

WHITE PAPER



Big Data Integration with Open ERP





The power of real-time data is unleashed with Open ERP. USA Odoo's Open ERP's data management capabilities, when seamlessly integrated, drive and connect enterprise resources with best-of-class analysis platforms.

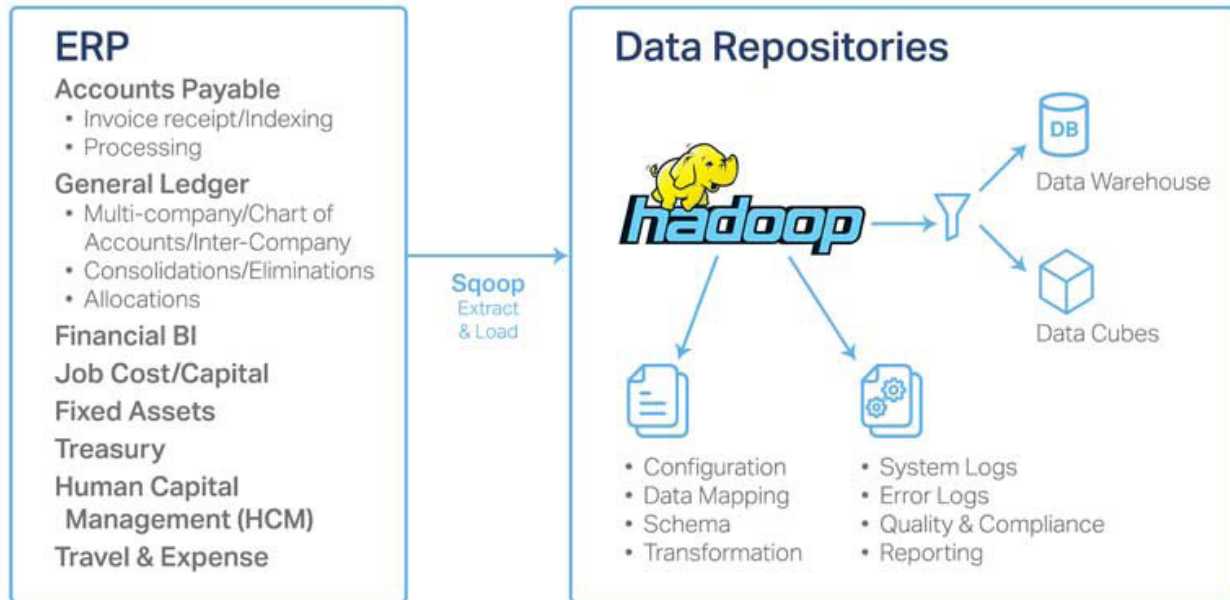
Data Integration Architecture

The USA Odoo community has devised a number of tools that manage Data Integration. Odoo provides an open API, and shares freely through a number of standard web service protocols. [Click here](#) for advanced details.

For the purposes of this white paper, we will offer several options. To determine the best approach, we would need to understand more about the "Data Lake" and if there is a proposed enterprise solution, such as Hadoop or an Enterprise Services Bus in the works.

Hortonworks Hadoop Data Lake

In this case, we are relying the Hadoop ecosystem to do the heavy lifting, and provide tools that will talk to USA Odoo’s underlying PostgreSQL database. The Hadoop tool for this is Sqoop, and it uses a JDBC connection to make the connection with USA Odoo.



Any needed “ETL” activities would be done in Hadoop, using PIG scripting and processing. Hadoop (and the extended ecosystem) are made for this data storage, ETL and integration.

The Hadoop approach may meet the needs that are diagrammed in the “DATA REPOSITORIES” diagram above.

On the USA Odoo side, we would use the USA Odoo module “Import_ODBC” to manage record and object changes. This will ensure that USA Odoo captures “Row/field” level changes, but by going through this module, we can ensure integrity. USA Odoo’s database should not be modified directly, as it may cause corruption problems.

Odoo Data Integration Modules

In this case, we are assuming little about the “Data Lake” and its ETL capabilities. We can do the data integration in USA Odoo with modules that map incoming data to the object models. We can map “like” or “unlike” schemas, using Python and submapping.



These tools are part of USA Odoo’s ecosystem, and work exceptionally well. They have provided the foundation of mapping to a number of external systems, such as E-Commerce, Content Management, Enterprise Service Buses, and similar systems.

By using this approach, we are assured of all data being transformed and validated by USA Odoo itself, ensuring schema cleanliness and row/field level changes.

This approach will work well bi-directionally with data warehousing technologies or other systems. However, maintaining multiple point-to-point integrations can be time consuming.



Kettle, Talend or ESB

In this case, we are considering using a “middleware” approach to data integration. Pentaho Kettle, Mule or Zato ESBs, and Talend are all examples of solutions.

When acquiring other companies, and integrating their systems quickly, this approach has some key advantages. It would allow the absorption of other systems rapidly, keeping newly acquired child companies intact, and operating, but not conflicting with the overall Enterprise architecture.

With all middleware approaches, we must consider the nuances of messaging, validation, and acknowledgement issues. They all offer high performance and lower long-term maintenance costs.

The tradeoff is that implementations are challenging and expensive.

For ESB’s, we would use the web service APIs previously referenced, and therefore, USA Odoo would see updates and changes natively and act appropriately.

For “online ETL” technologies, we would build the validation rules to ensure clean data updates to USA Odoo, and again, validate changes natively, rather than allowing direct writing to the database.

This approach will work well bi-directionally with data warehousing technologies or other systems. The initial cost and effort are higher for all systems, but long-term maintenance is lower.

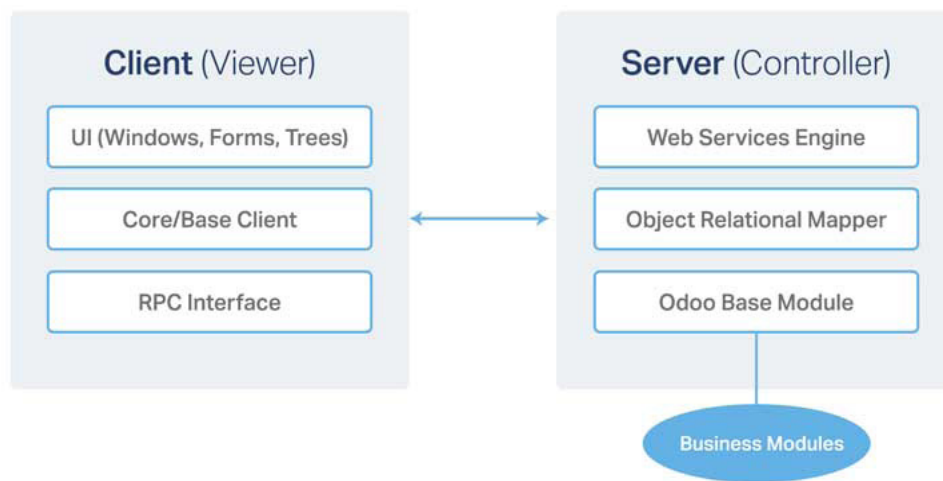
Open Source Integrators has experience with large ESB integrations for a Utility company, where reliability and safety were key, as well as Talend and Kettle ETL migrations of large USA Odoo data sets.

Services Architecture

USA Odoo has delivered a services architecture (XML-RPC and JSON-RPC) from the beginning. This approach has allowed it to support multiple clients (desktop, mobile, and thin), and a variety of integrations with external services.

The most popular services approach is to use XML-RPC. Please see the [documentation here](#) for examples of using this with Python, Ruby, PHP and Java.

By using this approach, developers are accessing USA Odoo the same way as the Client Tier within the Application Server.



This means that external system or program will be considered another client, and therefore, this connection will be managed in a way that goes with the grain of the software. All functionality and business logic will exist through this API.

Given the variety of preferences in the community, add-on modules exist that extend USA Odoo for REST type architectures. These tools are less utilized, given the power and functionality of the XML-RPC approach.

Finally, USA Odoo is using a structured, extendable schema in PostgreSQL. Clients do not access PostgreSQL directly, and will utilize the XML-RPC components for any “CRUD” type actions. This will maintain integrity and durability of the USA Odoo system.



The Power of Data Integration and Open ERP

It's all about connecting your people by bringing disparate systems and data together. Best practice open source data integrations will empower your executives to simplify and increase confidence in their decisions. Open Source Integrators will arm your teams to make better, quicker and more accurate decisions to increase the efficiency of the organization and the satisfaction of your stakeholders.



open source
integrators.com
AN URSA COMPANY

Visit OSI today to see how the right ERP can revolutionize your business.



opensourceintegrators.com



contact@opensourceintegrators.com



480-462-OPEN