Timely data onboarding provides a competitive advantage. But data onboarding is hard. Lumada Data Services accelerates the process so you spend less time onboarding and more time analyzing and taking advantage of data.

DATASHEET

Gain Open Connectivity, Automation, Faster Adoption and Low Maintenance

High-quality data is the foundation of the information-driven enterprise. But challenges persist to unlock the potential of that data: It is complicated, siloed and distributed. Consequently, companies do not take advantage of most of their data today. According to IDC, only about 2.5% of all data is analyzed\(^1\): The rest is lost and wasted.

But the greatest waste of all is the time spent simply looking for the right data. The average data analyst spends up to 80% of his or her time searching, cleaning and organizing data\(^2\). This statistic might explain why half of them leave their organization within three years.\(^3\)

If you want to take advantage of all your data and keep your data analysts, you need to make it easy to access the right data at the right time and in the right context.

Data Onboarding and Integration With Lumada Data Services

Hitachi’s Lumada Data Services uses open connectivity and automation to let anyone involved with data speed up the time to value of their data. By reducing the number of tools and eliminating manual coding, data onboarding and integration happen faster. As a result, ML algorithms can be tuned with timely production data to deliver more accurate predictions. Organizations get to use their data to drive better business outcomes.

Open Connectivity To Avoid Tool Proliferation

Hitachi’s array of 200+ connectors across edge, cloud and mainframe make it easy to connect to and onboard a wide variety of data sources. And blending data gives a more holistic view, resulting in deeper customer insights. Develop or reuse custom connectors as needed.

- **Packaged connectors:** Plug into a wide range of data stores with flexibility and insulation from change.
- **Operational technology (OT) and IT sources:** Use the rich library of prebuilt components to prepare and blend data from relational sources, big data stores on the premises or in the cloud, enterprise applications, edge devices, sensors and more.
- **SDKs and marketplace:** Rapidly assemble custom connectors to a variety of new and legacy applications by using Lumada’s APIs and SDKs.

Automation To Escape Low-Level Coding

Dataflow templates reduce the effort needed to develop and manage data onboarding routines across diverse data sources. Significantly reduce your development time and cost by templatizing repeatable data workflows and replace dedicated data pipelines (see Figure 1).

- **Template driven:** Use a template-based approach to rapidly onboard data sources into Hadoop and other data stores via metadata injection feature set.
- **Execute edge to cloud:** Seamlessly switch between execution engines, such as Spark and the native engine of Hitachi Vantara’s Pentaho software, to fit data volume and transformation complexity.
- **Low maintenance:** Spot check data in flight with immediate access to analytics, including charts, visualizations and reporting, from any data-prep step.

Intuitive Interface and Intelligent Tagging To Shrink Learning Curves

Improve productivity using drag-and-drop development. Cleanse, blend, enrich and...
tag data at scale with business definitions to curate data (see Figure 2).

- **Drag-and-drop user interface:** Use the intuitive designer to simplify the creation of data pipelines. Take advantage of code-free data transformation design, which results in faster productivity versus hand coding.

- **Intelligent tagging:** Discover data rapidly with ML-driven discovery and automated tagging, which reduces the risk of manual processing errors.

- **Community:** Use crowd-sourced ratings and reviews to search for and navigate to data faster, so you can spend more time analyzing the best available data.

---

**Lumada Data Services**

Lumada Data Services from Hitachi comprises a set of composable software services that help manage increasingly complex data ecosystems with an intelligent data foundation. Its composable web services for analytics, data flows, cataloging and data lake management work together through shared metadata, supported by intelligent orchestration and policy engines. These services can be deployed in the cloud, in the data center or at the edge, providing an agile data fabric and self-service data access for an enterprise DataOps practice.

---

1. IDC Analyst Connection, sponsored by Hitachi Vantara, “Achieving Intelligent Data Operations”, June 2019,
2. CrowdFlower Data Science Report, 2016
3. Software had the highest job turnover rate of any industry in 2017, TechRepublic, March 2018

---

**Read this datasheet** to see how our unique metadata injection capability accesses data sources on the fly and injects metadata into transformation templates to accelerate complex data onboarding projects.