HUMBOLDT STATE UNIVERSITY

HSU Document Converter

Guide for Students to Convert Image Files to Accessible (Text-Based) Files
Conversion Best Practices were written by Sean Keegan, Stanford University

The HSU Document Converter is a web-based service that will convert documents into a range of accessible formats. The service can be used to convert inaccessible documents such as image-only PDF files and JPG pictures into accessible (text-based) formats. The HSU Document Converter is intended to complement, not replace, the alternate media services that are provided by the Student Disability Resource Center (SDRC). The quality of the conversion is dependent upon the quality of the original document. If you are a student registered with the SDRC, are eligible to receive e-text services, and are having difficulty effectively reading a document after using the HSU Document Converter, please contact the let Student Disability Resource Center for assistance. The following best practices identify simple methods to prepare the file before converting in order to achieve a high-quality output.

Step 1 : Scan Your Document to Create Image-only PDF file

PDF and image-based files will be processed using optical character recognition (OCR) to create a text-based version of the document.

- If scanning the document:
  - Ensure the scanned image is free from smudges, dark marks, highlighted text, or artifacts in the image. These will affect the accuracy of the OCR process.
  - Minimize the effects from skewing. If the image is presented at an "off-angle", the accuracy of the OCR process will be lower resulting in a lower quality text version.

Step 2 : Upload Your Document to the HSU Document Converter

1. HSU Document Converter[1]
2. Select "File" as your source
3. Select "Browse…" to search for the file you would like converted
4. Find the file and select "Open". The file name appears to the right of the "Browse…" button
5. Select "Upload". This may take a few seconds
6. Select the output format. If converting to text, select "Accessibility conversion"
7. Select the format desired for the conversion
   - The HSU Document Converter will convert image-based document into MS Word, RTF, and text files. [With some image-based documents, you may achieve better results if you convert initially to Tagged PDF and then copy and paste the text from the Tagged PDF into MS Word. This may result in a better reading experience and may remove non-essential content.]
8. Enter your HSU email address (must use HSU email for this service)
9. Select "Submit"

---

[1] HSU Document Converter: http://www.humboldt.edu/disability/servicesavailable-sensus

Step 3 : Receive Converted file in your HSU email In-Box

The converted file will be sent to your HSU email in-box. Conversions that are between text-based formats will generally be completed in less than an hour. Conversions that involve converting image-based documents into text-based documents (e.g., scanned PDF to tagged PDF) or the creation of audio files may take longer.

Step 4 : Read the Document with Text-to-Speech Software (e.g., Read&Write)