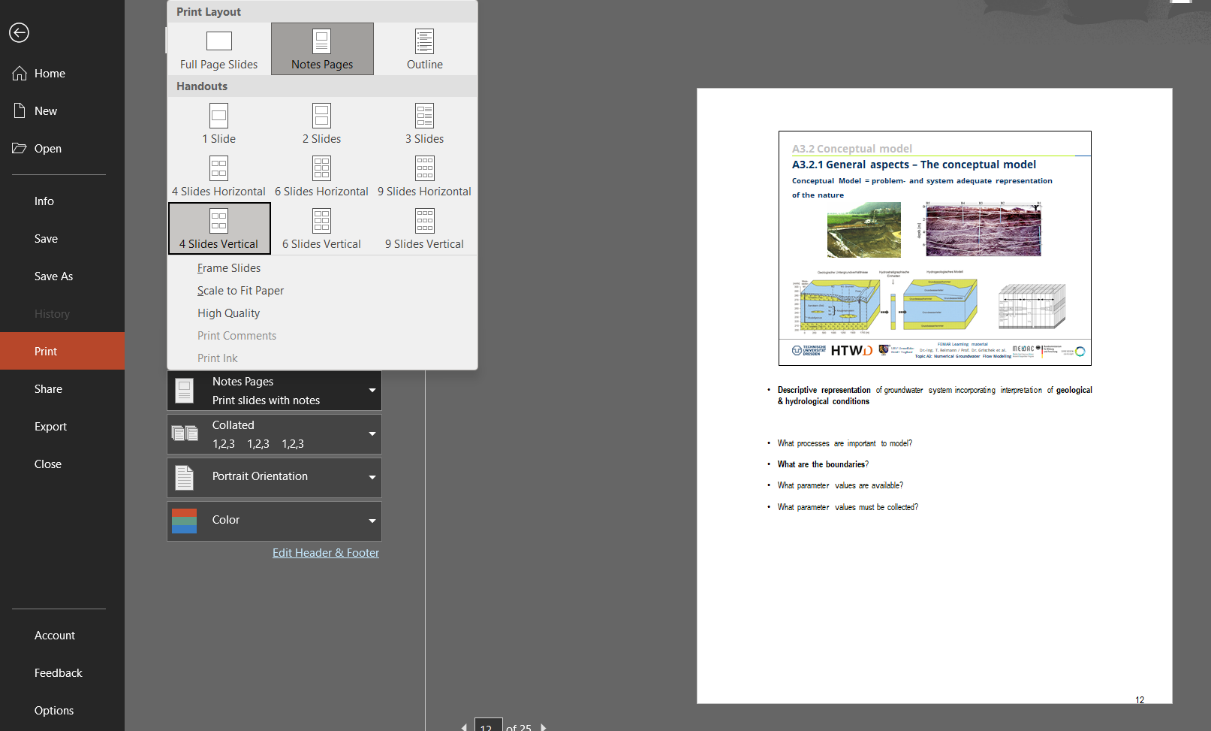
# **Steps to convert a Presentation with the notes to a Markdown file**

#### ****Part 1: Save the Presentation as a PDF****

1. **Open the Presentation**: Launch the PowerPoint file you want to convert.
2. **Save as PDF**:
   * Press Ctrl + P (or use the Print option).
   * Select a **PDF Writer** as the printer.
   * Under "Print Layout," select the **Notes** option.
   * Click **Print** and save the file to your desired location. This will generate a .pdf version with the notes included.

#### ****Part 2: Convert the PDF to a Markdown File Using Marker****

1. **Set Up Your Environment**:  
   * Install Python and Anaconda if not already installed.
   * Open **Anaconda Prompt** and create a new conda environment with a more widely supported Python version, such as 3.10 or 3.11. (We have to install the PyTorch package, and PyTorch is often not immediately compatible with the very latest Python versions like 3.13, and because of this, pip will not be able to find a compatible package). Then activate this environment.:

conda create -n my\_new\_env\_311 python=3.11

conda activate my\_new\_env\_311

1. **Install Required Packages**:  
   * To locate where the env is stored:  
       
     conda info
   * To check the installed packages:  
       
     conda list
   * Verify or install NumPy (a prerequisite for PyTorch):  
       
     conda install numpy
   * Use the [PyTorch website](https://pytorch.org) to determine the correct installation command. For example:

pip install torch torchvision torchaudio

* + Install Marker:

pip install marker-pdf

1. **Set Up OCRMyPDF (Optional but Recommended)**:  
   * Install OCR tools for higher-quality output:
     + **Windows**: Use these commands:

winget install -e --id Python.Python.3.11

winget install -e --id UB-Mannheim.TesseractOCR

Download **Ghostscript** from: [Ghostscript Downloads](https://github.com/ArtifexSoftware/ghostpdl-downloads/releases).

* + - **Run OCRMyPDF Installation**:

pip install ocrmypdf

* + For non-Windows systems, follow the [OCRMyPDF documentation](https://ocrmypdf.readthedocs.io/en/latest/installation.html).

1. **Configure Tesseract Language Models**:  
     
   The marker application uses Tesseract OCR for text recognition. For Tesseract to work properly, it needs to know where its language models (located in the tessdata folder) are stored. For this, we will set the TESSDATA\_PREFIX environment variable, and point the application to the correct directory to access the language models, by following these steps ahead.  
   * Locate the tessdata Folder on Windows: If you installed Tesseract using the Windows installer (e.g., from UB Mannheim's Tesseract builds), the tessdata folder is typically located in:  
     C:\Program Files\Tesseract-OCR\tessdata
   * Open the directory where your marker project is installed. For example, in my case:  
     C:\Users\Heena\anaconda3\envs\my\_new\_env\_311\Lib\site-packages
   * Open a text editor (e.g., Notepad) and create a new file named local.env. Inside the file, add the following line:  
     TESSDATA\_PREFIX=C:\Program Files\Tesseract-OCR\tessdata
   * Save the file in the root directory of the marker project
2. **Run Marker**:  
   * For a single PDF file:

marker\_single /path/to/file.pdf /path/to/output/folder

* + For a folder containing multiple files:

marker /path/to/input/folder /path/to/output/folder