WHITE PAPER

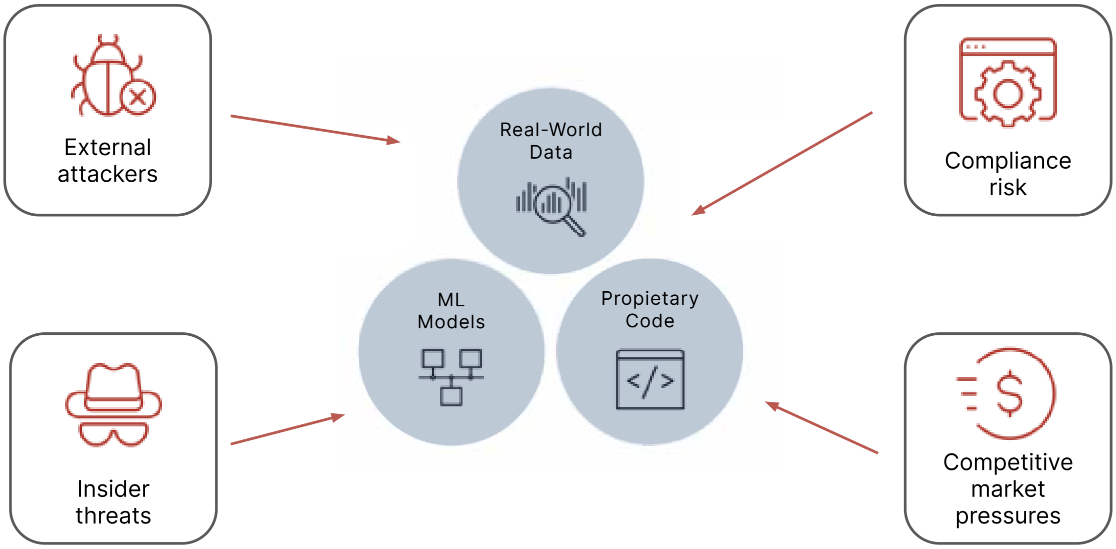
Anjuna Seaglass™ AI Clean Rooms [preview]

Accelerate AI with Trustworthy Multi-Party Collaboration

Data Collaboration Has Unlocked Potential

In today's enterprise environment, data often remains locked inside internal silos, even though Gartner reports that data collaborations with external partners result in triple the revenue and higher profit margins. 60% of data leaders cannot use a majority of their data due to privacy concerns, according to Bloor Research.

Today, valuable data collaborations in areas as varied as medical diagnosis, fraud detection, and customer engagement are blocked by data and privacy risks. The situation is net-negative: customers miss out on valuable insights, while enterprises are blocked from a new revenue stream.



Data risks are varied. External attackers are always looking for new ways to get valuable customer data to be sold or ransomed. Insider threats, including both malicious and well-meaning but compromised employees, challenge traditional security boundaries. AI adds to these risks by increasing the volume of data and exposing new attack surfaces for data pipeline compromise and model exploitation.

Compliance risks, for both security and privacy, are also increasing year-over-year. There has never been a worse time to be breached, yet the demand for data persists. The enterprise must grow, and those who stay in place will quickly fall behind their more agile and innovative competitors.

Why is traditional data collaboration difficult, and how can we improve?

Traditional Collaboration Methods are Painful and Ineffective

In legacy workflows, data collaboration is a manual process: data is exported and shared though spreadsheets in email, SFTP, ETL/ELT, or perhaps a data warehouse/lake. A data analyst then combines and processes the dataset, returning the results to all the collaborators. At every step of the way, the data is potentially at risk to be accidentally or maliciously exposed to outsiders. These breaches are extremely costly, between reputational damage and punitive fines: recent court cases in the US as well as privacy regulation like the EU GDPR demand strong and provable technical controls to protect customer data.

Various data security mitigations try to address this risk, with mixed results. Techniques like differential privacy and synthetic data create statistically-similar "fake data"; but fake data is not real data, resulting in biased or completely incorrect results, especially for ML models that require accurate data at both training and inference time. Anonymization and obfuscation also make it difficult to pinpoint exact matches across two or more data partners; for example, it is impossible to correlate the same fraud actor if the data partners are using different anonymization schemes. Ultimately, data scientists and AI systems are deprived of opportunity by data manipulation that destroys potentially-valuable features.

Besides accuracy, speed is critical. Approaches like zero-knowledge proofs and fully-homomorphic encryption are theoretically sound, yet practically unusable for streaming data for live decision-making and time-sensitive processing. In a different sense of speed, the ability to iterate quickly and ship your roadmap makes a difference. Data teams are hampered by working through layers of traditional obfuscation, and this leads to delayed projects, missed deadlines, and a huge opportunity cost.

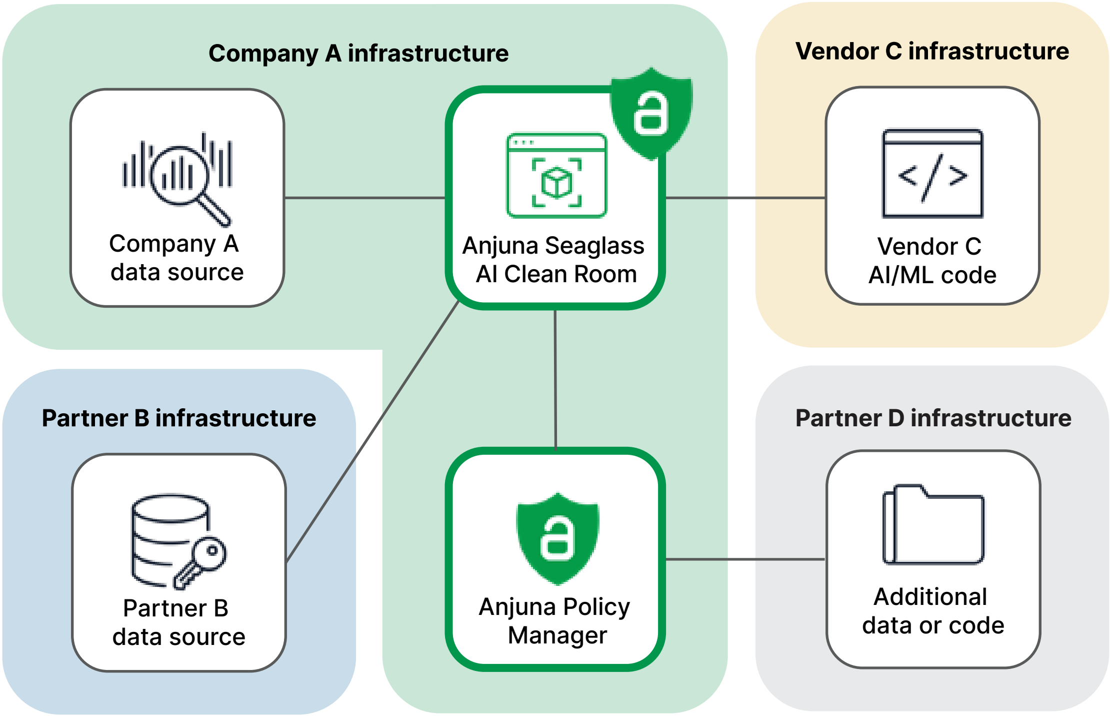
Despite these challenges, data security is not optional. Regulators, litigators, and your own customers will no longer accept mere organizational promises that data will be protected: there needs to be a technical control that provably ensures strong isolation of data in all states, while still enabling your data team to get things done.

At Anjuna, we noticed many of our existing customers were blocked and frustrated by traditional approaches to secure data collaboration. Whether they used analytics tooling, machine learning, or were early adopters of generative AI, the core challenges of data collaboration remained. That's why we built a Confidential Data Clean Room solution aligned to their needs: a "walled garden" that enables you to unlock valuable data and collaborate with partners, with verifiable protection against breaches.

What is a Confidential Data Clean Room?

A Confidential Data Clean Room uses Confidential Computing, or "secure enclaves", as a technical safeguard to protect sensitive data in collaborative use cases. Confidential Computing is a hardware-based technology that prevents unauthorized access and enables verification of the exact code running. The technology is enabled by next generation servers developed by companies including AMD, Intel, and NVIDIA, and it is available in all leading clouds, including AWS, Google Cloud and Microsoft Azure.

Anjuna Seaglass AI Clean Rooms, currently in Preview state, enables multiple parties to encrypt their data in a way that can only be decrypted within a "secure enclave". The resulting analysis outputs are aggregated results which do not reveal the underlying data.



The solution includes the Anjuna Policy Manager, an attestation-aware service that enables governance, ensuring the clean room only runs approved workloads that meet the security policies of all organizations involved. This enables true privacy over data inputs, models, processing, and results, as required by regulations like GDPR, CPRA, POPIA, and more.

The hardware-based guarantees of the Seaglass AI Clean Room eliminate the risk of data exposure from insider threats, highly-sophisticated attackers, and even rogue cloud service provider employees. This technical control enables Seaglass AI Clean Rooms to unlock your enterprise's sensitive data and AI/ML models for new collaborations.

What can you achieve with Anjuna Seaglass AI Clean Rooms?

- Deploy immediately into your preferred cloud (or your partner's cloud) and start processing data and models. Use business-focused web user interfaces, or integrate with APIs for automation and deeper technical integrations.
- Enable secure collaboration with partners by sharing both data and ML models, while maintaining tight access control and data security guarantees.
- Get full-fidelity insights on data instead of fighting with techniques like data masking, synthetic data, or obfuscations that create new risks from biased distributions.
- Earn customer trust by locking out insider access to their data, while still being able to use state-of-the-art machine learning for better, data-driven decisions.
- Provide access to private training datasets and pre-trained ML models to enable new monetization strategies without losing control of your intellectual property.
- Simplify and ensure compliance with security, privacy, and data sovereignty regulations.

Customer examples

A major North American bank wanted to unlock their first-party Level 3 data by using a data clean room, but existing market offerings didn't meet their data security requirements. By partnering with Anjuna, the bank was able to accelerate their own data clean room development by months, including replacing a complex synthetic data approach. The bank was able to meet banking and privacy compliance needs, reduce data use approval time by 90%, and enable new revenue by overcoming partner objections on sharing data.

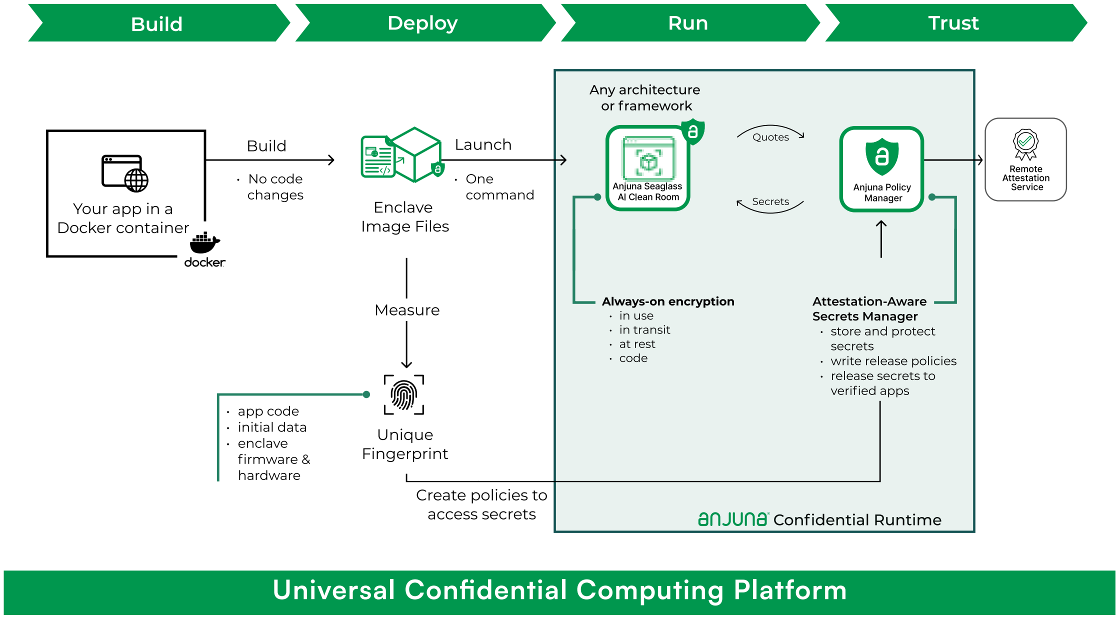
A specialist AI company for healthcare research spent months struggling to build their own cloud-native multi-party clean room, which delayed revenue growth. Seaglass AI Clean Rooms enabled custom AI model execution using AMD SEV-SNP in just days and set the course for future Confidential GPU use. This enabled AI partners to help healthcare providers without risking their valuable intellectual property investment in AI/ML.

A U.S. government agency wanted to combine real-time geospatial intelligence data with allied nations, but needed to ensure their security. If advanced persistent threat (APT) actors could access the combined data, the entire allied mission would be at risk. Using Seaglass AI Clean Rooms, the agency was able to quickly move their application and relevant mission data into a Confidential Data Clean Room. Instead of their old approach of using hours-old data, the agency and allies are now able to react with agility and adapt immediately to dynamic conditions.

How does Anjuna Seaglass AI Clean Rooms work?

Seaglass AI Clean Rooms builds on top of Anjuna Seaglass, the Universal Confidential Computing Platform, to secure your data for collaboration.

Inside look at Anjuna Seaglass AI Clean Rooms


Build
Deploy
Run
Trust
Your app in a Docker container
docker
Build
· No code changes
Enclave Image Files
Launch
· One command
Measure
Unique Fingerprint
· app code
· initial data
· enclave firmware & hardware
Create policies to access secrets
Any architecture or framework
Anjuna Seaglass AI Clean Room
Quotes
Secrets
Anjuna Policy Manager
Remote Attestation Service
Always-on encryption
· in use
· in transit
· at rest
· code
Attestation-Aware Secrets Manager
· store and protect secrets
· write release policies
· release secrets to verified apps
anjuna Confidential Runtime
Universal Confidential Computing Platform

1. Data providers use Anjuna's SDKs, APIs, or UI components to encrypt and upload data into the clean room.

2. The clean room runs your desired container workload and produces aggregated results.

3. Security and encryption policy is enforced by the Anjuna Policy Manager, ensuring that unauthorized workloads are never allowed to run.

Why choose Anjuna Seaglass AI Clean Rooms over alternatives?

Existing data collaboration solutions are based on synthetic data, redaction or tokenization, or brittle access control over shared infrastructure. These alternatives destroy valuable data, are difficult to implement and operate, and limit the types of insights available.

In comparison, Seaglass AI Clean Rooms offer the following advantages:

- Full power: many data clean rooms have very limited processing operations; for example, you can only find rows if they have exactly-matching ID fields. With Seaglass AI Clean Rooms, you can write any code, including AI/ML inference, to get the results you need.
- Full control: other data clean rooms are SaaS offerings, where the operator could potentially compromise your data. Seaglass AI Clean Rooms are deployed in your own infrastructure and never require remote access for outsiders.
- Policy-based verification: the unique capabilities of Confidential Computing are used by the Anjuna Policy Manager to ensure that only your approved workloads run - verifiable down to the hardware level.

Conclusion

The pains of security and privacy for traditional data collaboration are in the past. By using Anjuna Seaglass AI Clean Rooms, one of our earliest design partners (a major North American bank) cut out several months of engineering time, proved their business case, and are now scaling out their collaborations. This fresh yet simple Confidential Computing approach helps executives overcome risks in compliance and privacy and harness the power of their data. Let's accelerate your journey to unlock new data partnerships, better serve your customers, and bring in new lines of revenue too.

About Anjuna Seaglass

Anjuna Seaglass™ is the world's first Universal Confidential Computing Platform, capable of running applications in any cloud with complete data security and privacy. Anjuna Seaglass isolates workloads in a hardware-assisted Trusted Execution Environment that intrinsically secures data in every state – in use, at rest, and in transit – to create a zero trust environment.

Anjuna Seaglass AI Clean Rooms is currently in preview. Features and functionality may change as development progresses. For any questions or feedback, please contact Bobbie Chen, Senior Product Manager, at bobbie.chen@anjuna.io.

Get Started