

ALMATO just add  
digital

**Almato Bardioc**  
Semantic data platform

Release 7

# Inhalt

<b>1.</b>	<b>Einleitung</b>	<b>3</b>	<b>4.</b>	<b>Ontologie</b>	<b>20</b>
			4.1	Konzepte und Eigenschaften der OGIT Ontology	21
<b>2.</b>	<b>Übersicht</b>	<b>4</b>	4.2	OGIT Ontology Extensions	22
2.1	Herausforderungen	5			
2.2	Betriebssystem für Daten	5	<b>5.</b>	<b>Systemanforderungen und Betrieb</b>	<b>23</b>
2.3	Komponenten	6	5.1	Software as a Service / Platform as a Service	24
2.4	Funktionale Architektur	6	5.2	Hyperscaler	24
2.5	Digitale Souveränität	7	5.3	On-Premise	24
			5.4	System- und Integrationsanforderungen	25
<b>3.</b>	<b>Almato Bardioc</b>	<b>8</b>	5.5	Open Source	25
3.1	Access Manager	9	5.6	Zertifizierbarkeit	25
3.1.1	Datenschutz und Datensicherheit	9			
3.1.2	Identity and Access Management	10	<b>6.</b>	<b>Integrations- und Erweiterungsmöglichkeiten</b>	<b>26</b>
3.1.3	Access Token-Konzept	10	6.1	Knowledge Automation	27
3.2	Data Manager	11	6.2	Large Language Models / Retrieval-Augmented Generation	27
3.2.1	Scheduling	11			
3.2.2	Memory Management	11	<b>7.</b>	<b>Zusammenfassung</b>	<b>29</b>
3.3	Message Bus	12	7.1	Vergleich mit Graph Datenbanken	30
3.4	Knowledge Core	13	7.2	Alleinstellungsmerkmale von Almato Bardioc	31
3.4.1	Storage-Technologien	13			
3.4.2	Skalierbarkeit	14			
3.4.3	Historisierung	14			
3.5	Knowledge APIs	15			
3.5.1	APIs für fachliche Funktionen	15			
3.5.2	APIs für Echt- und Nahzeitanwendungen	15			
3.5.3	Connector SDK	16			
3.5.4	Funktionale Datenabfragen	16			
3.6	Desktop	17			
3.7	Graph Applications	18			

# 1. Einleitung

In der heutigen digitalen Ära erweist sich eine semantische Datenplattform als unverzichtbares Werkzeug, um komplexe Datenlandschaften zu verstehen und ihr volles Potenzial zu erschließen.

Almato Bardioc ist eine semantische Datenintegrations- und Verarbeitungsplattform, die über einen Zeitraum von 20 Jahren in Deutschland erforscht, konzipiert und entwickelt wurde. Sie integriert zahlreiche Open-Source-Technologien und vereint diese mit eigener Technologie zu einem vollständigen Betriebssystem für Daten. Die Plattform ist darauf ausgelegt, sehr große Datenmengen aus unterschiedlichen Quellen zu integrieren, zu verarbeiten und zu verwalten. Dabei werden komplexe Beziehungen über die Bedeutung und den Kontext der Daten gespeichert, sodass die Daten unabhängig von ihrer Quelle oder primären Verarbeitung ihre volle Bedeutung innerhalb der Plattform bewahren.

Bardioc stellt alle notwendigen Komponenten bereit, um skalierbare und revisionssichere Anwendungen zur Datenintegration, Datenverarbeitung, Datenanalyse sowie Datenvisualisierung auf Grundlage eines Datenpools zu entwickeln. Durch die bereits vorhandenen Anwendungen zur semantischen Datenanalyse in Echtzeit, kombiniert mit verschiedenen KI-Technologien, werden Unternehmen und Organisationen befähigt, fundierte Entscheidungen zu treffen und den vollständigen Wert ihrer Daten auszuschöpfen.

Bardioc bietet einzigartige Funktionen zur effizienten und wirtschaftlichen Skalierung für sehr große Datenmengen (mehrere hundert Petabyte) bei gleichzeitig hoher Abfrage- und Verarbeitungsgeschwindigkeit. Die effiziente Ausführung verteilter Algorithmen sowie die Zusammenführung von Ergebnissen werden vollständig automatisch von der Plattform gesteuert. Durch ausgeklügelte und geprüfte Sicherheitsmechanismen können granulare Datensicherheits- und Datenschutzkonzepte umgesetzt sowie höchste Compliance-Anforderungen erfüllt werden.

Bardioc eignet sich als Ontologie-basiertes Betriebssystem für Unternehmen und Organisationen aus Industrie und Verwaltung, um daten- und KI-gesteuerte Entscheidungen von hoher Relevanz zu treffen, die operative Effizienz auf ein neues Niveau zu heben oder schwer nachzuahmende Wettbewerbsvorteile zu erzielen. Durch die Abbildung von Wissen und Information innerhalb eines definierten Kontexts werden zudem die Agilität und Innovationsfähigkeit von Unternehmen und Organisationen unterstützt.

Bardioc ist das zentrale Betriebssystem für die Verwaltung und Speicherung sehr großer Datenmengen. Grundlage hierfür ist ein skalierbarer und per APIs zugänglicher semantischer Graph.

The background is a dark blue gradient with a network of glowing white and light blue nodes connected by thin lines, creating a sense of connectivity and data flow. The nodes are scattered across the frame, with some appearing brighter and more prominent than others. The overall aesthetic is modern and technological.

## **2. Übersicht**

## 2. Übersicht

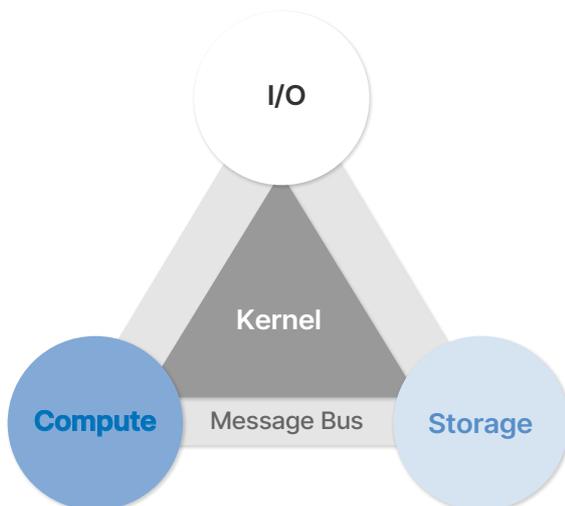
### 2.1 Herausforderungen

Viele Unternehmen und Organisationen arbeiten in komplexen Ökosystemen, in denen Daten über zahlreiche Plattformen, Abteilungen oder sogar externe Partner generiert und gespeichert werden. Bardioc ermöglicht die zentrale Integration verschiedener Datenquellen und bricht somit die herkömmlichen zweckgebundenen Datensilos auf. Erst dadurch kann ein umfassendes Verständnis der Datenlandschaft geschaffen werden, die wiederum Basis für eine funktionsübergreifende Zusammenarbeit ist. Dadurch können Unternehmen und Organisationen ganzheitliche und fundierte Entscheidungen treffen und kritische Herausforderungen bewältigen.

### 2.2 Betriebssystem für Daten

Funktional entspricht Bardioc einem Betriebssystem für Daten. Den Kern stellen Programme dar, die auf Daten zugreifen, Berechnungen durchführen sowie mit anderen Programmen kommunizieren. Das Äquivalent zum Dateisystem ist der Knowledge Core, bestehend aus Graph Datenbank, BLOB-Speicher und Zeitreihen-Speicher. Das Äquivalent zum Rechenkern sind Indizierung, Abfragen sowie die Möglichkeit, Code in Form von Programmen auf der Plattform auszuführen. Das Äquivalent zur Kommunikationsschicht ist der Message Bus, über den Komponenten und Programme untereinander sowie über verteilte Systeme hinweg kommunizieren können.

Ähnlich einem herkömmlichen Betriebssystem werden Zugriffsschutz, Programm- und Datenintegrität durch ein horizontal und granular implementiertes Berechtigungs- und Identitätssystem gewährleistet. Dieses Sicherheitssystem stellt darüber hinaus die Verschlüsselung von Daten zu jedem Zeitpunkt der Übertragung und Speicherung sowie die Verwaltung und Erkennung erlaubter Applikationen und Benutzer sicher.



Fundamentale Architektur von Betriebssystemen

#### Features:

- + Betriebssystem für Daten
- + In-Memory-Verarbeitung für hohe Performance
- + Kostenoptimierter Betrieb durch gezielte Speicherzuweisung

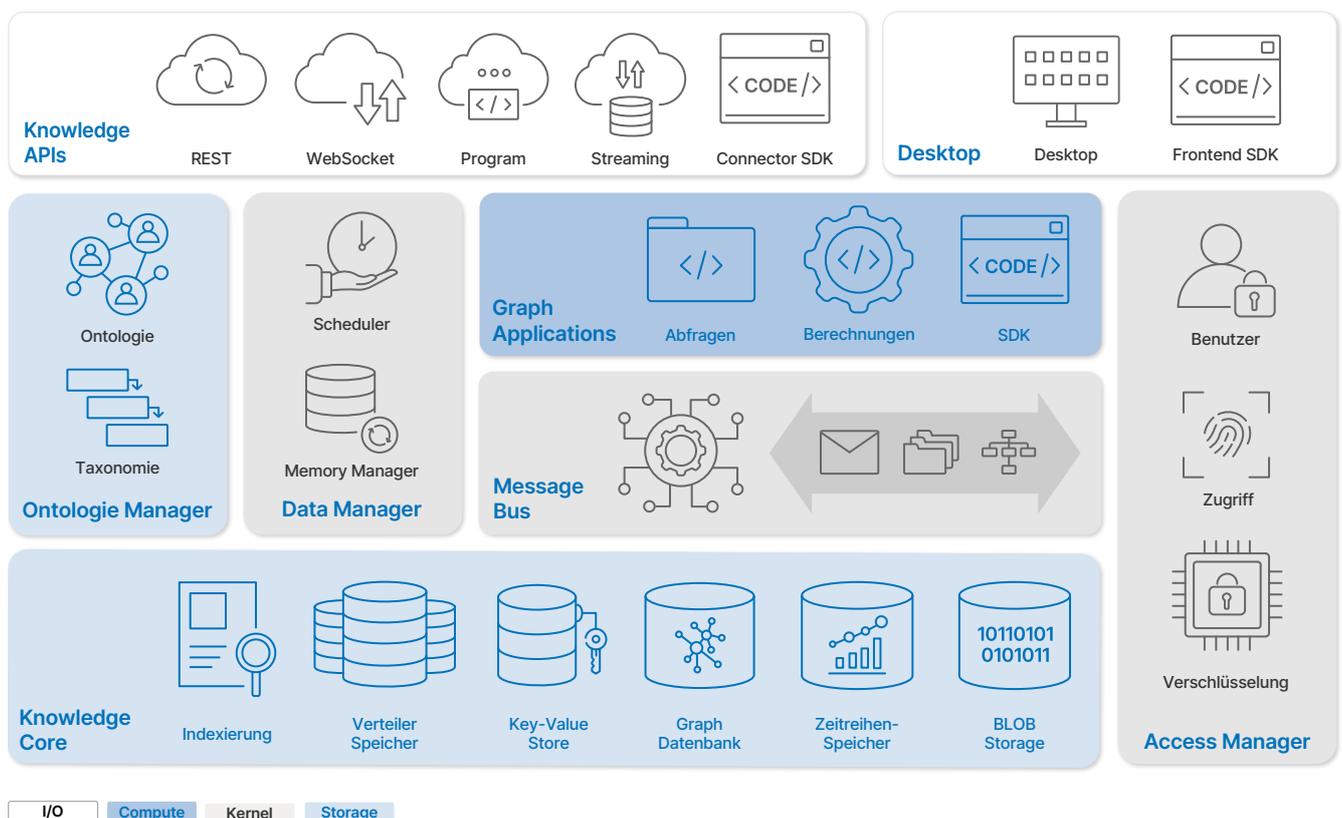
## 2.3 Komponenten

Bardioc besteht aus folgenden Produktbestandteilen und Komponenten:

- **Access Manager** als zentraler Access Control- und Security-Layer
- **Data Manager** übernimmt Scheduling- und Caching-Aufgaben und sorgt für eine kosteneffiziente Speicherverwaltung
- **Message Bus** integriert Storage- und Datenbanktechnologien und sorgt für einen hohen Datendurchsatz
- **Knowledge Core** als redundante und verteilte Datenhaltung
- **Knowledge APIs** als zentrales Connectivity Framework für den Zugriff auf die Daten
- **Desktop** als Framework zur Frontend-Entwicklung
- **Graph Applications** als individuelle Programme und Anwendungen auf Basis des Knowledge Graphen
- **OGIT Ontology** bildet den semantischen Kontext der Daten
- **OGIT Ontology Extensions** als kundenindividuelle Erweiterungen der Ontologie

## 2.4 Funktionale Architektur

Entsprechend der skizzierten fundamentalen Architektur für Betriebssysteme ist Bardioc als Betriebssystem für Daten konzipiert und weist folgende funktionale Architektur auf:



Funktionale Architektur von Bardioc

## 2.5 Digitale Souveränität

Zur Unterstützung der digitalen Souveränität im Rahmen der öffentlichen Verwaltung, aber auch in der Industrie, setzt Bardioc umfangreich auf Open Source Software (OSS). Der von Almato konzipierte Architekturansatz bietet ein Höchstmaß an Unabhängigkeit und Flexibilität. Durch die modulare Architektur sowie den Fokus auf Standards werden die Entwicklung von Programmen für die Plattform sowie der Betrieb der Plattform optimal durch das Know-how der Communities unterstützt.

Gleichzeitig wird mit Bardioc eine kommerzielle Lösung angeboten, die in ihrer innovativen Funktionsweise einzigartig am Markt ist. Sie ermöglicht Unternehmen und Organisationen, schrittweise eine semantische Datenintegrations- und Verarbeitungsplattform aufzubauen, die umfangreiche Optionen zur Erweiterung und Individualisierung bietet.



### **3. Almato Bardioc**

## 3. Almato Bardioc

### 3.1 Access Manager



Der Access Manager ist für Datensicherheit und Datenschutz innerhalb Bardioc zuständig und sorgt mit integriertem Identity and Access Management (IAM) sowie umfangreichen Verschlüsselungsmaßnahmen für die Einhaltung von Compliance-Anforderungen.

#### 3.1.1 Datenschutz und Datensicherheit

Datenquellen können persönliche oder sensible Informationen enthalten, deren Schutz gewährleistet werden muss. Die Implementierung von Datenschutz- und Sicherheitsmaßnahmen entsprechend rechtlichen und regulatorischen Anforderungen ist maßgebend, um Vertraulichkeit und Integrität der Daten im Knowledge Graphen zu gewährleisten. Dies betrifft besonders schützenswerte und sensible Projektdaten, die z. B. als Verschlusssache klassifiziert sind, der Geheimhaltung unterliegen oder nach dem Need-to-know-Prinzip verarbeitet werden müssen.

#### Datenschutz:

- + Software made in Germany
- + SaaS in Deutschem Rechenzentrum
- + Technologie für On-Premise-Betrieb verfügbar

- + **SSL/TLS** sowie **Authentifizierung** für verschlüsselte Datenübertragungen.
- + **Encryption at Rest** (Festplattenverschlüsselung) auf Basis AES256 zum Schutz ruhender Daten auf physischen Datenträgern.
- + **Audit-Logs** mit digitalen Signaturen zur Unterstützung regulatorischer Anforderungen (permanente Protokollierung aller Aktionen, nicht deaktivierbar, nicht löscherbar, nicht manipulierbar).

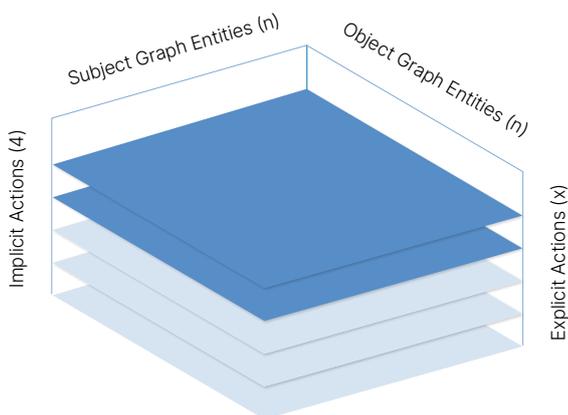
### 3.1.2 Identity and Access Management

Benutzern, Entwicklern sowie Anwendungen werden Identitäten zugewiesen, mit denen ihre Berechtigungen auf alle Daten und Ressourcen innerhalb der Plattform bestimmt werden und womit ein detailliertes Zugriffsmanagement gewährleistet wird. Die vom Access Manager zugewiesene Identität kann entweder ein OAuth-fähiger Account auf der Plattform sein oder sich zu bestehenden Single-Sign-on-Lösungen wie z. B. OneLogin oder internen Active-Directory-Lösungen verbinden.

- + **Identity and Access Management (IAM)** zur Bereitstellung von Accounts (Knowledge Core Accounts) auf Basis OAuth sowie zur Verwaltung und Kontrolle von Rollen und Rechten beim Zugriff auf Daten und Anwendungen innerhalb von Bardioc.
- + **Federation** mit anderen, unabhängigen IAM-Systemen oder X.509-fähigen Services zur Nutzer-Authentifizierung bei vertrauenswürdigen Identity Providern oder zum Zugriff auf fremde Ressourcen und Services per Single Sign-On via SAML.
- + **RBAC (Role-based Access Control)** sowie Implementierung eines **Need-to-know-Prinzips** für geschützten Datenzugriff auf Anwendungsebene.

### 3.1.3 Access Token-Konzept

Der Access Manager verfügt über ein ausgeklügeltes Access Token-Konzept. Alle Benutzer (Knowledge Core Accounts sowie über Federation verbundene Accounts) sowie alle Graph Applications werden mit Software-Tokens versehen. Ein Zugriff auf Daten im Knowledge Graphen ist nur in Kombination von Benutzer- und Application-Token möglich. Ein unberechtigter Zugriff von Benutzern auf den zentralen Datenpool ist somit nicht möglich.



Security Mesh Layer

Der Security Mesh Layer sorgt durch maximal effiziente Bit-Operationen für die Überprüfung der Berechtigung auf Einzeldaten, sodass diese ohne Beeinträchtigungen der Zugriffsgeschwindigkeit immer für jedes Datenelement durchgeführt werden können. Berechtigungsänderungen werden somit umgehend im Graphen wirksam, da diese nicht gecached werden. Dies ist insbesondere bei sensiblen Daten und Verschlusssachen von Bedeutung.

## 3.2 Data Manager



Der Data Manager ist für Scheduling und Caching innerhalb Bardioc zuständig.

### 3.2.1 Scheduling

Die Scheduler-Funktion analysiert alle Zugriffe auf Daten und berechnet bei verfügbaren Leerlaufzeiten der vorhandenen Infrastruktur die wahrscheinliche nächste Abfrage (**Preemption**) bereits vor, sodass die Berechnung deutlich schneller erfolgt als bei sequenzieller Bearbeitung. Durch Analyse der Daten und Zuweisung der geeigneten Speichermedien trägt der Scheduler zudem zur Kostenoptimierung beim Betrieb der Plattform bei.

### 3.2.2 Memory Management

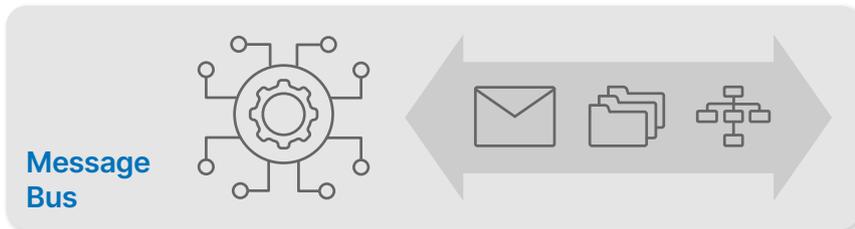
Das Memory Management des Knowledge Cores sorgt dafür, dass die Daten auf den für ihre Zugriffsmuster geeigneten Speichertechnologien gehalten werden. Veränderungen im Zugriffsmuster spiegeln sich somit in der Datenhaltung wider. Daten können im Speicher, auf schnellen Block Devices, langsamen Block Devices oder anderen Speichermedien gehalten werden.

Ein ausgeklügeltes System von Caches stellt sicher, dass die Verarbeitung soweit wie möglich im Arbeitsspeicher erfolgt, womit die Performance weiter optimiert wird. Durch die Implementierung von Copy-On-Write wird eine schnelle Initialisierung von Anwendungen, auch wenn sie große Datenmengen laden, ermöglicht. Diese Funktion sorgt auch dafür, dass Daten erst dann erneut geschrieben werden, wenn sie sich verändern. Eine De-Duplizierung zur Vermeidung unnötiger Redundanzen ist damit integraler Bestandteil des Systems.

Darüber hinaus stellt der Data Manager Funktionen bereit, die eine garantierte Echt- oder Nahzeitverarbeitung ermöglichen. Konventionell propagieren sich Nachrichten und Verarbeitungsschritte durch das System, wodurch ein Zeitverzug entsteht (eventual consistency). Durch die Skalierung und Priorisierung auf Abfrageebene lässt sich diese Latenz weitestgehend eliminieren.

## 3.3 Message Bus

Innerhalb Bardioc kommt eine massiv performante Message Streaming Platform zum Einsatz. Der Message Bus integriert die verschiedenen Storage- und Datenbanktechnologien von Bardioc und sorgt für einen hohen Datendurchsatz sowie niedrige Latenzen bei der Verarbeitung von Echtzeitdaten.



Der Message Bus von Bardioc stellt Schnittstellen und Connectoren (Producer, Consumer, Streams, Connect und Admin) zur Verfügung, die die nachfolgenden vier Kommunikationsfunktionen ermöglichen:

- + Kommunikation auf Systemebene zwischen den Komponenten des Datenbetriebssystems (Storage, I/O und Compute).
- + Kommunikation zwischen Graph Applications und dem Knowledge Core.
- + Kommunikation zwischen den Graph Applications untereinander (inklusive Chain of Responsibility, was den Betrieb gleicher Applikationen in unterschiedlichen Versionen ermöglicht).
- + Abonnement von Daten durch Graph Applications.

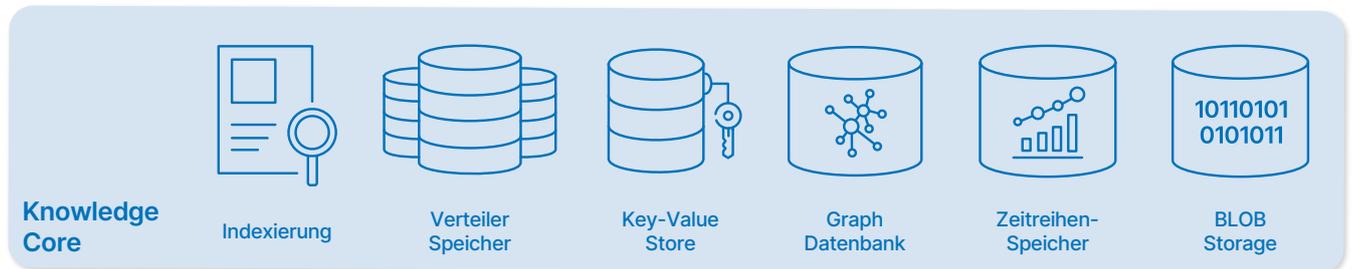
Jegliche Kommunikation innerhalb Bardioc wird somit über den Message Bus entsprechend den Berechtigungen des Access Managers geroutet. Darüber hinaus ergeben sich Vorteile durch den Message Bus in Bezug auf die Verarbeitung von Datenströmen im Graphen:

- + Eine **dynamische Datenaktualisierung** hilft bei der Integration von Datenquellen in den Daten-Stream, wodurch der Graph kontinuierlich und flexibel mit neuen Daten angereichert werden kann – auch in Echtzeit.
- + **Distribution der Datenänderungen** in die verschiedenen Datenbanken von Bardioc, sodass Datenänderungen zeitnah im Graphen reflektiert werden.
- + Durch die **ereignisgesteuerte Architektur** kann der Graph so gestaltet werden, dass er auf spezifische Ereignisse in den Daten-Streams reagiert, was dynamische und anpassungsfähige Datenverarbeitungs-Workflows ermöglicht.
- + Die **Orchestrierung** der Datenströme erleichtert die Integration und Transformation von Daten aus verschiedenen Quellsystemen.

Der Einsatz eines Message Backbones unterstützt zusätzlich die Skalierungsmöglichkeiten von Bardioc durch Parallelisierung und Lastverteilung.

## 3.4 Knowledge Core

Der Knowledge Core ist die eigentliche Datenhaltung, in der Graphdaten, Zeitreihendaten sowie BLOBs redundant und verteilt gespeichert, indiziert und für den effizienten Zugriff gecached werden.



### 3.4.1 Storage-Technologien

Der Knowledge Core nutzt und kombiniert nachfolgende Storage-Technologien, um sehr hohe Anforderungen an Skalierbarkeit, Performance sowie effiziente Ressourcennutzung sicherzustellen:

- + Skalierbare **Graph Datenbank** zur Speicherung und Abfrage von semantischen Graphen mit hunderterten Milliarden Knoten und Kanten.
- + Skalierbare und hoch verfügbare **NoSQL Datenbank** zur verteilten Datenspeicherung.
- + **Zeitreihen-Speicher** (TSDB) mit speziell optimiertem Datenmodell für hohe Abfragegeschwindigkeit durch In-Memory-Fähigkeit sowie mit integrierter Statistik-Funktion.
- + Skalierbarer und hoch verfügbarer **BLOB Storage** (Binary Large Object) zur Speicherung von unstrukturierten Daten und Objekten (Fotos, Videos, Container etc.).
- + Skalierbarer und hoch verfügbarer **Key-Value Store** mit hoher Abfragegeschwindigkeit durch In-Memory-Fähigkeit.
- + Verteilte und mandantenfähige **Indexierung**.

Eine besondere Ausprägung von Zeitreihen sind Log Files. Der Knowledge Core unterstützt hier mehrere Log-Einträge zum exakt gleichen Zeitpunkt sowie eine zeitliche Normierung von Log-Einträgen (z. B. auf definierte Zeitintervalle) für statistische Auswertungen.

### 3.4.2 Skalierbarkeit

Mit zunehmender Anzahl an Datenobjekten und Versionen steigt der Ressourcen-Bedarf (Speicher, CPU und Bandbreite) der Plattform an. Um die Skalierbarkeit und Datenintegrität des Systems zu gewährleisten, sind effiziente Speicherlösungen und -strategien erforderlich. Die lineare Skalierbarkeit von Bardioc auch für sehr große Datenmengen im Bereich von mehreren hundert Petabyte sowie hohe Nutzerzahlen werden unter anderem durch folgende Mechanismen erreicht:

- + **Distributed Nodes** mit Verteilung der Daten auf mehreren Knoten innerhalb von Clustern zur Vermeidung von Single Point of Failures.
- + Unterstützung von **Replikation** und **verteilten Rechenzentren** zur Umsetzung von Failover- und Disaster-Recovery-Strategien.
- + **Lineare Skalierung** von Lese- und Schreibdurchsatz durch Erweiterung von Ressourcen ohne Downtime oder Anwendungsunterbrechung.
- + **Sharding- und Partitioning-Algorithmen** zur horizontalen und vertikalen Partitionierung von Daten sowie zur besseren Lastverteilung bei sehr großen Datensätzen.
- + Optimierte **Indexmechanismen** für kurze Antwortzeiten bei komplexen Datenabfragen.
- + Einsatz von **MapReduce** als Programmiermodell zur Verarbeitung und Erzeugung großer Datensätze auf einem Cluster.

Mit Bardioc können umfangreiche Datenanalysen automatisch innerhalb des semantischen Graphen verteilt werden. Dadurch wird verhindert, dass alle abgefragten Daten zunächst an das Analyseprogramm übertragen werden.

### 3.4.3 Historisierung

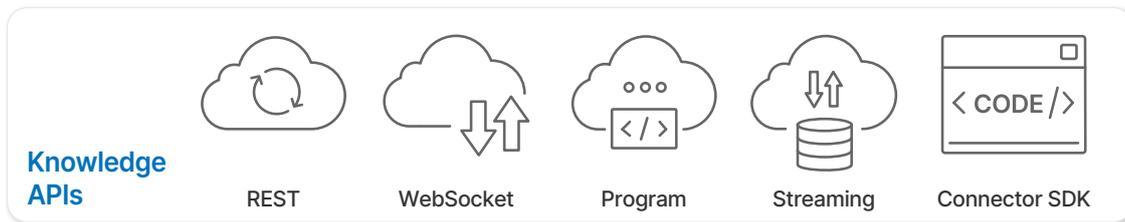
Eine semantische Datenplattform, die historische Datenobjekte effektiv verwalten soll, benötigt Mechanismen, die Abfragen im gesamten Kontext zu einem bestimmten historischen Zeitpunkt des Graphen ermöglichen. Datenreihen mit unterschiedlicher Gültigkeit oder zeitlicher Einordnung lassen sich mit Bardioc komfortabel verwalten und abfragen. Hierzu wurde eine Historisierungsfunktion der Daten entwickelt, die folgende Eigenschaften bietet:

- + Unterstützung der **zeitlichen Abfragefähigkeit im Gesamtkontext** des Graphen durch Zeitstempel und Versionierungskonzepte.
- + **Versionierung** von Objekten (Knoten, Kanten, Eigenschaften) im Graphen.
- + **Nachvollziehbarkeit** der Veränderung von Datenobjekten im Zeitverlauf zur Unterstützung von Regulierungsanforderungen.

Historisierung ist eine Grundvoraussetzung für Compliance, Revisionssicherheit sowie Erklärbarkeit und damit Voraussetzung für die Nutzung einer Datenplattform in jeglichem regulierten Umfeld.

## 3.5 Knowledge APIs

Mit den Knowledge APIs steht ein umfangreiches Connectivity Framework zur Verfügung, um Daten zu lesen, zu suchen, zu schreiben oder Berechnungen auf den Daten durchzuführen.



### 3.5.1 APIs für fachliche Funktionen

Für eine optimale Effizienz können fachliche Funktionen über zwei verschiedene technische APIs aufgerufen werden:

- + **REST API:** Die REST API stellt eine Schnittstelle bereit, die ohne Zustand (stateless) agiert und aufgerufen werden kann. Jeder einzelne Aufruf steht für sich allein, wird separat identifiziert und kann genau einen Aufruf auf funktionaler Ebene ausführen.
- + **WebSocket API:** Die WebSocket API stellt eine Schnittstelle bereit, die eine offene Verbindung zwischen einer Anwendung und der Plattform erzeugt. Die Verbindung wird zu Beginn autorisiert und unterliegt regelmäßigen Erneuerungen der Verschlüsselungstokens. Je Session sind grundsätzlich mehrere verschiedene funktionale Aufrufe der Plattform möglich.

### 3.5.2 APIs für Echt- und Nahzeitanwendungen

Darüber hinaus verfügt Bardioc über zwei Zugriffsmethoden, die speziell für die Entwicklung von Echtzeitanwendungen entwickelt wurden:

- + **Streaming API:** Mit der Streaming API wird analog zur WebSocket API eine dauerhafte Verbindung zum System hergestellt. Allerdings werden Abfragen im System registriert und das System führt diese ausschließlich unter bestimmten Konditionen durch. So können z. B. Veränderungen in den Daten an eine Anwendung automatisch weitergeleitet werden oder bei bestehenden Konditionen automatisch Berechnungen durchgeführt werden.
- + **Program API:** Mit der Program API können Programme, die auf Daten im Knowledge Graphen zugreifen, in Binär-Form (z. B. als .jar) an das System übergeben werden. Durch eine Analyse des Datenzugriffs der Programme können die Programme zu denjenigen Ressourcen der Plattform transportiert werden, auf denen die entsprechenden Daten zu finden sind. So können mit sehr kurzen Zugriffszeiten Berechnungen auf lokalen Daten ausgeführt werden, die dann über die Plattform automatisch wieder zusammengesetzt werden. Damit lassen sich Berechnungen auf sehr großen Datenbasen in hoher Geschwindigkeit ausführen.

### 3.5.3 Connector SDK

Um eine Transition zwischen tabellarischen Daten in relationalen oder objektorientierten Data Warehouse-Lösungen mit einem semantischen Graphen herzustellen, ist eine Umsetzung und Synchronisierung zwischen Tabellen und Graphen notwendig. Diese erzeugt Herausforderungen bei der Abfrage und Schreibperformance und benötigt eine besondere Art der Konfliktbehandlung (Conflict Management).

Das Connector SDK der Knowledge Platform stellt den Rahmen, die Dokumentation sowie vorgefertigten Skelett-Code bereit, um eine solche bi-direktionale Verbindung zwischen traditionellen Datenquellen und dem semantischen Graphen zu ermöglichen.

Die Entwicklung von Schnittstellen wird auf diese Weise im Vergleich zu der sonst üblichen zeitaufwändigen Enterprise Application Integration selbst bei komplexen Datenstrukturen um ein Vielfaches verkürzt.

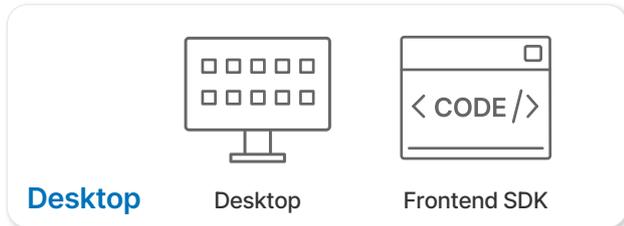
### 3.5.4 Funktionale Datenabfragen

Neben den oben genannten technischen Wegen lassen sich verschiedene funktionale Zugriffe auf die Daten über die Knowledge APIs ansteuern:

- + **Punktuelle Abfragen über Open Graph Protocol (OGP):** OGP ist das einfachste Protokoll, um den Knowledge Graphen abzufragen, wenn genau ein Knoten gesucht oder abfragt werden soll. Den Kanten des Knoten kann einzeln mit separaten API-Aufrufen gefolgt werden. OGP eignet sich für wenig komplexe Datenanfragen.
- + **Graph-Traversierungen über Gremlin:** Zur Implementierung echter paralleler Algorithmen, Abfragen und Suchen auf dem Graphen wird Gremlin unterstützt. Gremlin ermöglicht sowohl komplexe Suchen über ganze Wege im Graphen als auch die Parallelisierung von Abfragen über die Fähigkeiten der Plattform.
- + **Graph-Abfragen mit GraphQL zur tabellarischen Ausgabe:** GraphQL steht als Abfragesprache für Graphdaten zur Verfügung, um semantische Daten in Tabellen, Matrizen oder Vektoren zu wandeln, sodass diese statistisch oder mit anderweitigen, für lineare Daten optimierten Auswertungswerkzeugen, bearbeitet werden können.

## 3.6 Desktop

Desktop ist die Laufzeitumgebung von Bardioc für Anwendungen mit graphischer Benutzeroberfläche. Der Desktop integriert das Rechte- und Rollenkonzept des Access Managers über Token und steuert die Kommunikation mit dem Graphen.



Es stehen verschiedene Standardanwendungen zur Verfügung, die über den Desktop geladen werden können:

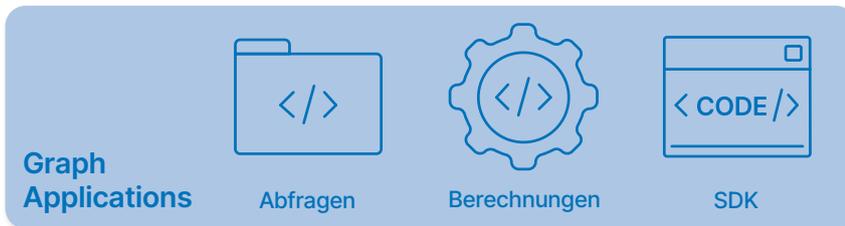
- + **Punktuelle Abfragen über Open Graph Protocol (OGP):** OGP ist das einfachste Protokoll, um den Knowledge Graphen abzufragen, wenn genau ein Knoten gesucht oder abfragt werden soll. Den Kanten des Knoten kann einzeln mit separaten API-Aufrufen gefolgt werden. OGP eignet sich für wenig komplexe Datenanfragen.
- + **Profile Manager** zur Verwaltung von Knowledge Core Accounts.
- + **Graph Explorer** zur Visualisierung von Graphdaten und zur Anzeige von Meta-Informationen.
- + **Ontology Manager** zur Verwaltung und Bearbeitung der Ontologie.

Über das Frontend SDK können individuelle Anwendungen erstellt werden, die ebenfalls dem Rechte- und Rollenkonzept des Access Managers unterliegen. Das Frontend SDK stellt folgende Features für Applikationen zur Verfügung:

- + **Authentifizierung und Autorisierung** gegenüber dem Access Manager.
- + **Messaging System** zur Kommunikation mit dem Graphen sowie zwischen den Applikationen.
- + **Error Handling** zur standardisierten Fehlerbehandlung.
- + **Responsive Design Templates** für die komfortable Erstellung von Oberflächen inkl. automatischer Anpassung der Anwendungen an das jeweilige Ausgabegerät.
- + **Pipeline zum Deployment** von Frontend-Anwendungen auf Bardioc (z. B. als JVM oder EVM).

## 3.7 Graph Applications

Bardioc ist das Betriebssystem für Daten und stellt mit dem Knowledge Core und dem Desktop sowohl die Rechenplattform als auch die grafische Oberfläche zur Verfügung, die man von einem solchen Betriebssystem erwarten kann. Auf Ebene des Graphen werden verschiedene Funktionen nach außen geführt bzw. bereitgestellt, die für Nutzer von semantischen Plattformen typischerweise von Interesse sind.



So können zum Beispiel über den Machine Learning Service eigene DeepLearning Netze trainiert und verwendet werden. Auch auf der graphischen Ebene stehen grundlegende Applikationen zur Verfügung, die das Verwalten von Profilen, Komponenten der Plattform sowie der Ontologie ermöglichen.

Standardmäßig sind folgende Graph-Algorithmen integriert:

---

### Traversal Recipes

---

#### Between Vertices

---

#### Centrality

Degree Centrality  
 Betweenness Centrality  
 Closeness Centrality  
 Eigenvector Centrality  
 PageRank Centrality

---

#### Collections

---

#### Connected Components

Small graph traversals  
 Small graph scalability  
 Large graphs

---

#### Cycle Detection

---

#### Duplicate Edge Detection

---

#### Duplicate Vertex Detection

---

#### Moving an Edge

---

#### Element Existence

---

#### If-Then Based Grouping

---

#### Looping

Conditional Looping with Max Depth  
 Emitting Loop Depth  
 Optional Loop Depth

---

---

**Operating on Dropped Elements**

---

**Pagination**

---

**Recommendation**

---

**Shortest Path**

---

**Traversal Induced Values**

---

<b>Tree</b>	Lowest Common Ancestor Maximum Depth Time-based Indexing
-------------	--

---

<b>OLAP traversals with Spark on YARN</b>	Approach Prerequisites Running the job Explanation Additional configuration options
---	---

---

**Anti-Patterns**

---

**Long Traversals**

---

**Unspecified Keys and Labels**

---

**Unnecessary Steps**

---

**Unspecified Label in Global Vertex lookup**

---

**Steps Instead of Tokens**

---

**has() and Traversal Arguments**

---

Mittels eines SDKs lassen sich sowohl Graph-Module als auch Frontend-Anwendungen zu Datenvisualisierung, Datenmanagement sowie Verwaltung erzeugen. Graph-Module sind Programme, die auf dem Graphen laufen und vom Data Manager verwaltet werden. Bibliotheken des SDKs erleichtern die Verwendung von Funktionen wie Identifikation, Verteilung, verteiltes Rechnen, Programmkommunikation sowie Security und beschleunigen so die Entwicklung individueller Anwendungen auf den eigenen Daten.

The background features a complex, abstract design. It consists of several overlapping, semi-transparent layers. The primary colors are various shades of teal and blue. A prominent feature is a grid of small, light-colored squares that appears to be draped or curved across the scene, creating a sense of depth and movement. The overall effect is a textured, digital landscape.

# 4. Ontologie

# 4. Ontologie

Zentrales Element des Datenmodells von Bardioc ist die Ontologie, die im Rahmen der Entwicklung der Plattform definiert und seither evolutionär weiterentwickelt wird. Die Ontologie bildet den semantischen Kontext bei der Strukturierung, Integration und Analyse der Daten ab.

## 4.1 Konzepte und Eigenschaften der OGIT Ontology

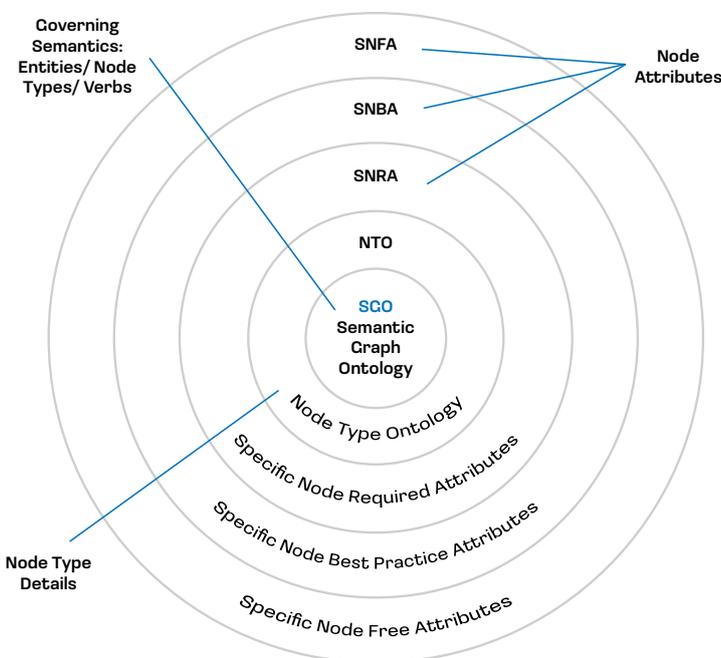
Die von Almato veröffentlichte OGIT Ontology definiert und organisiert die Beziehungen zwischen verschiedenen Datenentitäten, indem sie eine informationstechnisch geprägte Sicht auf die Welt auf Basis von Geschäftsprozessen und Personen abbildet.

Die OGIT Ontologie von Bardioc steht als Open Source zur Verfügung. Darüber hinaus sind individuelle Erweiterungen und Evolution des ontologischen Fachwissens um private Ontology Extensions möglich.

Zur evolutionären Weiterentwicklung sowie Pflege von Ontologie und Taxonomie steht ein definierter Prozess zur Verfügung. Diese werden in einem Open-Source-Prozess gepflegt und sind jederzeit öffentlich verfügbar. In diesem Prozess ist dokumentiert, wer für welche Ontologie-Erweiterungen verantwortlich ist und wie Verknüpfungen zwischen solchen Sub-Ontologien durch das Ontologie-Board beschlossen und genehmigt wurden.

- + Abbildung von **Knoten** (Datenentitäten, Node, Vertex) als grundlegende Bausteine der Ontologie zur Repräsentation realer Objekte, Konzepte oder Ereignisse, die im Datensystem modelliert werden. Beispiele könnten Kunden, Personen, Produkte, Transaktionen, Prozesse oder Standorte sein.
- + Abbildung von **Eigenschaften** (Label, Attribut, Beschriftung) zur Beschreibung spezifischer Merkmale. Beispielsweise könnte eine Kundenentität Attribute wie Name, Adresse, Geburtsdatum und Kundennummer haben.
- + Abbildung von **Kanten** (Verbindungen, Link, Edge, Beziehung, Relationship) zur Verbindung von Datenentitäten sowie zur Repräsentation der Beziehung zwischen den Entitäten. Die Kanten bilden die Grundlage für die semantische Analyse und ermöglichen die Navigation durch das Datennetzwerk. Beispiel ist die Beziehung zwischen Kunden und Bestellungen, die anzeigt, welcher Kunde welche Bestellung aufgegeben hat.
- + Abbildung von **Brückenknoten** (Bridge Node), um unterschiedliche semantische Datennetze miteinander zu verbinden.
- + Über **Hierarchien** können Entitäten organisiert werden, um komplexe Datenstrukturen abzubilden. Durch Kategorisierung und Klassifikation können Produkte zum Beispiel in Kategorien und Unterkategorien organisiert werden.
- + Mit Hilfe von **Regeln** und **Constraints** können in der Ontologie feste Regeln und Einschränkungen definiert werden, um die Datenintegrität sicherzustellen. Zum Beispiel kann festgelegt werden, dass jede Bestellung mindestens einen Kunden haben muss oder ein Geburtsdatum immer in der Vergangenheit liegen muss.

- + Durch die **semantische Anreicherung** können Daten mit Hilfe der Ontologie um zusätzliche Kontextinformationen angereichert werden. Dies erhöht das Verständnis der Daten und verbessert die Analysemöglichkeiten. Hierfür können externe Datenquellen integriert werden oder KI-Technologien wie NLP (Natural Language Processing) verwendet werden.
- + Da verschiedene Datenquellen oft unterschiedliche Terminologien und Konzepte verwenden, um ähnliche Dinge zu beschreiben, entstehen sogenannte semantische Lücken. Die OGIT Ontology sorgt für eine **semantische Heterogenität**, indem sie ein gemeinsames Vokabular für Beziehungen und deren Bedeutung innerhalb des Graphen definiert.
- + Durch die Verwendung offener Standards zeichnet sich die OGIT Ontology durch eine hohe **Interoperabilität** und **Integrierbarkeit** durch Daten aus verschiedenen Quellen aus.
- + Die OGIT Ontology bietet eine **dynamische Anpassbarkeit** durch evolutionäre Weiterentwicklung. Neue Anforderungen an das Datenmodell oder neue Datenquellen lassen sich so einfach integrieren, ohne dass der bestehende Kontext des Modells verloren geht.

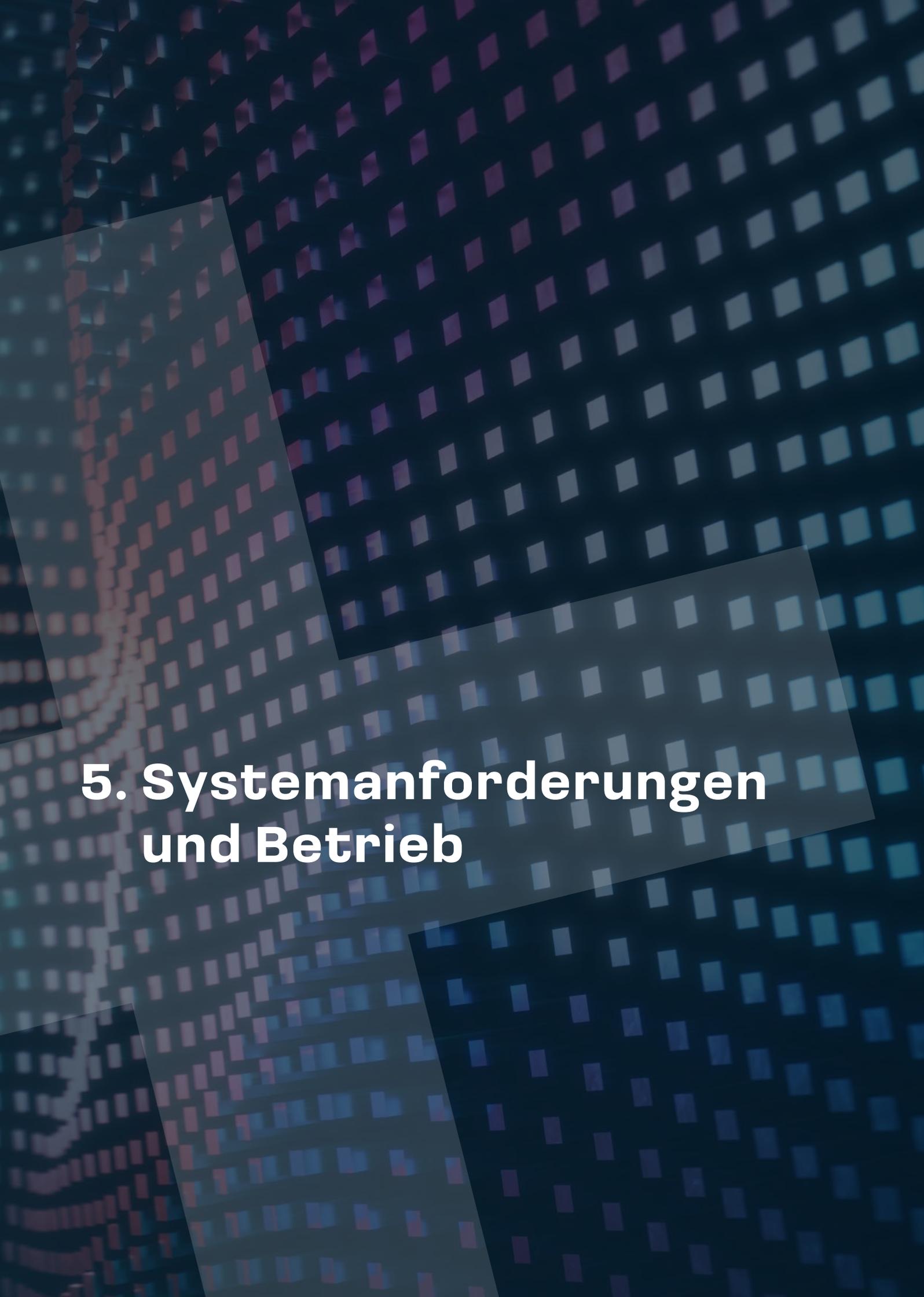


Meta-Modell der OGIT Ontology

## 4.2 OGIT Ontology Extensions

Die Ontologie kann durch fachspezifische Erweiterungen ergänzt werden. Dabei können auch bereits vorhandene Ontologien aus der eigenen Umgebung oder standardisierte und öffentlich verfügbare Ontologien eingebunden werden. Die OGIT Ontology Extensions gewährleisten, dass sich diese Domänen-spezifischen Sprachen unverändert einbinden lassen, ihre eigene Evolution vollziehen können und dennoch vollkommen in den semantischen Kontext der OGIT Ontology integriert und mit anderen Schichten oder auch anderen OGIT Ontology Extensions verbunden sind.

Auf diese Weise lassen sich bereits ausgearbeitete und als Standard anerkannte Ontologien oder andere semantische Strukturen in den Datenpool von Bardioc einbeziehen, ohne den Pflegeaufwand zu erhöhen. Gleichzeitig lassen sich die neu hinzugefügten Daten mit den bereits vorhandenen Daten verbinden, sodass keinerlei Limitationen bei der semantischen Interpretation sowie der Datenabfrage entstehen.



# **5. Systemanforderungen und Betrieb**

# 5. Systemanforderungen und Betrieb

Bardioc ist auf extreme Skalierung, hohe Sicherheit und Geschwindigkeit ausgelegt. Entsprechend anspruchsvoll sind die Anforderungen an die Infrastruktur und den Betrieb. Die Plattform kann in verschiedenen Betriebsmodellen verwendet werden.

## 5.1 Software as a Service / Platform as a Service

Almato betreibt Bardioc in konzerneigenen und geo-redundanten Data Centern der DATAGROUP am Standort Frankfurt am Main. Der RZ-Betrieb ist nach den Sicherheitsstandards der ISO 27001 sowie den Qualitätsstandards der ISO 20000 zertifiziert. Eine kontinuierliche Überwachung aller IT-Komponenten, Dienste und Kapazitäten sowie ein hoher Automationsgrad ermöglichen hohe Service Levels sowie einen wirtschaftlichen Betrieb.

Die Funktionen von Bardioc können als Software as a Service (SaaS) auf der mandantenfähigen, zentralen Plattform genutzt werden. Alternativ kann eine dedizierte Plattform in einer Enterprise Cloud oder Private Cloud in Form eines „Platform as a Service (PaaS)“-Modells bereitgestellt werden.

## 5.2 Hyperscaler

Bardioc ist kompatibel mit allen großen Cloud-Providern, insbesondere mit den Plattformen der großen Hyperscaler Amazon Web Services, Microsoft Azure sowie allen auf OpenStack basierenden Cloud-Plattformen. Für das Deployment der Plattform stehen entsprechende Referenzarchitekturen zur Verfügung.

## 5.3 On-Premise

Für Kunden mit hohem Schutzbedarf oder expliziter Make-Strategie steht die Technologie von Bardioc auch für den Betrieb im eigenen Data Center on-premise zur Verfügung. Eine On-Premise-Strategie sollte aufgrund der erforderlichen Infrastruktur- und Systemressourcen nur von Kunden in Erwägung gezogen werden, die sehr große Rollout-Anforderungen haben.

### Production ready:

- + Platform as a Service von Almato in Deutschem Rechenzentrum
- + Alle Hyperscaler (aws, Azure, OpenStack)
- + On-Premise-Betrieb im eigenen Rechenzentrum

## 5.4 System- und Integrationsanforderungen

Die System- und Integrationsanforderungen für On-Premise- oder Cloud-Installationen werden in kundenspezifischen Anforderungsworkshops ermittelt.

## 5.5 Open Source

Bardioc verwendet eine Reihe bewährter Open-Source-Komponenten, die sorgsam ausgewählt und spezifisch entsprechend ihrer Stärken zum Einsatz kommen. Eine Liste der Komponenten mit den dazugehörigen Lizenzen findet sich unter:

<https://bitbucket.org/almatoag/workspace/projects/KG>

## 5.6 Zertifizierbarkeit

Bardioc wurde in kundenindividuellen Umgebungen erfolgreich nach ISO 27001 sowie nach ISO 25051 zertifiziert. Für spezifische kundenindividuelle Installationen können Zertifizierungen bei Bedarf unterstützt werden.



## **6. Integrations- und Erweiterungsmöglichkeiten**

# 6. Integrations- und Erweiterungsmöglichkeiten

Bardioc kann vielseitig erweitert werden, um spezifische kundenindividuelle Anforderungen umsetzen zu können.

## 6.1 Knowledge Automation

Die Knowledge Automation Platform (siehe separate Produktbeschreibung) kombiniert eine auf künstlicher Intelligenz basierende Automatisierungs-Engine (Regel-Engine) mit einer auf Deep Learning basierenden Wissensmanagement-Engine und bietet damit ein Framework zur Formalisierung von Betriebswissen sowie zur nachhaltigen Automatisierung hochindividueller Aufgaben oder Prozesse.

Dank dieser einzigartigen Technologiekombination kann die Knowledge Automation Platform Aufgaben vergleichbar einem menschlichen Experten lösen und ist nicht auf die zwingende Einhaltung genau definierter Verfahren, Runbooks oder Skripte zur Automatisierung angewiesen.

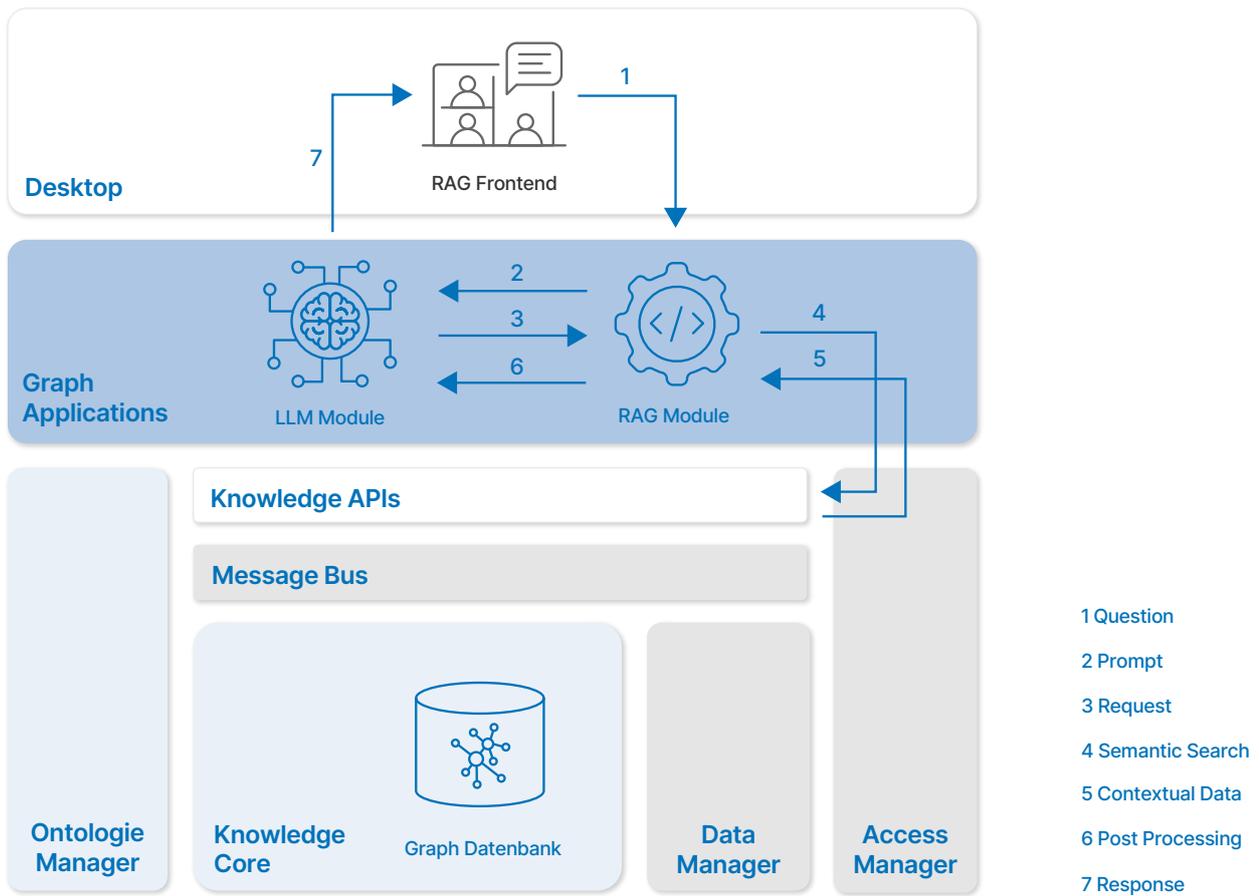
Mit Hilfe von Knowledge Acquisition Tools können komplexe Aufgaben in atomare Informationseinheiten, sogenannte Knowledge Items, zerlegt werden. Die Knowledge Automation Platform findet für gegebene Aufgaben die passende Lösung auf Grundlage des Kontexts des Problems sowie der verfügbaren Knowledge Items.

## 6.2 Large Language Models / Retrieval-Augmented Generation

In Kombination mit einem Large Language Model (LLM) kann Bardioc zu einem RAG-System (Retrieval-Augmented Generation) erweitert werden. Über das LLM können Anfragen an den Knowledge Graphen in natürlicher Sprache gestellt werden, wobei auch die Antwort benutzerfreundlich und verständlich formuliert generiert wird. Das RAG-System liefert eine konkrete Antwort auf Basis der im Knowledge Graphen gespeicherten relevanten Daten bzw. des dort gespeicherten strukturierten Wissens. Dadurch können nicht nur isolierte Fakten, sondern auch deren semantische Zusammenhänge und zeitliche Abhängigkeiten bei der Beantwortung berücksichtigt werden.

Knowledge-Graph-basierte RAG-Systeme bieten folgende Vorteile:

- + **Verbesserte Genauigkeit** der generierten Antworten durch die Nutzung strukturierten Wissens.
- + **Erweiterte Kontextualität** durch die Verknüpfung von Entitäten und deren Beziehungen.
- + **Höhere Erklärbarkeit der Antwort** durch bekannte Herkunft der Information, wodurch die Transparenz und Nachvollziehbarkeit insbesondere bei kritischen Anwendungen gesteigert wird.



Architektur für Retrieval-Augmentation Generation auf Basis von Bardioc

# 7. Zusammenfassung

# 7. Zusammenfassung

## 7.1 Vergleich mit Graph Datenbanken

Im Vergleich mit Graph Datenbanken bietet Almato Bardioc folgende Vorteile:

Eigenschaft	Graph Datenbanken	Almato Bardioc
Entwicklung und Betrieb der Lösung in Deutschland	—	+
Vollintegrierte semantische Datenplattform (Betriebssystem für Daten)	—	+
Integrierte Ontologie mit informationstechnisch geprägter Sicht auf Basis von Prozessen und Personen	—	+
Lineare Skalierbarkeit bis mehrere hundert Petabyte Datenvolumen	—	+
Performerer Umgang mit Super Nodes (stark vernetzte Knoten)	—	+
Hohe Geschwindigkeit bei komplexen Abfragen oder Transaktionen im Graphen	—	+
Effiziente Speicherverwaltung durch ausgeklügeltes Memory Management	—	+
De-Duplizierung zur Vermeidung unnötiger Datenredundanzen	—	+
Inhärente Systemsicherheit durch Token-basiertes Berechtigungskonzept für Benutzer und Applikationen	—	+
Wirksamkeit von Berechtigungsänderungen in Echtzeit	—	+
SDKs zur Entwicklung von Graph Applications, Frontend-Anwendungen sowie zur Integration von APIs	—	+
Historisierung für zeitliche Abfragefähigkeit im Gesamtkontext	—	+
Audit-Logs für höchste Compliance-Anforderungen	—	+

## 7.2 Alleinstellungsmerkmale von Almato Bardioc

- + Bardioc bietet **Security by Design**. Verschlüsselte Datenspeicherung und -übertragung, granulare Berechtigungen auf Datensatzebene sowie token-basierte Zugriffsmechanismen sind inhärente Designkriterien der Plattform und keine optionalen Add-ons.
- + Bardioc bietet lineare **Skalierungsmöglichkeiten** für sehr große Datenmengen. Viele Datenvorhaben starten als Pilotprojekte mit kleinen Datenmengen. Skalieren die Datenmengen schnell und kommen Anforderungen an verteilte Systeme, selbstständige und effiziente Ressourcen-Allokation ohne manuelle Eingriffe sowie Parallelisierung von Aufgaben hinzu, stoßen einfache Graph Datenbanken sehr schnell an ihre Grenzen.
- + Mit der **Historisierungsfunktion** von Bardioc wird die Grundvoraussetzung für Compliance, Revisionssicherheit sowie Erklärbarkeit geschaffen, Voraussetzungen für die Nutzung einer Datenplattform in jeglichem regulierten Umfeld. Gleichzeitig sorgt sie durch die automatische Zuweisung des effektivsten Speichermediums für eine hohe Wirtschaftlichkeit.
- + Bardioc verfügt mit der OGIT Ontology über eine **eigene Ontologie**, die die Integration weiterer Domänen-spezifischer Sprachen (DSL) sowie die Einbindung von Ontology Extensions unterstützt. Der definierte und transparente Prozess zur evolutionären Weiterentwicklung der Ontologie schafft die Voraussetzungen für einen integrierten semantischen Datenpool.
- + Durch die stringente Konzeption von Bardioc als **Betriebssystem für Daten** sowie die Bereitstellung umfangreicher APIs und SDKs können Programme und Anwendungen den Daten im Graphen zugeführt werden, anstelle die Daten aus dem Graphen extrahieren und den Programmen zuführen zu müssen.

**Almato AG**

Ein Unternehmen der DATAGROUP  
Theodor-Heuss-Straße 9  
70174 Stuttgart  
+49 711 3406-7810  
sales@almato.com  
almato.com

**Standorte**

Stuttgart  
Barcelona  
Bonn  
Neu-Isenburg  
Reutlingen