

Centific's Gen AI suite leverages Azure Open AI and other Azure Cloud services and is designed to cater to all needs of an enterprise. Enterprises looking for out of the box Gen AI solutions are having to evaluate several platforms and solutions for their various needs. Centific Suite ensures that all the users in an Enterprise have access to Gen AI to perform their roles better.

Centific GenAI suite is powered by our Intelligent Data Platform (IDP.ai) wherein configuration activities like Datasource connection, Data Fabric configuration with Metadata Mapping.

The suite includes:

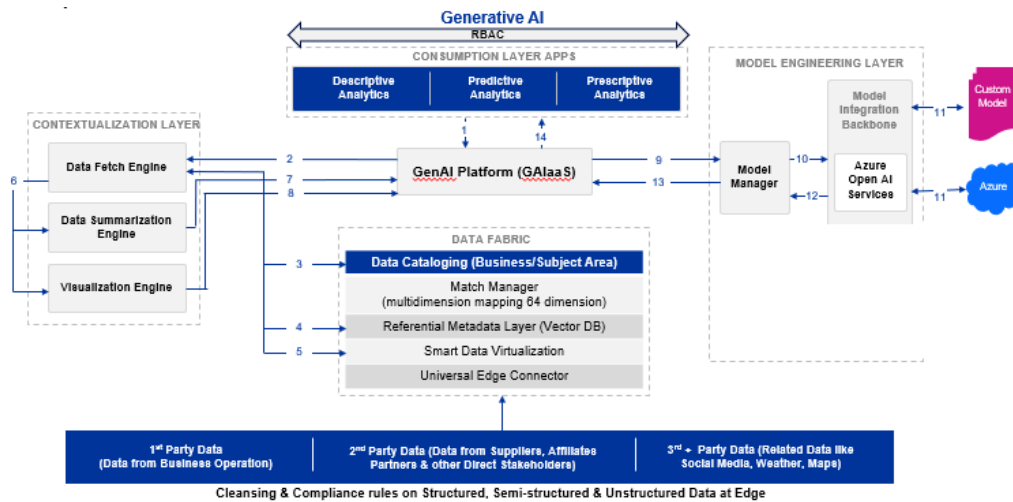
- 1) Talk2Data- Gen AI solution targeted for users managing the Business. Key KPIs around Demand Forecasting, Return on Investment, Cost Analysis & Budgeting, Business Planning can be done using Talk2Data. Talk2Data can also be used by Functional teams like HR, Finance, Legal etc. for information dissemination retrieval within the enterprise
- 2) Talk2Cloud- Gen AI solution tailored for IT Teams to track the FinOps, SecOps and CloudOps setup and suggest steps to improve the posture. It is a really powerful tool particularly if the number of subscriptions are high and decision making is decentralized
- 3) Talk2Code- Gen AI based Co-pilot solution to help Engineering teams be more productive. Helps identify gaps in coding standards, Requirements coverage and Test Coverage
- 4) Talk2Application- Gen AI solution for ISVs to manage multiple tenants and their subscription

The suite is offered as a **Managed solution** and will be deployed on the **customer's Azure subscription**.

Some of the services used in the Gen AI solution that will be consumed include:

- ACR
- Storage Accounts
- Azure Container Registry
- Azure Databricks
- Azure Cache for Redis
- Azure Kubernetes Service (AKS)
- Key Vault
- Azure Monitor
- Virtual Network
- Azure Private Link
- Azure SQL Database
- Azure Front Door
- API Management
- Azure Cosmos DB
- Azure OpenAI Service
- Azure AI services
- Microsoft Defender for Cloud
- App Service
- HDInsight(Spark)
- HDInsight(Kafka)

Technical Architecture



The components of Centific's Gen AI architecture are:

- 1) Universal Edge connector- PreBuilt connectors to connect to various structured , semi structured and unstructured data sources
- 2) Data Fabric- Ensures that only the metadata is extracted and Data is processed in real time through virtualization without migrating or creating copies
- 3) Model Engineering layer is provided by Open AI or custom model
- 4) Contextualization layer is where Business context is applied to align the output to Business KPIs
- 5) Consumption Apps are the presentation layer that can view the outcome based on Role Based access controls defined

Engagement Model (Implementation/Professional Services)

Implementation of Centific's Talk2 Suite happens through a Proof of Concept/Proof of value phase which is about 12 weeks long. This includes 3 steps:

- 1) Interaction with Business owners to identify use case for PoC and the appropriate Talk2 solution
- 2) 1-2 week planning and preparation phase to identify use case and KPI to be measured
- 3) 8-10 week PoC implementation phase with exit criteria being a clear demonstration of Proof of value realized
- 4) The duration and size of team for full implementation varies between 20-60 person months depending on complexity of use cases and datasets to be integrated

Deliverables

- 1) Deployed Solution on Azure

- 2) Demonstrated and Measured Improvements on KPIs agreed as part of use case definition (KPI before Gen AI and after Gen AI implementation)
- 3) Business Architecture and Technical Architecture documents
- 4) User training & handover to Technical and Business teams

Target Audience

- 1) Talk2Data – Line of Business owners, Functional team leaders (Head of HR, Head of Marketing, Head of Legal, Head of Finance etc.)
- 2) Talk2Cloud- CTO, IT Leader responsible for improving Azure consumption and posture
- 3) Talk2Code- Engineering Leaders responsible for custom development and improving Quality and Productivity Metrics of Engineering teams

Benefits/Value Proposition:

- 1) Suite of products for all enterprise needs- option to select based on Business needs
- 2) Centific’s Gen AI suite is based on the principal of Data Virtualization wherein all Data is processed at the source(edge) itself without the need to migrate to a centralized repository
- 3) Fully hosted on Azure – SaaS offering where customers pay based on consumption of Azure services
- 4) Centific provides professional services covering implementation, user training and application support
- 5) Centific also provides roadmap to graduate to open source proprietary (Custom) LLM solution based on Gen AI adoption levels within the enterprise



**Subject to funding approval from Microsoft and Centific*

Azure consumption Bundle

ACR	Small	Medium	Large (Enterprise)
Storage Accounts	Block Blob Storage, General Purpose V2, Hierarchical Namespace, LRS Redundancy, Hot Access Tier, 1,000 GB Capacity - Pay as you go, 500 x 10,000 Write operations, 499 x 10,000 Read operations, 10 x	Block Blob Storage, General Purpose V2, Hierarchical Namespace, LRS Redundancy, Hot Access Tier, 1,000 GB Capacity - Pay as you go, 500 x 10,000 Write operations, 499 x 10,000 Read operations, 10 x	Premium Block Blob Storage, Hierarchical Namespace, ZRS Redundancy, Hot Access Tier, 1 TB Capacity - Pay as you go, 1,000 x 10,000 Write

	10,000 Iterative Read operations, 10 x 100 Iterative Write operations, 1,000 GB Data Retrieval, 1,000 GB Data Write, 1,000 GB Index, 500 x 10,000 Other operations	10,000 Iterative Read operations, 10 x 100 Iterative Write operations, 1,000 GB Data Retrieval, 1,000 GB Data Write, 1,000 GB Index, 500 x 10,000 Other operations	operations, 1,000 x 10,000 Read operations, 10 x 10,000 Iterative Read operations, 10 x 100 Iterative Write operations, 1,000 GB Data Retrieval, 1,000 GB Data Write, 1,000 GB Index, 100 x 10,000 Other operations
Azure Container Registry	Standard Tier, 2 registries x 30 days, 10 GB Extra Storage, Container Build - 4 CPUs x 301 Seconds - Inter Region transfer type, 5 GB outbound data transfer from East US to East Asia	Standard Tier, 2 registries x 30 days, 10 GB Extra Storage, Container Build - 4 CPUs x 301 Seconds - Inter Region transfer type, 5 GB outbound data transfer from East US to East Asia	Standard Tier, 2 registries x 30 days, 10 GB Extra Storage, Container Build - 4 CPUs x 301 Seconds - Inter Region transfer type, 5 GB outbound data transfer from East US to East Asia
Azure Databricks	All-Purpose Compute Workload, Premium Tier, 1 D8AV4 (8 vCPU(s), 32 GB RAM) x 730 Hours, Pay as you go, 1.5 DBU x 730 Hours	All-Purpose Compute Workload, Premium Tier, 1 D8AV4 (8 vCPU(s), 32 GB RAM) x 730 Hours, Pay as you go, 1.5 DBU x 730 Hours	All-Purpose Compute Workload, Premium Tier, 1 D8AV4 (8 vCPU(s), 32 GB RAM) x 730 Hours, Pay as you go, 1.5 DBU x 730 Hours
Azure Cache for Redis	Premium tier; 1 Shard per Instance, 0 Additional Replicas per Shard, 1 P3 instances, 730 Hours, Pay as you go	Premium tier; 1 Shard per Instance, 0 Additional Replicas per Shard, 1 P3 instances, 730 Hours, Pay as you go	Premium tier; 1 Shard per Instance, 0 Additional Replicas per Shard, 1 P3 instances, 730 Hours, Pay as you go
Azure Kubernetes Service (AKS)	Standard; Cluster management for 3 clusters; 14 DC2es v5 (2 vCPUs, 8 GB RAM) x 730 Hours (Pay as you go), Linux; 3 managed OS disks – E15	Standard; Cluster management for 3 clusters; 14 DC2es v5 (2 vCPUs, 8 GB RAM) x 730 Hours (Pay as you go), Linux; 3 managed OS disks – E15	Standard; Cluster management for 3 clusters; 9 DC4es v5 (4 vCPUs, 16 GB RAM) x 730 Hours (Pay as you go), Linux; 3 managed OS disks – E2
Key Vault	Vault: 500 operations, 0 advanced operations, 0 renewals, 0 protected keys, 0 advanced protected keys; Managed HSM Pools: 0 Standard B1 HSM Pool(s) x 730 Hours	Vault: 500 operations, 0 advanced operations, 0 renewals, 0 protected keys, 0 advanced protected keys; Managed HSM Pools: 0 Standard B1 HSM Pool(s) x 730 Hours	Vault: 10,000 operations, 0 advanced operations, 0 renewals, 0 protected keys, 0 advanced protected keys; Managed HSM Pools: 0 Standard B1 HSM Pool(s) x 730 Hours

<p>Azure Monitor</p>	<p>Log analytics: Log Data Ingestion: 5 GB Daily Analytics logs ingested, 5 GB Daily Basic logs ingested, 1 months of Interactive Data Retention, 0 months of data archived, 100 Basic Log Search Queries per day with 1 GB data scanned per query, 0 GB of Log Data Exported per day, Platform Log Data Processed per day: 0 GB with Destination to Storage or Event Hub and 0 GB with Destination to Marketplace Partners, 0 Search job Queries per day with 0 GB data scanned per query; Managed Prometheus: 0 AKS nodes in cluster, 0 Prometheus metrics per node, 3 seconds of Metric collection interval, 0 Average daily Dashboards users, 7 Dashboards, 50000 Data samples queried per dashboard, 25 promql alerting rules, 25 promql recording rules; Application Insights: 3 months Data retention, 0 Multi-step Web Tests; 6 resources monitored X 1 metric time-series monitored per resource, 0 Log Alerts at 5 Minutes Frequency, 0 Additional events (in thousands), 0 Additional emails (in 100 thousands), 0 Additional push notifications (in 100 thousands), 0 Additional web hooks (in millions)</p>	<p>Log analytics: Log Data Ingestion: 5 GB Daily Analytics logs ingested, 5 GB Daily Basic logs ingested, 1 months of Interactive Data Retention, 0 months of data archived, 100 Basic Log Search Queries per day with 1 GB data scanned per query, 0 GB of Log Data Exported per day, Platform Log Data Processed per day: 0 GB with Destination to Storage or Event Hub and 0 GB with Destination to Marketplace Partners, 0 Search job Queries per day with 0 GB data scanned per query; Managed Prometheus: 0 AKS nodes in cluster, 0 Prometheus metrics per node, 3 seconds of Metric collection interval, 0 Average daily Dashboards users, 7 Dashboards, 50000 Data samples queried per dashboard, 25 promql alerting rules, 25 promql recording rules; Application Insights: 3 months Data retention, 0 Multi-step Web Tests; 6 resources monitored X 1 metric time-series monitored per resource, 0 Log Alerts at 5 Minutes Frequency, 0 Additional events (in thousands), 0 Additional emails (in 100 thousands), 0 Additional push notifications (in 100 thousands), 0 Additional web hooks (in millions)</p>	<p>Log analytics: Log Data Ingestion: 10 GB Daily Analytics logs ingested, 10 GB Daily Basic logs ingested, 1 months of Interactive Data Retention, 0 months of data archived, 1000 Basic Log Search Queries per day with 1 GB data scanned per query, 0 GB of Log Data Exported per day, Platform Log Data Processed per day: 0 GB with Destination to Storage or Event Hub and 0 GB with Destination to Marketplace Partners, 0 Search job Queries per day with 0 GB data scanned per query; Managed Prometheus: 9 AKS nodes in cluster, 5 Prometheus metrics per node, 1 seconds of Metric collection interval, 2 Average daily Dashboards users, 7 Dashboards, 50000 Data samples queried per dashboard, 25 promql alerting rules, 25 promql recording rules; Application Insights: 3 months Data retention, 0 Multi-step Web Tests; 6 resources monitored X 1 metric time-series monitored per resource, 0 Log Alerts at 5 Minutes Frequency, 0 Additional events (in thousands), 10 Additional emails (in 100 thousands), 10 Additional push notifications (in 100 thousands), 5 Additional web hooks (in millions), {0} Voice calls and 200 SMS in United States (+1)</p>
<p>Virtual Network</p>	<p>East US (Virtual Network 1): 500 GB Outbound Data Transfer; East US (Virtual Network 2): 10 TB Outbound Data Transfer</p>	<p>East US (Virtual Network 1): 500 GB Outbound Data Transfer; East US (Virtual Network 2): 10 TB Outbound Data Transfer</p>	<p>East US (Virtual Network 1): 5000 GB Outbound Data Transfer; East US</p>

			(Virtual Network 2): 10 TB Outbound Data Transfer
Azure Private Link	5 Endpoints X 730 Hours, 100 GB Outbound data processed, 100 GB Inbound data processed	5 Endpoints X 730 Hours, 100 GB Outbound data processed, 100 GB Inbound data processed	5 Endpoints X 730 Hours, 100 GB Outbound data processed, 100 GB Inbound data processed
Azure SQL Database	Single Database, vCore, General Purpose, Provisioned, Standard-series (Gen 5), Locally Redundant, 1 - 10 vCore Database(s) x 730 Hours, 1100 GB Storage, LRS Backup Storage Redundancy, 0 GB Point-In-Time Restore, 0 x 5 GB Long Term Retention	Single Database, vCore, General Purpose, Provisioned, Standard-series (Gen 5), Locally Redundant, 1 - 10 vCore Database(s) x 730 Hours, 1100 GB Storage, LRS Backup Storage Redundancy, 0 GB Point-In-Time Restore, 0 x 5 GB Long Term Retention	Single Database, vCore, General Purpose, Provisioned, Standard-series (Gen 5), Zone Redundant, 1 - 8 vCore Database(s) x 730 Hours, 1000 GB Storage, ZRS Backup Storage Redundancy, 0 GB Point-In-Time Restore, 12 x 50 GB Long Term Retention
Azure Front Door	Azure Front Door Premium - Base instance included, 5 GB Data Transfer Out to Client, 5 GB Data Transfer In to Origin, 0 x 10,000 Requests	Azure Front Door Premium - Base instance included, 5 GB Data Transfer Out to Client, 5 GB Data Transfer In to Origin, 0 x 10,000 Requests	Azure Front Door Premium - Base instance included, 2 TB Data Transfer Out to Client, 2 TB Data Transfer In to Origin, 0 x 10,000 Requests
API Management	Premium tier, 1 base unit x 730 Hours, 0 additional units	Premium tier, 1 base unit x 730 Hours, 0 additional units	Premium tier, 1 base unit x 730 Hours, 0 additional units
Azure Cosmos DB	Azure Cosmos DB for NoSQL (formerly Core), Standard provisioned throughput (manual), Always-free quantity disabled, Single Region Write (Single-Master) - East US (Write Region), 12,000 RU/s x 730 Hours, 1,000 GB transactional storage, Analytical storage disabled, 2 copies of periodic backup storage, Dedicated Gateway not enabled	Azure Cosmos DB for NoSQL (formerly Core), Standard provisioned throughput (manual), Always-free quantity disabled, Single Region Write (Single-Master) - East US (Write Region), 12,000 RU/s x 730 Hours, 1,000 GB transactional storage, Analytical storage disabled, 2 copies of periodic backup storage, Dedicated Gateway not enabled	Azure Cosmos DB for NoSQL (formerly Core), Standard provisioned throughput (manual), Always-free quantity disabled, Single Region Write (Single-Master) - East US (Write Region), 20,000 RU/s x 730 Hours, 1,000 GB transactional storage, Analytical storage disabled, 2 copies of periodic backup storage, Dedicated Gateway not enabled
Azure OpenAI Service	Language Models, GPT-4-32K, 5000 x 1000 prompt tokens, 5000 x 1000 completion tokens	Language Models, GPT-4-32K, 100000 x 1000 prompt tokens, 10000 x 1000 completion tokens	Language Models, GPT-4-32K, 80000 x 1000 prompt tokens, 40000 x 1000 completion tokens

<p>Azure AI services</p>	<p>Azure AI Language, Pay as you go, Standard, AI Services for Language: 4000 x 1,000 text records, Text Analytics for health: 100 x 1,000 text records, Custom question answering and prebuilt question answering: 1000 x 1,000 text records, Custom named entity recognition and Custom text classification: 100 x 1,000 text records, 100 training hours, 100 endpoint hosting models, Conversational Language Understanding and Orchestration Workflow: 100 x 1,000 text records, 100 Advanced training hours</p>	<p>Azure AI Language, Pay as you go, Standard, AI Services for Language: 4000 x 1,000 text records, Text Analytics for health: 100 x 1,000 text records, Custom question answering and prebuilt question answering: 1000 x 1,000 text records, Custom named entity recognition and Custom text classification: 100 x 1,000 text records, 100 training hours, 100 endpoint hosting models, Conversational Language Understanding and Orchestration Workflow: 100 x 1,000 text records, 100 Advanced training hours</p>	<p>Azure AI Language, Pay as you go, Standard, AI Services for Language: 10000 x 1,000 text records, Text Analytics for health: 100 x 1,000 text records, Custom question answering and prebuilt question answering: 1000 x 1,000 text records, Custom named entity recognition and Custom text classification: 100 x 1,000 text records, 100 training hours, 100 endpoint hosting models, Conversational Language Understanding and Orchestration Workflow: 100 x 1,000 text records, 100 Advanced training hours</p>
<p>Microsoft Defender for Cloud</p>	<p>Microsoft Defender for Cloud by Resource: 1 Plan 1 servers x 730 Hours, 1 Plan 2 servers x 730 Hours, 2 Container vCores x 730 Hours, 0 additional container image scans, 4 App Service nodes x 730 Hours, 2 SQL Database servers on Azure, 0 SQL Database servers outside Azure x 730 Hours, 0 MySQL Instances, 0 PostgreSQL Instances, 0 MariaDB Instances x 730 Hours, Cosmos DB 2 x100 RU/s x 730 Hours, 1 Storage accounts x 730 Hours with 0 million total coverage of transactions across each storage account and 10 GB storage scanned for malware, 2 Key Vault(s), 2 Subscription(s)</p>	<p>Microsoft Defender for Cloud by Resource: 1 Plan 1 servers x 730 Hours, 1 Plan 2 servers x 730 Hours, 2 Container vCores x 730 Hours, 0 additional container image scans, 4 App Service nodes x 730 Hours, 2 SQL Database servers on Azure, 0 SQL Database servers outside Azure x 730 Hours, 0 MySQL Instances, 0 PostgreSQL Instances, 0 MariaDB Instances x 730 Hours, Cosmos DB 2 x100 RU/s x 730 Hours, 1 Storage accounts x 730 Hours with 0 million total coverage of transactions across each storage account and 10 GB storage scanned for malware, 2 Key Vault(s), 2 Subscription(s)</p>	<p>Microsoft Defender for Cloud by Resource: 1 Plan 1 servers x 730 Hours, 1 Plan 2 servers x 730 Hours, 2 Container vCores x 730 Hours, 0 additional container image scans, 4 App Service nodes x 730 Hours, 2 SQL Database servers on Azure, 0 SQL Database servers outside Azure x 730 Hours, 0 MySQL Instances, 0 PostgreSQL Instances, 0 MariaDB Instances x 730 Hours, Cosmos DB 2 x100 RU/s x 730 Hours, 1 Storage accounts x 730 Hours with 0 million total coverage of transactions across each storage account and 10 GB storage scanned for malware, 2 Key Vault(s), 2 Subscription(s)</p>

App Service	Premium V2 Tier; 2 P3V2 (4 Core(s), 14 GB RAM, 250 GB Storage) x 730 Hours; Linux OS; 0 SNI SSL Connections; 0 IP SSL Connections; 0 Custom Domains; 0 Standard SLL Certificates; 0 Wildcard SSL Certificates	Premium V2 Tier; 2 P3V2 (4 Core(s), 14 GB RAM, 250 GB Storage) x 730 Hours; Linux OS; 0 SNI SSL Connections; 0 IP SSL Connections; 0 Custom Domains; 0 Standard SLL Certificates; 0 Wildcard SSL Certificates	Premium V2 Tier; 2 P3V2 (4 Core(s), 14 GB RAM, 250 GB Storage) x 730 Hours; Linux OS; 0 SNI SSL Connections; 0 IP SSL Connections; 0 Custom Domains; 0 Standard SLL Certificates; 0 Wildcard SSL Certificates
HDInsight(Spark)	NA	Spark Component: 2 A4m v2 (4 cores, 32 GB RAM) Head nodes x 730 Hours, 5 A4m v2 (4 cores, 32 GB RAM) Worker nodes x 730 Hours, 5 A4m v2 (4 cores, 32 GB RAM) Edge nodes x 730 Hours, Enterprise Security Package x 730 Hours	Spark Component: 2 E64i v3 (64 cores, 432 GB RAM) Head nodes x 730 Hours, 5 E64i v3 (64 cores, 432 GB RAM) Worker nodes x 730 Hours, 5 A4m v2 (4 cores, 32 GB RAM) Edge nodes x 730 Hours, Enterprise Security Package x 730 Hours
HDInsight(Kafka)	NA	Kafka Component: 2 A4m v2 (4 cores, 32 GB RAM) Head nodes x 730 Hours, 5 E64i v3 (64 cores, 432 GB RAM) Worker nodes x 730 Hours, 3 A5 (2 cores, 14 GB RAM) Zookeeper nodes x 730 Hours, 5 A4m v2 (4 cores, 32 GB RAM) Edge nodes x 730 Hours, 1 Standard disks	Kafka Component: 2 E64i v3 (64 cores, 432 GB RAM) Head nodes x 730 Hours, 5 E64i v3 (64 cores, 432 GB RAM) Worker nodes x 730 Hours, 3 A5 (2 cores, 14 GB RAM) Zookeeper nodes x 730 Hours, 5 A4m v2 (4 cores, 32 GB RAM) Edge nodes x 730 Hours, 2 Premium disks