# SIX DESIGN TIPS IN INFORMATICA DATA QUALITY

INFORMATICA DATA QUALITY DESIGN TIPS:
TOP SIX DESIGN TIPS FOR IDQ

BY SALLY MCCORMACK

**The design phase of the Data Quality Management Process is typically the third step in a five step iterative process. In this phase, the developer will take the business rules as defined by the data steward in phase two (Rule Definition) and convert them into data quality useful goals.**

In an upcoming article "Tips for Establishing Data Quality Rules", we will see how the use of collaboration tools pre-built into the software makes the process work more smoothly. In phase two, the developer already has a good head start on what the data steward identified inside the Informatica Data Quality tool. The Developer can then build on top of what the Data Steward has already defined.

## OUTLINED BELOW ARE SIX DIFFERENT TIPS & TRICKS FOR THE DESIGN PHASE OF THE **DATA QUALITY MANAGEMENT PROCESS.**

**1**

### INFORMATICA DATA QUALITY DESIGN TIP: CONSISTENT NAMING AND CODING STANDARDS

When designing rules in Informatica Data Quality, the developers and data stewards will see the same rules. Therefore, it is important to develop consistent naming and coding standards. For example, both the data steward and the developer will understand what "rule_" means while not everyone will understand what "mplt_" means. Therefore, mapplets should be named rule_ if they are used in both the Analyst and Developer tools.

It is a best practice to include anchors and descriptions in DQ mappings for faster modifications, readability and comprehension. Any metadata changes to a source or target object in a mapping (e.g. a new column) can quickly be made if anchor transformations exist immediately following the source object and immediately preceding the target object. These can simply be pass-through transformations. All objects should also have descriptions. The description will be displayed in the Analyst tool and can be useful to data stewards for understanding the data anomalies when viewing profiles. This can also help remind the data steward what rule was applied and whether or not it needs to be updated.

## ② INFORMATICA DATA QUALITY DESIGN TIP: CONSIDER THE ENVIRONMENT

Data Quality or PowerCenter? That is the question – and it almost sounds like something from the BI version of Shakespeare's Hamlet. However, it is a valid question when developing data quality rules and processes. We need to consider the environment that will be leveraged. Will all development be done in Informatica Data Quality (IDQ), or will we need to integrate with PowerCenter? PowerCenter, for example, should be leveraged for improved performance and scalability, reliability, or as part of an ETL process. If there isn't a performance impact, then IDQ can be leveraged alone and integration with PowerCenter is not needed.

## ③ INFORMATICA DATA QUALITY DESIGN TIP: REUSE

Reduce! Reuse! Recycle! These are three common words we hear often in reference to protecting the environment around us. We can use the same approach when it comes to designing in Informatica – design for reuse. Many of the same rules can be leveraged across data domains and verticals.

> **A few examples of rules that can be used anywhere are; uppercasing values and trimming blank spaces.** These reusable rules should be located in a shared location. Having reusable rules will allow you to design multiple mapplets which can then be placed within one mapping. This decreases the level of complexity in the mapping…. and ultimately saves time and reduces errors.
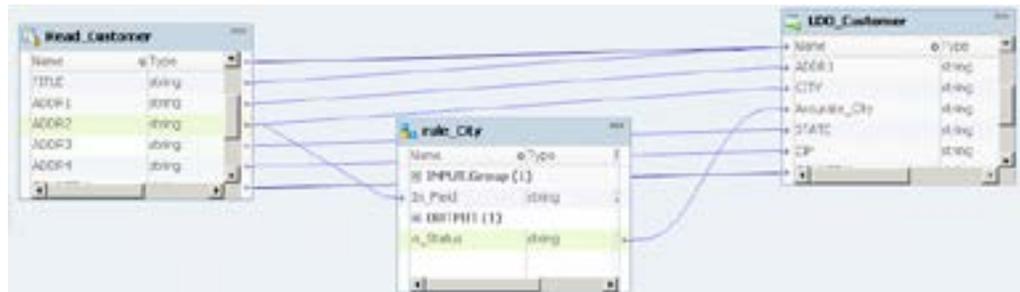
## ④ INFORMATICA DATA QUALITY DESIGN TIP: PRE-FAB VERSUS NEW CONSTRUCTION

When buying a house, it is usually easier to purchase a house that is already built (a.k.a. Pre-Built construction) versus starting from the ground up and waiting months for new construction. Mapplets and other objects are very similar. Pre-built mapplets, content sets, reference data, etc. should be leveraged when possible. Informatica Data Quality provides core accelerators that provides easy and pre-built solutions to common data quality issues.

## 5    INFORMATICA DATA QUALITY DESIGN TIP: LDO'S

Leverage Logical Data Objects (LDO's) during design. LDO's are virtual mappings that allow you to apply filters and can be used in multiple profiles where the LDO is the source object. LDO's allow you to join data together, join multiple tables and include them in one profile, exclude columns from being shown in the profile, filter out records, rename columns, etc. so you don't have to do that in your physical data object. For example, you can rename "ADDR2" to "CITY".



## 6    INFORMATICA DATA QUALITY DESIGN TIP: REFERENCE TABLES

Reference tables should also be leveraged to standardize data. Reference tables enable validation, parsing, enrichment, and enhancement of the data. These are managed by the data steward in Informatica Analyst. Informatica Data Quality keeps an audit trail of metadata and data changes to reference tables. Rather than hardcoding values into the code for a particular rule, a reference table can be used to look up the values instead.

If you'd like more tips & tricks regarding Informatica Data Quality, simply reach out to us through email.

# PARTNER WITH US

**Do you have a current Data Quality initiative? How is your Data Quality project progressing?** We'd like to hear from you. Take a few minutes to send us an email and let us know about your project. If you'd like further assistance with your current project, feel free to contact us directly. We'd be happy to talk with you and develop a proposal around your project.