



CloverETL®

**CASE STUDY**

# **DATA QUALITY & ADDRESS VALIDATION**

**CloverETL** provides a fast-growing logistics company with a data quality solution to replace a team of 30+ workers dedicated to manual address validation and cleansing – an automated platform that allows them to scale throughput by several orders of magnitude.



# DATA QUALITY & ADDRESS VALIDATION

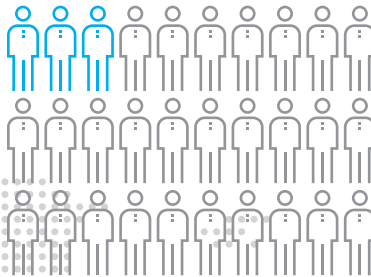


A logistics company's automated CloverETL-based solution validates, geo-locates, and repairs 80 to 90% of addresses instantly, while also providing self-learning tools that help the client's team manually work through exceptions quickly and efficiently.

0 → 85 → 90%

**AUTOMATION COVERAGE**  
Grows with self-learning

**WORKFORCE SAVED**  
Only the exceptions require human interaction



## ACHIEVEMENTS

- Removed a hard-to-scale bottleneck of 30+ workers needed for manual address validation
- Implemented a scalable address validation and cleansing framework customizable to country-specific rules for future expansion to new markets
- Minimized human interactions down to 1/10th – a figure that's still decreasing with the system's self-learning capability

## THE GOAL

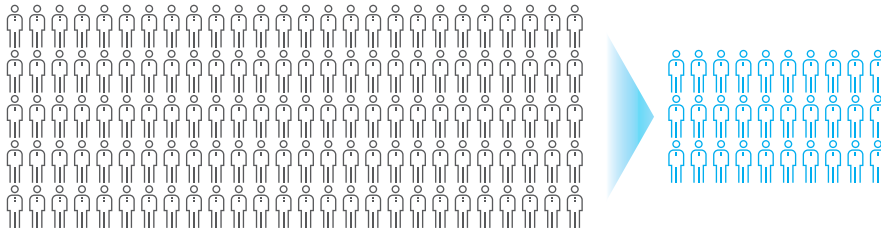
Solving the address quality problem bolsters the company's planned large-scale expansion by clearing the way for a full utilization of their transportation management software, which relies heavily on quality address data for an effective optimization of logistics processes.

## THE SITUATION BEFORE

For a logistics company, expansion to new territories means more address data, notoriously difficult to validate and interpret. Particularly in Emerging Markets, companies are facing vast regional differences in both address structure and rules, and must contend with the availability of tools and databases, or lack thereof, to process them.

### A data quality team tasked with manually verifying and cleaning the data

The logistics company originally had assembled a data quality team tasked with manually verifying and cleaning the data, working shifts to meet overnight delivery deadlines. However, the ever-growing need for manpower to process more data with a quick turnaround became a bottleneck that blocked the company's ability to expand further.



The CloverETL-based automated solution removed a hard-to-scale bottleneck of 30+ workers

## FINDING A SOLUTION WITH THE CLIENT

We started working with the client after their development team had already taken a look at CloverETL data integration. Typically software is developed in-house for this particular client, and as such, they hadn't used a data integration platform before. In adopting CloverETL, they were able to immediately see how our approach complemented their style of working, programming thinking, and skills.

During the initial Proof of Concept project, roughly a week or so, our services team used CloverETL to explore various validation and cleansing techniques, including the use of external libraries and services such as AddressDoctor, HERE maps, Google Maps, and Baidu.

In that short time, we focused on finding ways to solve the core address quality pains, then presented clear results to bring meaningful value to the client. We demonstrated how CloverETL could orchestrate and combine multiple third-party services, as well as created a skeleton of a user interface that the company's teams would use to interact with the solution.



## MOVING FORWARD

After several planning workshops and deep requirements analyses, our services team architected an extensible framework that enabled additional functionality beyond what was originally envisioned. We were able to define and deliver a solution of much greater value and build off of our experience in providing forward-looking solutions. By building a solution on CloverETL, we offered the company not only a useful framework, but also a scalable approach for them to consider and tackle new data-related challenges.

## DELIVERY

### CloverETL Professional Services

The project was implemented by our services team during a roughly five-month period.

### CloverETL Server Corporate + custom Web-based application

The solution is powered by a single instance CloverETL Server automating all the validations with a custom-built user interface for human interactions.

## CHALLENGES

- Address data typically are missing required elements which causes problems in route optimizations
- Geo-coding was needed, as couriers were wasting time on finding exact locations (e.g. a front desk of a large building complex)
- Legal and regulatory challenges in specific countries had to be considered within the solution
- Difficulties with address structures and validation rules varying from country to country

## LEARNING

- The scope of a project changes as the customer gains understanding of what the technology and implementation team can do for them.
- Changes occur naturally in creative processes such as this one and lead to better outcomes than initially envisioned. Successful projects manage changes effectively – even suggesting, rather than avoiding them along the way.
- Getting the whole team, including management, business users, and IT on a training helps them understand not only the current project, but also the potential benefits of deploying CloverETL in other areas.

## THE NITTY-GRITTY

The project was delivered as a framework that can be expanded to support additional countries with their specific validation rules and external lookup and geo-coding services. For example, because AddressDoctor does not support certain countries, our consultants had to switch over to alternative providers in certain cases.

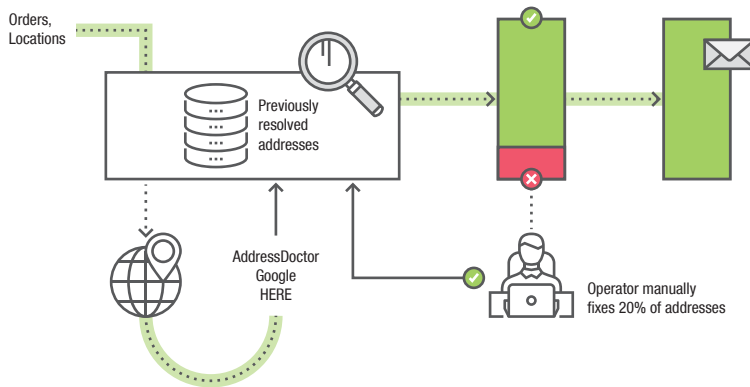
The framework consists of numerous CloverETL transformations and automation jobflows that form the core rule engine. Rules can be easily added or modified without the need for coding. This future-proofs the framework for further expansions without costly change requests.

## CLOVERETL




**CloverETL offers a range of benefits that make this a successful project:**

- Support for connecting to Web Services (public and private APIs used extensively)
- Enabling easy connectivity with all of the company's data (DBs, Flat Files, XML, JSON)
- Publication of CloverETL jobs as a web service, enabling integration between the core solution (CloverETL) with a custom-built user interface via CloverETL Launch Services
- Scale-Out option for handling larger volumes of data, incorporating additional territories, and providing a robust SLA-driven service to the business
- Simple management of complex business rules
- Out-of-the-box Scheduling, Monitoring, Orchestration, and Automation
- Licensing to fit both the budget of the project and the need for future expansion

### CloverETL® AUTOMATED ADDRESS VALIDATION & GEO-CODING



Using local address validation sources and services	Low quality addresses (plus non-latin characters)
Automated validation of addresses against Baidu, Google and HERE maps validators including scoring of single validator results	Near real-time processing
Unique geo coordinates methodology to meet legal specifications	Easily extensible solution that can grow with the size of the customer's business (growing amounts of transactions as well as business in new countries)

<b>PLATFORM</b>	 <b>CloverETL Corporate Server</b> single instance	 <b>Microsoft SQL Server</b> system storage	 <b>Grails Web App</b> custom user interface
-----------------	--	---	--

**North America and LATAM**  
2111 Wilson Blvd.  
Suite 320  
Arlington, VA 22201  
USA

+1 (703) 259 8585

**EMEA and APAC**  
8 Devonhurst Place  
Heathfield Terrace, Chiswick  
London, W4 4JD  
United Kingdom

+44 (0) 203 789 2070

**Germany (DACH)**  
Katharina-Paulus-Straße 8  
65824 Schwalbach  
Frankfurt  
Germany

+49 6196 5869 555

**HQ**  
Vinohradská 174  
2<sup>nd</sup> floor  
130 00 Prague 3  
Czech Republic

+420 277 003 200

