

# SPEC SHEET • ECCENCA CORPORATE MEMORY 26.1

## Functional Architecture

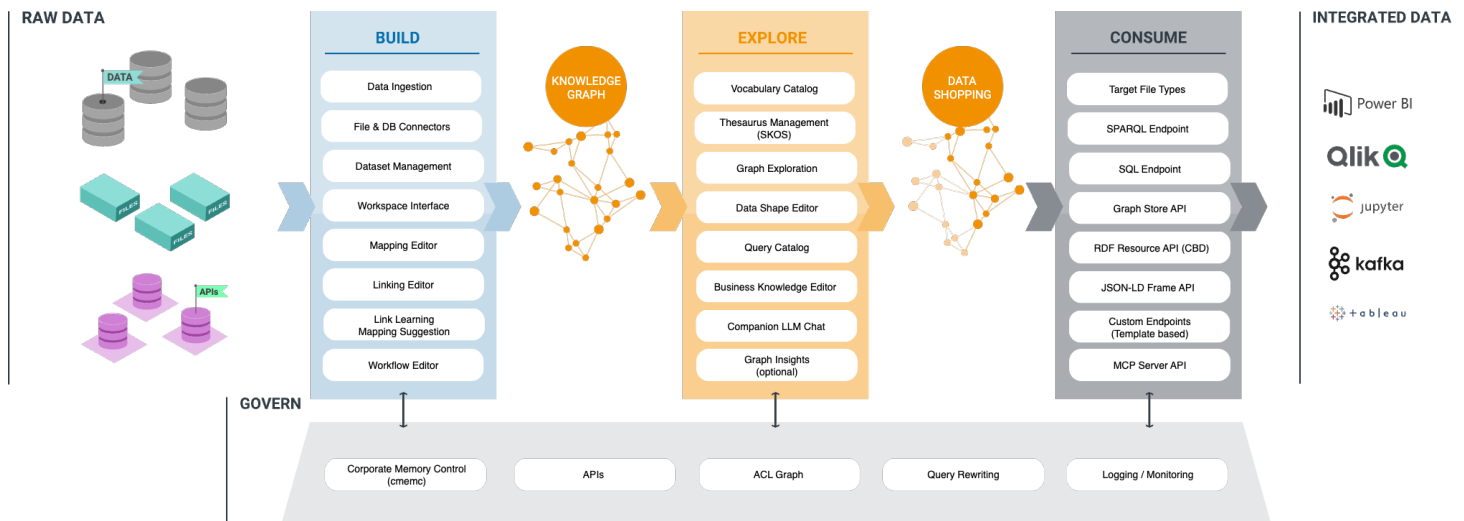


Figure 1. Functional Architecture

## Build

### Workflows

- Compose workflows and workflows-of-workflows based on mapping, linking and other operators
- Python SDK to build custom workflows operators
- Turn workflows into APIs by visual configuration, and use them via REST API or `cmemc workflow io` command
- Project individual permissions (group assignments)

### Data Linking

- Automatic and manual rule learning with integrated editor (drag & drop)
- Engineering roundtrip allowing refinement and validation of learned rules
- The optional Link-Rules module supports management and approval workflows for rules
- Use custom built transformation operators

## Capabilities through our Python Plugins Image

- Authenticate to services and provide the OAuth2 access token for other tasks.
- Combine CSV files with the same structure into one dataset.
- Convert prices between currencies.
- Update your Graph Insights snapshots within Corporate Memory.
- Send queries to a GraphQL endpoint and save the results in a JSON dataset.
- Create unique International Registration Data Identifier (IRDI).
- Query and extract data from Jira with JQL.

### Data Mapping

- Hierarchical mappings
- Mapping suggestions and automatic generation of mappings
- Value translations from inline or excel based mapping tables
- Table and query based JDBC access
- Visual transformation rule builder
- Use custom built transformation operators
- Mapping Creator for visual drag & drop and LLM-assisted mapping generation
- Target schema guided mapping

### Data Manipulation

- Execution of extraction, linkage, cleansing and integration rules on arbitrary datasets
- Python SDK to build custom transformation operators
- Support for different writing strategies
- REST APIs for synchronous or asynchronous executions
- Dataset generation via SPARQL queries
- Export of file-based datasets as e-mail attachments
- XSLT operations on XML documents

- Process data structures with the jq expression language.
- Send and receive messages from Apache Kafka.
- Start processes on pods on your Kubernetes cluster.
- Interact with Large Language Models (Execute Instructions, Create Embeddings).
- Retrieve data from a Logpoint Security information and event management (SIEM) system.
- Loop over the output of a task and start a sub-workflow for each entity.
- Send messages to Mattermost channels and users.

- Extract and process data from your Nextcloud instance.
- Convert numbers between binary, octal, decimal and hexadecimal representation.
- Extract and process data from your Office365 instance.
- Set or overwrite parameters of a task.
- Extract text and tables from PDF documents.
- Store and search for embedding vectors in a Postgres vector store.
- List and upload project resources.
- Validate your Knowledge Graphs based on tests generated from SHACL shapes.
- Perform reasoning tasks and validate OWL consistency.
- Send or receive data from your organization's Salesforce account.
- Generate SHACL shapes from an instance data graph.
- Split a text file into parts of the same filesize
- Get files and process output via secure shell protocol.
- Create Universally Unique Lexicographically Sortable Identifiers (ULID) in transformations.
- Create universally unique identifiers (UUIDs) versions 1, 3, 4, 5, 6, 7 and 8 in transformations.
- Validate graph and data structures.
- Get workflow reports and status information.
- Load and parse YAML documents.

## Explore

### Business Knowledge Editor

- Visually explore and edit knowledge graphs
- Save, load and share explorations
- Fine-tune the representation with SHACL Shapes
- Complex individual search filter settings

### Thesaurus Module

- Thesaurus project catalog to create, remove and manage metadata of thesaurus projects
- Thesaurus editor to create, remove and manage metadata of concept schemes together with concept trees and concept metadata

### Companion LLM Chat

- Conversational AI interface
- Ask questions about your graphs
- 10+ tools available to the internal chat and via MCP server API

### Charts Catalog

- Use Apache eCharts to visualize data
- Create your data views in the query catalog and configure a chart
- Support for assisted form-based chart definitions as well as an advanced mode that allows full customization of the eCharts JSON declarations

### SPARQL Query User Interface

- Save and reuse queries as well as add metadata
- Use placeholders in queries
- Preview and export query results
- Use multiple query catalog graphs

### Exploration Module

- Tree, list and resource views
- Management of knowledge bases
- Visualization of ontologies / vocabularies and instance level data
- Inline authoring and detailed edit view
- Add new resources and properties
- Custom authoring interfaces based on SHACL Shapes
- Download result tables in different formats (Excel, CSV, JSON)

### Workspace Configuration

- Create, remove and manage workspace configurations in the GUI
- Inherit from other configurations or start from scratch
- Workspace configuration is represented in RDF and can be exported and imported
- Query based badges for visual annotations

### Graph Insights (optional)

- Visually explore your graphs' contents and relationships in an aggregated representation
- Flexible aggregated class and resource overview, details on demand

### Further Functions and Interfaces

- Vocabulary management
- User management and access control
- Search and facet based filtering

## Consume

### Data Discovery, Exploration and Visualization

- JSON-LD Framing with SPARQL 1.1 construct queries
- On-the-fly log level configuration
- SPARQL 1.1 query endpoint with server-side optimization (pagination, sorting, variable expansion, filtering, aggregation, triple-store optimized search)
- SPARQL endpoint service description
- Flexible, template-based custom APIs

### PowerBI Connector

- Custom data connector to consume data from the knowledge graph directly from Microsoft PowerBI Desktop

### MCP Server

- Integrate with 10+ tools via our MCP server API

## Govern

### Corporate Memory Control CLI (cmemc)

Various command groups to manage and automate eccenca Corporate Memory:

- Admin – Import bootstrap data, backup/restore workspace or get status.
- Config – List and edit configs as well as get config values.
- Dataset – List, create, delete, inspect, up-/download or open datasets.
- Graph – List, import, export, delete, count, tree or open graphs.
- Project – List, import, export, create, delete or open projects.
- Query – List, execute, get status or open SPARQL queries.
- Vocabulary – List, (un-)install, import or open vocabs / manage cache.
- Workflow – List, execute, status or open (io) workflows.

### Access Control Management

- Based on RDF access control model (graph-based, context-based)
- Authentication via keycloak (OAuth2 protocol, OpenID Connect, SAML 2.0), refer to <https://www.keycloak.org/>
- Integration with external user management systems (e.g. LDAP, Active Directory)

## eccenca Corporate Memory Core Components

### eccenca Build

Component for integration of multiple databases, provides the eccenca Build capabilities.

### eccenca Explore

Single-page web application for administration of knowledge graphs, provides the eccenca Explore capabilities as well as semantic middleware for data access management, provides the eccenca Consume capabilities.

### eccenca Graph Insights (optional)

Visual graph exploration and analytics component, provides the eccenca Explore Graph Insights capabilities.

### Corporate Memory Control

Command line interface `cmemc` for the automation and integration of eccenca Corporate Memory.

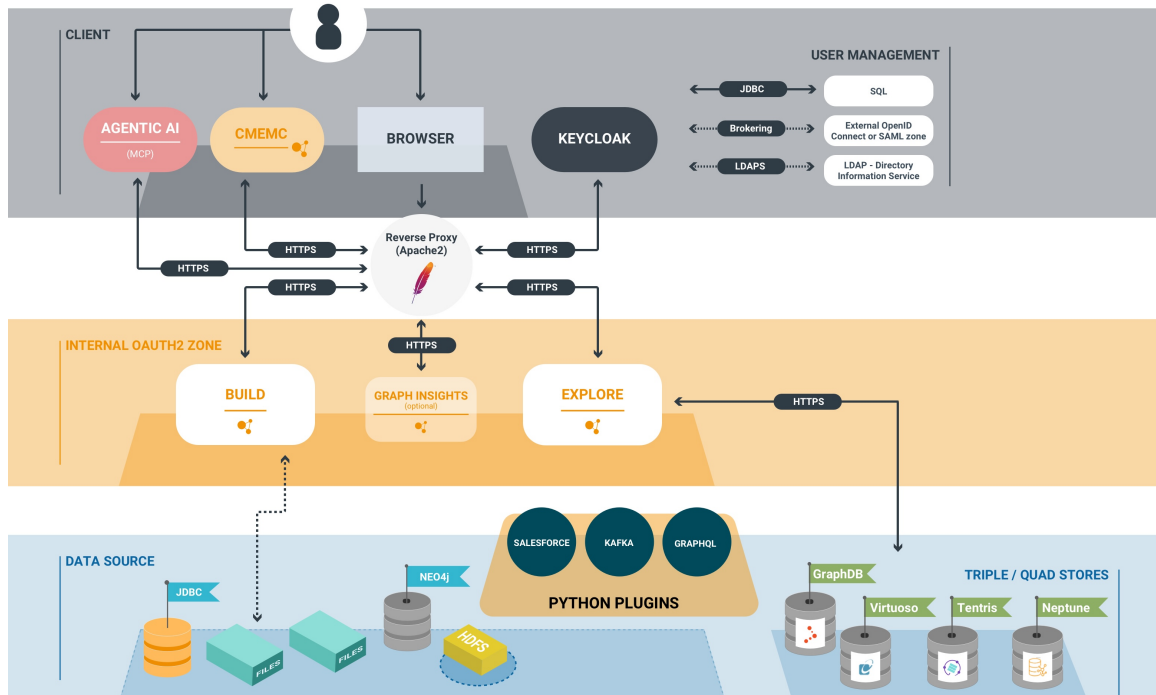


Figure 2. System Architecture

## Standards and Dataset Support

### Supported Dataset File Types

- .txt, .csv, .xlsx, .json, .jsonl (JSON-Lines), .yaml, .xml, .rdf, .orc, .avro, .parquet (all also zip-compressed)
- multi-csv-zip
- binary files support (e.g. images, pdf, etc.)

### Supported Dataset Endpoints / APIs

- SPARQL, SQL, JDBC, Hive, MariaDB, Microsoft SQL Server, MySQL, Neo4j, PostgreSQL, SQLite, Snowflake, Trino, pgvector
- Excel on Google Drive
- Excel on Microsoft OneDrive (Office 365/SharePoint)
- REST APIs, MCP Server API

### Supported Standards

- RDF 1.1 Turtle, XML Syntax, N-Quads, N-Triples
- OAEI Alignment Format
- JavaScript Object Notation (JSON)
- JSON-LD 1.0
- Extensible Markup Language (XML) 1.1
- Hypertext Transfer Protocol – HTTP/1.1
- OAuth 2.0 Authorization Framework
- SPARQL 1.1 Query Language, Federated Query, Update, Query Results JSON Format, Graph Store HTTP Protocol
- SPARQL Query Results XML

## Deployment Options

- Deployment via Docker or Kubernetes
- Various roll-out options (single instance to multiple instances, HA options)
- Helm charts and Docker Compose configuration templates

## Further Information

- Refer to <https://documentation.eccenca.com/26.1/deploy-and-configure/requirements/> for system requirement details.
- Refer to <https://documentation.eccenca.com/26.1/deploy-and-configure/installation/> for more details on deployment and sample configurations.