



unilog

CASE STUDY

Created Information as a
Service Offering for Over
3.5 Million Product SKUs

unilog

Unilog is a global technology company specializing in enterprise e-commerce solutions and product content services in the business to business (B2B) marketplace.

Unilog's flagship product, CIMM2, comprises numerous features including product information management (PIM), a content management system (CMS), an e-commerce shopping cart, and intuitive search and filtering capabilities. The software also provides data analytics capabilities.

Unilog's clients who use the CIMM2 software system are small to mid-size industrial, electrical, and plumbing distributors.

EXECUTIVE SUMMARY



Duplicate data identification that previously took 4 hours is now automated and done **in 10 minutes**



Automated data prep has created much faster rates of bringing enriched data into the marketplace



Direct revenue has been generated from a service that matches customer's product data against Unilog's rich catalog

Unilog's Information as a Service Offering: Master Product Catalog

One of the critical aspects of Unilog's business is providing product content to its customers – specifically the images, detailed descriptions, categorizations, taxonomies, and other product information that distributors need to populate catalogs, portals, and websites to more effectively sell products to their respective customers.

To accomplish this for one specific customer, Unilog created a Master Product Catalog – a comprehensive product data set of 3.5 million SKUs.

This Master Product Catalog provides insights on the customer's entire product database, including 1) item status, 2) security authorization for item access, 3) classification and taxonomy assignment, and 4) merged item information from industry and individual manufacturers' data sources.

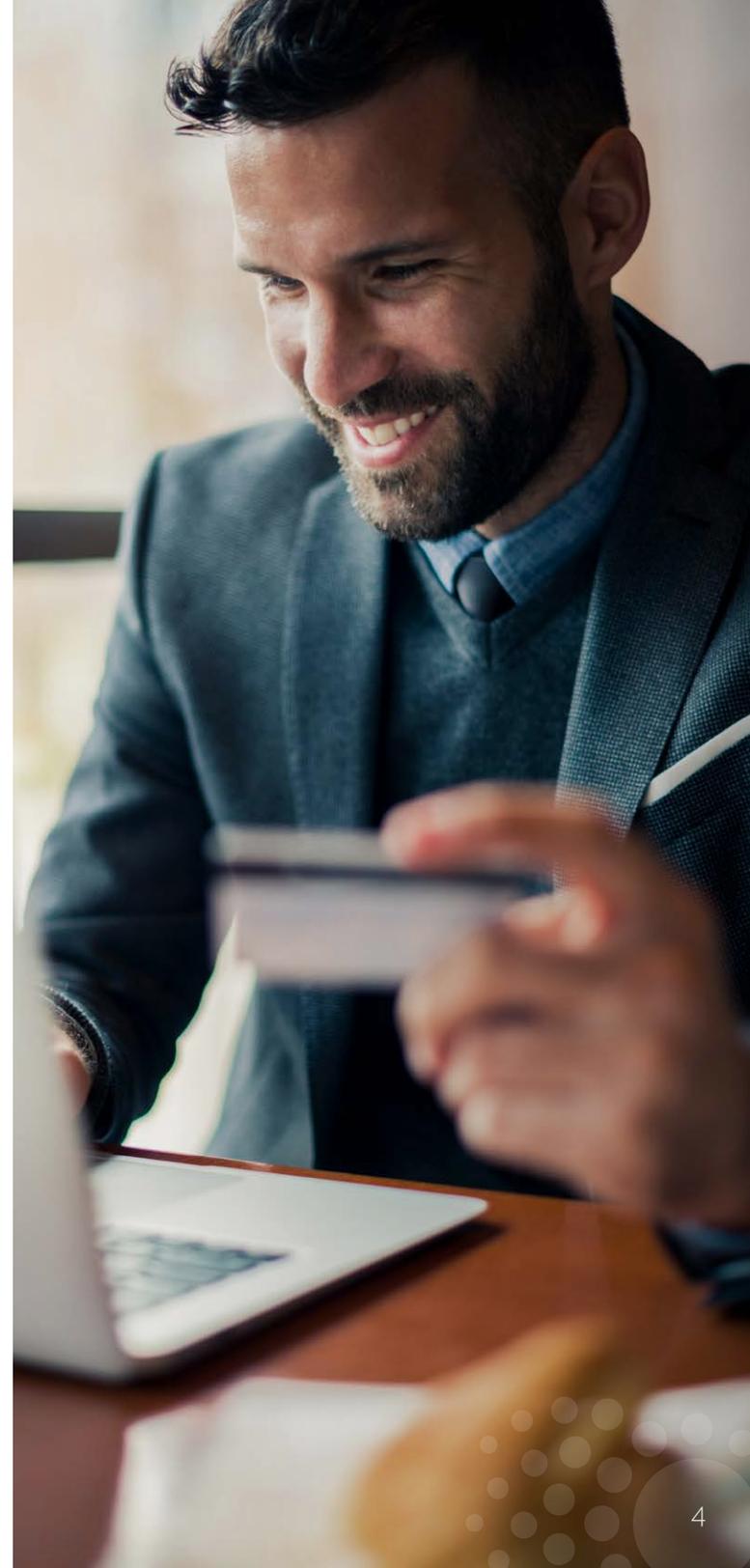
On a daily basis, Unilog's Content Subscription group is responsible for:

- 1 **Loading, combining, and analyzing data** from a data warehouse where multiple manufacturers upload their data, Unilog's own master data, and master taxonomy (in Excel)
- 2 **Using dynamic data prep rules** to validate and de-duplicate data
- 3 **Publishing the results**, which eventually land in Unilog's PIM system

Considering that the product catalog is constantly expanding, the Content Subscription group is tasked with inserting approximately 100,000 additional product items every month into the Master Product Catalog. Rapidly onboarding new product information enables Unilog to meet SLAs with its clients.

This ongoing product expansion requires Unilog to constantly analyze and reconcile the differences between its own production database and the industry data warehouse where multiple manufacturers upload pertinent product data, including country of origin, basic description, UPC codes, weights, and measurements/dimensions.

In many cases, this reconciliation also includes enrichment by associating images and technical documents, applying proper categorization, and providing longer descriptions to the data so that the information is ready for use by distributors.



“

We needed to solve the problem of the team having to use Excel spreadsheets which would blow up as the size of our data became too large.

- NOAH KAYS,
DIRECTOR OF CONTENT SUBSCRIPTIONS

Enterprise Requirements and Technical Challenges

Unilog initially used Excel to address these duties. This approach was adequate, at first, when its customer's product set was merely 100,000 SKUs. However, as the product data grew to 3.5 million SKUs, the group reached Excel's limitations and the process became increasingly unwieldy.

Additionally, the data changed daily. The combination of data frequency and the numerous Excel files required Unilog to divide and analyze the data into more manageable segments which contributed to increasing the time required for data analysis.

Moreover, between using desktop tools and the VLOOKUP function in Excel, processes would take hours to complete.

Searching for the Right Solution: The Criteria and Why Paxata

Given the limitations Unilog experienced with Excel and its strong desire to automate its processes to accelerate and improve its efficiency, the company initiated a search for a new data prep solution. Unilog felt the solution needed to meet three criteria:

- 1 **Have a user-friendly interface** comparable to Excel in terms of look, feel and usability, which would positively drive internal adoption
- 2 **Include a licensing model** based on data volume rather than number of users
- 3 **Offer a turn-key, cloud-based application** so that all members could share data and analyses

Unilog first evaluated Talend and other Extract, Transform, and Load (ETL) tools, but found their older user interfaces would hurt adoption. Some command line base was also needed, which would require a developer.



Why Paxata

Ultimately, Unilog selected Paxata as its solution of choice, based on the following capabilities and features found in Paxata:

- + **Ease of use:** Excel-like interface to drive adoption internally
- + **Speed:** Ability to extract data from its database and into the data prep tool quickly
- + **Licensing model** based on data volume not the number of users
- + **Core capabilities:** Shaping functions such as pivot and de-pivot, filtergrams to immediately highlight missing values, and lookups for reconciling the reporting view with raw transactions for a complete view of parts, manufacturers, and performance levels
- + **Turnkey, fully-managed SaaS application** running on Amazon Web Services (AWS)
- + **Cloud-based software** that allows all data analyses to be shared between all members of the team

The Results

Prior to Paxata, the Excel-based process used by Unilog was highly manual and time-consuming. It was difficult to identify and repair missing detailed records, and individuals without SQL knowledge were unable to analyze the database tables. Moreover, the process lacked repeatability, as all refreshes were performed daily.

Paxata provided Unilog with a simple-to-use interface to access the underlying transactions for immediate data validation. This enabled Unilog's Client Services team to gain visibility into its customer's entire Master Product Catalog by using interactive product matching and validation. This aspect was especially appealing to users in marketing, sales, and client-facing teams with minimal technical skills.

Also, in order to work with Excel's maximum row limitations when identifying duplicate entries, Unilog previously needed to split its 3.5 million Master Product Catalog into four spreadsheets, perform VLOOKUPS against each, and then reconcile the data. Unilog is now able to bypass this cumbersome and time-consuming process because Paxata allows the team to work within the entire data set on a single, large, powerful interface.

Furthermore, Unilog automated the daily and weekly analysis of fresh data within the Master Product Catalog. The repeatability of Paxata's solution further drove Unilog's ROI, enabling automatic refreshes of the data and the creation of the corresponding Paxata AnswerSet to publish internally as well as programmatically into Unilog's PIM system for consumption by its customer's own BI tools.

“

Paxata met our immediate need as far as manipulating the data and running reports, and the more we are exposed to it, the more we learn how to embed it into our processes.

– NOAH KAYS,
DIRECTOR OF CONTENT SUBSCRIPTIONS

Specifically, Paxata delivered the following results for Unilog:

- + Identifying duplicate entries previously involved utilizing 4 different spreadsheets, a process that took 3-4 hours. The process is now automated and can be performed in 5-10 minutes. In one specific use case, this has freed up an entire resource.
- + The AnswerSet produced by Paxata is now Unilog's new record of standard and the single source of truth, ensuring consistent information for analysis.
- + Automation empowers Unilog to divert hours previously spent on maintaining workbooks to other, more strategic initiatives.
- + Unilog is actively selling a service to its 160+ customers centered on matching products against its database and is generating direct revenue as a result.

Unilog's Use of Paxata

Sources	TEXT
[Item Master Table v.20] + [ITEM_DETAIL_REPORTING_VIEW v.35] + [Master_Taxonomy_Hierarchy v.1]	2,597,972
[Item Master Table v.20] + [ITEM_DETAIL_REPORTING_VIEW v.35]	555,883
[Item Master Table v.20]	53,790

Net / New parts missing from analytics view

LEVEL_CATEGORY	LEVEL_CATEGORY_NAME	LEVEL_CATEGORY_CODE	LEAF_NODE_NAME	LEAF_NODE_CODE	Clean MFG and PART
1112130000	Cut Resistant Gloves	1112130000	ShovelsandGloves9903		
1112110000	Chemical Resistant Gloves	1112110000	HeavyWeightSafetyGloves		
1112110000	Chemical Resistant Gloves	1112110000	HeavyWeightSafetyGloves		
1114240000	High Visibility Safety Vests	1114240000	HighVisSafetyVest		
1114240000	High Visibility Safety Vests	1114240000	HighVisSafetyVest		
1114240000	High Visibility Safety Vests	1114240000	HighVisSafetyVest		
1114240000	High Visibility Safety Vests	1114240000	HighVisSafetyVest		
1116120000	Hard Hats	1116120000	HeavyWeightSafetyHats		
1116120000	Hard Hats	1116120000	HeavyWeightSafetyHats		
1116120000	Hard Hats	1116120000	HeavyWeightSafetyHats		
1116120000	Hard Hats	1116120000	HeavyWeightSafetyHats		
1116120000	Hard Hats	1116120000	HeavyWeightSafetyHats		
1116130000	Hard Hats	1116130000	HeavyWeightSafetyHats		

Excel VLookups = 4 hrs
Paxata Lookups = 3 min

SQL Desktop Load - 2 hrs
Paxata Load - 7 minutes

Automation - addresses daily updates

“

The Unilog team is highly anticipating Paxata's new Intelligent Ingest capability.

Our team is especially excited as it will enable us to quickly understand unknown file types, and allow us to preview, validate the file content, and select the right fields across multiple data sources all in the same step. This augments our own internal efforts to make large amounts of varied data available to all levels of business and data analysts.

- NOAH KAYS, DIRECTOR OF CONTENT SUBSCRIPTIONS

Future Plans

Many of Unilog's customers are currently running legacy enterprise resource planning (ERP) software, and therefore lack complete manufacturing or product information, such as Universal Product Codes (UPCs). Unilog's information services provide the enriched information to these customers.

With Paxata's new Intelligent Ingest capabilities, Unilog is able to bring data into its marketplace much faster – and therefore accelerate its billable services. While Paxata is currently rolled out to one third of the company, Unilog is continuing to refine and leverage the Paxata AnswerSet within other projects, including new billable professional services to its customers.

Companies around the globe rely on Paxata to get smart about information. Paxata is the pioneer that intelligently empowers all business consumers to transform raw data into ready information, instantly and automatically, with an enterprise-grade, self-service data preparation application and machine learning platform. Our Adaptive Information Platform weaves data into an information fabric from any source and any cloud to create trusted insights. Business consumers use clicks, not code to achieve results in minutes, not months. With Paxata, Be an Information Inspired Business.

Paxata is headquartered in Redwood City, California with offices in New York, Ohio, Washington D.C., and Singapore.



Paxata Headquarters 305 Walnut Street Redwood City, CA 94063 1-855-9-PAXATA paxata.com

