

Doc Number: Korea JTC1/SC2: k2653_281_word_to_png_transfer_manual.txt

Date: 2026.06.11.

Author: SHIN, SangHyun

본 스크립트는 Microsoft Word 문서(*.docx, *.doc) 내에 포함된 모든 테이블(표)을
고해상도 PNG 이미지로 자동 변환하여 저장하는 도구입니다.

1. 주요 목적 (Purpose)

Word 문서에 작성된 테이블을 개별 이미지 파일로 변환하여 다른 문서나 웹페이지,
시스템 등에서 쉽게 활용할 수 있도록 고해상도 PNG 포맷으로 내보냅니다.

2. 주요 특징 (Key Features)

* ****외부 종속성 최소화****: LibreOffice나 Pandoc 등 외부 프로그램이나 바이너리 없이,
Python 패키지만을 사용하여 Windows 환경의 MS Word를 제어하고 변환을
수행합니다.

* ****고해상도 래스터화****: 기본 400 DPI 해상도로 변환하여 텍스트 및 선이 깨지지
않는 깔끔한 화질을 보장합니다. (원하는 DPI 지정 가능)

* ****여백 자동 크롭 (Autocrop)****: 테이블 주위의 무의미한 흰색 여백을 자동으로
추적하여 잘라내며, 잘라낸 뒤 상하좌우에 정확히 1mm 폭의 흰색 여백을 추가하여
안정적인 구도를 확보합니다.

* ****다중 페이지 테이블 병합 (Stitching)****: 여러 페이지에 걸쳐 길게 나누어진
테이블인 경우, 페이지들을 세로 방향으로 이어 붙여(수직 스티칭) 하나의 온전한 단일
이미지로 저장합니다.

* ****16열 한자 표 특화 레이아웃 보정****:

* 테이블의 열 개수가 16개인 경우 한자 표로 인식하여 다음의 서식을 자동으로
최적화합니다.

* 열 너비 강제 조정 (홀수 번호 열: 넓게, 짝수 한자 열: 좁게)

* 단락 공백 제거 및 한 줄 간격(wdLineStyleSingle) 설정

* 셀 수직 가운데 정렬 및 행 높이 고정 (1행: 16pt, 2행 이후: 30.5pt)

* 모든 테두리 제거 (음영이 연속적인 띠 형태로 나타나도록 유도)

* ****자동 출력 디렉토리****: Word 파일이 있는 폴더 하위에 `output_tables/<파일명>/`
경로를 생성하여 결과물을 일괄 저장합니다.

* ****파일명 기준 자동 탐색 및 확장자 보정****: 전체 경로 대신 파일 이름만 입력해도
현재 작업 디렉토리(CWD)와 스크립트가 위치한 디렉토리에서 대상 파일을 자동으로
찾아내며, 확장자(`.docx`/`.doc`)가 생략된 경우에도 자동으로 추가하여 실행합니다.

3. 요구 사양 및 실행 환경 (Prerequisites)

* ****OS****: Windows (Microsoft Word 프로그램이 반드시 설치되어 있어야 함)

* ****Python****: 3.10 +

* ****필수 Python 패키지****:

```
``bash
```

```
pip install pywin32 pymupdf pillow
```

```
````
```

\* `pywin32`: MS Word COM 객체 제어용

\* `pymupdf`: PDF 문서를 고해상도 이미지로 변환(래스터화)용

\* `pillow`: 이미지 스티칭, 크롭 및 PNG 저장용

---

## ## 4. 파이프라인 (Execution Pipeline)

1. **\*\*Word 문서 열기\*\***: COM 객체를 통해 백그라운드(비화면 모드)에서 지정한 Word 문서를 엽니다.
2. **\*\*테이블 복사 및 임시 문서 생성\*\***: 개별 테이블을 복사해 새로운 임시 문서에 넣고 여백을 제거(최소화)한 뒤 최대 크기 페이지로 지정합니다.
3. **\*\*PDF 변환\*\***: 임시 문서를 최고 화질(Print) PDF로 변환하여 임시 저장합니다.
4. **\*\*고해상도 이미지 변환\*\***: PyMuPDF를 사용하여 지정된 DPI 배율로 PDF를 래스터화합니다.
5. **\*\*수직 병합\*\***: 2페이지 이상 분할되어 있으면 Pillow를 사용해 세로로 이어 붙입니다.
6. **\*\*배경 크롭 및 여백 생성\*\***: `ImageChops`로 투명/흰색 여백 영역을 자동으로 깎아내고 상하좌우 1mm 여백을 부여합니다.
7. **\*\*최종 저장 및 정리\*\***: PNG 파일 저장 후 임시 생성했던 PDF 파일을 완전히 제거합니다.

---

## ## 5. 사용 방법 (Usage)

### ### 기본 변환 실행

명령 프롬프트(cmd) 또는 PowerShell을 열고 아래 명령어를 실행합니다.

**\*\*1) 파일 이름과 확장자를 입력하여 실행 (스크립트 폴더 또는 현재 실행 폴더에 파일이 있는 경우)\*\***

```
``powershell
python word_to_png_transfer.py "파일명.docx"
``
```

*\*(참고: 확장자 `.docx` 또는 `.doc`를 생략하고 `python word\_to\_png\_transfer.py "파일명"`으로 입력해도 폴더 내에서 파일을 자동으로 찾아 실행합니다.)\**

**\*\*2) 전체 경로를 지정하여 실행\*\***

```
``powershell
python word_to_png_transfer.py "D:\경로\파일명.docx"
``
```

### ### 출력 해상도(DPI) 지정 실행

해상도를 변경하려면 `--dpi` 옵션을 사용합니다. (기본값: 400)

```
``powershell
python word_to_png_transfer.py "파일명.docx" --dpi 600
``
```

### ### 주의 사항

- \* 경로 및 파일명에 공백(띄어쓰기)이 있는 경우에는 반드시 경로 전체를 **\*\*큰따옴표(")\*\***로 감싸주어야 오류가 발생하지 않습니다.
- \* 실행 도중 MS Word가 백그라운드에서 동작하므로 프로그램이 강제 종료되지 않도록 주의하십시오.