IBM Knowledge Accelerator for Cross Industry

Accelerate the implementation of data governance, enterprise vocabularies, and compliance with regulatory and industry standards to create comprehensive industry content for catalog-generated management.
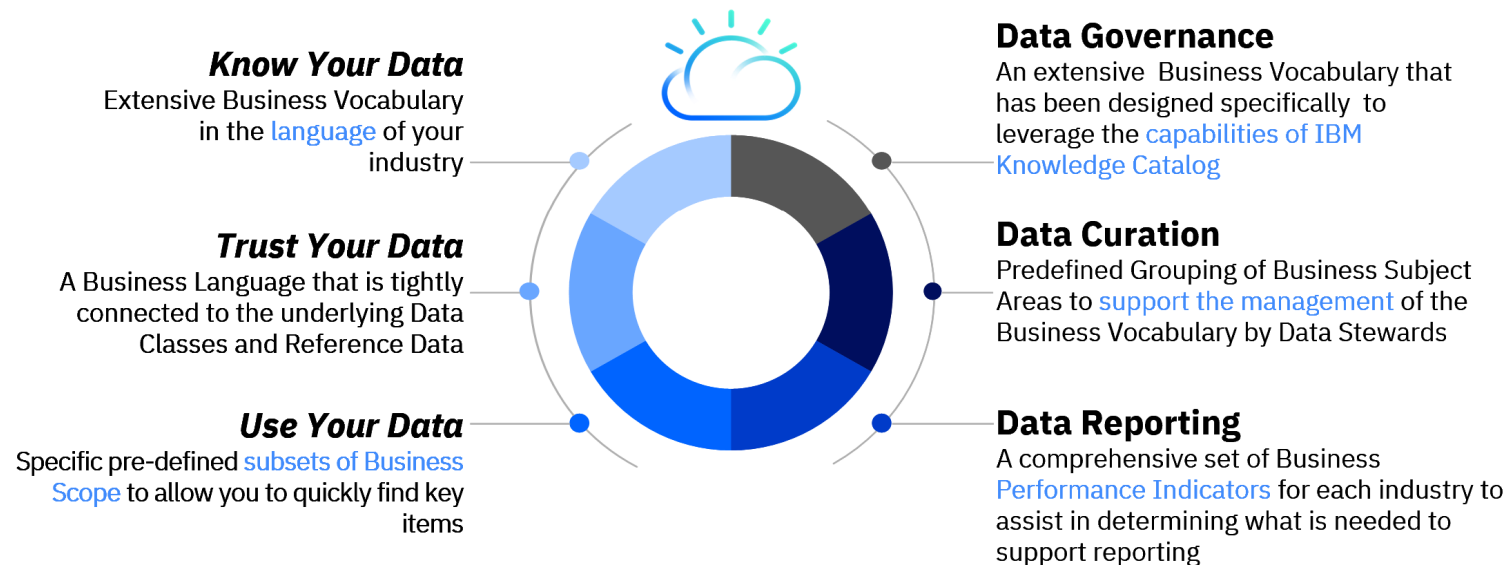
IBM®

Table of contents

Business Context

Enterprise knowledge management requires a trusted, industry-oriented business vocabulary that captures the business meaning and context for the industry concepts, practices, and assets related to business operations.

Enterprise Knowledge Management

Requires a trusted, industry-oriented Business Vocabulary

Know Your Data
Extensive Business Vocabulary in the language of your industry

Trust Your Data
A Business Language that is tightly connected to the underlying Data Classes and Reference Data

Use Your Data
Specific pre-defined subsets of Business Scope to allow you to quickly find key items

Data Governance
An extensive Business Vocabulary that has been designed specifically to leverage the capabilities of IBM Knowledge Catalog

Data Curation
Predefined Grouping of Business Subject Areas to support the management of the Business Vocabulary by Data Stewards

Data Reporting
A comprehensive set of Business Performance Indicators for each industry to assist in determining what is needed to support reporting

However, building a business vocabulary is more than creating a word list without definitions or business context. It takes time to create a usable business vocabulary, lots of time. It also requires a depth of industry experience and a breadth of operational business knowledge that usually is beyond a single person's capability. Thus, building a business vocabulary takes time and takes a team.

Best Practice Governance, Compliance, and Enterprise Search

An enterprise business vocabulary provides great value for organizations because it centralizes the governance of enterprise information and assets. Activities such as compliance risk management are supported by a business vocabulary that connects regulatory concepts to the enterprise data footprint and the data policies and rules that surround it. Additionally, a well-governed enterprise business vocabulary enables enterprise search capability to provide a true "everything to everyone" view of the business.
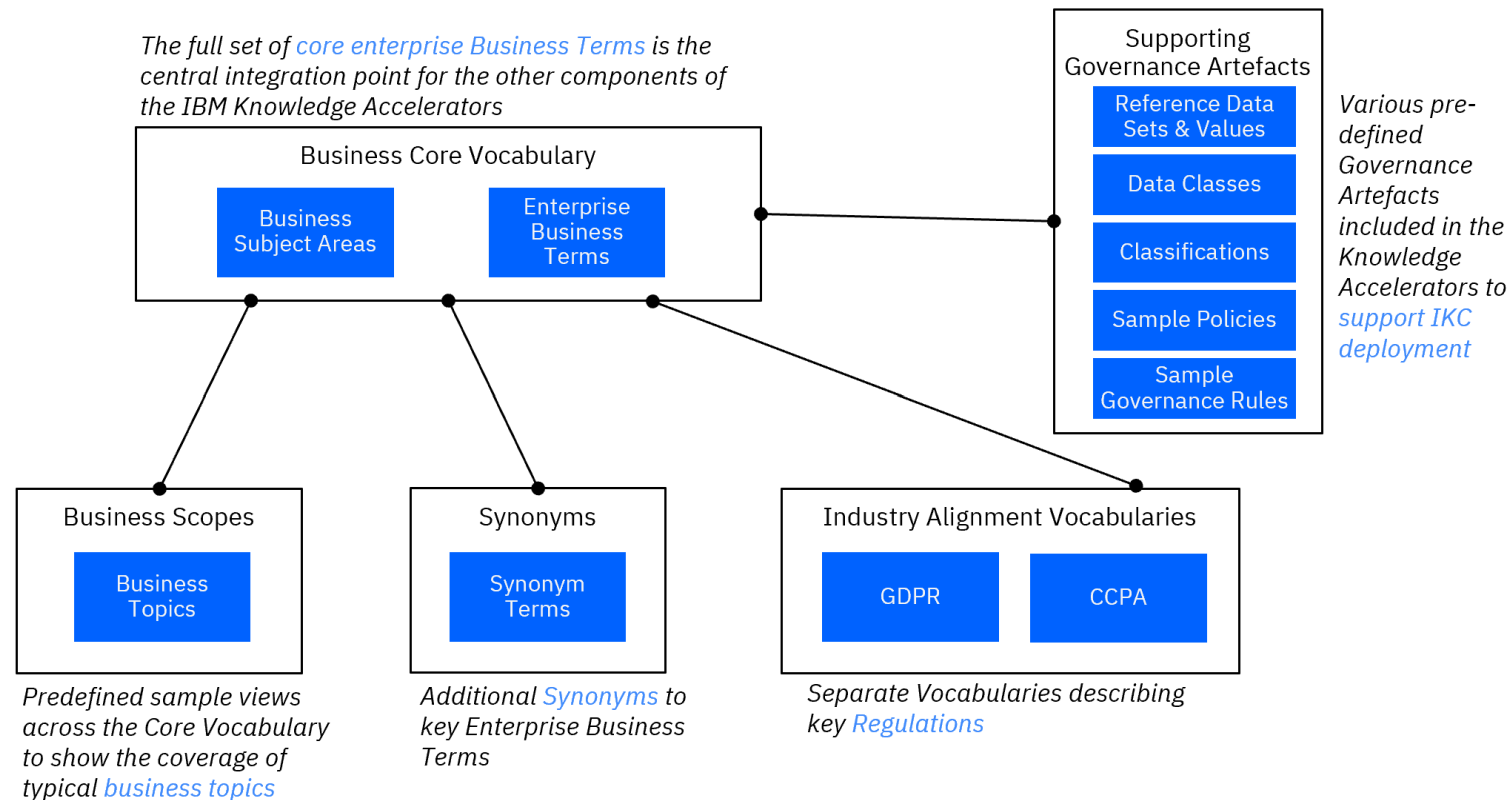
In addition, an enterprise business vocabulary is an important basis for the overall semantic layer of a Data Fabric ecosystem.
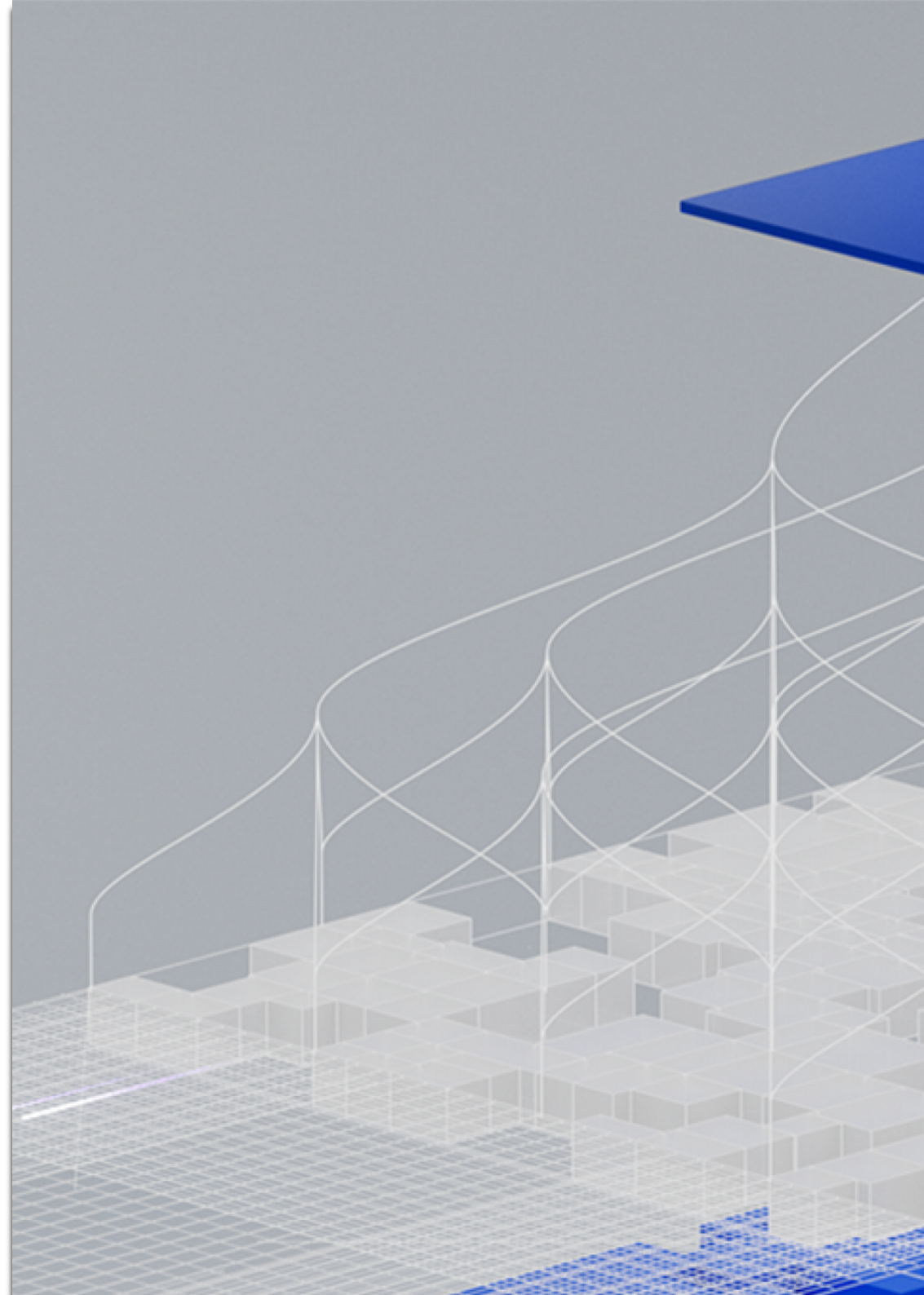
IBM Knowledge Accelerator for Cross Industry (IBM KACI)

IBM® Knowledge Accelerator for Cross Industry (KACI) is a vocabulary that is imported to [IBM Knowledge Catalog](). The comprehensive, multilayer business vocabulary includes thousands of inter-related terms that are grouped by intuitive subcategories.

The terms in these subcategories (Business Core Vocabulary, Business Scopes, Synonyms and Industry Alignment Vocabularies) are used by different business users to establish, understand, navigate, and use an enterprise vocabulary.

The full set of core enterprise Business Terms is the central integration point for the other components of the IBM Knowledge Accelerators

Business Core Vocabulary

- Business Subject Areas
- Enterprise Business Terms

Supporting Governance Artefacts

- Reference Data Sets & Values
- Data Classes
- Classifications
- Sample Policies
- Sample Governance Rules

Various pre-defined Governance Artefacts included in the Knowledge Accelerators to support IKC deployment

Business Scopes

- Business Topics

Predefined sample views across the Core Vocabulary to show the coverage of typical business topics

Synonyms

- Synonym Terms

Additional Synonyms to key Enterprise Business Terms

Industry Alignment Vocabularies

- GDPR
- CCPA

Separate Vocabularies describing key Regulations

Find out more about the IBM Knowledge Accelerators at: https://www.ibm.com/products/knowledge-accelerators

IBM KACI subcategories

The IBM Knowledge Accelerator for Cross Industry is represented as a category of business terms within IBM Knowledge Catalog. The components within IBM KACI are represented by the following subcategories:

- Business Core Vocabulary

Defining industry concepts in plain business language, the Business Core Vocabulary contains business terms that have properties and can be related to each other. Business terms can include a term hierarchy and are organized by logical business categories. Clearly defined business terms help standardization and communication within an organization and are a foundation for governance of enterprise information assets.

- Business Scopes

Business Scopes provide groupings of related Business Core Vocabulary terms that address a specific business use case. These are typically at a finer grain than the business core vocabulary categories and will help business users to focus on the terms required to support a particular business need. The business scopes are available for separate download and import.

- Synonyms

Synonym terms represent a word with the same meaning and are synonyms of Core Business Vocabulary terms.

- Industry Alignment Vocabularies

Industry Alignment Vocabularies include terms that align key elements from industry regulations to the terms within the Business Core with regulatory coverage and compliance (GDPR, CCPA).

Business Core Vocabulary

IBM Knowledge Accelerator for Cross Industry includes a comprehensive vocabulary of business terms. These are organized into a set of business categories to enable users to quickly identify the terms relevant to their area of interest. For example, the business category Communication, contains over 200 terms related to communication metadata.

Activity	Equipment	Payment
Communication	Location	Person
Customer	Organization	Product

Business Scopes available for import

The IBM Knowledge Accelerator for Cross Industry has three business scopes available for separate download and import. These can be imported individually to address one or more business use cases. Each Business Scope contains a set of business terms with supporting reference data sets and data classes. Multiple scopes can be imported with common terms being shared across the scopes. Users can also import the full KACI vocabulary to supplement any scope vocabulary content already imported. The following business scopes are available for separate download and import:

Data Privacy
A Business Scope that groups terms for Data Privacy, addressing the personal information of individuals and the processing of that information.

Contact Center
A Business Scope that groups terms describing the characteristics of the contact center communication, people and systems involved, contact preferences and documentation.

Weather Insights
A Business Scope that groups terms for weather-related terms information. The scope covers weather observations and forecasts terms and data related terms that may be provided as part of weather data products by specialized weather information service providers. This includes everyday weather patterns and severe weather occurrences.

Industry Alignment Vocabularies

IBM KACI includes Industry Alignment Vocabularies which are provided to guide mapping from regulations to the Business Core Vocabulary content. Each supported regulation is defined using subcategories and terms below the Industry Alignment Vocabularies top-level category.

Industry Alignment Vocabularies delivered with IBM KACI:

CCPA (California Consumer Privacy Act of 2018)

GDPR (EU General Data Protection Regulation)

IBM Knowledge Accelerators subcategories

The "Knowledge Accelerators" category contains related artifacts such as Data Classes, Reference Data, Policies, and Rules organized in subcategories.

Data Classes

IBM KACI extends the set of data classes that are provided in IBM Knowledge Catalog with several data classes describing types of data commonly found in data assets across multiple industries.

The existing IBM Knowledge Catalog data classes and the new data classes that are provided by the Knowledge Accelerators are all combined in a new category hierarchy based on topics such as Demographics, Finance, and Location. This subcategorization allows them to be selectively included in different metadata enrichments to improve matching accuracy. As the data classes are pre-mapped to terms in the Business Core Vocabulary, metadata enrichment also automatically assigns the business term when it matches data in an asset to a data class. Data classes that are relevant to personal data also have a classification of either Personal Information (PI) or Sensitive Personal Information (SPI) depending on the sensitivity of the data it describes. This classification helps in the identification of personal data in data assets and supports the application of data protection rules to control data access.

Reference Data

IBM KACI includes almost 40 reference data sets with 100's of reference values. This reference data, where applicable is mapped to the core business vocabulary. In IBM Knowledge Catalog, reference data sets can be used to generate Data Classes, which are used to profile the data in the Data Catalog. Try the guided demo of IBM Knowledge Catalog to explore the capabilities.

Policies

IBM KACI provides a sample selection of governance policies in the areas of Data Privacy and AI Governance. The policies are further broken down into policy subcategories, which contain individual policies that illustrate how the organization can define its own policies in these areas. The Organization might also use this category structure to create policies in other areas such as Sustainability, or Diversity. Policies are descriptive rather than enforceable in IBM Knowledge Catalog.

Rules

IBM KACI provides a sample selection of governance rules in the areas of Data Privacy and AI Governance. These rules express how the organization intends to implement their policies. Governance Rules can be related to one or more governance policies. They can also be related to business terms where such terms are clearly applicable to the rule. Like policies, Governance Rules are not enforceable. This sample that is provided illustrates how the organization can define its own expanded set of rules.

Data Privacy and Protection

IBM Knowledge Accelerator for Cross Industry contains business terms with classifications to guide the identification of personal information and sensitive personal information. By using the metadata enrichment tool in IBM Knowledge Catalog, the business terms can be assigned to imported data assets to identify the assets that contain personal data.

IBM KACI includes a **Data Privacy Scope** that contains a set of Business Core Vocabulary terms to accelerate the discovery and governance of personal information and the activities that are related to processing such information.

The Business Terms in the Data Privacy scope are categorized to align with key data privacy topics:

- Personal Data
    - Affiliations
    - Basic Personal Information
    - Beliefs, Views & Opinions
    - Biometrics And Neurodata
    - Criminal & Offence Activity
    - Demographics
    - Education & Skills
    - Finance
    - Health
    - Human Resources
    - ID - Government
    - ID - Others
    - Online Activities
    - Online Behavior
    - Personal Communication
    - Personal Location
    - Personal Preference Information
    - Schedules
    - Telephony & Video

- Personal Data Management

The scope is not intended to be a definitive standard for data privacy and protection requirements as the classifications of Personal Information (PI) and Sensitive Personal Information (SPI) vary in different legal jurisdictions. However, the business terms in the scope include suggested PI and SPI classifications.

The Data Privacy Scope complements specific regulations such as GDPR or CCPA and includes Business Core Vocabulary terms that are related to these regulations in the following categories.

- Regulatory Alignment
    - CCPA Related Terms
    - GDPR Related Terms

The Data Privacy Scope helps accelerate the building of the organization's data privacy taxonomy and the creation of an inventory of data assets that contain personal information. This scope is available for separate download and import.

Data Privacy Regulations

IBM Knowledge Accelerator for Cross Industry contains key concept terms from two leading data privacy regulations. These regulatory terms are related to the business terms within the Business Core Vocabulary, which helps data users identify enterprise information that is relevant to these regulations. This coverage can help the organization understand what components need to be considered including consumer rights, personal data types, processing activities, processing purposes, and roles.

GDPR

IBM KACI supports the General Data Protection Regulation (GDPR) and provides an industry-specific vocabulary that can help the organization discover and govern privacy data. It can help organizations ensure that their enterprise data architecture is able to provide the necessary data artifacts to report on data protection issues. It can also help to determine and define which type of personal data the organization uses.

CCPA

IBM KACI also supports the California Consumer Privacy Act (CCPA). Building on the foundations put in place with GDPR, IBM KACI identifies key terms within the CCPA regulation, which are then mapped to the Business Core Vocabulary terms.

AI Governance

IBM KACI includes a set of policies relating to governance of the use of Artificial Intelligence by the organization. The AI governance policies are organized into a set of subcategories that describe the main process flow of an AI project.

- **AI Scoping and Planning Policies** relating to the scoping and planning of AI projects including assignment of responsibilities, establishing the use case, and identifying potential ethical issues and risks.
- **AI Data Collecting and Organizing Policies** relating to the collection and organization of data for use in AI projects. It is critical that such data is appropriate to the use case and does not introduce bias into the model.
- **AI Building and Training Policies** relating to the building and training of AI models by the organization. This includes mitigating risks and bias, and the systematic recording of assessment results.
- **AI Validating and Deploying Policies** relating to the validation of machine learning models against their original context as well as the predetermined performance thresholds prior to deployment.
- **AI Monitoring and Managing Policies** relating to the ongoing monitoring and management of deployed machine learning models used by the organization.

This sample policy set is provided for illustrative purposes only. It shows how the organization can define its own expanded set of policies in this and other areas.

Industry Led Governance

While a business vocabulary is often defined as a way to provide a business view to connect with more technical artifacts, or as the starting point for defining the business scope to be used to guide the identification of a subset of logical model elements, the business vocabulary has significant value in its own right.

A properly defined business vocabulary that reflects the needs and terminology of the various users can be used to:

- act as a common language across the different business users across the organization
- provide a reference point when aligning new or acquired businesses into the overall organization
- provide the Data Steward with a comprehensive and scalable template with which to grow their business vocabulary, addressing both the needs of the central IT as well as the needs of the individual groups of Self-Service users

- provide a basis for the identification of gaps and overlaps between different projects or activities that the enterprise may be engaged in
- leverage the AI/ML capabilities of IBM Knowledge Catalog. Where relevant, the business terms in the IBM Knowledge Accelerators are integrated with the Data Classes provided out of the box with IBM Knowledge Catalog and the Knowledge Accelerators. This ensures a higher coverage and accuracy of the Auto assignment of Data Assets to Terms, thus improving the integration of the Knowledge Accelerators with the governed assets
- broaden the current coverage of the AI-driven Data Discovery process. The IBM Knowledge Accelerators include sample Reference Data Sets and Reference Data values that, when customized as required, can be used as a basis for the creation of additional Data Classes. This results in a broader range of data classes available to underpin the Data Discovery process.

Data Lake, Lakehouse and watsonx.data

The [data lake](#) and the more recent lakehouse have emerged as the possible mechanism to enable organizations to define, manage and govern the use of various big data technologies. This represents an evolution of big data towards mainstream use in an enterprise and the associated focus on management of such assets.

Many of the same traditional imperatives for the use of IBM KACI also exist when organizations deploy a lakehouse

- the need to establish a common cross enterprise set of assets for use by the business, with such assets fully integrated with the overarching layer of business terms
- the use of business-friendly terminology to ensure a suitable basis for self-service access by the business users of the lakehouse

- the use of classifications and tags on terms to assist business and technical users when searching and navigating the vocabulary
- the need for consistent and scalable structures of the vocabulary to ensure common understanding of the lakehouse assets by the business and technical users
- the need to enforce a common governance layer around the lakehouse.

IBM KACI contains a number of components which are ideally suited to supporting deployment to a lakehouse architecture. IBM Knowledge Catalog enables organizations to govern and protect their data assets in a [watsonx.data](#) lakehouse.

Learn More

IBM Knowledge Accelerator Metamodel

IBM Knowledge Accelerators are designed based on a common metamodel specification described in the components section of the IBM Documentation.

IBM Knowledge Accelerators for Cloud Pak for Data Components
https://www.ibm.com/docs/SSQNUZ_5.2.x/ka/components/compnts.html

IBM Knowledge Accelerators for Cloud Pak for Data as a Service Components
https://dataplatform.cloud.ibm.com/docs/content/ka/components/compnts.html

IBM Documentation

IBM Knowledge Accelerator for Cross Industry documentation is available on the following sites:

IBM KACI for Cloud Pak for Data
https://www.ibm.com/docs/SSQNUZ_5.2.x/ka/ka-cpd.html

IBM KACI for Cloud Pak for Data as a Service
https://dataplatform.cloud.ibm.com/docs/content/ka/ka-cpd.html