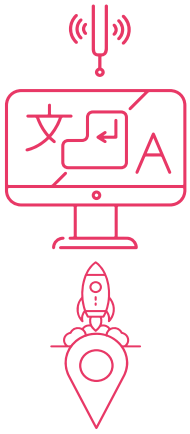
AI data security through in-house fine-tuned language models

Alpha CRC addresses the growing risk of IP leakage by developing in-house fine-tuned language models, which anonymize internal documents for training, continuously update the models, and enforce strict access controls on private servers. This approach ensures data sovereignty, regulatory compliance, and robust security, thereby enhancing efficiency and customer trust while maintaining a competitive edge.

Tools



LLMs, Launchpoint, Anvil

Outcomes

- **Limited risk** of data exposure to third parties
- **Enabled** complete customization of the LLM
- **Improved efficiency** on a wide range of localization services

Use Case description

With AI tool uptake on the rise, the risk of IP leakage is a growing fear for many companies. Alpha CRC has developed a set of AI tools through the creation of in-house language models. This is performed by anonymizing internal documents for training, continuously updating the models, and enforcing strict access controls on private servers.

The solution ensures data sovereignty, customization for specific needs, regulatory compliance, and robust security measures. Consequently, Alpha CRC not only fortifies its data security but also gains a competitive advantage through improved efficiency and reinforced customer trust.

In-house language model security

Challenge

Alpha CRC handles sensitive data for various clients. With the increasing sophistication of cyber threats and the need for compliance with various data protection regulations (like GDPR, HIPAA), Alpha CRC is required to ensure data security and privacy while maintaining efficient data processing and customer service operations.

The use of third-party language processing tools poses a risk of data breaches and leaks, as sensitive information could be inadvertently exposed to external entities.

Solution

To address these security concerns, Alpha CRC decided to fine-tune in-house language models (LLMs) tailored to its specific needs. The process involves:

1. Collecting a large dataset of company-specific documents and communications that are thoroughly anonymized to remove any sensitive information.
2. Training a base language model on this dataset to understand the company's terminology, workflows, and communication styles.
3. Continuously fine-tuning the model with updated data to improve its accuracy and relevance to the company's evolving needs.
4. Implementing strict access controls and monitoring systems to ensure that only authorized personnel can use and train the model.
5. Running the language models on secure, private servers within the company's own infrastructure to prevent unauthorized external access.