

Modern Data Management with Boomi

Accelerate business outcomes by enabling more people to gain insights from data.

463 exabytes

of data created each day by 2025¹

45%

of data scientists' time is spent on data preparation tasks, including loading and cleaning data.²

Boomi a leader in Enterprise iPaaS

In 2020, Boomi was positioned as a Leader in the Gartner Magic Quadrant for Enterprise iPaaS for the seventh consecutive year.³

Find and leverage your data

Data is the new rocket fuel for every organization — but only if users can find and leverage the information that they need to drive business outcomes. This wealth of data is growing as enterprises progress on their digital journeys, but a large portion of enterprise data is unknown, dormant or underused. Even as enterprises are producing, capturing and storing more data than ever before, many potential users can't find critical information and don't have the tools that they need to turn data into actionable insights.

For too many organizations, the move to big data technologies has created barriers to information for business and data analysts. Worse, data opacity and analytics complexity put the burden on IT to provide data extracts because organizations in the business lack the required source system access and coding skills to independently explore the data.

To enable a broader community of data consumers, a new approach is required — one that enables efficient access to data, no matter where it resides, for everyone with a business need.

Boomi Data Catalog and Preparation (DCP) software on Dell EMC PowerEdge servers can change the game for organizations working toward becoming more data driven. Together they enable more people in your organization to derive business value from practically any combination of enterprise data sources, without learning to program.

Boomi Data Catalog and Preparation

Boomi Data Catalog and Preparation (DCP) is a platform in which both IT and the business analyst community can find value. IT can easily add Boomi DCP to a new or existing big data cluster as easily as adding a traditional node to Hadoop. IT staff responsible for data governance can configure the right data access for the right users, using DCP role-based security integrated with existing source systems controls.

Business analysts using DCP can create powerful data transformations and summaries, with basic relational data structure and common SQL commands knowledge. These jobs are then run using highly scalable Hadoop services to produce new data artifacts. The artifacts are suitable for use with many popular reporting and visualization tools, both inside and outside of the Hadoop ecosystem.

Boomi DCP is easy for IT to install and manage, easy for data governance management, and easy for business analysts to use. This combination enables organizations that are investing in big data technologies to safely expand the end-user base, provide more value, and improve the ROI of any data-driven initiatives.

¹ The Next Platform, "[New Applications, More Users Drive Bright Outlook for HPC](#)," November 2019.

² Datanami, "[Data Prep Still Dominates Data Scientists' Time, Survey Finds](#)," July 2020.

³ Boomi, [Boomi Positioned as a Leader in the 2020 Gartner Magic Quadrant for Enterprise iPaaS for the Seventh Consecutive Year](#), September 2020.

Dell Technologies and Boomi

Dell Technologies acquired Boomi in November 2010 to help customers bridge traditional environments with the cloud. Boomi has since become the fastest growing acquisition in Dell Technologies' history, helping businesses around the world and across industries address their most daunting integration challenges.

Learn more

[Read the guide](#)

[Infohub](#)

delltechnologies.com/analytics

[Modern Data Management with Boomi](#) describes an end-to-end use case that demonstrates a typical Boomi DCP workflow. The use case example addresses the different types of problems that can be solved using the Boomi DCP toolset and shows how the data discovery, data cataloging and data transformation capabilities of Boomi DCP can be combined into an integrated business solution.

It demonstrates how someone with a basic understanding of data structures and table-joining can use GUI-based wizards and minimal coding to transform data. The output of a data transformation job definition is highly efficient SQL code that can be immediately run or scheduled for later on a Cloudera® cluster. Thus, Boomi DCP expands capabilities and capacity for business users who would otherwise have to use spreadsheets or single-threaded Python® or R on workstations to generate value from data.

For added flexibility, Boomi Data Catalog can be used stand-alone and/or with Boomi Data Prep and Hadoop. The software can be added to a new or existing cluster as easily as adding a node to an Apache® Hadoop® cluster. Using the integrated Cloudera wizard to extend a few Hadoop services to the Boomi node is all that is required to prepare to install the platform. Boomi DCP comes with its own installation scripts that are used once the cluster node is prepared.

Adding Boomi DCP to a Hadoop cluster adds significant functionality, providing a simplified extract, transform, load (ETL) experience. The all-in-one tool has significant value for organizations trying to leverage their Hadoop investments. The Boomi DCP framework enables IT to focus on the management of business-critical Hadoop investments while supporting the needs of business data analysts.

Dell Technologies engineers combined a Dell EMC PowerEdge R640 Server with Boomi DCP software, however, Boomi DCP can run on other server models including the PowerEdge R740/xd as needed to simplify ordering, maintenance and planning operations. The [Dell Technologies Cloudera CDP Data Center on Dell EMC Hardware Reference Architecture](#) includes a list of Dell EMC PowerEdge servers that are suitable for Boomi DCP with a Hadoop cluster.

If you're considering implementing Boomi DCP with Hadoop in your environment, you don't have to go it alone. [Dell Technologies Services](#) — ranging from consulting and education to deployment and support — are available when and where you need them.

