

BUSINESSOBJECTS

UNIVERSAL DATA CLEANSE (UDC)

Universal Data Cleanse

- Adds new parsing classification capabilities that are industry- and region-specific
- Customizes rules that build on dictionaries
- Loads new data standards from BusinessObjects Data Insight XI
- Supports mixed language product description

UDC Supports Organizational Initiatives

- Mergers and acquisitions
- Spend analysis
- Master data management
- International data processing
- Global coding systems
- Unstructured and semistructured data processing

SEE ALL COMPANY DATA IN THE SAME LIGHT

Universal Data Cleanse (UDC) is an option of the BusinessObjects™ Data Quality XI solution suite from Business Objects, an SAP company. UDC greatly eliminates the knowledge shadows cast by noncustomer or regional-specific data duplication or discrepancies. With UDC, customers access an automated data standardization solution from Business Objects that helps identify abnormalities not even imagined by those processing the data. With UDC, your enterprise can now obtain a single, reliable, panoramic view of products, supply chains, and other key information.

Every company struggles with the myriad product and customer data that more often than not is represented in a variety of ways across disparate systems or departments. For example, a product name such as “LIGHT BULB-120 VOLT 100 WATT” may be entered multiple times with different capitalizations, spelling, numbers, abbreviations, and hyphenation – resulting in costly inefficiencies and lost time for the organization, which is forced to cross-check databases, find the defective data, and weed it out to ensure effective business operations. With UDC, your organization is reassured that its data – regardless of type or where it’s stored – is clean, reliable, and actionable.

DATA CLEANSING TAILORED TO YOUR BUSINESS

Typically, organizations feed all types of information to business intelligence (BI) or operational applications, which further process the data for activities such as analysis, inventory management, invoicing, or decision-making. Until now, only customer data could easily be cleansed and standardized. UDC extends the data cleansing capabilities beyond your customer information to include any type of data your organization needs to process. Industries such as financial, pharmaceutical, or manufacturing can now process their custom data sets and get a 360-degree view of products, supply chains, and vendors. An add-on option for BusinessObjects Data Quality XI, UDC offers the following capabilities:

- Process bulk loads – incorporate product specifications coming directly from business analysts
- Add new classifications – create industry or region-specific identifiers
- Add custom output – easily communicate results in your local language
- Store product data specifications in repository – support the master data management (MDM) goals of your organization
- Use preloaded dictionaries – colors, sizes (in English, French, German, Italian, Portuguese, Spanish), and weights (U.S. and Metric)

UDC USE-CASE SCENARIOS

Product Data Parsing and Standardization

One typical use-case scenario of UDC is to support structured and semistructured data processing. A typical goal for a manufacturer, for example, may be more consistency between the descriptions of items on the Web site and items in the catalog. To achieve consistency, the company would need to parse out the descriptions of each item and understand the relationships between them so it can take out any duplicate entries and consolidate to a consistent format. An example of parsing product data follows.

UDC Can Be Extended to Support Industry Standards

- United Nations Standard Products and Services Code® (UNSPSC)
- Export Control Classification Number (ECCN)
- North American Industry Classification System (NAICS)
- Electronic Commerce Code Management Association Open Technical Directory (eOTD)
- International Standard for the Classification of Product and Services (eCl@ss)
- International Organization for Standardization (ISO)
- National Stock Number Codes (NSN) – military
- And More!

Customer input string: Glove ultra grip profit 2.3 large black synthetic leather elastic with Velcro Mechanix Wear

This graphic shows a complex input text and parsing routine. Product data parsing is achieved by breaking the input down into smaller pieces, assigning a meaning to the pieces, matching the pieces to the rule, and creating output from the rule matches.

Figure 1: Complex Input Text and Parsing Routine

Custom output fields	Standardization
Product Category	Glove
Size	Large
Material	Synthetic Leather
Trademark	Pro-Fit 2.3 Series
Cuff Style	Elastic Velcro
Palm Type	Ultra-Grip
Color	Black
Vendor	Mechanix Wear

International Data Parsing and Standardization

Another typical use-case scenario of UDC is processing international data. For an increasing number of organizations, it's important to be able to create intelligent parsing standardization processes – regardless of what language the original data is in. Equally important is the ability to present data in the local language and in the appropriate fields. With UDC, not only are the fields named appropriately, but your users also can modify the rules and new names to optimize the process for your organization's specific region or language.

Customer input string: Bernaś Bronisław s. Józefa

This graphic depicts the ability of UDC to adjust output parsing fields for unique regional attributes.

Figure 2: Use UDC to Adjust Output Parsina Fields

English fields	Western parsing		Polish fields	Polish parsing
Extra	Bernas	Middle name is now son of	Nazwisko	Bernaś
First_name	Bronislaw		Imie	Bronislaw
Mid_name	s.		Pokrewienstwo	syn
Last_name	Jozefa		Rodzic	Józef Bernaś
Gender	Unassigned		Plec	Strong male

COLLABORATES WITH BUSINESSOBJECTS DATA INSIGHT

The UDC option in BusinessObjects Data Quality XI seamlessly integrates with the entire data quality product suite, allowing our customers to take advantage of additional functionality. Most notably, the collaboration between the UDC option and BusinessObjects Data Insight XI – our data profiling tool – greatly enhances process efficiencies between business analysts and IT users. Using these products in tandem allows your business analysts to conduct ongoing contextual analysis of input and output, and directly add, modify, or delete product classifications. Those changes can then be directly loaded into UDC for your IT users to incorporate into the data cleansing cycle. This means your IT users no longer need to guess at discerning the appropriate business rules – instead, they can work far more effectively.

PROVIDING THE MOST COMPLETE DATA QUALITY APPROACH

With UDC and Data Quality XI from Business Objects, business and IT users now have the power to manage data quality standards across deployments – in a way that's consistent, predictable, and easy to use.

businessobjects.com



© 2008 Business Objects. All rights reserved. Business Objects owns the following U.S. patents, which may cover products that are offered and licensed by Business Objects: 5,555,403; 5,857,205; 6,289,352; 6,247,008; 6,490,593; 6,578,027; 6,831,668; 6,768,986; 6,772,409; 6,882,998; 7,139,766; 7,299,419; 7,194,465; 7,222,130; 7,181,440 and 7,181,435. Business Objects and the Business Objects logo, BusinessObjects, Business Objects Crystal Vision, Business Process On Demand, BusinessQuery, Crystal Analysis, Crystal Applications, Crystal Decisions, Crystal Enterprise, Crystal Insider, Crystal Reports, Desktop Intelligence, Inxight, the Inxight Logo, LinguistX, Star Tree, Table Lens, ThingFinder, Timewall, Let there be light, Metify, NSite, Rapid Marts, RapidMarts, the Spectrum Design, Web Intelligence, Workmail and Xcelsius are trademarks or registered trademarks in the United States and/or other countries of Business Objects and/or affiliated companies. All other names mentioned herein may be trademarks of their respective owners. Part # WS464-B