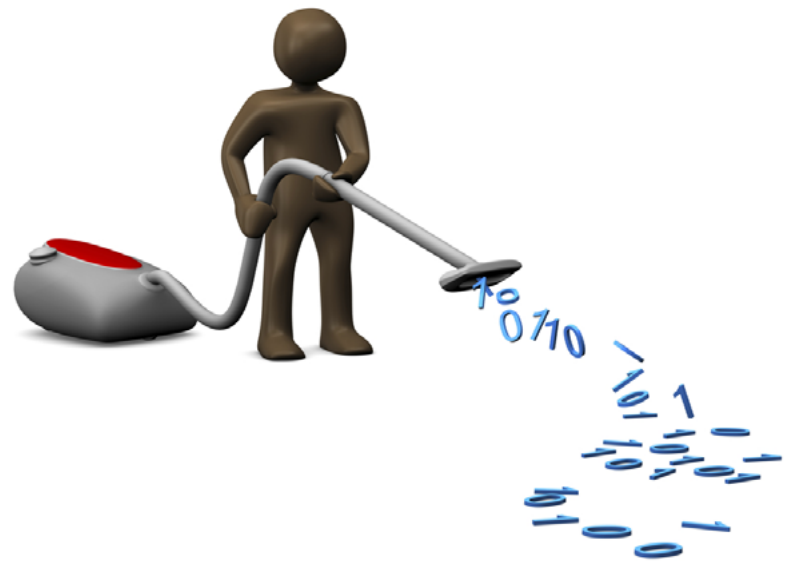## BellaDati ETL Summary

BellaDati enables to combine data from different data sources, using 100+ ready made connectors while enriching, cleansing, and transforming the data.

Data cleaning

BellaDati has advanced features for data cleaning, that is major part of the ETL process. It deals with detecting and removing errors and inconsistencies from data in order to improve the quality of data. Data quality problems are present in single data collections, such as files and databases, e.g., due to misspellings during data entry, missing information or other invalid data. When multiple data sources need to be integrated, e.g., in data warehouses, federated database systems or global web-based information systems, the need for data cleaning increases significantly. This is because the sources often contain redundant data in different representations. In order to provide access to accurate and consistent data, consolidation of different data representations and elimination of duplicate information become necessary.

Simple data cleaning for standard volume of data can be performed during upload process. For big data with more demanding cleaning functions BellaDati ML Studio shall be used.

BellaDati features for step nu 1 in data cleaning - Data analysis

BellaDati offers wide range of functions and high flexibility for Data analysis. In order to detect which kinds of errors and inconsistencies are to be removed, a detailed data analysis is required. In addition to a manual inspection of the data or data samples, BellaDati offers the range of readily available functions that should be used to gain metadata about the data properties and detect data quality problems. Developers and power users can as well use the development environment within our ML Studio to select internal and external data sources and write their own scripts for data cleaning.

BellaDati features for Step nu 2 in data cleaning: Definition of transformation workflow and mapping rules

Belladati Advanced Analytics Tool And Platform supports the cleaning of the data from high number of data sources, with high degree of heterogeneity and the "dirtiness" of the data where a large number of data transformation and cleaning steps must be executed.

ETL (Extraction, Transform and Loading) and data cleaning in BellaDati
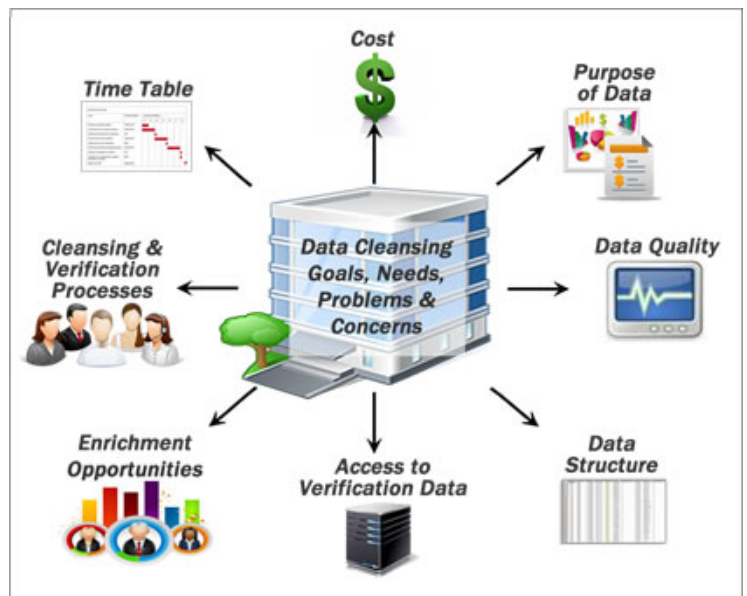
It supports for example:

- a schema translation that is used to map sources to a common data model
- a relational representation that is typically used for data warehouses
- correction of single-source instance problems and preparation the data for integration
- schema/data integration and cleaning multi-source instance problems, e.g., duplicates.

The schema-related data transformations as well as the cleaning steps are specified by declarative query and mapping language to enable automatic generation of the transformation code. In addition, it is possible to execute user-written cleaning script and special purpose features during a data transformation stage.

BellaDati supports multilevel transformation – stageing, joining of dataset and metadata transformation.
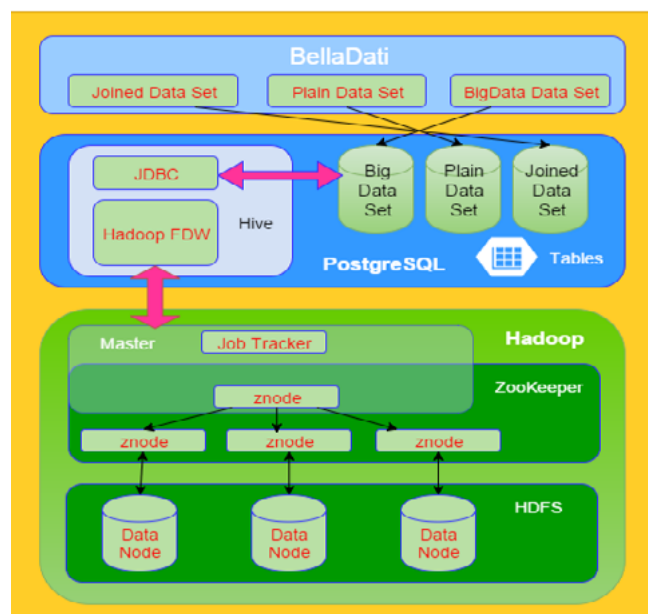
After inconsistencies are removed, the transformed data replace the dirty data in Belladati database. For data warehousing, the cleaned data are available.

BellaDati supports Revert changes in job in own database.Database connections can be used with all jobs. This enables commits and rollbacks on a job level.

Big Data cleaning

BellaDati provides Distributed data transformation and computing as well as load balancing. BellaDati provides load balancing of data within transformations and over multiple cluster nodes when using transformation clustering. BellaDati is highly distributable and is using this approach with heavy data processing and transformation. It means, that the transformation can be optionally executed in parallel mode and distributed over BellaDati transformation nodes cluster.

BellaDati is Agile BI, Pure Web Application, Complete BI Solution for Manufacturing

mongoDB

SAP HANA

facebook

salesforce

hadoop

Cloud or On-Premise
Deploy in the cloud or install On-Premise. All you need is web-browser or mobile device.

Analytics Platform
with API and SDK for creating Embeded Data Solutions. Deliver BellaDati as Your Brand.

Mobile BI
Pull reports out of your pocket on- or off-line with BellaDati Mobile for iOS and Android. Filter, share and comment or harvest new data on the go.

Predefined BellaApps
Take the fastest paths. Leverage pre-built dashboards, reports and measures to cut down costs and time.

BIG Data and Predictions
In Memory Real time Statistics and Prediction with SAP HANA.

NO-SQL/SQL biz data integration and analysis
from more than 100 databases, APIs, Cloud Solutions and Social Networks.

Social network for Data Discovery
Realtime collaboration on the insights. Search, share, copy, export, get notified on data changes.

BellaDati Cloud

BellaDati On-Premise

Ask for more use cases...

United States
+ 1(415) 800 3540

Asia
(+65) 3159-0549

Europe
+420 702 023 558

BellaDati