

XLSCOUT

AI-Powered IP Analytics at Global Scale — Secure, Compliant, and Production- Ready on AWS

How XLScout built a 4-region, SOC 2 and GDPR-compliant AI platform for patent intelligence — serving attorneys and universities across 3 continents — without building a dedicated DevOps team.

394

Containers in production

4

AWS Regions — US, Japan & Germany

SOC2SOC 2 & GDPR Compliance Via
Trustcloud**24x7**Support — Incidents identified &
Resolved in minutes



At a Glance

The foundation behind XLScout's global AI infrastructure

COMPANY

XLSCOUT

394 containers running AI-powered patent analytics workloads across production EKS clusters

INDUSTRY

AI-Powered IP Analytics / LegalTech

Multi-region AWS infrastructure spanning US East, US West (Oregon), Japan, and Germany (eu-central-1) to serve a global user base across 3 continents

CLOUD

AWS — US East, US West (Oregon), Japan (ap-northeast-1), Germany (eu-central-1)

Proprietary domain-tuned patent LLMs with Para-Embed™, Para-RAG™, and Para-Rank powering IP analytics on cloud-agnostic infrastructure across AWS and Azure

COMPLIANCE

SOC 2 (via TrustCloud); GDPR; AWS KMS Encryption at Rest

SOC 2 audit readiness achieved — controls documented across 20+ domains with TrustCloud and DuploCloud

CUSTOMER SINCE

April 2023

GDPR compliance support for EU operations, including data residency in Germany (eu-central-1)

DUPLOCLOUD SERVICES

DevSecOps Platform; Managed Ops; Kubernetes & EKS; Compliance Engineering

Multiple EKS cluster upgrades to AL2023 AMIs completed with zero production downtime

AWS-consolidated infrastructure replacing fragmented Azure and GCP workloads



The Challenge

Scaling a global AI platform without a large internal infrastructure team

XLScout's mission is to make patent intelligence accessible and AI-powered — giving attorneys, researchers, and universities the tools to analyze IP landscapes, search prior art, and draft patent claims faster than ever before. Their flagship product, Patdigger, depends on a sophisticated stack of machine learning inference services, large-scale search infrastructure, and globally distributed data stores.

As the platform grew to serve customers across the US and Japan, the infrastructure footprint expanded accordingly — from a single AWS region to a multi-region deployment spanning three continents. Managing Kubernetes clusters, database infrastructure, compliance audits, patching cycles, and cost optimization across this footprint required a level of DevOps depth that a fast-growing AI company cannot easily staff in-house.

XLScout needed a DevSecOps platform and dedicated engineering partner that could run their cloud infrastructure end-to-end — delivering the compliance posture and operational reliability their enterprise and institutional clients demand, while keeping pace with the team's AI product roadmap.



The Solution

DuploCloud DevSecOps platform + Managed Ops across a multi-region AI stack

XLScout's AI platform is built on a proprietary suite of domain-tuned patent and IP models — purpose-built for the complexity of patent language, legal standards, and technical documentation that general-purpose LLMs cannot reliably handle. Core production workflows rely on fine-tuned LLMs trained specifically for IP intelligence, complemented by open-source models where domain specialization and customization matter most. The AI infrastructure is cloud-agnostic, deployed across both AWS and Azure to support global enterprise customers — with DuploCloud managing XLScout's AWS footprint across four regions (US East, US West, Japan, and Germany).

The GenAI architecture is organized around three proprietary components: **Para-Embed™** — a custom embedding system for semantic retrieval over patents, standards, technical documents, litigation data, and licensing corpora; **Para-RAG™** — a retrieval-augmented generation layer that grounds patent drafting, analysis, and SEP (Standard Essential Patent) mapping in verified source documents to minimize hallucination risk; and **Para-Rank** — a relevance scoring and prioritization engine for patents and standards. The Standigger product combines Para-Embed, Para-RAG, and AI-driven analysis with a human expert review loop for enterprise SEP workflows. XLScout's platform also incorporates multi-modal capabilities — processing patent images, flowcharts, sequence diagrams, and standard specification structures alongside text. No RAG orchestration frameworks such as LangChain or LlamaIndex are publicly disclosed; the architecture appears to be custom-built end-to-end.

The operational outcomes are measurable: XLScout reports up to a **60% reduction in manual workload** for SEP review workflows and a **75% improvement in section/subsection extraction accuracy** in standards analysis use cases. Across the platform, customers benefit from accelerated patent drafting and prior art search, higher-precision patent-to-standard mapping, reduced false positives, and AI-assisted portfolio monetization and licensing intelligence — all grounded in retrieval-first generation to maintain accuracy on extremely large patent corpora.



DUPLOCLOUD DEVSECOPS PLATFORM

Full-stack cloud management across 3 AWS regions — EKS, Aurora, DocumentDB, ElastiCache, API Gateway, and Elasticsearch — provisioned and secured via DuploCloud's abstraction layer with least-privilege IAM, KMS encryption, and built-in network isolation per tenant.

PROPRIETARY PATENT AI — PARA-RAG™, PARA-EMBED™ & PARA-RANK

XLScout's AI stack is built on proprietary domain-tuned models running on cloud-agnostic inference infrastructure across AWS and Azure. DuploCloud manages the AWS layer — providing the EKS, storage, and networking infrastructure on which XLScout's Para-Embed™ retrieval, Para-RAG™ generation, and Para-Rank relevance scoring workloads run at production scale.

MANAGED OPS & EKS LIFECYCLE

DuploCloud's Managed Ops team operates XLScout's infrastructure — handling EKS cluster upgrades to AL2023 AMIs, SSM Patch Manager cycles, auto-scaling policy tuning, and ECR lifecycle management — all coordinated during maintenance windows with zero production disruption.

SOC 2 & GDPR COMPLIANCE ENGINEERING

DuploCloud partnered with XLScout through their full SOC 2 audit cycle — using TrustCloud as the GRC platform and gathering evidence across 20+ controls including encryption, RBAC, DR testing, vulnerability scanning, and environment separation. DuploCloud also supported GDPR compliance for EU operations, with data residency enforced in Germany (eu-central-1).

MULTI-REGION AWS ARCHITECTURE

XLScout's global user base is served from US East, US West (Oregon), Japan (ap-northeast-1), and Germany (eu-central-1) — with DocumentDB, Aurora, and Elasticsearch replicated across regions and cross-region connectivity managed through DuploCloud's network infrastructure.

OBSERVABILITY & COST OPTIMIZATION

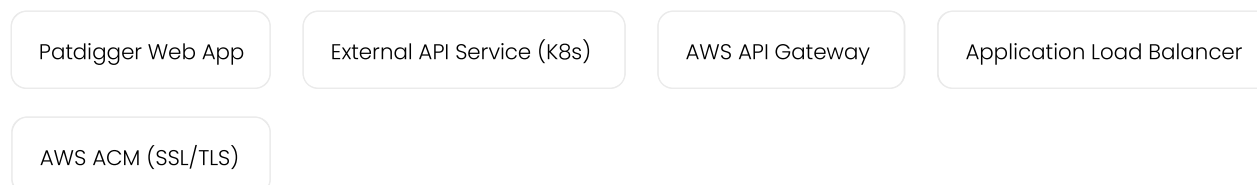
CloudWatch-based monitoring, Grafana dashboards, and right-sizing exercises surface actionable cost and performance signals. DuploCloud continuously reviews instance types, Fargate scaling policies, backup retention, and ECR image retention to keep infrastructure lean.



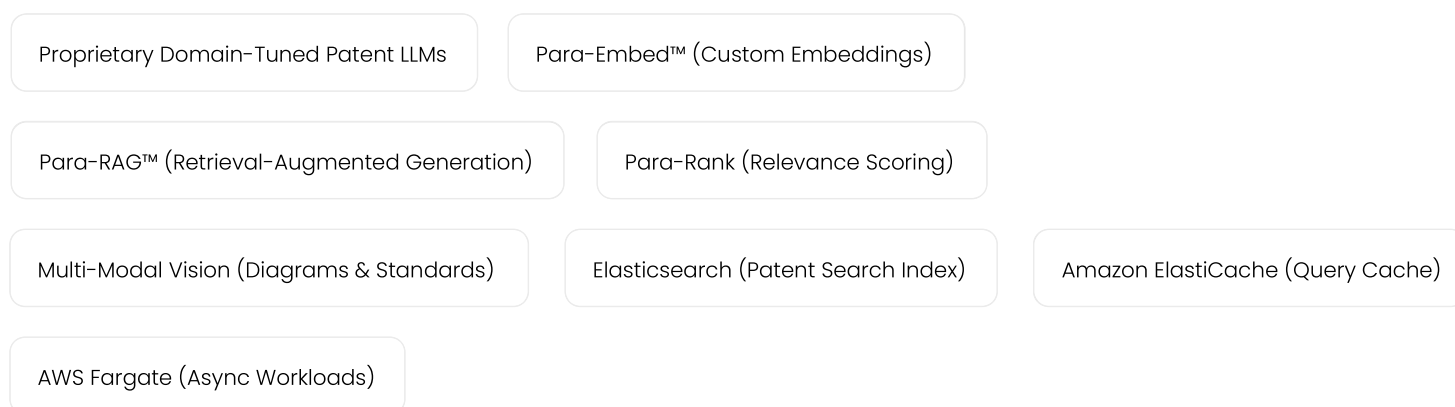
Architecture Overview

Multi-region AI inference and IP analytics infrastructure on AWS

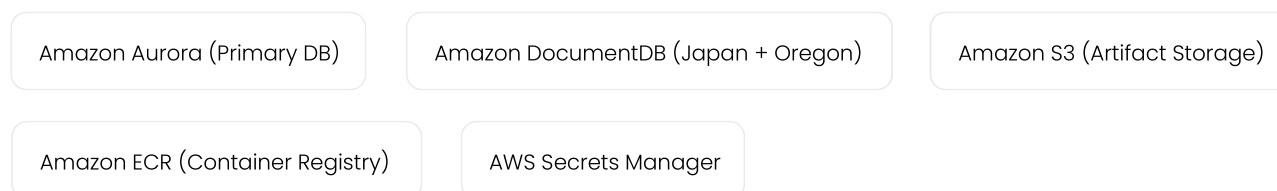
CLIENT & API LAYER



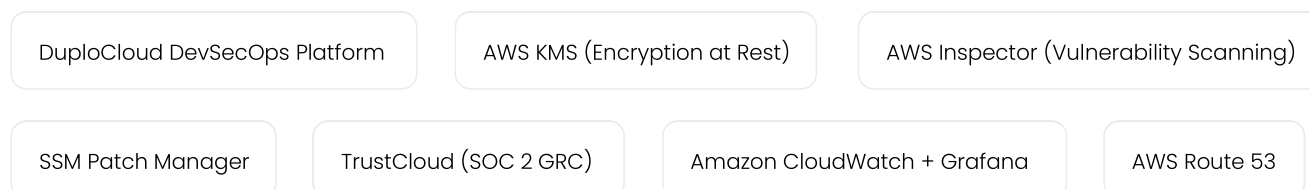
AI / ML PROCESSING



DATA & STORAGE LAYER



SECURITY, COMPLIANCE & OBSERVABILITY





The Results

Production-grade AI infrastructure with compliance built in from day one

Global Reach: 4 AWS Regions Across 3 Continents, Zero Downtime Upgrades

XLScout's platform operates across US East, US West (Oregon), Japan, and Germany (eu-central-1) — with DocumentDB and Aurora data stores serving regional workloads. Multiple EKS cluster upgrades to AL2023 AMIs were completed during scheduled maintenance windows without any production outage, ensuring uninterrupted service for global users.

4 Regions

SOC 2 & GDPR Compliance Achieved

DuploCloud worked with XLScout through the complete SOC 2 evidence cycle — gathering documentation for encryption, RBAC, DR testing, environment separation, patch management, and vulnerability scanning via TrustCloud. DuploCloud also supported GDPR compliance for XLScout's EU operations, with data residency enforced in Germany (eu-central-1) to meet European regulatory requirements.

SOC 2 + GDPR ✓

24×7 Support — Critical Incidents Resolved in Minutes

DuploCloud's Managed Ops team provides around-the-clock monitoring and engineering coverage for XLScout's production infrastructure. Critical infrastructure incidents, if any, are identified and resolved within minutes through proactive alerting and dedicated on-call support — ensuring enterprise SLAs are met at all times.

24×7 Coverage

394 Containers — Proprietary AI Workloads Running at Scale

XLScout runs 394 production containers across 11 isolated tenants on DuploCloud-managed EKS — hosting the Patdigger API, Para-RAG™ and Para-Embed™ inference services, and multi-modal vision pipelines. Right-sizing and auto-scaling policies keep infrastructure costs lean as AI workload demand scales across a global IP analytics customer base.

394 Containers



AWS Gen AI Competency Highlights

How XLScope's platform demonstrates AWS Gen AI Competency capabilities

Proprietary Patent AI — Para-RAG™, Para-Embed™ & Domain-Tuned LLMs

- Para-Embed™ delivers semantic retrieval over patents, standards, litigation, and licensing corpora at enterprise scale.
- Para-RAG™ grounds patent drafting and SEP analysis in verified source documents — minimizing hallucination risk on complex IP workflows.
- Para-Rank scores and prioritizes patents and standards relevance; Standigger combines these with a human expert review loop.
- Multi-modal capabilities process patent images, flowcharts, and sequence diagrams alongside text.
- DuploCloud manages the AWS infrastructure layer supporting cloud-agnostic AI workloads across 4 regions

Secure, Compliant AI Infrastructure on AWS

- SOC 2 audit preparation supported across 20+ controls with TrustCloud GRC integration.
- GDPR compliance for EU operations — data residency enforced in Germany (eu-central-1).
- AWS KMS encryption at rest for all data stores; SSL/TLS via ACM for in-transit.
- AWS Inspector + SSM Patch Manager for continuous vulnerability management.
- 11 isolated EKS tenants enforce environment and workload separation.

Multi-Region Resilience & Operational Excellence

- 4 AWS regions (US East, US West, Japan, Germany) spanning 3 continents with regional data stores.
- Multiple EKS cluster upgrades to AL2023 AMIs — zero downtime, coordinated maintenance windows.
- Aurora replica auto-scaling and ElastiCache tuned for AI query workloads.
- DuploCloud Managed Ops provides 24/7 engineering coverage for production operations.