SCS Case Study

Creating a Configurable ETL Process that can be Controlled by Account Service Personnel

Robust and Flexible Solution to Handle at Least 99% of Foreseeable Data Formats

Solution Snapshot

> **Business Situation:** Employees of companies to which our client provides service, require access to a website where reports concerning services can be viewed. The scope of information available to a given employee must be restricted to the data relevant to the employee's position in the company.
>
> **Multi-Point Solution:** SCS designed a solution that eliminated programmer involvement in creating a huge number of diverse ETL processes and instead created a configurable ETL process that can be controlled by account service personnel.
>
> **Benefits:** At the start of the project, the client was faced with creating and managing 300+ ETL processes with the knowledge that this number would be constantly changing as their client base changed. At the end of the QA experience, the client was well satisfied that the design was robust and flexible enough to handle at least 99% of foreseeable data formats.

The Client:

Our client provides many services, including water, hygiene and energy technologies and services to the food, energy, healthcare, industrial and hospitality markets, to many corporate customers that each have large numbers of locations to be serviced.

The Challenge:

Employees of companies to which our client provides service, require access to a website where reports concerning services can be viewed. The scope of information available to a given employee must be restricted to the data relevant to the employee's position in the company. For example, a store manager should have access only to data concerning services rendered at the store they manage, while a district manager should have access to all information pertaining to stores within his/her district, and only that district.

To accomplish this, our client needed to be able to import employee lists and company hierarchy information from each of the large companies it services. Unfortunately, most of these companies have stock formats they use to supply this type of information to their vendors. This left our client with the daunting task of developing and maintaining several hundred ETL programs to deal with all the data formats presented across their ever-changing client portfolio.

The Solution:

By taking advantage of the service representatives' familiarity with Excel, Superior Consulting Services (SCS) designed a solution that eliminated programmer involvement in creating a huge number of diverse ETL processes. Instead, SCS created a configurable ETL process that can be controlled by account service personnel.

As clients migrate to online reporting, the account representative looks at their client's personnel and hierarchy reports and 'tells' the ETL where to find the data it needs to properly configure data access restrictions for each employee of that company.

Solution Details

Service representatives add a new company to the ETL process by adding the company's name and the name of the directory where data to be imported can be found to an Excel spreadsheet that sits in a specific location monitored by the ETL. This is the Master configuration file.

In the directory that the representative named in the Master configuration file, they place a company configuration file along with the files containing employee and hierarchy data from the client. The company configuration file is nothing more than another Excel file where they specify the configuration of the client data on various pre-defined tabs according to SCS supplied documentation.

The SCS-developed ETL solution reads the Master configuration to determine the locations where data to be imported might reside. It uses the company configuration file in each directory to look for new data to import and understand the specific format supplied by that company.

The spreadsheet used to configure how to read the client supplied data files allows the representative to specify the names of the files being supplied and their basic format (for example CSV, TSV or XLSX), including:

- The columns to be used as input
- How to recognize rows that should be ignored
- 'Regular expression' validations to perform on the data before importing
- Data transformations—the transformations are needed, for example, to standardize name formats

The ETL archives successful data imports and reports detailed descriptions of trouble when a problem is encountered so the representative can either contact the client to fix the problem or change the configuration as needed.

Dynamic reports allow authorized people to examine the data being imported at each stage in its transformation—from a client-defined format to a 'standard' format where a dedicated ETL can easily apply the updates. These reports allow examination of raw 'customer format' data, normalized data and transformed data.

The Result:

After a month of requirements gathering and design by an SCS designer, that designer developed a QA-ready product in two months with the help of one client ETL programmer and another SCS consultant. Client involvement as well as 50+ pages of documentation insured our client retained detailed knowledge necessary to maintain the system over time.

The QA process involved training representatives to use the tool and then use it to import some of the highest priority and most diverse client report formats. After one month of testing, the

representatives were well trained, the ETL was thoroughly exercised and the system was ready for production.

At the start of the project, the client was faced with creating and managing 300+ ETL processes with the knowledge that this number would be constantly changing as their client base changed.

At the end of the QA experience, the client was well satisfied that the design was robust and flexible enough to handle at least 99% of foreseeable data formats. Moreover, since SCS created configurations for the most diverse known formats during testing, the client's personnel had templates for most types of formats that further simplified creation of new configurations as they completed initial rollout and their client base changed over time.

Contact

Superior Consulting Services, LLC
350 West Burnsville Parkway, Suite 550
Burnsville, MN 55337
952-890-0606
www.teamscs.com