

# Oracle Customer Data Librarian

**Axel Broecker**

International Atomic Energy Agency **Vienna**

## **Keywords**

*Field report, Customer Data Librarian, Data Quality Management (DQM), Merge, Trading Community Architecture (TCA)*

## **Introduction**

The International Atomic Energy Agency (IAEA) is the world's centre of cooperation in the nuclear field. It was set up in 1957 as the world's "Atoms for Peace" organization within the United Nations family. IAEA works with its Member States and multiple partners worldwide to promote safe, secure and peaceful nuclear technologies. Three main "pillars" or areas of work underpin the IAEA's mission: Safety and Security; Science and Technology; Safeguards and Verification. These are reflected in the Agency's departmental structure: it is made up of the Departments of Nuclear Sciences and Applications, Nuclear Energy, Nuclear Safety and Security, Safeguards and Technical Cooperation (TC).

Until 2011 the IAEA had over 60 different information systems, few of which communicated with one another, and many of which presented data in forms that were difficult to compare. Hence, in 2009 the IAEA began the adoption of Oracle E-Business suite. This would be carried out in four phases or plateaus. Plateau 1, comprising the areas of finance, programme and project implementation, asset management and procurement, went live in 2011. Plateau 2, implemented in 2012-2013, covered the areas of programme and project budgeting and forecasting, as well as the management of the IAEA's contacts: its suppliers, meeting participants and so on.

This field report discusses how IAEA uses Oracle Applications for its contacts management, covering around 250,000 persons and organizations imported from several legacy systems into the central repository of TCA. The legacy systems have either been decommissioned or integrated.

The main solution objectives were

- An improved set of business processes for the acquisition, management, quality control, ownership and sharing of contacts data across the Agency
- A holistic view of contact information and related contact activities to all authorized Agency staff
- An integrated master contact database capturing all relevant information about contacts in AIPS.

The following features of Oracle Customer Data Librarian and Supplier Data Steward were used to achieve the objectives:

- A common TCA model
- Powerful DQM searches to identify potential duplicates in real time and in batch mode
- Merge capabilities for elimination of duplicates and marking of non-duplicates

It was necessary to add some extensions to ensure full support for business processes.

### **Common TCA model**

The core concept of TCA is a *party*. A party is either a person or an organization. Relationships can be maintained between different parties. A party may have a supplier and/or a customer account assigned. The parties with which IAEA deals include:

- Individuals and organization as suppliers
- Individuals as publication customers
- Organization as publication or donor customers
- Meeting participants
- Research institutes, training institutions, fellowship placement institutions
- External project stakeholders, including Technical Cooperation project counterparts, research contract investigators, national liaison officers, national liaison assistants, national coordinators under a cooperation agreement.

During data migration about 165,000 person parties and about 73,000 organization parties were imported into TCA. They represent the history of the last 20 years. Where possible they were matched against existing data created in TCA during Plateau 1.

Currently about 190,000 person parties and about 78,000 organization parties are maintained, of which about 20% are suppliers and about 1% are customers. Roughly 400 users work with this data. Now that all this information is stored centrally, duplicates can be identified and managed.

Parties are accessible to all users and are maintained by them using Oracle Customer Online.

Party creation is the start of the party lifecycle. In this cycle parties can become external team members in Oracle Projects *e.g.* as soon as they have become involved in one of IAEA's more than 2,600 technical cooperation projects. They can also become suppliers if, for example, IAEA pays them travel expenses to attend a meeting on nuclear safety. A party may also become a customer when it orders publications or laboratory samples.

IAEA identified one shortcoming in the TCA data model relating to suppliers: with the exception of employees, TCA only supports organization suppliers. The vast majority of IAEA's suppliers are external persons. A customization was implemented to help with that.

TCA provides a powerful tool, Data Sharing and Security or DSS, to restrict access rights on party definitions. DSS enables the management of who can create, modify, and delete information. In the IAEA, persons representing governments (for example) can be changed only by one dedicated organizational unit. Setting up DSS rules made it possible to implement this requirement.

The protection had some shortcoming regarding the granularity of the protection level. It is possible to control permissions on the level of an entity – normally a page – but not for specific attributes. For instance, the IAEA asked that party names be protected from being updated by the majority of users when the party is a supplier. Nonetheless, those users should be allowed to change other attributes such as dates of birth. On the other hand, descriptive flexfields and UDA cannot be protected by DSS. Both requirements were addressed through a customization.

## DQM Search

Data Quality Management provides various features that work together as a powerful and flexible data transformation and matching tool. These features are integrated with TCA and can be used by any Oracle E-Business Suite application to search and identify duplicates.

For IAEA it is important to be able to find existing parties – especially persons – when using party search pages. Even when a party is created on party creation pages, a duplicate check is executed automatically. When similar parties are found, the user is presented with an option to choose one of those or continue with the creation process. This has been achieved by applying DQM rules. The rules were defined to provide sufficient search possibilities by configuring IAEA specific search parameters tailored to specific user groups. To be able to use custom attributes defined as UDA (such as passport number) custom transformation procedures were added. It is not possible to search by supplier number, therefore some customization was developed.

Challenges arise from data quality in general (such as typing errors) but also in translating names from other alphabets like Chinese or Cyrillic into the Latin alphabet. Common examples for the first challenge are Tomm instead of Tom or Newten instead of Newton. The Russian name Sergej which is sometimes spelled Sergey or Sergei represents the second challenge. Certain abbreviations of names like Bill instead of William are also common. These lead to situations where search parameters and existing data might not match.

DQM has built-in capabilities to identify similar names using fuzzy logic. These rules can be configured, and they allow the calculation of a matching score. This score indicates how closely the identified record matches the search parameter. Records below a threshold score are not shown. These built-in features covered most of IAEA's needs with a limited development effort.

DQM had some shortcomings regarding the usage of non-English characters in names, for example Jose is not identified as a variation of José. To ensure that both versions are found when searching for a party a custom translation function was developed.

Related to this, it was important to the IAEA that multiple names could be assigned to a party. Organizations all across the world work with the Agency. Frequently they are referred to in the English language representation of their name (Institute of Nuclear Fusion), but sometimes also in their national language name (Instituto de Fusión Nuclear) or their abbreviation (IFN). IAEA needs to have all those names available for search. This has been achieved by using standard functionality.

Another means of identifying duplicates is the usage of System Duplicate Identification (SDI) batches. These batches compare existing parties based on specific DQM rules with each other and identify potential duplicates. For those, merge requests are automatically created.

SDI batches can be used to proactively identify potential duplicates. This ensures a higher degree of data quality and reduces misdirected transactions.

Defining DQM rules for this purpose follows a different logic than those for DQM rules used on screens. To avoid a large number of potential duplicates, rules should be designed with less sensitivity to deviations in name spellings. There are many common names like John Smith or Mohamad Abdul. When spelling differences are taken into account, a large number of potential duplicates will be identified.

To avoid this, it is highly desirable to use additional attributes that might help to identify a person. Date of birth, passport number or email addresses are good options. Address information is not as useful because many persons in contact with the Agency work in changing locations, depending on their assignments.

## Merge

The Party Merge and Account Merge functionalities are part of the Data Quality Management (DQM) tools provided by TCA. Supplier merging capabilities come with AP.

Different applications across the Oracle E-Business Suite have data associated with parties, suppliers and customer accounts. Any references to parties and accounts are updated when parties or accounts are merged. The target party replaces the source party in all affected entities. This ensures data integrity across the Oracle E-Business Suite.

For IAEA it is important that the merge consistently update references in

- Applications Object Library for parties used as supplier contacts
- E-Business Tax for suppliers
- Grants Proposal
- Human Resources
- iSupplier Portal
- Marketing
- CRM Foundation – Resource Manager
- Payables
- Payments
- Projects
- Receivables
- Sourcing
- Purchasing
- TCA

Once a merge has been completed it cannot be undone. It must therefore be ensured that the parties to be merged are identical. A merger and acquisition activity does not necessarily imply a merge activity in Oracle E-Business suite. From a technical point of view they are not duplicates.

Due to the difficulties in the spelling of names, language-specific characters and alternate names, a number of potential duplicates has accumulated, necessitating the use of Oracle Customer Data Librarian.

IAEA wanted UDA and DFF to be merged as well. A customization was developed to achieve this.

### **Axel Broeker**

Internal Atomic Energy Agency  
PO Box 100  
A-1400 Vienna  
Austria

+43(0)1 260024653  
[a.broeker@iaea.org](mailto:a.broeker@iaea.org)  
[www.iaea.org](http://www.iaea.org)