



# Data Extraction Overview

## Global Treasury Intelligence (GTI)

**November 2023**

# Contents

- A background to spend data
- Example source systems
- Extraction process
- Data mapping
- Data extraction guidelines
- Data quality



# A background to spend data

## What is meant by spend data?

Spend Data refers to the accounts payable data, usually captured in the ERP system that the client provides to RobobAI for processing during the project. This typically includes purchase order, invoice and purchasing card data and any other supporting data including third party payables data. This data can also include transactional data and vendor master data in the accounts payable system.

## What is Accounts Payable (AP) Spend data?

- Invoice entered for final payment
- Invoice receipt
- Invoice payment
- PO receipt

## What types of data attributes /transactions/spend should be included?

- Invoice data supported by PO/Invoice without supporting PO
- Purchasing Card data
- A set timeframe agreed with RobobAI e.g. Your Fiscal Calendar
- The ERP column / field that determines Intra/Inter-Company spend
- The ERP column / field that determines if spend is in scope/out of scope? If applicable
- The ERP column / field that determines if suppliers are internal or external. If applicable
- The ERP column / field that determines if spend is direct or indirect or any other internal high level classification. If applicable
- Travel Expenses from travel management company or system.
- Any other key requirements for spend analysis or reporting or data enrichment that you want RobobAI to process.

## What types of transactions/spend should be excluded?

- Payroll
- Any spend that cannot be affected by Procurement e.g. Banking fees

# Example source systems

## What is a Source System?

The standalone system (as per contract) that the client will access / build an extract in, to provide / report the required spend data to RoboBAI.

The creation of a complete set of spend data may require input from multiple files or systems to extend the Accounts Payable data

## Example of different Source Systems (Oracle ERP vs SAP ERP)

invoice_number	invoice_line_description	Part_Number	Vendor_No	gl_account_desc	po_unit_price	po_qty	source_system
13124651	Computer	546	4564	Computer H/w	542	1	Oracle
6456464	Nut	64654	56465	Misc	0.254	565	Oracle
16354641	Nut	646	464	Misc	0.2547	262	Oracle
646464	Bolt	654	464	Misc	528	652	Oracle
46549	IT & T Services		6464	Building	2650	58	SAP
6494	Hardware	646	464	Commodity	0.235	521	SAP
654965	Quote no. x55	464	64687	Commodity	0.25	2323	SAP
4646	Quote no. x55	451	6897496	Commodity	235	25	SAP
45454	Literature	2532	4987496	Adv	54	1	SAP

ORACLE ERP as 1st

SAP ERP as 2nd source system

Example of the structure of the extract file that needs to be generated from the ERP system for RoboBAI to classify and visualise in dashboards

Invoice File					GL File		PO File			Supplier File	
invoice_number	invoice_line_number	Vendor_No	invoice_line_description	invoice_line_amount_ex_tax	gl_account_code	gl_account_desc	po_number	po_line_number	po_line_description	Vendor_Name	Street_Address
1313613	1	1254	Computer	250	125	IT	25122	1	Laptop	Computer r us	1, Market st. Market land

# Extraction + data mapping process

The following process is recommended for data extraction:

## 1. Define Business Specifications

## 2. Develop Mapping Document:

- This document should include all the fields of the RobobAI Data Extraction file and mapped (if used) to the client ERP fields
- This document should be the basis for development of the extraction script itself
- Provide finalized and approved document to RobobAI for review to ensure critical data elements are included in the extract files (Recommended) and critical values reconciled post data extraction
- Data mapping changes should be tracked accordingly in the Mapping Document by the Project Team

## 3. Define Technical Specifications

4. Generate Test Files over a short timeframe e.g. 1 month and upload to RobobAI Platform using a safe delimiter like Pipe "|" or carat "^"

5. Generate the final files from production environment of your ERP system with the complete set of the data that is in scope only after all the issues from test files are resolved

## Data Field List - Column Descriptions

### 1. Column 'B' - 'Robobai\_Column'

Defines the standard Data columns used by Robobai to develop any dashboards / reporting

### 2. Column 'C' - 'Grouping'

Defines the type of the source data

### 3. Column 'D' - 'Column Definition'

Definition of each column available in 'Robobai\_Column'

### 4. Column 'E' - 'AK'

Defines required columns for configuration of respective dashboards available from column 'L' to Column 'AR'

# Data extraction guidelines

The following guidelines are recommended :

1. Data files are always preferred to be in .csv or .txt format
2. Data delimiter to be used should not be present within descriptions in order to avoid any data shifting. Suggested delimiter is caret (^)
3. Encoding of the data file is always recommended to be UTF-8
4. All date columns should be in same format across all the files
5. File names are need to be descriptive enough in terms of source (invoice, po, material), time frame (q1, jan), date of extraction or month of extraction, etc.
6. File format should be constant throughout the project lifecycle in terms of file type, data header names, their sequence & format of the dates

# Data quality drive business performance

Data quality is paramount. Investing in data improvement not only provides the opportunity for immediate gains, but it also enables you to create compounding value for the future.

## Data transformation

Extracted data might not always be in the format you need. Your data may need transformation to standardise data across different ERPs. This could involve mapping categories, currencies, units of measurement, and more.

## Data consolidation

If you're dealing with spend data from multiple ERPs, you'll likely want to consolidate the data into a single repository for analysis and reporting. This might involve using ETL (Extract, Transform, Load) tools to transform and load the data into a data warehouse or a central database.

## Data Cleansing and Validation

Data quality is crucial. Cleanse and validate the extracted data to ensure accuracy. This includes identifying and rectifying inconsistencies, missing values, duplicates, and other data quality issues.

RobobAI provide [Managed Services](#) to support all of your data needs. Our data engineers, IT specialists, and ERP experts can improve the extraction process.

For technical support please contact  
[support@robobai.com](mailto:support@robobai.com)