Data Blending

Things you can do with the right data blending tool

- Combine data – structured, unstructured & streaming sources.

- Gain truer, more comprehensive insights.

- Simplify data prep and reduce overall prep time.

- Gain control over data via error handling & recovery, validation, and pipeline management options.

- Democratize data - Enable self-service data preparation and analysis.

- Build actionable data sets with a few simple clicks.

- Customized data pipelines to fit business needs.

Overview

Data blending allows you to blend or mix data in any way you need to comprehensively answer complex business questions. It doesn't matter if data resides in the cloud or in your database, streams from the IoT or social media, or is shared or bought through third-party sources such as partners or data brokers. You can blend it all easier and faster using automated data blending toolsets than you can through traditional means requiring specialized talent, labor intensity, and more time than any company can spare.

With the right data blending toolsets, your organization can access, integrate, cleanse, and consume data in record time and with a high degree of accuracy and security.

Not all data blending tools are created equal, however. Some lock you into proprietary technologies while others use open-stack based technologies and components that don't limit your options. This is why the open source movement is so strong in big data today.

The best strategy is to select data blending toolsets that offer the most interoperability options and the most flexibility in analytics while also delivering the fastest and safest automated blending capabilities.

Data Blending Is the Process of...

1. Gather data from various sources

2. Combine useful data into a functioning data set

3. Eliminate unnecessary data

4. Enable heterogeneous data to coexist in a common data repository

5. Gain new insights

Data Blending Solution

Accelerate your time-to-insight with our unique data blending process. Our open source based process replaces a traditional ETL approach with a modern self-service driven approach. Our data blending approach provides data quality controls, including baked-in security, superior workload control via pause, quick resume, recover and other data pipeline controls.

Completely user-friendly with no special skills required, users can easily execute everything through an intuitive user interface. With a few clicks, you can blend diverse data sets irrespective of format or the physical location of your data.

Specifically, our comprehensive toolset does the following:
- Combines disparate data stores & formats
- Builds large actionable datasets easily
- Enables real-time decision making
- Drives accurate strategies - validate data
- Provides UI-driven, ready-to-use data transformations
- Provides large-scale, distributed & iterative ML algorithms
- Allows reusability of scripts as user defined functions in data pipelines
- Simplifies the design of complex workflows using a 'drag-and-drop' approach
- Monitors and measures data quality against business requirements
- Eliminates errors at the onset and reduces error remediation time
- Provides numerous pipeline management options like pause, stop, recover, resume, cleanup, schedule, partial execution and more…
- Empowers all: business analysts, data stewards and data scientists

Impetus is focused on creating big business impact through Big Data Solutions for Fortune 1000 enterprises across multiple verticals. The company brings together a unique mix of software products, consulting services, Data Science capabilities and technology expertise. It offers full life-cycle services for Big Data implementations and real-time streaming analytics, including technology strategy, solution architecture, proof of concept, production implementation and on-going support to its clients.

To learn more, visit www.impetus.com or write to us at inquiry@impetus.com.