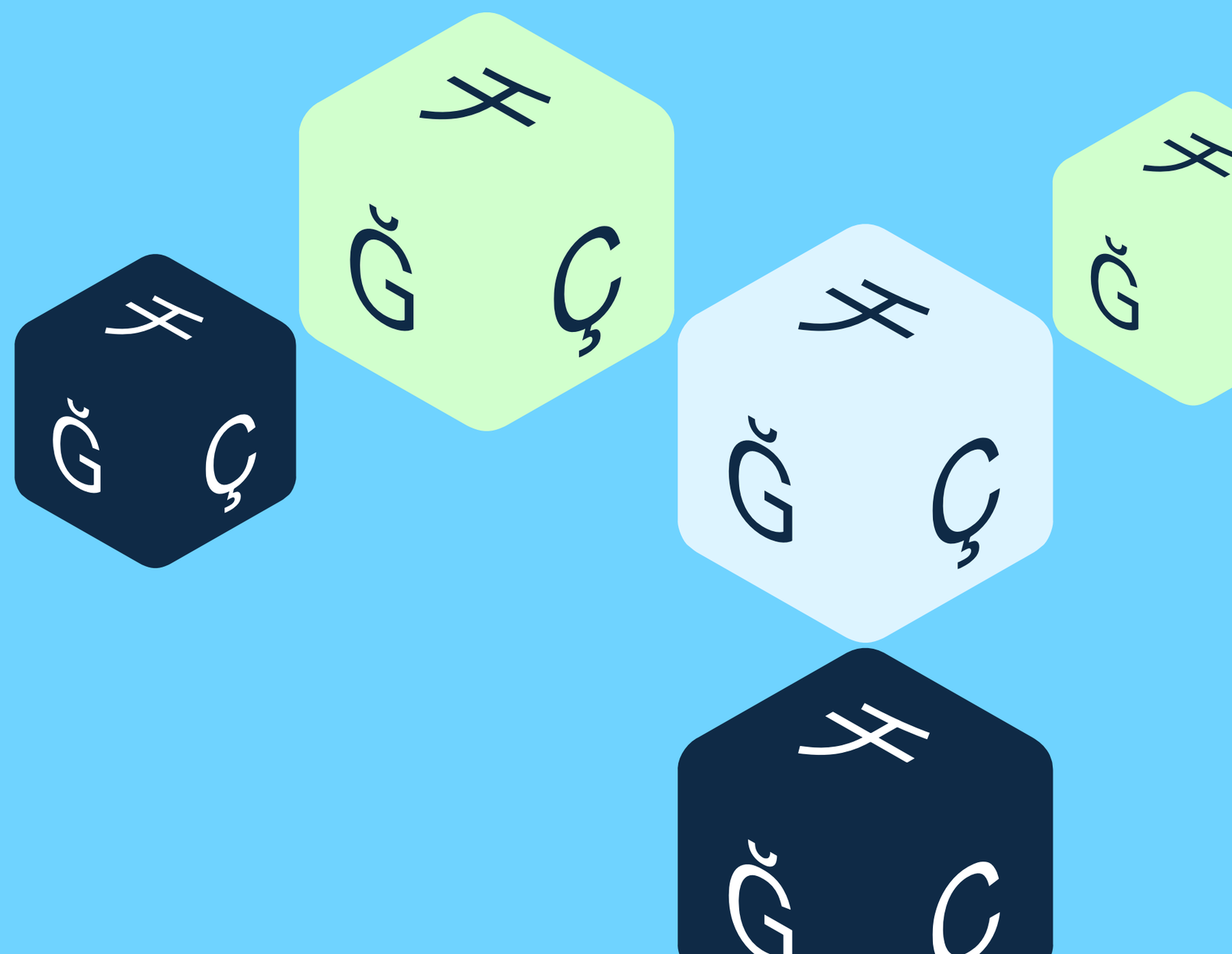
DeepL

Report

Security and data residency whitepaper:

The hybrid infrastructure model

Architectural safeguards for AWS and private data center integration



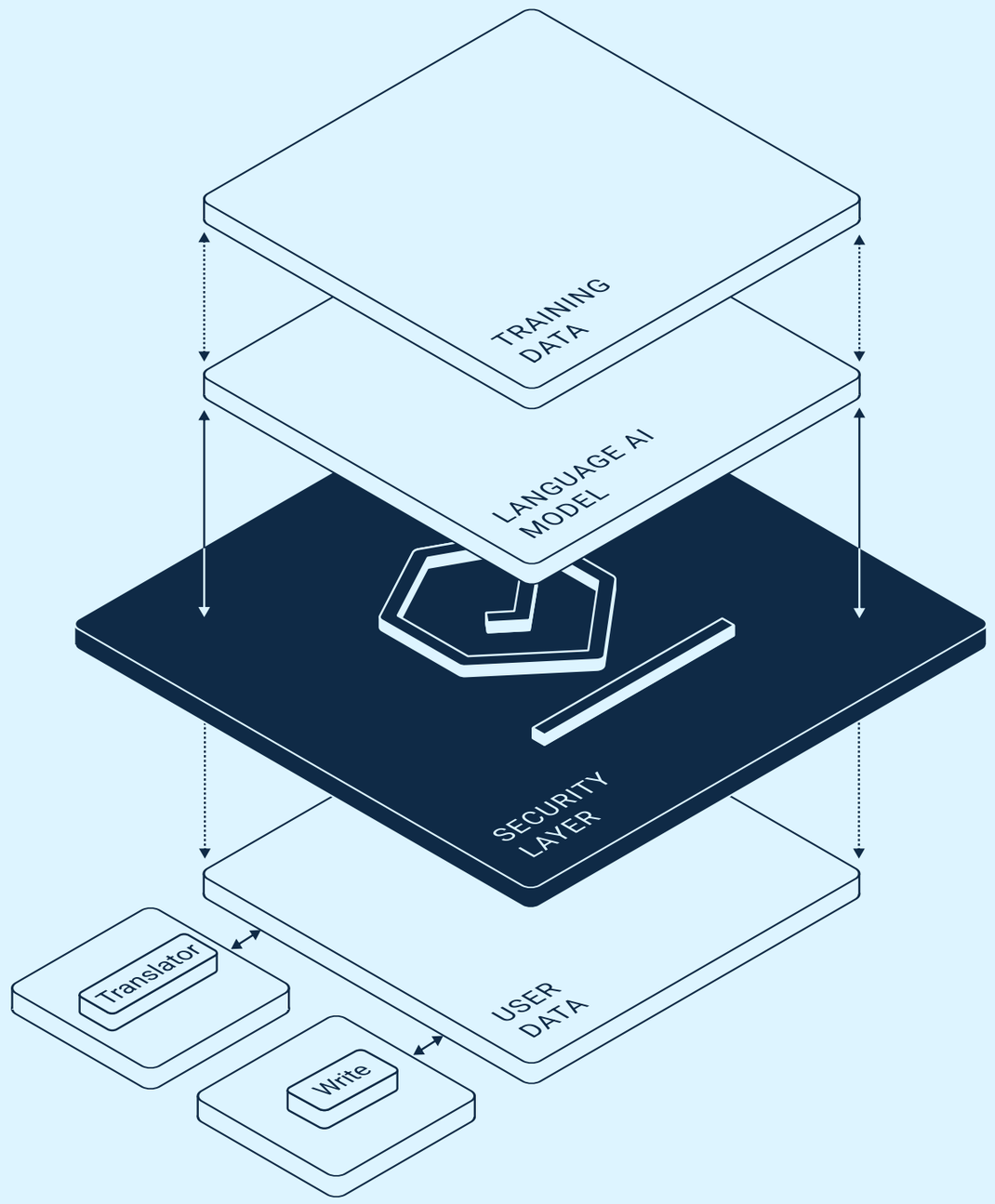
Executive summary

Transitioning to a hybrid infrastructure requires a security model that matches the rigor of your on-premises data centers. Our architecture is engineered specifically for privacy-sensitive organizations, combining the scalability of AWS with the sovereign control of our private data center infrastructure.

By taking a hybrid approach, we can better serve customers–for example by utilising AWS expertise in running extremely robust database services while keeping data secured, enabling us to focus on providing more reliable and effective services and products for customers.

Our commitment to security is backed by industry-standard certifications, ensuring that as you move to a hybrid model, your compliance posture remains uninterrupted.

- **Compliance standards:** Fully aligned with, and certified to, **SOC2 Type II** and **ISO 27001**
- Compliance with the **BSI Cloud Computing Compliance Controls Catalogue (C5)** is maintained for specific in-scope services and regions, including AWS Stockholm (eu-north-1).



Hybrid security architecture

The following diagram and narrative outline the high-level flow of our infrastructure, demonstrating the integration between our secure edge, AWS Stockholm, and our private data center environments.

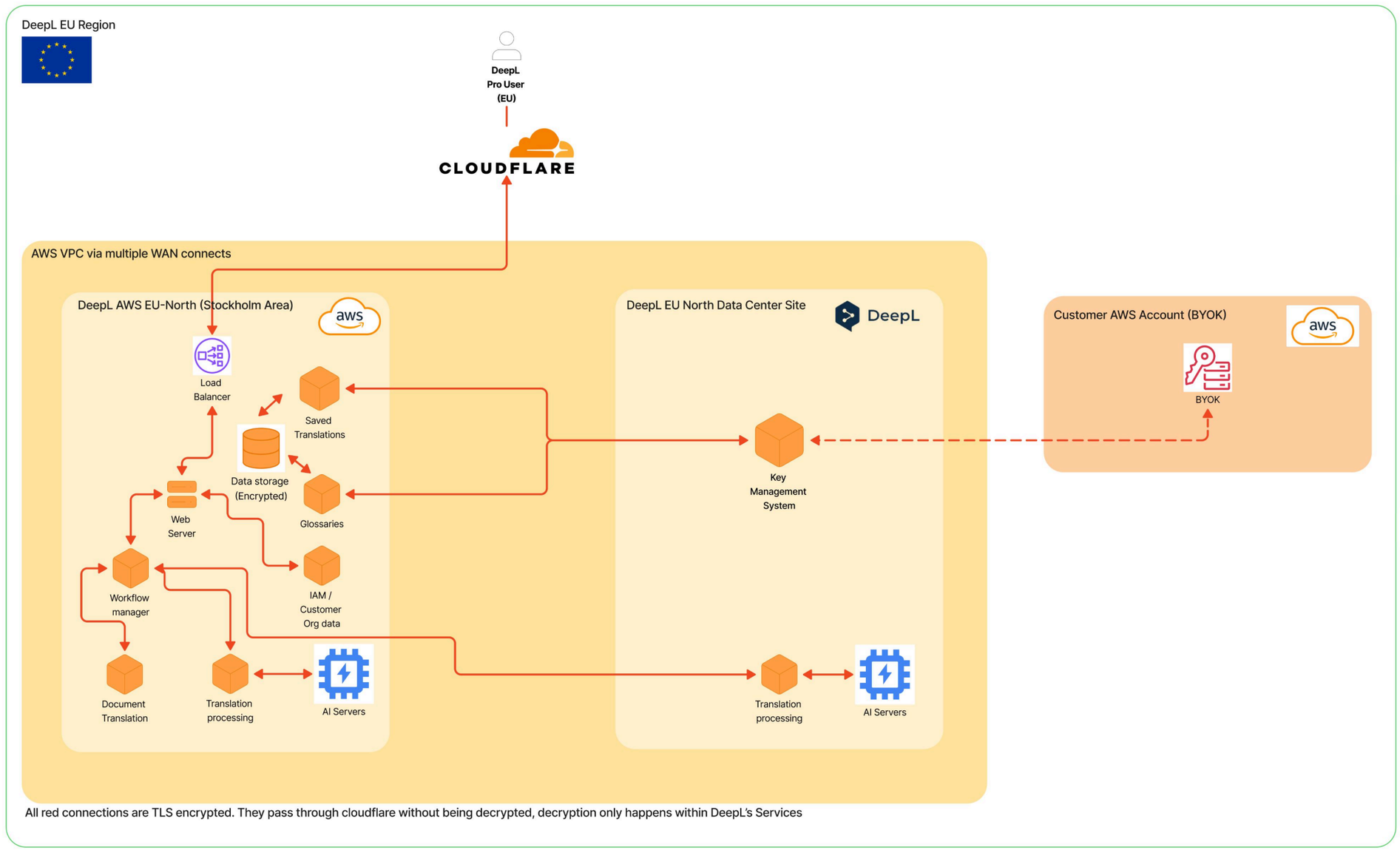


Figure 1

As illustrated in **Figure 1**, requests first pass through our CDN, Cloudflare which provides DeepL with a layer of security by filtering out malicious bot requests and attempts. Once filtered, requests, still encrypted, are passed through to our servers. Static web content such as images and javascript are served by Cloudflare to improve page load speeds for all customers.

Requests are decrypted by our load balancers, before being passed to down-stream services for processing using our internal TLS encryption layer.

Translation requests are assessed, for example by language pair, before being sent to our translation engines. Translation engines reside in both AWS and on-premises environments, enabling us to use of compute capacity as efficiently as possible.

Requests passing from one data center to another are encrypted in low-latency, high-speed virtual-private links (TLS Application layer + MACSec fallback).

When customers make use of glossaries or saved translations, this content is encrypted (envelope encryption) using our internally managed KMS service. By default this uses DeepL-managed encryption keys, hosted in our on-premises service. BYOK options are also available, which rely on a customer providing their keys via their AWS account.

This envelope-encrypted data is itself held on encryption-at-rest storage devices. Encryption at rest is provided by AWS, using AWS keys and KMS.

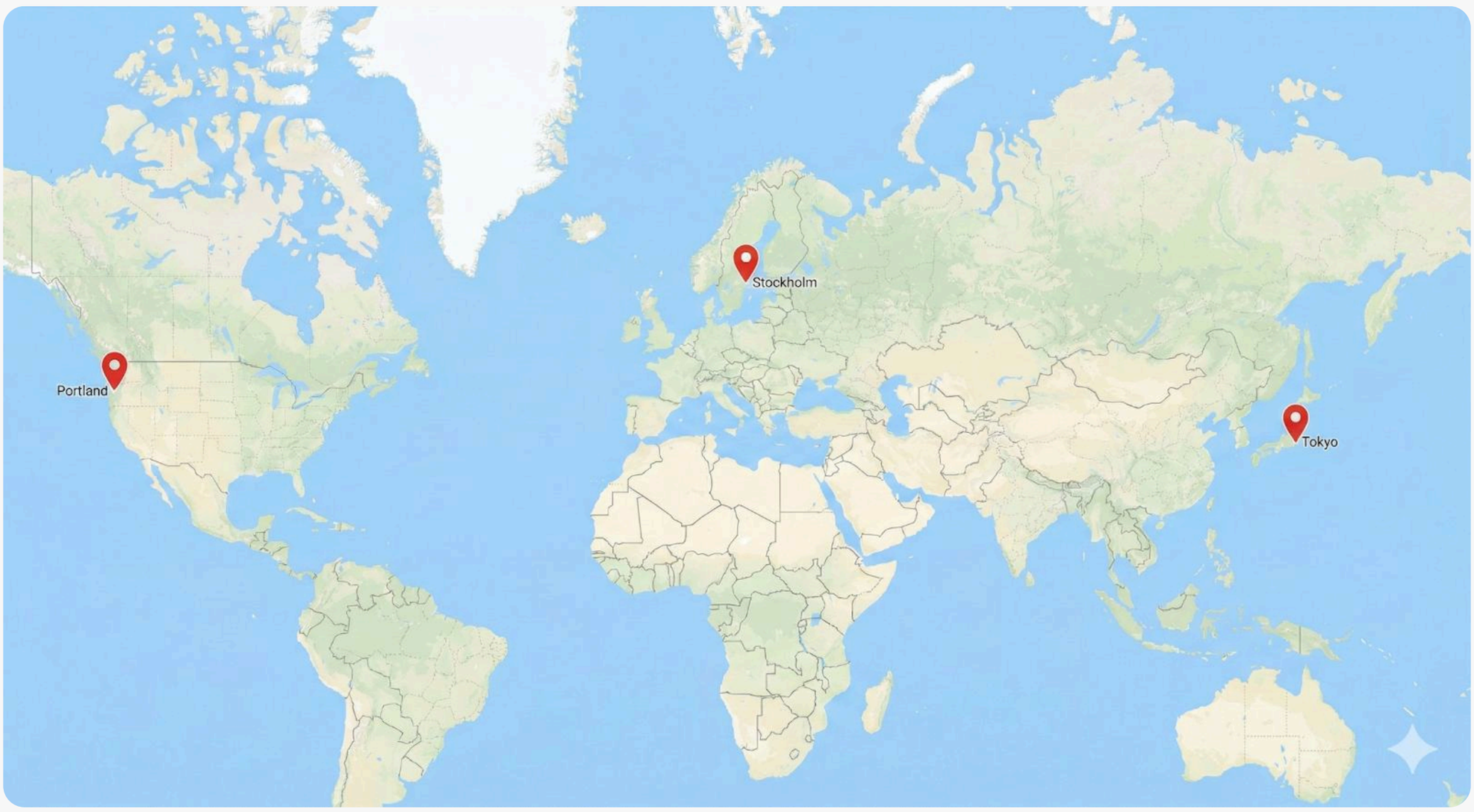


Figure 2

Figure 2 shows the regions that DeepL is using in the global setup. Oregon and Tokyo are the AWS-based regions for our regional customer services. Stockholm is our hybrid region, combining AWS with our data center site powering European regionality.

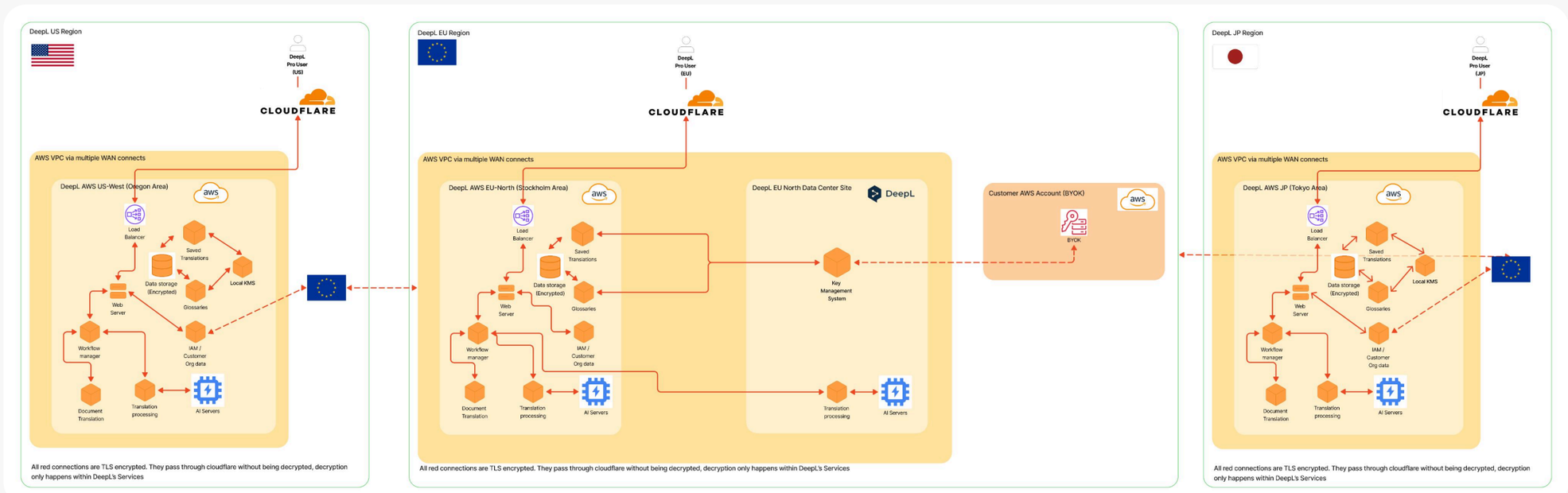


Figure 3

As illustrated in **Figure 3**, DeepL's regional deployments incorporate the whole stack and ensure the customer's interactions are handled within the preferred region. In Europe only, keys are maintained outside AWS. Bring your own key is available in all regions and relies on AWS. Configuration of the IAM (authentication, licenses) is implemented in a one-way sync from the European site.

Guaranteed data residency

We offer a data residency guarantee via paid add-on to simplify compliance and satisfy internal audit requirements. For our customers who have opted for data residency:

- **Region-locked processing via paid add-on:** Content from customers can be processed and stored exclusively within their chosen region.
- **Hybrid sovereignty:** While we utilize AWS regions within the EU (such as Stockholm), encryption key material is stored outside AWS within EU-based data centers via our agnostic KMS, unless a customer chooses **Bring Your Own Key,** which relies on an AWS feature.
- In other regions, key material is stored on AWS using our provider agnostic KMS service.
- For customers who do not purchase the data residency add-on, DeepL retains the right to process content from any of our regions.

Hardened perimeter and zero-knowledge edge

We utilize a "Zero-Trust" approach to global traffic. While we leverage Cloudflare for its world-class availability, your content remains a "black box" to the edge provider.

- Encrypted pass-through: Cloudflare manages bot detection and DDoS mitigation, but **cannot decrypt** your translation requests.
- TLS encryption in transit: Every request is protected via TLS encryption (represented by the **red lines** in the diagram) from your terminal to our secure environment.

Ephemeral "in-flight" data handling

For API, Website, and App interactions, we adhere to a strict policy of **Data Minimization.**

- **Zero residency:** Text operations are processed "in-flight." We do not store your source text or translations on disk during standard operations.
- **Just-in-time decryption:** Data is only decrypted at the point of entry to our services and re-encrypted on exit.

Data sovereignty (KMS and BYOK)

We bridge the gap between cloud flexibility and on-premises control through our proprietary key management approach.

- **Agnostic KMS:** Our Key Management Service (KMS) is independent of AWS. In Europe, it resides in our private data centers. This ensures that even when utilizing AWS infrastructure, the "root of trust" stays within our audited physical control.
- **Bring Your Own Key (BYOK):** Maintain ultimate authority by importing your own organization keys directly from your **AWS KMS.**
- **Document-level protection:** For stored content, we utilize an “envelope” encryption method where the client application holds the key, ensuring content is double-encrypted. Once at the object layer and once at the storage layer (encrypted at rest) and inaccessible to unauthorized parties. Document translations have a similar double encryption but using document-level, symmetric encryption rather than envelope encryption.

Security matrix

Feature	Protection Mechanism	Security Benefit
Data in transit	TLS 1.2+	TLS is recognised as the standard method of encrypting data in transit, and is used throughout the world by institutions with the highest standards.
Edge protection	Non-Decrypting CDN	Our CDN is used to accelerate the delivery of static assets - e.g. website images and javascript to improve performance. It also provides us with the first layer of defense against Bot and DDoS attacks. Our CDN has no visibility of what is inside the encrypted connections to our services.
Saved data	Agnostic KMS + optional BYOK	In addition to the regular encryption-at-rest of our storage, we employ a layer of encryption above this enabling us to maintain control of keys and cryptographic material, both on-prem and on cloud.
Processing	In-flight transient decryption	We decrypt traffic at the latest possible point, and continue to work on pushing this further into our own services.

Glossary

AWS
Amazon Web Services

BSI
Bundesamt für Sicherheit
in der Informationstechnik

BYOK
Bring Your Own (Encryption) Key

GDPR
General Data Protection Regulation (EU)

DDoS
Distributed Denial-of-Service

KMS
Key Management System

TLS
Transport Layer Security

VPC
Virtual Private Cloud

WAN
Wide Area Network

IAM
Identity and Access Management

DeepL SE
Maarweg 165,
50825 Cologne, Germany
info@deepl.com