Choose Hitachi Content Intelligence to:
- Transform data into valuable business information.
- Surface new insights from relevant data.
- Mitigate risk by improving data quality and discovery.

DATASHEET

Hitachi Content Intelligence: Data Transformation and Exploration Solution

Be in the Know

Enterprises are putting increasing importance on data as we now live in a 24/7 connected world that is fueled by information. New content is coming into the enterprise at breakneck speed in a host of formats, on top of the mountains of legacy data it is already maintaining. The challenge for businesses is how to disseminate this data and make it work in their favor. What if you could aggregate all of your organizational data (internal or external), classify and categorize it, and allow your knowledge workers to independently explore it? Imagine the benefits to the organization when data is used to determine trends, discover new opportunities, and make informed decisions from aggregated and related data sets. Imagine continuously optimizing resource utilization by preventing data from going dark.

Organizations must address the challenges of exploring and discovering relevant, valuable and factual information across a growing number of data producers and siloed repositories. They need a solution that surfaces rapid insights from data that is both big and deep.

Hitachi Content Intelligence enables organizations to turn multistructured data into valuable business information. Content Intelligence aggregates that data to create a centralized information hub for your workforce to explore, discover and surface actionable business insights quickly.

Connect, Transform, Enrich and Act

Hitachi Content Intelligence automates the extraction, classification, enrichment and categorization of data residing on both Hitachi Vantara and third-party repositories, located on-premises and in clouds, and across heterogeneous data repositories (internal and external). This approach drastically reduces time spent searching for what is needed, or recreating what already exists. Additionally, Content Intelligence delivers:

- Guided data exploration based on automated classification and categorization.
- Immediate visibility to all of your data by unifying data access across disparate locations and data types.
- Important insights by transforming your data into valuable business information.
- Managed access to sensitive data with granular access controls and security system integrations.
- The right data to the right person at the right time, using personalized and self-service features and user experiences.

An Innovative and Intelligent Approach to Understanding Data

Hitachi Content Intelligence is a data recommendations solution, built upon the feature-richness, scalability and extensibility of an open-core architecture. Integrating, packaging and extending open source ensures that Content Intelligence has a mature and broadly adopted code-base that is already in use in many industries, verticals and in many applications. As a result, our solution appeals to a very active developer community. They can use the extensible nature of the product to integrate it with your other services, or to create new solutions and user experiences tailored for your audience.

Content Intelligence takes advantage of the collocation of its data and indexes with its query services, to generate recommendations from enterprise data. The recommendations are based on user requests. These may come from the solution's first applicable use case, Search, or through a user experience where Content Intelligence is integrated into an existing application or used to create a custom application.

How It Works

As shown in Figure 1, Hitachi Content Intelligence can connect to and index data that resides on Hitachi Content Platform (HCP), HCP Anywhere, remote CIFS/NFS shares, databases, HDFS, cloud repositories and a wide variety of other structured, semi-structured and unstructured data sources. Once data repositories are connected, Hitachi Content Intelligence supports multiple processing workflows with access to a comprehensive library of data analytics, extraction, transformation and enrichment stages that can be conditionally applied as your data is processed. When data processing is complete, the results can be stored in multiple locations with access

controls, to ensure that only authorized users have access to specific data.

Hitachi Content Intelligence delivers purpose-built interfaces to perform the necessary functions of solution configuration, content analytics and transformations, and access to personalized results in the following ways:

- IT administrators use the administrative console to monitor, configure and integrate the solution into existing IT services.
- Content managers and data architects use the workflow designer to create connections to data repositories, design processing pipelines to extract and blend content, and define how and where the processed results are stored.
- End users have access to a search and discovery interface delivered with and powered by Content Intelligence. This application allows interaction with the personalized results using data categorization, natural language searches and data refinements, providing an intuitive and user-friendly experience.
- Storage administrators can perform enhanced performance monitoring and analytics of Hitachi Content Platform with Hitachi Content Monitor. A feature of Hitachi Content Intelligence, this software monitors multiple HCP clusters at scale, with customized visualizations and dashboards to improve capacity planning of storage, compute and networking resources.

Architectural Benefits

- **Flexible deployment options:** Choose from bare metal, software only or in-the-cloud deployment.
- **Highly scalable:** Based on an open-core clustered architecture, Content Intelligence is designed to support dynamic cluster sizes that can burst to 10,000+ instances.
- **Safe and secure:** Maintain the sensitivity of data by ensuring only authorized users have access to specified results. This capability integrates into existing directory services, such as lightweight data access
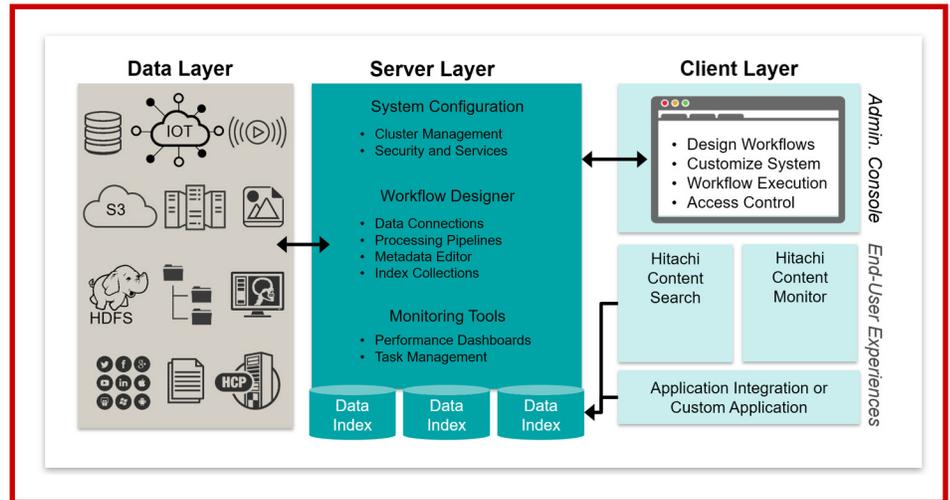

Figure 1. Hitachi Content Intelligence Architecture

TABLE 1. SYSTEM REQUIREMENTS

Requirements per Hitachi Content Intelligence Instance	
Operating System	A 64-bit Linux distribution running Docker v1.13.1 or later
Processor	4 cores (minimum), 8 cores (recommended)
Memory	24GB (minimum), 32GB (recommended)
Disk Space	50GB (minimum), 500GB (recommended)
Platform	• Hitachi Content Intelligence can be deployed on physical x86 servers. • Hitachi Content Intelligence can be deployed as a virtual machine using VMware's Open Virtualization Format template. • Hitachi Content Intelligence can be deployed in a hosted environment, such as Amazon Web Services.

protocol (LDAP) compatible servers, Microsoft® Active Directory®, OpenLDAP and 389 Directory Server.

- **Personalized results, granular controls:** Tailor result-set views to the user or group accessing the data, define document-level security controls, and apply role-based access rules.
- **Highly available, flexibly deployed:** Physical, virtual or hosted, Content Intelligence is a clustered solution that can be stretched geographically and architected to tolerate failures, based on the importance of the data being processed.
- **Extensible:** Create custom data connections, transformation and enrichment stages specific to the data you're working with. Content Intelligence provides

a fully featured software developer's kit. Its RESTful application programming interface support enables creation of custom connections and stages as well as integration of the solution into existing organizational tools and/or the creation of new applications.

Summary

As the market leader in digital transformation, Hitachi Vantara is uniquely positioned to help enterprises to incorporate digital discovery into their organization. Our Hitachi Content Intelligence solution effectively exposes insights for faster decision-making, productivity improvements and business risk reduction.

Hitachi Vantara