# oerforge.convert

Content Conversion Utilities for OERForge

## Overview

oerforge.convert provides functions for converting Jupyter notebooks (.ipynb) and Markdown files to various formats, managing associated images, and updating a SQLite database with conversion status. It supports batch and single-file conversion, image extraction and copying, and database logging.

## Functions

### setup\_logging

def setup\_logging()

Configure logging for conversion actions. Logs to log/export.log and the console.

### query\_images\_for\_content

def query\_images\_for\_content(content\_record, conn)

Query the database for all images associated with a content file.

**Parameters** - content\_record (dict): Content record dictionary. - conn: SQLite connection object.

**Returns** - list[dict]: List of image records.

### copy\_images\_to\_build

def copy\_images\_to\_build(images, images\_root=IMAGES\_ROOT, conn=None)

Copy images to the build images directory. Returns a list of new build paths.

**Parameters** - images (list[dict]): List of image records. - images\_root (str): Destination directory for images. - conn: SQLite connection object (optional).

**Returns** - list[str]: List of copied image paths.

### update\_markdown\_image\_links

def update\_markdown\_image\_links(md\_path, images, images\_root=IMAGES\_ROOT)

Update image links in a Markdown file to point to copied images in the build directory.

**Parameters** - md\_path (str): Path to the Markdown file. - images (list[dict]): List of image records. - images\_root (str): Images directory.

### handle\_images\_for\_markdown

def handle\_images\_for\_markdown(content\_record, conn)

Orchestrate image handling for a Markdown file: query, copy, and update links.

**Parameters** - content\_record (dict): Content record dictionary. - conn: SQLite connection object.

### convert\_md\_to\_docx

def convert\_md\_to\_docx(src\_path, out\_path, record\_id=None, conn=None)

Convert a Markdown file to DOCX using Pandoc. Updates DB conversion status if record\_id and conn are provided.

**Parameters** - src\_path (str): Source Markdown file path. - out\_path (str): Output DOCX file path. - record\_id (int, optional): Content record ID. - conn: SQLite connection object (optional).

### convert\_md\_to\_pdf

def convert\_md\_to\_pdf(src\_path, out\_path, record\_id=None, conn=None)

Convert a Markdown file to PDF using Pandoc. Updates DB conversion status if record\_id and conn are provided.

**Parameters** - src\_path (str): Source Markdown file path. - out\_path (str): Output PDF file path. - record\_id (int, optional): Content record ID. - conn: SQLite connection object (optional).

### convert\_md\_to\_tex

def convert\_md\_to\_tex(src\_path, out\_path, record\_id=None, conn=None)

Convert a Markdown file to LaTeX using Pandoc. Updates DB conversion status if record\_id and conn are provided.

**Parameters** - src\_path (str): Source Markdown file path. - out\_path (str): Output LaTeX file path. - record\_id (int, optional): Content record ID. - conn: SQLite connection object (optional).

### convert\_md\_to\_txt

def convert\_md\_to\_txt(src\_path, out\_path, record\_id=None, conn=None)

Convert a Markdown file to plain TXT (extracts readable text). Updates DB conversion status if record\_id and conn are provided.

**Parameters** - src\_path (str): Source Markdown file path. - out\_path (str): Output TXT file path. - record\_id (int, optional): Content record ID. - conn: SQLite connection object (optional).

### batch\_convert\_all\_content

def batch\_convert\_all\_content(config\_path=None)

Batch process all files in the content table. For each file, checks conversion flags and calls appropriate conversion functions. Organizes output to mirror TOC hierarchy.

**Parameters** - config\_path (str, optional): Path to \_content.yml config file.

## CLI Usage

python convert.py batch  
python convert.py single --src <source> --out <output> --fmt <format> [--record\_id <id>]

## Requirements

* Python 3.7+
* Pandoc (for docx, pdf, tex conversions)
* nbconvert
* markdown-it-py
* SQLite3

## See Also

* [Pandoc Documentation](https://pandoc.org/)
* [nbconvert Documentation](https://nbconvert.readthedocs.io/en/latest/)

## License

See LICENSE in the project root.