
Case Study: Mixed Data Sets with Chinese



Executive Summary

The Dataactics solution was selected for use in an Asian Trade Development Council to replace an outdated manual process of maintaining data. Dataactics was chosen in part because of its unique ability to handle all variations of Chinese characters within the same data set, and also because of its high level of reporting, tracking and reviewing changes made to the data, ensuring low risk. The solution has reduced the time taken in making data available from months to days.

Speedy and accurate capturing of visitors' registration information from trade fairs and data from other sources is critical to make available the company information collected for marketing and business matching use.

However data analysts within the Trade Development Council were taking months to import, de-duplicate and clean data. Data management was done manually taking up the resources of skilled staff. This costly and laborious process was slowing down the entire organization.

The Customer

The Trade Development Council organises 30 world-class international trade fairs annually, attracting some 500,000 visitors. It maintains a databank of nearly 9 million records collected from various countries. The data is used for general marketing purposes and to support specific business matching activities. It is critical that the information captured in the databank be up-to-date, and maintained to a high quality standard.

The Solution

The Dataactics solution replaced the old de-duplication module which was no longer able to cope with new business and market conditions within the Trade Development Council's Data Hygiene Process.

Dataactics provided a highly configurable de-duplication solution with sophisticated logic to automate the whole data cleansing process, providing good quality data quickly and cost effectively.

The Challenge

The major source of company information is from the 30 world-class international trade fairs organised by the Trade Development Council annually.

Dataactics effectively handled mixed data sets including Chinese and other Asian languages providing the Trade Development Council with the ability to track the changes to the data and manually review lower confidence matches by way of the Dataactics Master Record Manager. The solution has reduced the time taken in making accurate data available from months to days.

UK Headquarters

9 Heron Avenue, Heron Road, Belfast BT3 9LF
Tel: +44 (0) 28 9073 8854 Fax: +44 (0) 28 9045 0665
Email: info@dataactics.com www.dataactics.com

USA Headquarters

One South Dearborn Suite 2100, Chicago, IL 60603
Tel: +1 312 212 4363 Fax: +1 312 212 4401
Email: info@dataactics.com www.dataactics.com

“We needed a solution that would yield immediate benefits in accelerating the accessibility of accurate data. The team at Dataactics was not fazed by handling different character sets and their solution will significantly reduce the time we need to make essential information available very quickly.” **Trade Development Council,**

The Benefits

- Provides a user friendly web interface with good response times to speed up the de-duplication process where human intervention is unavoidable.
- Provides improved data security.
- Reduces the turnaround time needed to de-duplicate new company data from external sources from weeks to days.
- Automates the importing of new company data into the database, releasing existing skilled resources for higher value work.
- Performs white and black lists via a Master Record Manager module, ensuring the right data is retained/de-duplicated.
- Facilitates the outsourcing of the data hygiene work to third parties to further speed up the data handling process.

UK Headquarters

9 Heron Avenue, Heron Road, Belfast BT3 9LF
Tel: +44 (0) 28 9073 8854 Fax: +44 (0) 28 9045 0665
Email: info@dataactics.com www.dataactics.com

USA Headquarters

One South Dearborn Suite 2100, Chicago, IL 60603
Tel: +1 312 212 4363 Fax: +1 312 212 4401
Email: info@dataactics.com www.dataactics.com