



## remote data integration framework



### An ETL framework to continuously collect and centralize data from hard-to-reach sources.

Cilantro fills a unique niche by collecting hard-to-reach data. Traditional ETL tools typically require some kind of API which is exposed to the outside world. Basil takes a different approach: self-updating agents are installed near the sources and regularly push new data, requiring no remote access for continuous operation.

### Use cases

Cilantro is an ideal match for applications which need up-to-date data from outside sources that cannot be directly accessed by the application. E.g. sources located at a customer, a supplier or a subsidiary whose systems are not part of the private company network and are not accessible from outside.

Data collected by Cilantro is readily available in the Core Warehouse, which can be accessed using well-known methods, like SQL. Or Cilantro can push it to where it is needed, e.g. in an event stream or data lake.

### Functionality



#### Collect data from diverse sources

Sources can be (SQL) databases, flat files, spreadsheets, APIs... on any location inside or outside your organization, on-prem or cloud-based, wherever a Cilantro agent can be deployed to access them.



#### Transform data into a uniform model

Collected data is automatically conformed to a master data model and stored in the Core Warehouse, allowing consuming applications to be blissfully unaware of the many original source formats.



#### Synchronize at any frequency

Relevant changes to the source data are detected and pushed to the backend in near real-time.



#### Easily integrate hundreds of sources

No coding required to add a new source. A self-service workflow allows third-party system administrators to install the agent.

### Features

#### Easy to define new source types

New source types are defined only once, with minimal effort, using a low-code approach. Cilantro supports most common storage technologies out-of-the-box and the architecture is strongly focused on reuse.

#### Privacy by design

Agents sync only relevant source data and can redact sensitive data before it leaves the source system. The Cilantro backend provides smart retention policies which reconcile reliability with GDPR compliance.

#### Batteries included

Cilantro comes with tools to remotely monitor and control large numbers of agents.



## Security

Cilantro was audited by an independent, certified third party. Agents are assumed to run in a zero-trust environment, limiting the blast radius when compromised. Data is end-to-end encrypted, in transit and at rest.

## Devops automation

While Cilantro's backend can run on any cloud or server, an instance can be created or re-deployed to AWS using an entirely code-driven, automated process. Due to automation and centralized configuration, error-prone system administration tasks are very limited, leading to a very low total cost of ownership.

## Architectural overview

