Introduction

PHP package that transforms spreadsheet data into database structures and content, with a schema model and diagrams

Anonymous · 10/10/2026

From Spreadsheets to Databases Seamlessly

A PHP package that transforms spreadsheet data into database structures and content with minimal effort.

Overview

Derafu ETL provides a streamlined solution for converting data between spreadsheets and databases. With a clean, fluent API, it simplifies complex data integration tasks through a pipeline architecture.

```
$pipeline = new Pipeline();
$result = $pipeline
    ->extract('data.xlsx')      // Extract data from a spreadsheet.
    ->transform($rules)         // Apply transformations rules (optional).
    ->load('database.sqlite')   // Load into a database.
    ->execute()
;
```

Key Features

- **Extract** data from various sources (XLSX, ODS, CSV, databases).
- **Transform** data with customizable rules.
- **Load** data into different target systems.
- **Bidirectional** conversion between spreadsheets and databases.
- **Schema management** with automatic table creation and structure updates.
- **Data visualization** capabilities with schema export to Markdown, D2, and more.
- **Extensible** architecture for custom source and target systems.

Installation

Install via Composer:

```
composer require derafu/etl
```

Requires PHP 8.5 or later and `doctrine/dbal` 4.5 or later. If you are coming from a previous version, see Upgrading.

Quick Start

Command Line

The quickest way to use Derafu ETL is through the command line:

```
php app/console.php derafu:etl data.xlsx database.sqlite
```

This extracts data from `data.xlsx` and loads it into a new SQLite database on `database.sqlite`.

Example

Run the example used in tests with:

```
php app/console.php derafu:etl tests/fixtures/spreadsheet-data.xlsx
```

This will create a `spreadsheet-data.sqlite` in the current directory.

PHP Code

```
use Derafu\ETL\Pipeline\Pipeline;

$pipeline = new Pipeline();
$result = $pipeline
    ->extract('data.xlsx')  // Load data from a XLSX.
    ->transform()           // This will use default transformations.
    ->load([                // You can specify the configuration for Doctrine.
        'doctrine' => [
            'driver' => 'pdo_sqlite',
            'path' => 'database.sqlite',
        ]
    ])
    ->execute();            // This is will run the process.

echo "Rows loaded: " . $result->rowsLoaded();
```

Understanding ETL Pipelines

An ETL pipeline consists of three main steps:

1. **Extract**: Read data from a source (e.g., spreadsheet).
2. **Transform**: Apply rules and transformations to the data.
3. **Load**: Write the transformed data to a target (e.g., database).

Derafu ETL provides a clean interface for each step while handling the complex details behind the scenes.

More than just move data to a target

Export Database Schema to Markdown

```
use Derafu\ETL\Database\DatabaseManager;
use Derafu\ETL\Schema\Target\MarkdownSchemaTarget;

$manager = new DatabaseManager();
$database = $manager->connect('database.sqlite');

$target = new MarkdownSchemaTarget();
$markdown = $target->applySchema($database->schema());

file_put_contents('schema.md', $markdown);
```

Generate Database Diagram

```
use Derafu\ETL\Database\DatabaseManager;
use Derafu\ETL\Schema\Target\D2SchemaTarget;

$manager = new DatabaseManager();
$database = $manager->connect('database.sqlite');

$target = new D2SchemaTarget();
$d2 = $target->applySchema($database->schema());

file_put_contents('schema.d2', $d2);
```

Last updated on 10/10/2026