Elastic Privacy Datasheet

Elastic AI

Customers are responsible for making their own independent assessment of the information in this document. This document: (a) is for informational purposes only, (b) represents current Elastic product offerings, services, and practices, which are subject to change without notice, and (c) does not create any commitments or assurances from Elastic and its affiliates, suppliers, or licensors. The responsibilities and liabilities of Elastic to its customers are controlled by Elastic agreements, and this document is not part of, nor does it modify, any agreement between Elastic and its customers.

Product Summary: Elastic AI

Elastic AI encompasses the artificial intelligence and machine learning capabilities built into the Elastic Platform.

These offerings are designed to let organizations add semantic search, embeddings, reranking, and agentic AI experiences to their data without independently sourcing, hosting, or managing machine learning models. Because different components process customer data in different ways, the sections below distinguish between models that run entirely within Elastic-controlled infrastructure and features that may route customer content to Elastic Managed LLMs or, where the customer configures it, to trusted third-party model providers.

- **Product type**: Software/ AI features and managed inference service
- **Deployment model(s)**:
  - **On-premises (self-managed)**: for models that can run on your own GPU, CPU (for example, Jina On-Prem) or machine learning (for example, ELSER, E5, Elastic Rerank)
  - **Cloud (SaaS)**: Elastic Cloud Hosted and Elastic Cloud Serverless
  - **Hybrid**: including self-managed clusters connected to EIS through Cloud Connect

Components covered by this Datasheet

- **Built-in NLP models:** Pre-trained models shipped by Elastic that require no fine-tuning, including ELSER (Elastic Learned Sparse EncodeR) for English-language semantic search, E5 for multilingual semantic search, Elastic Rerank, language identification, and the Jina family of embedding, reranker, and reader models.
- **Elastic Inference Service (EIS):** A managed inference service that runs Customer selected AI models for ingest, search, and chat, without the customer deploying or scaling machine learning nodes. EIS also provides Elastic Managed LLMs used by AI features such as Agent Builder, the AI Assistant, Attack Discovery, Automatic Import, and Search Playground.
- **Elastic Agent Builder:** An AI conversational platform for creating agents that answer questions and take actions over Elasticsearch data using natural language, combining LLM reasoning with built-in and custom tools, and exposing tools and agents to external clients through an MCP server, A2A server, and REST APIs.

Data Processed in Elastic AI

The data processed by Elastic AI depends on the component in use and the customer's configuration. Elastic primarily provides the platform and models; customers determine what data they submit for inference and what agents are permitted to access.

Categories of personal data necessary to access and use Elastic AI:

- Admin and user identifiers, credentials, permissions, session cookies, and activity logs. The processing of such data is necessary for product deployment, configuration, administration, management, secure access, role-based access control (RBAC), and auditable use.

Additional categories of personal data processed may include, but are not limited to, data contained within:

- Inference inputs and outputs. Any data the customer submits to a model for embedding, sparse encoding, reranking, or generation, and the results returned. Depending on the customer's use case, inputs can contain personal data.
    - Examples: search queries; passages and documents submitted for embedding or reranking; text, and, for multimodal Jina v5 omni models on EIS, images, audio, video, and documents such as PDFs submitted for embedding; the vector embeddings and relevance scores returned.
- Agent Builder conversations and context. Natural-language prompts, the LLM responses returned, and the data retrieved from Elasticsearch indices by the tools an agent invokes to ground its answers.
    - Examples: user questions typed into Agent Chat; content retrieved from indices the agent's tools are permitted to query; generated ES|QL and tool call parameters; conversation history.
- Data exchanged through the MCP and A2A servers. When external clients (for example, Claude Desktop, Cursor, VS Code, or LangChain applications) connect to Agent Builder tools, the requests they send and the index data returned in tool results are processed. Access is scoped to the privileges of the API key used.

Typical Purpose(s) of Processing Customer Data with Elastic AI

In their specific deployments, customers have sole discretion to determine the actual purpose(s) of processing any customer data they choose to feed to Elastic AI. In general, Elastic AI is designed and intended for use cases such as:

- To perform semantic search over English-language content using ELSER, which expands text into weighted sparse token representations that capture learned associations rather than exact keyword matches.
- To perform multilingual and cross-domain semantic search using dense vector embedding models such as E5 and the Jina embedding models, storing embeddings in dense_vector or semantic_text fields for vector similarity search.
- To improve relevance ordering of search results through reranking models, including Elastic Rerank and the Jina reranker models.
- To generate multimodal embeddings across text, images, audio, video, and documents in one shared vector space using the Jina v5 omni models on EIS.
- To run inference for ingest, search, and chat as a managed service through EIS, without deploying or scaling machine learning nodes in the customer's environment.

- To build and operate AI agents that answer questions and take actions over Elasticsearch data using natural language, grounding responses in the customer's own indexed data.
- To expose Agent Builder tools and agents to external systems through the MCP server, A2A server, and REST APIs, and to import tools from external MCP servers.
- To power AI features such as the AI Assistant, Attack Discovery, Automatic Import, and Search Playground through Elastic Managed LLMs.

Product Usage Data

Elastic automatically collects certain information related to the use of Elastic AI features. Such Product Usage Data is processed as described in the Elastic Product Privacy Statement. In on-premises and certain hybrid deployment models, customers may restrict or disable Product Usage Data processing in the product's configuration settings (please refer to the applicable product Documentation). However, such Product Usage Data collection and processing is indispensable to the proper functioning of the product in Cloud (SaaS) deployment models.

Access to Customer Data

- **Access by Customers:** Customers retain full control over the data they submit for inference and the data their agents can reach. Inference endpoints, models, and Agent Builder tools are managed through Kibana and the inference APIs. Role-based access controls are integrated into Elastic deployments, Kibana, and Agent Builder, allowing administrators to define roles and API keys that limit which indices, tools, and features a user or external client can access, enforcing least-privilege principles.
  For Agent Builder's MCP server, tools execute with the scope assigned to the API key presented by the connecting client. Elastic recommends restricting API keys to only the specific indices the tools need, setting API key expiration dates, and using read-only privileges so agents cannot modify data. This allows customers to control precisely what data is exposed to external MCP clients.

- **Access by Elastic**: For Elastic Cloud offerings, a limited number of Elastic employees have privileged access to the production environment solely for platform management, maintenance, and support. This access strictly adheres to the principles of least-privilege and need-to-know and is regularly reviewed and modified as necessary. Elastic does not have access to data within self-managed deployments that are not Cloud connected.
  - For inference processed through EIS, ELSER and Jina requests are processed entirely within Elastic inference infrastructure. Requests to Elastic Managed LLMs and, where the customer configures external inference, to third-party embedding or generative models, may involve additional processing by the respective model providers, which can operate in different cloud platforms or regions.
  - Elastic has implemented centralized logging, encompassing proxy logs, access logs, Elasticsearch logs, and Auditbeat logs to record all access to customer data and the systems on which it resides. Elastic's internal teams actively develop and implement

detections for suspicious internal account activity and unauthorized access, including file integrity monitoring and account takeover indicators.

Sub-Processing and Third-Party Model Providers

- **Elastic-hosted models.** ELSER, E5, Elastic Rerank, language identification, and the Jina models made available through EIS are run within Elastic-operated inference infrastructure.
- **Elastic Managed LLMs.** On Elastic Cloud Hosted and Elastic Cloud Serverless, AI features can use Elastic Managed LLMs running on EIS with no additional API key management. These managed LLMs may be delivered using trusted third-party model providers acting as sub-processors under Elastic's agreements. See our external sub-processors for more information.
- **Customer-configured external inference and connectors.** Customers may connect Agent Builder and other AI features to their own third-party model providers (for example, OpenAI, Anthropic, Amazon Bedrock, or Google Vertex) using external inference endpoints or Generative AI connectors, or to self-managed and local LLMs. Where a customer configures such a provider, customer inputs and outputs are transmitted to that provider under the customer's own arrangement, and that provider's data handling terms apply to that processing. Customers should review the relevant provider's retention and privacy terms.

Processing Locations

Elastic Cloud deployments are hosted on certified cloud platforms managed by industry-leading IaaS providers that include Amazon Web Services (AWS), Google Cloud Platform (GCP), and Microsoft Azure. Customers can select the geographic region where they want to host their deployments.

The Elastic Inference Service is available in selected AWS and GCP regions. Inference requests sent through EIS are routed to the nearest available region, regardless of where the customer's Elasticsearch deployment or Serverless project is hosted. Depending on the model, request processing may involve Elastic inference infrastructure and, in some cases, trusted third-party model providers. ELSER and Jina requests are processed entirely within Elastic inference infrastructure, whereas large language models or third-party embedding models may involve additional processing by their respective providers, which can operate in different cloud platforms or regions. Refer to the EIS region and hosting documentation for the current list of supported regions.

Compliance with Privacy Regulations

Elastic captures, processes, stores, and protects Customer Personal Data in accordance with the applicable customer Data Processing Addendum, and the information outlined in this Privacy Datasheet. Our Trust Center serves as a comprehensive resource for information on our privacy practices and compliance efforts.

- **Data Ownership:** Customer data remains the property of the customer. Elastic only processes this data for the purposes specified in the customer's agreement, and never sells customer data to third parties. Customer inputs to Elastic-hosted models are not used to train Elastic's models except as permitted under the customer's agreement.
- **GDPR Compliance:** Elastic Cloud features are designed to support compliance with the General Data Protection Regulation and other global data protection and privacy laws. This includes prioritizing the security of personal data through effective technical and organizational measures, offering a GDPR-compliant Data Processing Addendum. Our dedicated privacy team at Elastic oversees GDPR compliance.
- **Data Privacy Framework (DPF):** Elasticsearch, Inc. is certified under the EU-U.S. Data Privacy Framework, UK Extension to the EU-U.S. Data Privacy Framework, and the Swiss-U.S. Data Privacy Framework as set forth by the U.S. Department of Commerce.
- **Cross-Border Data Transfer:** Elastic legalizes transfers of personal data across the jurisdictions where we operate by relying on current Standard Contractual Clauses, and through our participation in the Data Privacy Framework for transfers to the U.S. and from there onwards. Additionally, we implement robust supplementary measures to protect data during transfers, such as encryption in transit and at rest, protocols for challenging public authority requests, and providing customers with the option to select regional servers for hosting.

Data Portability

Customers retain full control over the data they ingest into their own clusters, including the vector embeddings and enriched fields produced by Elastic AI models. Data stored in Elasticsearch can be retrieved using the REST API and through various Elasticsearch clients for common programming languages, which supports customers in fulfilling data subject access requests. Agent Builder configurations, tools, and conversations are accessible through Kibana and the Agent Builder REST APIs.

Retention and Deletion

Elastic provides tools and capabilities that enable customers to manage their data retention and deletion policies effectively.

- Embeddings, enriched fields, and other inference outputs stored in the customer's indices are retained under the retention and index lifecycle policies the customer defines within their own environment, and can be deleted by the customer at any time.
- For inference processed through EIS, ELSER and Jina requests are processed within Elastic inference infrastructure to return a result; Elastic minimizes retention of inference content consistent with the principle of data minimization and the transient nature of inference.

- Where a customer configures an external model provider, the retention of inputs and outputs by that provider is governed by the customer's arrangement and that provider's terms. Customers should review those terms, including any options for zero-retention or limited-retention processing.
- Agent Builder conversations and configurations persist in the customer's deployment and can be managed and deleted by the customer.
- Logstash can be used to perform data transformation, including anonymization and pseudonymization, which helps achieve data minimization goals and reduces data security risks. For data subject deletion requests, Elastic provides capabilities to tag data for retention under an exception, or to permanently delete data using permissible deletion and de-identification techniques.

Security

The security of your data also relies on you keeping your deployments configured securely and maintaining the confidentiality of your login credentials and API keys (please see our Security FAQ for more information). Elastic supports a comprehensive defense-in-depth security model designed to protect customer data throughout its entire lifecycle: in transit, at rest, and in memory, as well as through robust key management procedures.

- **Encryption:** Customer data is encrypted at rest using AES-256 and in transit via TLS 1.2 or higher. Elastic implements strict encryption key management procedures.
- **Access Controls:** Elastic maintains technical, logical, and administrative controls to restrict data access solely to authorized users. These controls include multi-factor authentication, strong password strength standards, and the use of Virtual Private Networks (VPNs) for administrative access. Role-based access controls are integrated into Elastic deployments, Kibana, and Agent Builder, allowing granular control over user permissions and over which indices and tools agents and API keys can access.
- **API key scoping for AI features:** Access to the Agent Builder MCP server requires an API key with the appropriate Kibana application privileges. Elastic recommends scoping API keys to only the required indices, applying read-only privileges, and setting expiration dates so that data exposed to agents and external clients follows the principle of least privilege.
- **Logging and Monitoring:** Elastic has implemented centralized logging, encompassing proxy logs, access logs, Elasticsearch logs, and Auditbeat logs to record all access to customer data and the systems on which it resides. Elastic's Threat Detection and Response team utilizes automated workflows to detect and investigate suspicious internal account activity and unauthorized access, including file integrity monitoring and account takeover indicators.
- **System Updates and Patches:** Elastic systems are regularly updated and deployed based on the latest operating system kernels. Appropriate patches are applied promptly whenever a Common Vulnerability and Exposure (CVE) is identified in any component software.
- **Incident Detection and Response:** Elastic has implemented and continuously updates detection rules for suspicious activity and unauthorized access. These detections are part of automated workflows that alert the Threat Detection and Response team, triggering analyst investigations. Elastic uses Elastic Security internally, which includes Endpoint Detection and Response

capabilities and integrates threat intelligence feeds to support intrusion detection. It supports automated responses through integrations with third-party SOAR platforms.

- **Secure Software Development Framework (SSDF):** Elastic maintains a secure software development framework based on NIST 800-218, guiding the process to securely design, develop, deploy, track, and maintain all Elastic software, ensuring a “secure by design” and “secure by default” approach.
- **Third-Party Vendor Review:** Elastic partners with major IaaS providers (AWS, GCP, Azure) and conducts rigorous reviews of their security and compliance standards, including SOC 2 audits and ISO 27001 and 27701 certifications, as part of its third-party risk management program. Third-party model providers used to deliver Elastic Managed LLMs are subject to this program.
- **Penetration Testing:** Independent third parties conduct annual application and network penetration tests against Elastic Cloud, at a minimum. Elastic also hosts a public Bug Bounty Program for continuous researcher testing.
- **Employee Training:** All Elastic employees are required to complete comprehensive information security and data protection/privacy training upon hire and at least annually thereafter.