Weights & Biases Service Definition Document

G-Cloud 15 Lot 2b

Version: 2.0

Date: January 2026

Provider: Weights & Biases LLC

Service Name: Weights & Biases AI Development Platform

---

1. Executive Summary

Weights & Biases is an end-to-end AI development platform for building, evaluating, and monitoring foundation models and AI agents. Provides serverless reinforcement learning, agentic workflow tracing, LLM evaluation, AI guardrails, and model governance. Supports SaaS, dedicated cloud, and self-managed deployments with ISO 27001 certified security controls.

The platform addresses the complete AI lifecycle from foundation model training and reinforcement learning through agentic workflow development, evaluation, and production monitoring. Government organisations gain unified visibility, reproducibility, and governance for next-generation AI systems whilst maintaining full compliance with UK data protection and security standards.

---

2. Service Overview

2.1 Service Description

Weights & Biases provides a unified AI development platform designed for organisations building foundation models, large language model

applications, and AI agents. The service addresses critical challenges including model fine-tuning, reinforcement learning, agentic workflow observability, LLM evaluation, AI safety guardrails, and production monitoring.

Primary Use Cases:

- Foundation model fine-tuning and RLHF (Reinforcement Learning from Human Feedback)
- AI agent development with serverless reinforcement learning
- Agentic workflow debugging and multi-turn execution tracing
- LLM application evaluation and prompt optimisation
- AI safety with automated guardrails and content filtering
- Model governance with complete lineage and audit trails

2.2 Deployment Models

SaaS (Cloud-Hosted): Fully managed service hosted on enterprise-grade cloud infrastructure (AWS, GCP, Azure) with 99.5% availability guarantee.

Self-Managed (On-Premises/Private Cloud): Customer-controlled deployment on organisation's own infrastructure, providing complete data sovereignty and customised security configurations.

Dedicated Cloud: Isolated cloud deployment for customers requiring dedicated infrastructure whilst maintaining managed service benefits.

---

3. Service Features

The platform provides ten core capabilities positioning W&B as the unified AI development platform for the agentic era:

1. **End-to-end AI development: training, evaluation, deployment, monitoring** - Unified platform spanning the complete AI lifecycle from initial experimentation through production operations, eliminating the need for multiple point solutions.
2. **Foundation model fine-tuning and RLHF with experiment tracking** - Comprehensive tooling for fine-tuning large language models and foundation models using supervised learning and reinforcement learning from human feedback (RLHF), with automated tracking of all training parameters, metrics, and artefacts.
3. **Serverless reinforcement learning for AI agent post-training** - Fully managed, serverless infrastructure for training reliable AI agents using reinforcement learning. This capability,

unique to W&B, enables post-training optimisation of agent behavior without infrastructure management overhead.

4. **Agentic AI tracing with multi-turn workflow observability** - Detailed execution tracing for AI agent systems, capturing multi-turn conversations, tool calls, decision points, and workflow trajectories. Supports Model Context Protocol (MCP) for comprehensive agent debugging.
5. **LLM evaluation framework with custom and pre-built scorers** - Systematic evaluation of large language model outputs using both pre-built evaluation metrics and custom scoring functions. Enables comparison across prompts, models, and configurations before production deployment.
6. **AI guardrails for content moderation and prompt safety** - Automated safety controls including prompt injection detection, harmful content filtering, and output validation. Ensures LLM applications meet organisational safety and compliance requirements.
7. **Model registry with governance, lineage, and version control** - Centralized repository for all AI artefacts including models, datasets, and configurations. Provides complete lineage tracking, approval workflows, and audit trails for regulatory compliance.
8. **Prompt versioning, playground testing, and A/B comparison** - Comprehensive prompt management including version control, interactive testing playground, and systematic A/B comparison of prompt variations to optimize model outputs.
9. **SSO, role-based access control, and audit logging** - Enterprise security controls including Single Sign-On (SSO) via SAML 2.0, granular role-based access control (RBAC), and comprehensive audit logging of all system actions.
10. **SaaS, dedicated cloud, or self-managed on-premise deployment** - Flexible deployment options supporting fully managed SaaS, dedicated cloud infrastructure, or self-managed on-premises installations to meet data sovereignty and compliance requirements.

---

4. Service Benefits

Government organisations gain ten key benefits aligned with modern AI development requirements:

1. **Accelerate AI agent development with unified training and evaluation** - Eliminate tool fragmentation by consolidating

agent training, evaluation, and monitoring in a single platform, reducing development cycles and procurement complexity.

2. **Reduce agent failures using reinforcement learning post-training optimisation** - Improve AI agent reliability and task success rates through systematic reinforcement learning, ensuring agents behave appropriately in production environments.
3. **Debug complex agentic workflows with multi-turn execution tracing** - Rapidly identify and resolve issues in multi-step AI agent workflows by examining complete execution traces including tool calls, decision points, and intermediate outputs.
4. **Ensure LLM safety with automated guardrails and content filtering** - Protect against prompt injection attacks, harmful content generation, and inappropriate outputs through automated safety controls that operate before deployment.
5. **Improve model quality faster through systematic prompt experimentation** - Accelerate LLM application development by testing and comparing prompt variations systematically, identifying optimal configurations based on measurable criteria.
6. **Maintain governance with complete model lineage and audit trails** - Meet regulatory and internal compliance requirements by maintaining comprehensive records of all model versions, training data, and approval decisions.
7. **Deploy confidently using comprehensive LLM evaluation before production** - Reduce production incidents by systematically evaluating model outputs against quality criteria before deployment, with quantitative evidence of performance.
8. **Collaborate securely across teams with role-based access controls** - Enable multi-team collaboration on sensitive AI projects whilst maintaining appropriate access boundaries and audit visibility.
9. **Maintain data sovereignty with flexible on-premise deployment options** - Meet strict data residency requirements through self-managed deployments that provide complete control over data location and infrastructure.
10. **Standardize AI operations across foundation models and agents** - Establish consistent development, evaluation, and deployment practices across diverse AI initiatives, from traditional ML through agentic AI systems.

---

5. Technical Specifications

5.1 System Requirements

Supported Operating Systems:

- Windows 10 and above
- macOS 10.14 and above
- Linux (Ubuntu 18.04+, CentOS 7+, RHEL 7+)

Browser Compatibility:

- Chrome (latest two versions)
- Firefox (latest two versions)
- Safari (latest two versions)
- Edge (latest two versions)

Network Requirements:

- HTTPS connectivity for web interface
- Outbound HTTPS for SDK communication
- Configurable for private networking and IP allowlists

5.2 Service Interface

Users access Weights & Biases through multiple interfaces. They use the W&B Python SDK and other language SDKs integrated into training scripts. This logs metrics, artefacts, and metadata automatically.

The W&B web application visualises and manages results. It provides dashboards, experiment tracking, reports, and collaboration tools.

Users also interact via the W&B CLI for authentication and project management. APIs enable automation and integration with ML frameworks, CI/CD systems, and data platforms.

5.3 Integration Capabilities

The platform integrates with:

- **ML Frameworks:** PyTorch, TensorFlow, Keras, Scikit-learn, XGBoost, LightGBM
- **Deep Learning:** Hugging Face Transformers, JAX, MXNet
- **LLM Frameworks:** LangChain, LlamaIndex, OpenAI API
- **Agent Frameworks:** Model Context Protocol (MCP), AutoGPT, LangGraph
- **Reinforcement Learning:** OpenAI Gym, Stable Baselines, Ray RLlib
- **Orchestration:** Kubeflow, MLflow, Airflow
- **Cloud Platforms:** AWS SageMaker, Google Vertex AI, Azure ML

- **Version Control:** Git, GitHub, GitLab, Bitbucket

---

6. Security and Compliance

6.1 Security Certifications

Weights & Biases maintains the following independently audited certifications:

- **ISO/IEC 27001:2022** - Information security management
- **ISO/IEC 27017:2015** - Cloud security controls
- **ISO/IEC 27018:2019** - Cloud privacy controls
- **ISO/IEC 27701** - Privacy information management
- **SOC 2 Type II** - Security, availability, and confidentiality
- **HIPAA** - Healthcare data protection compliance

6.2 Compliance Standards

The platform supports compliance with:

- **UK GDPR** - Data protection and privacy requirements
- **NIST 800-53** - Security and privacy controls for information systems
- **NIST SP 800-61 Rev. 2** - Incident handling framework

6.3 Data Protection

Encryption:

- Data in-transit: TLS 1.2 or higher
- Data at-rest: AES-256 encryption

Data Residency:

- Customer data isolated by organisation and workspace
- Self-Managed deployments provide complete data sovereignty
- SaaS deployments support region selection for data residency requirements

Data Export:

Users can export data from W&B through several supported mechanisms. They can use the W&B APIs and SDKs to download runs, metrics, system logs, and associated metadata.

Users can also export models, datasets, and other files through W&B Artefacts. Artefacts can be downloaded via the SDK, the CLI, or directly through the web UI. If artefacts are stored in customer-managed object storage, users may also retrieve the data directly from that underlying storage system.

For Self-Managed deployments, users retain full access to the underlying databases and storage backends, allowing complete data extraction using standard database and storage tools.

6.4 Access Controls

Authentication Methods:

- Multi-factor authentication (MFA) via Okta
- Single Sign-On (SSO) using SAML 2.0
- OAuth-based identity federation (including Google Apps)
- Email and password with mandatory MFA
- API keys for programmatic access

Authorisation:

- Role-based access control (RBAC) with granular permissions
- Organisation and workspace-level access controls
- Segregation of duties for administrative actions
- Comprehensive audit logging of all access and changes

6.5 Network Security

Enterprise and Self-Managed deployments support additional controls. Customers can restrict network access using private networking, IP allowlists, and secure ingress configurations to meet internal security requirements.

---

7. Service Management

7.1 Support Levels

AI Solutions Engineers provide enablement on best practices and features to achieve specific workflows.

Solutions Architects deliver guidance and support for deployment and maintenance at scale.

Support Engineers assist via virtual methods including email, Slack, and Teams.

7.2 Support Response Times

Support response times vary by plan tier and issue priority:

Standard/Standard Plus Plans:

- Critical (P0): 4 business hours
- High (P1): 12 business hours
- Medium/Low (P2/P3): 24 business hours

Premium Plan:

- Critical (P0): 1 business hour
- High (P1): 2 business hours
- Medium/Low (P2/P3): 4 business hours

7.3 Service Availability

SaaS Deployments: The service will be available at least 99.5% of the time in any calendar month.

Self-Managed Deployments: Availability depends on the customer's infrastructure design. W&B provides guidance and best practices to support highly available configurations.

7.4 Onboarding and Offboarding

Product documentation is publicly available online. Onboarding and offboarding documentation is typically provided as a document in formats such as PDF, Word, or Google Docs.

---

8. Change and Configuration Management

8.1 Change Management Process

Our change management process operates through two categories: Normal and Emergency changes. Normal changes follow a decentralised approval model where teams authorise changes within defined quality, security, and reliability standards. Emergency changes address critical incidents with retroactive documentation where necessary, subject to subsequent review.

All changes require risk and impact assessment, peer review, and segregation of duties between initiator and approver. Security team approval is mandatory for platform security changes. Testing occurs in non-production environments, with post-implementation monitoring and

validation. We maintain a fix-forward approach with documented remediation plans for each change.

8.2 Configuration Management

Configuration management utilises infrastructure-as-code via Terraform, version-controlled in Git with approval-based workflows. We employ Golden Images for VM standardisation, security-reviewed before production deployment. Continuous monitoring and automated drift detection operate through Wiz, with configurations baselined against CIS benchmarks for AWS, Azure, GCP, EKS, GKE, and AKS.

Change coordination uses Jira for tracking, GitHub for code changes, Notion for reviews, and Slack for emergency communications.

Our processes align with ISO 27001:2022 controls (A.8.32, A.8.9, A.5.20) and SOC 2 Type II requirements. Monthly security reviews ensure ongoing compliance and continuous improvement. Terraform state files are backed up with immediate rollback procedures available for critical issues.

8.3 Component Lifecycle Tracking

Service components are tracked throughout their lifecycle using infrastructure-as-code via Terraform, version-controlled in Git. Golden Images standardise VM configurations. Wiz provides continuous monitoring and automated drift detection against CIS benchmarks for AWS, Azure, GCP, and Kubernetes platforms. Monthly security reviews ensure compliance.

All changes undergo mandatory risk and impact assessment before implementation. Security team approval is required for platform security changes. Changes are peer-reviewed with segregation of duties between initiator and approver. Testing occurs in non-production environments. Post-implementation monitoring validates security controls remain effective. Emergency changes receive retroactive security review.

---

9. Vulnerability Management

9.1 Vulnerability Assessment

We assess threats through continuous automated scanning of infrastructure, dependencies, source code, and cloud environments using Docker CVE scans, Wiz, SAST tools, Trufflehog, and Socket scans. Quarterly third-party penetration testing supplements automated detection. Vulnerabilities are prioritised using CVSS scoring, assessing business impact, exploitability, and compliance obligations.

Patch deployment follows risk-based timelines: Critical vulnerabilities remediated within 30 days, High within 60 days, Medium within 90 days.

Threat intelligence sources include National Vulnerability Database, vendor security advisories, penetration test findings, bug bounty programmes (HackerOne, Bugcrowd), cloud security alerts, and security research communities. All findings are aggregated in Linear for tracking and remediation management.

9.2 Vulnerability Management Process

Vulnerabilities are normalised, categorised by severity (Critical, High, Medium, Low), and logged with source tags. Security engineers review findings, prioritising based on business impact, exploitability, and compliance obligations. Remediation progress is tracked through status workflows. Dashboards monitor trends by severity and source, with metrics including mean time to remediate (MTTR). Quarterly reviews assess trends and improve processes.

---

10. Incident Management

10.1 Incident Detection

We identify potential compromises using RunReveal and Sentinel One to detect security events through continuous monitoring of logs, alerts, and system activity. Multiple detection sources include automated scanning, employee reporting, customer notifications, and bug bounty programmes.

10.2 Incident Response

Our incident management follows NIST SP 800-61 Rev. 2 framework with pre-defined playbooks for common scenarios including data breaches, ransomware, and system compromises. Incidents are categorised by severity (P0-P3) with defined response procedures.

Users report incidents immediately to security@wandb.com. Employees, customers, vendors, and bug bounty programmes serve as detection sources. The Security Incident Response Team (SIRT) triages all reports via private Slack channels.

Upon detecting potential compromises, the Security Incident Response Team (SIRT) assembles immediately via private Slack channels. We implement containment and eradication measures as quickly as possible, which may include taking systems offline, removing access, or applying patches.

10.3 Incident Notification

We provide incident reports to impacted customers within 72 hours of confirmed impact. Post-incident debriefs occur within one week. External communications are issued solely by the CISO or authorised designee. Internal incident documentation remains confidential.

10.4 Standards Compliance

W&B's incident management processes conform to:

- **NIST SP 800-61 Rev. 2** - Computer Security Incident Handling Guide
- **ISO/IEC 27001:2022** - Controls A.5.24, A.5.25, A.5.26, A.5.27, A.5.28, A.6.8
- **ISO/IEC 27017:2015** - Cloud security incident management
- **SOC 2 Type II** - Incident response and management controls

---

11. Protective Monitoring

11.1 Monitoring Processes

We identify potential compromises using RunReveal and Sentinel One to detect security events through continuous monitoring of logs, alerts, and system activity. Multiple detection sources include automated scanning, employee reporting, customer notifications, and bug bounty programmes.

Upon detecting potential compromises, the Security Incident Response Team (SIRT) assembles immediately via private Slack channels. We implement containment and eradication measures as quickly as possible, which may include taking systems offline, removing access, or applying patches.

Response times: employees report incidents immediately to security@wandb.com. Impacted customers receive notification within 72 hours of confirmed impact. Post-incident reviews occur within one week.

11.2 Standards Compliance

Protective monitoring processes conform to:

- **SOC 2 Type II** - Security monitoring, logging, and incident detection
- **ISO/IEC 27001:2022** - Monitoring controls (A.8.15 Logging, A.8.16 Monitoring activities)
- **ISO/IEC 27017:2015** - Cloud-specific security monitoring

- **ISO/IEC 27018:2019** - Privacy monitoring requirements

---

12. Secure Development Practices

W&B demonstrates adherence to secure software development best practices through independently audited certifications:

- **ISO/IEC 27001:2022** - Secure development lifecycle controls (A.8.25, A.8.26, A.8.27, A.8.28, A.8.29, A.8.30, A.8.31, A.8.32)
- **SOC 2 Type II** - Secure development and change management controls
- **ISO/IEC 27017:2015** - Cloud-specific security development practices

---

13. User Authentication

13.1 Service Access Authentication

Users authenticate through multiple methods to access W&B services. We support email and password-based authentication with multi-factor authentication (MFA) via Okta for enhanced security.

Enterprise customers use Single Sign-On (SSO) through SAML 2.0 identity federation, enabling integration with existing providers including Google Apps and other OAuth-based identity providers.

For programmatic access, users authenticate via API keys for SDK, CLI, and automation workflows. Service and bot accounts also authenticate using API keys.

13.2 Management Access Authentication

W&B authenticates management access through:

- Multi-factor authentication (MFA) via Okta for enhanced security on administrative accounts
- Identity federation using Single Sign-On (SSO) with SAML 2.0 and OAuth-based identity providers
- Email and password-based authentication with mandatory MFA for privileged accounts
- API keys for SDK, CLI, and programmatic management access with restricted permissions

All management actions are subject to role-based access controls (RBAC), segregation of duties, and comprehensive audit logging. Security team approval is required for platform security changes.

13.3 Management Interface Access Restrictions

Management interfaces require authentication via Okta SSO with multi-factor authentication. We enforce role-based access controls (RBAC) with segregation of duties between initiators and approvers for changes.

Administrative access uses identity federation through SAML 2.0 or OAuth-based providers, API keys for programmatic access, and email/password with MFA. All management actions are logged with comprehensive audit trails.

Support channels restrict access to authorised personnel only. Internal support uses dedicated Slack channels and Zoom with access controls. Customer support access is granted on a need-to-know basis with session logging.

Enterprise deployments support IP allowlists, private networking, and additional access restrictions.

13.4 Access Control Testing

W&B tests access controls through multiple mechanisms:

- **Quarterly penetration testing** (every 3 months) by third-party security firms
- **Monthly security reviews** of configurations and access controls
- **Annual SOC 2 Type II audits** independently verifying access control effectiveness
- **Annual ISO/IEC 27001:2022 audits** assessing implementations against standard requirements

Continuous automated vulnerability scanning monitors access control configurations on an ongoing basis through Wiz for cloud security and infrastructure monitoring.

---

14. Service Constraints

The service operates within the following constraints:

Resource Limits: Usage quotas apply based on subscription tier. Rate limiting prevents excessive API usage that could impact platform performance.

Integration Dependencies: Some integrations require specific versions of third-party frameworks and libraries. Compatibility matrices are maintained in product documentation.

Network Requirements: Internet connectivity required for SaaS deployments. Self-Managed deployments can operate in air-gapped environments with appropriate configuration.

Browser Compatibility: Latest two versions of major browsers supported. Older browser versions may experience degraded functionality.

---

16. Support and Documentation

16.1 Documentation

Comprehensive public documentation available at docs.wandb.ai including:

- Getting started guides
- API reference documentation
- Integration tutorials
- Best practices guides
- Troubleshooting resources

16.2 Support Channels

- Email: support@wandb.com
- Slack Connect (for Enterprise – Premium Support customers)
- Microsoft Teams integration (for Enterprise – Premium Support customers)

16.3 Training and Enablement

- Online training materials and tutorials
- Webinars and workshops
- Custom training sessions available upon request

---

17. Service Roadmap and Updates

W&B continuously develops new features and improvements. Regular updates are deployed to SaaS infrastructure with minimal disruption. Customers self-managing Weights & Biases are recommended to maintain the latest supported release of W&B Server, with public details available at https://docs.wandb.ai/release-notes/server-releases

Feature roadmap communicated through:

- Product changelog
- Customer newsletters
- Executive business reviews (Enterprise customers)
- Annual customer summit

---

18. Contact Information

Sales Enquiries:
Email: sales@wandb.com
Website: wandb.ai

Technical Support:
Email: support@wandb.com

Security Issues:
Email: security@wandb.com

Legal and Compliance:
Email: legal@wandb.com

Company Address:
Weights & Biases LLC
400 Alabama Street
San Francisco, CA 94110
United States

End of Service Definition Document

This document provides a comprehensive overview of the Weights & Biases AI development platform for UK Government procurement evaluation. All information is accurate as of January 2026 and subject to update in accordance with product development and regulatory requirements.