

DATA, DATA CLEANING

# How to Choose the Right Data Quality Tool [06/2018 UPDATE]

JUNE 25, 2018 · 4 MINUTE READ

Data cleaning is part of a greater effort to **achieve the highest data quality possible in used in business decisions and operations**. It requires organizational effort and participation throughout a business and when done correctly, can help to provide valuable insights and analytics for decision making. A few additional benefits associated with data cleaning include:

- Streamlined business practices
- Increased productivity
- Faster sales cycle
- Better analytics

Given the **ever growing quantity of data for many businesses**, automation is required in **data cleaning**. The right data tool can fill in these gaps and manage a number of issues automatically before they have a chance to become truly problematic. This can ultimately help businesses to become more efficient and more profitable in their efforts.

Choosing the right data cleaning tool for your organization is essential to getting the most utility for your investment. To help in your decision making, this post answers the following:

- [What are some criteria for choosing the right data cleaning tool?](#)
- [What are the key features in data quality tools?](#)

- What are some major tools that have been developed for data cleansing?

## Criteria for Choosing a Data Quality Tool

The right data cleaning practices can have a **huge positive impact across an organization**, so it's worthwhile to take the time to choose the right tools to support it. In the case of large or complicated datasets, outsourcing the entire process to a third party could also be considered.

Some of the criteria that should be included when choosing a tool are:

**Price:** Is it a subscription or one-time fee? Are there addons that will cause the price to inflate?

**Support:** A strong support team can be a big factor in decision making.

**Usability:** Not only in terms of analytical uses/IT users that are working for setup/implementation, but will business users need this?

**Scalability:** Whether or not your tool will be able to keep up as your data sources grow and evolve; and how easy it will be to make upgrades and changes down the line

**Features:**

- **Auditing capabilities:** Being able to see when and where changes were made to a record is important for internal and external auditing and compliance concerns.
  - **Compatibility/integrations:** Having a tool that can work with all the data sources that your business utilizes for daily activities.
  - **Cloud vs on-premise:** A cloud based option opens up many more choices for smaller businesses with limited hardware resources.
  - **Metadata support:** Metadata is important for avoiding 'insight gaps' where valuable data is that could be used for analysis becomes separated from data scientists and other business users
- **Compatibility with different sources:** How many, and what sources, can data be taken

from? How long does it take to run any processes or to prepare for them?

- **Batch processing capabilities:** Being able to program ahead of time regular data cleaning practices (?) can help to ensure the ongoing quality of your data

## Considerations for Different Sized Businesses

The size of your business will play a major role in helping you to choose the right tool. There are three general categories that will have differing needs:

1. **Small businesses with 10 employees or less:** Businesses of this size generally do not have a need for extensive data cleaning tools.
2. **Medium businesses with 10-100 employees:** At a midsize level, businesses begin to encounter an interesting problem where there's enough data to need the tools and effort to keep it clean, but putting together an entire team isn't realistic. Subsequently, it is important to choose a robust tool that can help to fill in the 'gaps'.
3. **Large businesses with 100-500 employees:** At this level, the volume of data going in and out of an organization will generally mandate a dedicated team to ensure data quality. However, choosing a high quality tool can help to simplify their jobs and allow them to focus on key quality related tasks.

## Common Functionalities of Data Quality Tools

No matter what tool you ultimately choose for your organization, there are several common functionalities that can be found in a wide range of tools:

- **Data profiling:** Scanning through data to find patterns, missing values, character sets and other essential characteristics. This will enable to tool to later identify data as irregular.
- **Data elimination:** The removal of duplicate data and also data that doesn't meet the desired profile.
- **Data transformation:** For erroneous data that is valuable, it can be transformed into 'good' data through correcting typos, standardization, and normalizing numeric values to fall between minimum and maximum values.

- **Data standardization:** Putting data into a common format for easier analysis.
- **Data harmonization:** Similar to standardization, this practice takes data from a range of sources and puts them into a common format. Unlike standardization, which is about conformity, harmonization is about consistency.

## Data Quality Tools Overview

Every day the number of data cleaning tools available on the market grows. Some of common vendors include:

SHOW  ENTRIES

SEARCH:

SHOW  ENTRIES

SEARCH:

Name	Founded	Status	Number of Employees
OpenRefine	2012	Open source	N/A
Trifacta Wrangler	2012	Private	11-50
TIBCO Clarity	1997	Private	1,001-5,000
IBM Infosphere Quality Stage	1911	Public	10,001+
Foxtrot	2014	Private	11-50
Symphonic Source Cloudingo	2010	Private	11-50
Quadient Data Cleaner	2014	Public	1,001-5,000
Data Ladder	2006	Private	11-50
Winpure	2003	Private	11-50
Nmondal Solutions Datamartist	2008	Private	2-10

Showing 1 to 10 of 10 entries

◀ Previous Next ▶

Showing 1 to 10 of 13 entries

◀ Previous Next ▶

Choosing a data quality tool can seem intimidating, but with some careful research and the advice of a trusted 3rd party, it can ultimately be one of the most effective methods of achieving high quality data.

Interested in topics like knowledge management, artificial intelligence (AI), and machine learning? Be sure to follow [our blog](#) where we regularly posts update on topics like these and more. Need an AI tool or vendor to solve a specific problem? Check out our [use case library](#) and find the perfect solution.

### [Featured image source](#)

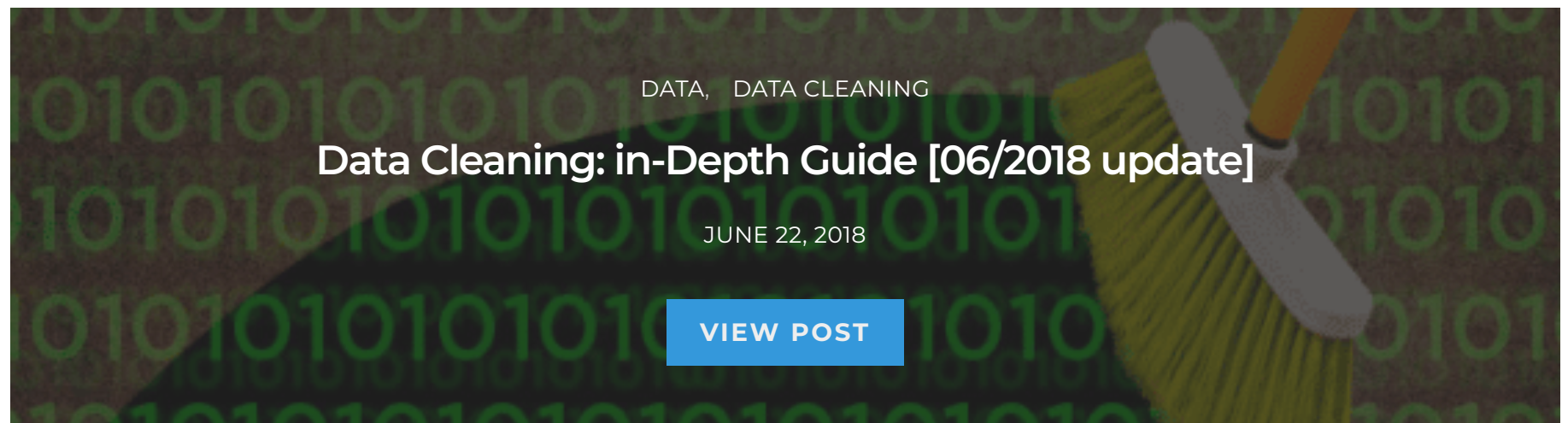
 LIKE

 TWEET



### Related Topics

[#ANALYTICS](#) [#DATA](#) [#DATA CLEANING](#) [#TOOLS](#)



### YOU MAY ALSO LIKE



DATA, DATA VISUALIZATION

## 34 Data Visualization Tools for Better Business

MAY 31, 2018

DATA, DATA VISUALIZATION

## Improving Understanding with Data Visualizations

JUNE 4, 2018

DATA, DATA CLEANING

## Data Cleaning: in-Depth Guide 106/2018



### LEAVE A REPLY

YOUR EMAIL ADDRESS WILL NOT BE PUBLISHED. REQUIRED FIELDS ARE MARKED \*

YOUR COMMENT \*

NAME \*

EMAIL \*

WEBSITE

POST COMMENT

Enter Keyword

SEARCH

## CATEGORIES

AGI

AI COMPANY

AI CONSULTING

AI STARTUPS

AI VENDORS

AIQ

ALPHABET

ANALYTICS

ANALYTICS

APPLIEDAI

ART

ARTIFICIAL INTELLIGENCE

AUTOMATED MACHINE LEARNING

AUTONOMOUS DRIVING

**BLOCKCHAIN CONSULTING**

---

**CARLABS**

---

**CHATBOT**

---

**CONSULTING**

---

**CONTENT AUTOMATION**

---

**CREDIT SCORING**

---

**CUSTOMER SERVICE**

---

**DATA**

---

**DATA CLEANING**

---

**DATA INTEGRATION**

---

**DATA SCIENCE**

---

**DATA  
VISUALIZATION**

---

**DECEPTION TECHNOLOGY**

---

**DRUG DISCOVERY**

---

**DYNAMIC PRICING**

---

**E-COMMERCE**

---

**ETL**

---

**FACEBOOK**

---



**FINTECH**

---

**FRAUD DETECTION**

---

**GUIDELINES**

---

**HEALTHCARE**

---

**HR**

---

**INFORMATION SECURITY**

---

**INTELLIGENT CALL ROUTING**

---

**IT**

---

**KNOWLEDGE MANAGEMENT**

---

**LEAD GENERATION**

---

**MANUFACTURING ANALYTICS**

---

**MARKETING**

---

**MARKETING ANALYTICS**

---

**NONPROFIT**

---

**PERSONALIZED CARE**

---

**PREDICTIVE MAINTENANCE**

---

**PREDICTIVE SALES**

---

**RECOMMENDATION SYSTEM**

---

**ROBOTIC PROCESS AUTOMATION (RPA)**

---

**SALES**

---

**SALES ANALYTICS**

---

**SELF CHECKOUT SYSTEM**

---

**SURVEY&REVIEW ANALYTICS**

---

**UNCATEGORIZED**

---

**VERITONE**

---

**VOICE RECOGNITION**

---

**WEBSITE PERSONALIZATION**

---

## **RECENT POSTS**

How to Choose the Right Data Quality Tool [06/2018 UPDATE]

---

Data Cleaning: in-Depth Guide [06/2018 update]

---

Top 10 Knowledge Management Tools [06/18 update]

---

Fraud Detection: In-Depth Guide [06/18 update]

---

15 Fraud Detection Software: A Complete List [06/18 update]

---

## **APPLIEDAI**

ABOUTUS

We democratize Artificial Intelligence.

We are building a transparent marketplace of companies offering B2B AI products & services.

We allow companies to look beyond marketing speak, beyond huge marketing budgets to understand how they can use AI in their businesses and evaluate AI services objectively.

VISIT WEBSITE



EXPLORE APPLIEDAI

---

USE CASE GUIDES

---

AI STARTUPS AND SOLUTION PROVIDERS

---

PLAN YOUR AI JOURNEY

---

VENDOR SELECTION

---

ABOUT

---

BLOG

---

ABOUT US

---

CONTACT US

---



