|  |  |
| --- | --- |
| Title | Architecture Glossary |
| **Owner** | Architecture Board |
| **Version Number** |  |
| **Document Id** |  |
| **Primary Audience** | All staff including self-employed planners and contractors |
| **Document Location** |  |
| **Next Review Date** | This glossary is regularly updated |

|  |  |  |
| --- | --- | --- |
| **Term** | Description | See also |
| AAA | Authentication, Authorization and Auditing |  |
| ACID | Data Transaction Properties: Atomicity, Consistency, Isolation, and Durability | SQL |
| API | Application Programming Interface | ReST, GraphQL |
| Architecture | A high-level abstraction of the system, that defines structure, functions, components and interactions, i.e. how it is put together. |  |
| AVRO | Binary-encoded JSON-based data format, the preferred format for serialising record data for streaming. Also used in Big Data, although uses a row-based storage approach | Parquet, CSV |
| BASE | Basic Availability, Soft State, Eventual Consistency | CAP, ACID |
| BI | Business Intelligence |  |
| Big Data | Too large or complex data to be dealt with using conventional means |  |
| Blueprint | The plan/specification for building the system, i.e. manufacturing/fabricating instructions, e.g. functional and non-functional specifications based on the requirements | Design, Architecture |
| BPM | Business Process Management |  |
| CAP | A design trade-off in distributed system design: Consistency, Availability, Partition tolerance | BASE, ACID |
| CBOR | Concise Binary Object Representation, based on JSON | JSON |
| CCPA | California Consumer Privacy Act | GDPR |
| CDC | Change Data Capture |  |
| CDM | Common Data Model |  |
| CI | Configurable Items, as found in a CMDB | CMDB |
| CICD | Continuous Integration and Continuous Deployment | GIT |
| CMDB | Configuration Management Database | CI |
| CORS | Cross-Origin Resource Sharing |  |
| COTS | Commercial Off the Shelf, i.e. commonly obtainable |  |
| CRUD | The 4 standard data operations: Create, Read, Insert and Delete | SQL |
| CSS | Cascading Style Sheet |  |
| CSV | Comma-delimited String Value row-oriented data file format | Parquet, Iceberg |
| DAD | Data Architecture Document | HLD |
| Data Fabric | A methodology and architectural approach to break down siloed data by typically virtualizing all data and so make it accessible to the right consumers, and to manage its life cycle |  |
| Data Lake | An unstructured store of the enterprise's raw and unprocessed data. | Data Warehouse, Date Lakehouse |
| Data Lakehouse | A store of structured, semi-structured and unstructured enterprise data with easy access to all data, transactioning and indexing. | Data Lake, Data Warehouse |
| Data Mart | A partition of the enterprise's data warehouse that specializes in a particular aspect of the business | Data Warehouse |
| Data Mesh | A methodology to fluidize data access. Not ready for use yet. | Data Fabric |
| Data Pipeline | The route that data follows from create, collect, ingest, store, transform, compute and consumption, in either a streamed or a batched mode. This does not include discrete approaches like event-based integrations (ReST API etc...). | ETL, ELT |
| Data Warehouse | A structured store of the enterprise's data. Structured data can be in a table schema of facts (aggregates) and dimensions (details) for convenient analytics, or in a normalized schema for operational systems. | Data Mart, Data Lake |
| Delta Lake | An additional layer to make Data Lakes ACID compliant, and offers schema enforcement and time travel | ACID, Data Lake |
| DORA | Digital Operational Resilience Act from the EU to drive digital operational resilience in financial organizations. | GRPD, DSAR, CCPA |
| DSAR | Data Subject Access Request | GRPD, CCPA, DORA |
| ESB | Enterprise Service Bus |  |
| ETL | Extract Transform Load data process on a batched basis | ELT, Data Pipeline |
| ELT | The process to Extract Load streamed data and then Transform it | ETL, Data Pipeline |
| Framework | A foundational set of tools, libraries and best practices, i.e. how to build it, as can be shown with e.g. a functional decomposition. |  |
| GAB | Group Architecture Board | Architecture, HLD, DAD |
| GDPR | General Data Protection Regulation | DSAR, CCPA |
| GIT | The de facto code control and versioning system | CICD |
| GraphQL | Hierarchical Graph Query Language | ReST |
| gRPC | Cross-platform, high-performance network protocol |  |
| HIPAA | Data compliance regulation in healthcare | PCI/DSS |
| HLD | High-Level Design document | DAD |
| HMAC | Hash-based Message Authentication Code |  |
| HTML | Hypertext Mark-up Language | CSS, SSL |
| ICEBERG | High-performance data file format for analytics and data lakes | CSV, Parquet, AVRO, ORC |
| Infrastructure | Physical components and resources that constitute the system |  |
| IT | Information Technology |  |
| JSON | JavaScript Object Notation for schema-less data interchange | YAML |
| Least privilege | A security principle that limits user access to the minimum resources and permissions needed to perform their tasks |  |
| LLM | Large Language Model (AI) | RAG |
| MDM | Master Data Management, which creates a single source of truth for business-critical data |  |
| Message Queue | Inter-component asynchronous communication method |  |
| Methodology | A set of methods and principles for building a system, i.e. how to use the tools to build the system, such as Agile and DevOps |  |
| Model | A particular implementation of an architecture, describes the working parts of the system, i.e. how it works, as can be illustrated with e.g. a data model |  |
| MPP | Massively Parallel Processes |  |
| MVP | Minimum Viable Product |  |
| NLP | Natural Language Processing | RAG |
| NoSQL | A schema-less database paradigm that holds semi-structured data such as emails, JSON, XML & HTML and other document files, which is harder to retrieve and manipulate. NoSQL databases scale horizontally. | SQL |
| Observability | The ability to measure the current holistic state of a system by examining the data collected across multiple layers of the system's architecture. |  |
| ORC | An Optimized Row Columnar data file format, used in Big Data | CSV, Iceberg, Parquet, AVRO |
| Parquet | A columnar data file format, used in Big Data, especially nested data | CSV, Iceberg, ORC, AVRO |
| PCI/DSS | Data compliance regulation | HIPAA |
| PII | Personally Identifiable Information | GDPR |
| Pipeline | Data: The route from creation, ingestion, transformation and loading through to consumption |  |
| RAG | Retrieval-Augmented Generation, an architectural approach for improving accuracy of an LLM in AI | LLM |
| RCS | Rich Communication Service, successor to SMS |  |
| RDP | Remote Desktop Protocol, for accessing and viewing Windows desktops |  |
| RDBMS | Relational Database Management System |  |
| ReST | Representational State Transfer | API, GraphQL |
| RCO | Recovery Point Objective | RTO |
| ROPA | Record of Processing Activities: Record of What, Why and How data is processed, and Who owns the data. |  |
| RTO | Recovery Time Objective | RCO |
| SAN | Storage Area Network |  |
| SASE | Secure Access Service Edge, a network architecture that combines VPN and WAN capabilities with cloud-based security. | Zero Trust |
| Scalability | The ability of a system to handle an increased workload without losing performance or compromising availability. |  |
| Schema | The blueprint that defines the logical structure and organization of the data |  |
| Schema-on-Read | The data schema is created and imposed as data is read from a data lake | Data Lake |
| Schema-on-Write | The structured data schema is created before data is ingested into a data store, typically an RDBMS | RDBMS |
| SIEM | Security Information and Event Management |  |
| SQL | Search Query Language, a scripting language used for creating statements that create, read, insert and delete data from structured data databases. | NoSQL, CRUD |
| SSL | Secure Sockets Layer | HTML |
| System | All the related parts and rules that form a unified solution |  |
| TOGAF | The Open Group Architecture Framework |  |
| UML | Unified Modelling Language |  |
| Unstructured Data | Schema-less, format-less data that cannot easily be queried, manipulated or analysed and is stored in its native format, e.g. documents, audio, video, social media posts. Requires specialist tools. |  |
| Vector Search | NLP technique to find nearest matching strings, images, audio, video, and other unstructured data sets | NLP |
| YAML | A human-readable (as opposed to JSON) data serialization language | JSON |
| Zero Trust | A Security framework that assumes that every interface can be a hostile threat | SASE |
| XML | Extensible Mark-up Language, a data serialization language |  |