



Trifacta v4: Extends Enterprise Data Wrangling to Any User, Any Data, Any Cloud

Industry leader’s latest release dramatically enhances user experience, improves scale and performance and extends connectivity and cloud deployment support

Trifacta, the global leader in data wrangling, released Trifacta v4. The latest release expands upon Trifacta’s award-winning approach to data wrangling, with capabilities specifically designed to work for more users, more diverse data sources and within more cloud environments.

Trifacta v4 features the general availability of Builder, a new menu-driven workflow to guide users through data wrangling steps. The latest release includes the general availability of the Photon Compute Engine improving the scale of data that users are able to wrangle on-the-fly directly within the Trifacta application. Photon provides an optimized engine for data sets that don’t require parallel processing within Trifacta’s Intelligent Execution architecture. The v4 release also expands support for customers deploying Trifacta in cloud environments such as Amazon Web Services, Google Cloud Platform and Microsoft Azure, while extending the ability of users to directly connect to a variety of enterprise data sources, including Microsoft SQL Server, MySQL, Oracle, PostgreSQL and Teradata.

<p>Enhanced User Experience</p> <ul style="list-style-type: none">Builder GAPattern ProfilingFuzzy JoiningColumn LineageSelf-Service Operationalization	<p>Improved Performance & Scale</p> <ul style="list-style-type: none">Photon Compute Engine GASpark 2.0	<p>Cloud Deployment & Data Source Connectivity</p> <ul style="list-style-type: none">Relational (Oracle, MySQL, Teradata)Cloud (AWS, Microsoft Azure, Google Cloud Platform)API Connectivity
--	---	---

“ We’re seeing tremendous demand for solutions that can put data preparation capabilities into the hands of business users, where the requirements and desires of analytic outcomes are best understood. Trifacta has established itself in the fast-growing self-service data preparation market and is continuing to build meaningful differentiation into their product as evidenced by the v4 release. Making the process of data wrangling easier and faster for a wider set of sources and deployment environments is critical to enterprise adoption. ”

STEWART BOND
Research Director, IDC

“ At Nordea Bank, we are constantly striving to improve the timeliness, accuracy and level of trust in our data to internal and external stakeholders. Trifacta v4 will enable us to involve our business subject matter experts more efficiently than ever before. This has allowed us to fundamentally reduce time to market and cost of managing data while demonstrably increasing the quality of our data products. ”

ALASDAIR ANDERSON
Executive Vice President of Data Engineering, Nordea Bank

Enhanced User Experience

A core focus area of the v4 release is to enrich Trifacta's unique data wrangling user experience by offering a new workflow for building data preparation steps. The addition of Builder to the Trifacta interface augments the ability of users to wrangle data without



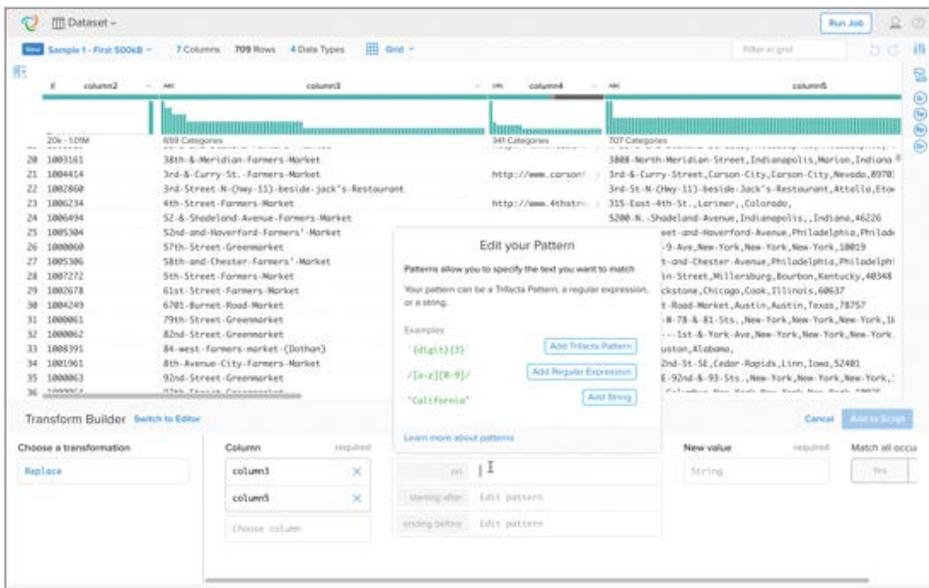
Builder Highlights

New menu-driven workflow for building wrangling steps

Intelligently breaks down the elements of each wrangling operation

Provides extensive descriptive guidance, examples and intelligent suggestions

the need to utilize scripts. Builder is designed to guide users through complex data wrangling tasks, providing greater ease-of-use whether simply selecting a suggested transform or using drop-down menu options to build wrangling steps from scratch. With Builder, the process of preparing data is dramatically simplified by intelligently breaking down the steps of each wrangling task to enhance how non-technical users handle common and complex data.



Builder screenshot

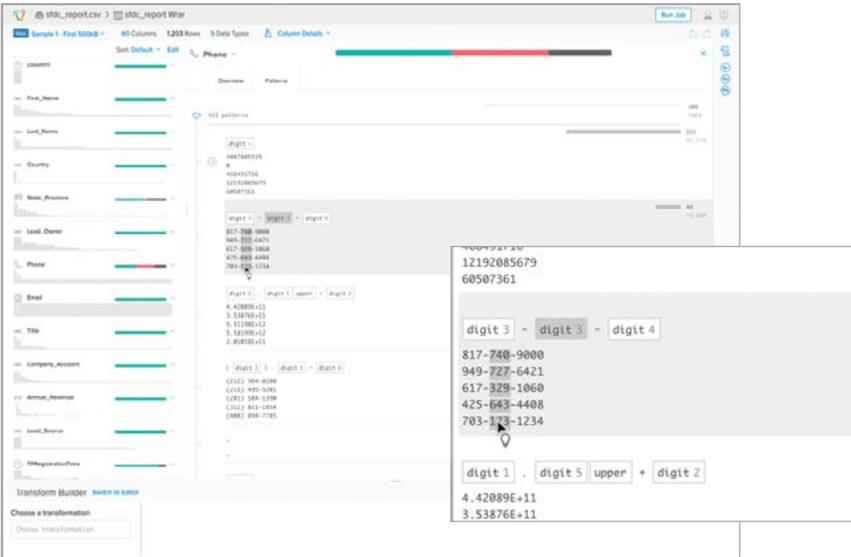
The v4 release also expands upon Trifacta's blend of data visualization and machine learning to guide users through common data wrangling tasks. With pattern profiling, users visualize common and anomalous text patterns that are automatically detected within each column. The addition of fuzzy join allows users to blend together disparate data sources with similar values but non-exact matches. v4 also features the debut of column lineage, a breakthrough visual technique to expose the lineage of how each attribute or column within a dataset originated. With operationalization, v4 allows end users to set and manage end-to-end data wrangling workflows in a completely self-service process.

“ At Sanofi, a key corporate strategy is improving our processing of data across technical groups to provide more concise treatment, improve operational efficiency and reduce security risks. Trifacta is a core part of our success because it gives the Infrastructure Management Team the ability to manage large, diverse data sets and wrangle them into the formats we need for analysis. We're excited about the release of v4 and especially how Builder will enable a broader set of users within Sanofi to intuitively prepare data in a simple, guided workflow. We hope to see more groups and departments use Trifacta moving forward for their data wrangling as we move to make it a service on our data analytics platforms. ”

JASON STOUTE
Senior Manager of Infrastructure
Architecture, Sanofi

“ Trifacta v4 represents our most significant release since the launch of the company. From the beginning, our goal has been to provide a self-service data preparation solution that helps customers connect their big data strategy to business value. As the leader in data wrangling, we're excited about the innovations v4 will deliver to the more than 3,500 companies using our products today. ”

ADAM WILSON
CEO, Trifacta



Pattern Profiling

Improved Performance & Scale

The latest version delivers greater performance and scale for working with data directly within the Trifacta application, and an optimized in-memory data processing engine for data sets that don't require parallel processing. The general availability of the Photon Compute Engine enables users to wrangle a 100x larger volume of data on-the-fly directly within the application while still maintaining the fluid experience and immediate feedback, both of which are core to Trifacta's user experience.

For files, Photon enables users to transform entire data sets completely on-the-fly within the application, and also integrates seamlessly with Trifacta's intelligent execution architecture, complementing existing data processing engines Spark and MapReduce. Photon was specifically built to underpin Trifacta and provides unmatched performance and scale for data wrangling use cases when compared to other interactive computing engines. As part of v4, Trifacta has also enhanced support for executing transformations at scale, leveraging the Spark data processing framework by adding support for Spark 2.0.



INTELLIGENT EXECUTION



“As an analyst, I spend much of my time exploring and refining data sets, running analysis, and examining the outcome to find the best solution to the business problem in front of me. The workflow is extremely important to my process. Delays and interruption can lead to hours of lost time on a project. With Trifacta, the data wrangling process is seamless, making it much easier for me to be productive and efficient. The addition of Photon improves upon what is already a great user experience by allowing us to interactively work with greater volumes of data while maintaining the same fluid workflow.”

MIKE RIEGLING
Supply Chain Data Analyst, PepsiCo

“We're seeing tremendous growth in the enterprise adoption of Microsoft Azure for critical analytics and business intelligence processes. A challenge customers mention to us is the need for a more effective process for cleaning and joining together diverse data. With Trifacta's added support and integration with Microsoft Azure Storage and Microsoft HDInsight as part of their v4 re lease, customers will now be able to accelerate these analytics processes with an industry-leading data wrangling solution for the cloud.”

TIFFANY WISSNER
Head of Big Data Marketing, Microsoft



Deploy on EC2

Support for S3 & Redshift as input/output sources



Deploy on GCE

Support for Cloud Storage & BigQuery as input/output sources

Processing via Google DataFlow



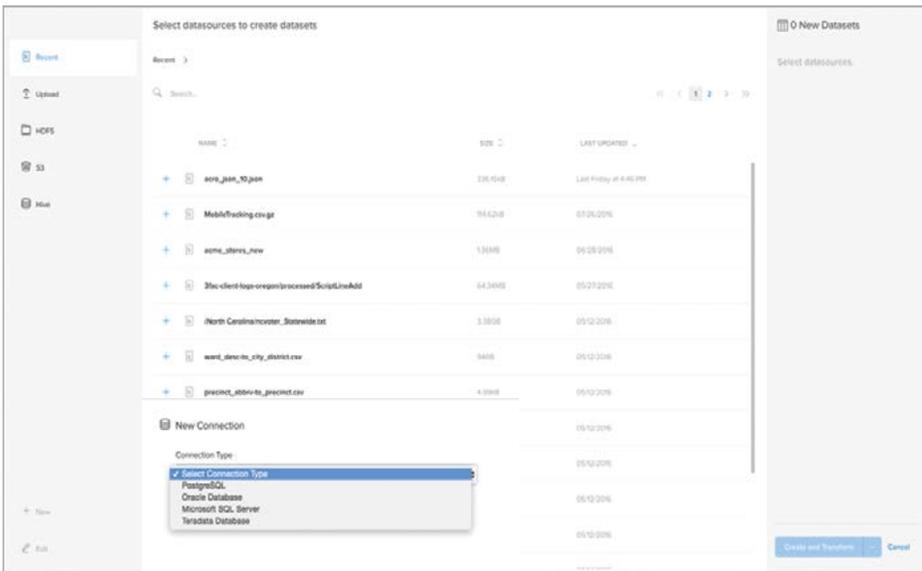
Deploy on Azure HDI

Support for Azure Blob Storage as input/output sources

Extended Cloud Deployment and Data Source Connectivity

With v4, customers benefit from expanded support for deploying Trifacta in the cloud through integrations with Amazon Web Services, Google Cloud Platform and Microsoft Azure. For Amazon Web Services, Trifacta provides integration with Amazon S3 and Redshift as input and output sources and deployment on EC2. Trifacta v4 also supports the Google Cloud Platform ecosystem with support for Google Cloud Storage and BigQuery as input and output sources, data processing via Google Dataflow and deployment on Google Compute Engine. The Microsoft Azure cloud platform is also supported in v4. Trifacta adds support for deployment on Microsoft Azure HDI and can integrate data from Azure Blob Storage.

Trifacta has also expanded support for creating live connections to common relational sources such as Microsoft SQL Server, MySQL, Oracle, PostgreSQL and Teradata. Unlike approaches that force customers to make copies of data prior to preparation, Trifacta creates a live connection, streaming in live data from external sources to incorporate directly into the wrangling process. v4 also includes the initial release of Trifacta's connectivity API, giving customers and partners the ability to seamlessly integrate Trifacta with external data and services.



Live connections screenshot

To meet Trifacta team members and see a live demo of the v4 release, visit Trifacta at Strata + Hadoop World in New York at booth 539 from September 26-29, 2016.

About Trifacta:

Trifacta, the leading data wrangling solution for exploratory analytics, significantly enhances the value of an enterprise's big data by enabling users to easily transform and enrich raw, complex data into clean and structured formats for analysis. Leveraging decades of innovative work in human-computer interaction, scalable data management and machine learning, Trifacta's unique technology creates a partnership between user and machine, with each side learning from the other and becoming smarter with experience. www.trifacta.com

Additional Resources:

[Learn more about Trifacta: trifacta.com](http://www.trifacta.com)

[Trifacta at Strata + Hadoop World NYC 2016: trifacta.com/strata-nyc-2016/](http://www.trifacta.com/strata-nyc-2016/)

[Read the blog on the v4 launch: https://www.trifacta.com/blog/our-most-significant-release-yet-for-any-user-any-data-and-any-cloud/](https://www.trifacta.com/blog/our-most-significant-release-yet-for-any-user-any-data-and-any-cloud/)

[Read more on the Trifacta blog: trifacta.com/blog](http://www.trifacta.com/blog)

[Follow us on Twitter: twitter.com/trifacta](https://twitter.com/trifacta)

[Become a fan on Facebook: facebook.com/trifacta](https://www.facebook.com/trifacta)

[Connect on LinkedIn: linkedin.com/company/trifacta](https://www.linkedin.com/company/trifacta)