverse proxy. | |
| Lucidchart: Cloud-based diagramming tool for architecture and collaboration. | Oracle Database: Enterprise-grade relational database management system. | |
| Maven: Build automation tool primarily for Java projects. | OutSystems: Low-code platform for rapid application development. | |
| Mocha: JavaScript testing framework for asynchronous testing. | PostgreSQL: Open-source relational database with advanced features. | |
| OWASP ZAP: Open-source tool for finding web application vulnerabilities. | Redis: In-memory key-value store for caching and real-time applications. | |
| pgAdmin: Administration and management tool for PostgreSQL databases. | Sauce Labs: Cloud platform for automated testing across browsers and devices. | |
| Playwright: Browser automation tool for end-to-end testing across browsers. | SQL Server: Microsoft's relational database management system. | |
| Podman: Container management tool, alternative to Docker. | Traefik: Cloud-native reverse proxy and load balancer. | |
| Postman: API testing tool for designing, testing, and documenting APIs. | | |
| Prometheus: Monitoring and alerting toolkit for time-series metrics. | | |
| Pulumi: Infrastructure as Code (IaC) tool supporting multiple programming languages. | | |
| Pytest: Python testing framework for unit and integration tests. | | |
| Selenium: Browser automation tool for web application testing. | | |
| Snyk: Security tool for identifying vulnerabilities in code and dependencies. | | |
| SQL Server Management Studio (SSMS): IDE for managing Microsoft SQL Server databases. | | |
| Talend: Data integration platform for ETL and data management. | | |
| Terraform: Infrastructure as Code tool for provisioning cloud resources. | | |
| TestRail: Test case management tool for organizing and tracking tests. | | |
| Vite: Modern JavaScript build tool for fast frontend development. | | |
| Visual Studio Code (VS Code): Lightweight, customizable IDE for multiple languages. | | |
| Webpack: JavaScript module bundler for building web applications. | | |
| WebStorm: IDE optimized for JavaScript and TypeScript development. | | |
| Zephyr: Test management tool integrated with Jira for QA workflows. | | |

Copyright © 2024 DRAUP. All Rights Reserved.

# Glossary – IT Tech Stack

| Tools | Platforms | Languages |
|---|---|---|
| AnyDesk: Remote desktop software for end-user support. | AWS (Amazon Web Services): Cloud platform with services like EC2, S3, and Systems Manager. | Bash: Scripting language for automating IT tasks in Unix environments. |
| BMC Remedy: ITSM platform for incident and asset management. | Azure: Microsoft's cloud platform with Virtual Machines, AKS, and Security Center. | JavaScript: Language for scripting and web-based IT tools. |
| Burp Suite: Penetration testing tool for web application security. | Cisco Secure: Suite of network security solutions. | PowerShell: Microsoft's scripting language for Windows automation. |
| Chef: Configuration management tool for automating infrastructure setup. | Fortinet: Network security platform for firewalls and threat protection. | Python: General-purpose language for IT automation and scripting. |
| CrowdStrike: Endpoint security platform for threat detection and response. | GCP (Google Cloud Platform): Cloud platform with Compute Engine and Security Command Center. | |
| Drata: Compliance automation tool for frameworks like SOC 2 and ISO 27001. | Google Workspace: Collaboration suite for productivity and communication. | |
| Elastic Security: SIEM platform for security analytics and threat detection. | Kubernetes: Container orchestration for cloud infrastructure. | |
| Freshdesk: Helpdesk software for customer and end-user support. | Linux: Open-source operating system for servers and endpoints. | |
| Jamf: Management tool for macOS and iOS devices. | macOS: Apple's operating system for desktops and laptops. | |
| Jira Service Management: ITSM platform for ticketing and service workflows. | Microsoft 365: Productivity suite with collaboration and endpoint management tools. | |
| Kali Linux: Linux distribution for security testing and penetration testing. | Palo Alto Networks: Cybersecurity platform for firewalls and threat prevention. | |
| Lansweeper: IT asset management tool for network inventory. | Windows: Microsoft's operating system for desktops and servers. | |
| Metasploit: Penetration testing framework for security assessments. | | |
| Microsoft Intune: Endpoint management for mobile and desktop devices. | | |
| Microsoft Sentinel: Cloud-native SIEM for threat detection and response. | | |
| Nessus: Vulnerability scanning tool for identifying security risks. | | |
| Okta: Identity and access management platform for single sign-on. | | |
| OpenVAS: Open-source vulnerability scanner for security assessments. | | |
| Opsgenie: Incident management and alerting tool. | | |
| PagerDuty: Incident response platform for IT operations. | | |
| Power Automate: Microsoft tool for automating workflows and processes. | | |
| PRTG: Network monitoring tool for infrastructure and performance. | | |
| Puppet: Configuration management tool for infrastructure automation. | | |
| SaltStack: Configuration management and orchestration tool. | | |
| ServiceNow: ITSM platform for service management and automation. | | |
| SolarWinds: Network and system monitoring tool for IT operations. | | |
| TeamViewer: Remote desktop software for end-user support. | | |
| Terraform: Infrastructure as Code tool for provisioning resources. | | |
| UiPath: Robotic Process Automation (RPA) platform for IT tasks. | | |
| Vanta: Compliance management tool for automating audits. | | |
| Wazuh: Open-source security monitoring and SIEM platform. | | |
| Zendesk: Helpdesk platform for customer and end-user support. | | |

Copyright © 2024 DRAUP. All Rights Reserved.

# Glossary – Data Science Tech Stack

| Tools | Platforms | Languages |
|---|---|---|
| Apache Airflow: Workflow orchestration for data pipelines. | Apache Iceberg: Open table format for data lakes. | Java: Language for data engineering and big data frameworks. |
| Apache Flink: Stream processing framework for real-time data. | AWS (Amazon Web Services): Cloud platform with SageMaker, Glue, and Redshift. | Julia: High-performance language for numerical and scientific computing. |
| Apache Kafka: Distributed streaming platform for data pipelines. | Azure: Microsoft's cloud platform with Synapse Analytics and Data Factory. | Python: Primary language for data science, ML, and automation. |
| Apache NiFi: Data integration tool for ETL and data flow automation. | Cassandra: Distributed NoSQL database for large-scale data. | R: Language for statistical analysis and visualization. |
| Apache Spark: Big data processing framework for batch and streaming. | Databricks: Unified analytics platform for data engineering and ML. | Scala: Language for big data processing with Spark. |
| Apache Superset: Open-source data visualization and dashboarding tool. | Delta Lake: Storage layer for data lakes with ACID transactions. | SQL: Query language for data manipulation and analysis. |
| BentoML: Model serving framework for deploying machine learning models. | DynamoDB: AWS NoSQL key-value and document database. | |
| caret: R package for machine learning and model training. | GCP (Google Cloud Platform): Cloud platform with BigQuery and Vertex AI. | |
| dbt: Data transformation tool for analytics workflows. | GitHub: Platform for code hosting and collaboration. | |
| Deequ: Data quality validation library for Spark. | GitLab: Platform for code hosting and CI/CD. | |
| DVC (Data Version Control): Version control for machine learning datasets and models. | MongoDB: NoSQL document database. | |
| Evidently AI: Tool for monitoring machine learning model performance. | PostgreSQL: Relational database with advanced features. | |
| ggplot2: R library for advanced data visualization. | Redshift: AWS data warehouse for analytics. | |
| Great Expectations: Data quality and validation framework. | Snowflake: Cloud-based data warehouse for analytics. | |
| Jupyter Notebook: Interactive environment for data analysis and coding. | Kubernetes: Container orchestration for ML and data pipelines | |
| Kubeflow: Machine learning workflow platform for Kubernetes. | | |
| Looker Studio: Google's data visualization and reporting tool. | | |
| Matplotlib: Python library for plotting and visualization. | | |
| Metabase: Open-source business intelligence and visualization tool. | | |
| MLflow: Platform for managing machine learning lifecycles. | | |
| NumPy: Python library for numerical computations. | | |
| Pandas: Python library for data manipulation and analysis. | | |
| Plotly: Interactive visualization library for Python and R. | | |
| Prefect: Workflow orchestration tool for data pipelines. | | |
| Prophet: Time-series forecasting tool by Meta AI. | | |
| PyCharm: IDE for Python development with data science features. | | |
| Pytorch: Machine learning framework for deep learning. | | |
| Qlik Sense: Business intelligence platform for data visualization. | | |
| RStudio: IDE for R programming and data analysis. | | |
| Scikit-learn: Python library for machine learning algorithms. | | |
| Seaborn: Python visualization library based on Matplotlib. | | |
| Seldon: Machine learning deployment and serving platform. | | |
| SimPy: Python library for discrete-event simulation. | | |
| SPSS: Statistical software for data analysis. | | |
| Statsmodels: Python library for statistical modeling. | | |
| Tableau: Data visualization and business intelligence platform. | | |
| Talend: Data integration platform for ETL/ELT. | | |
| TensorFlow: Machine learning framework for deep learning. | | |
| VS Code: Lightweight IDE for data science and programming. | | |
| Weights & Biases: Tool for experiment tracking in machine learning. | | |
| XGBoost: Gradient boosting library for machine learning. | | |

Copyright © 2024 DRAUP. All Rights Reserved.