



Computer Info sheet



Using Office for OCR

If you have an image of text (a pdf document or text that was scanned as a image) and you don't have a scanner with OCR¹ capability, you can use Word to convert it.

You need to do this in a few stages.

1) Convert the Image to a Tagged Image File Format (.tif .tiff)

The original picture must be in .tif format for this to work.

To create a .tif image.

- Open **Paint Shop Pro** (PSP) or a similar paint program.
- Open the existing image of the text.
 - If you have a .pdf, Select and Copy the text and paste it into PSP this will create an image.
- Use **Save As** to save and convert it as a .tif file.
- Close **PSP**

2) Use OCR in Microsoft Office Document Imaging

- Click the Windows **Start** menu, on the **Task Bar**
- Click **All Programs**
- Click **Microsoft Office**, then click **Microsoft Office Tools**,
- Select **Microsoft Office Document Imaging**.
- From the **File** menu choose **Open** and locate your .tiff image and open it.
- Click the **OCR** icon. (located on the Standard Toolbar or in the Thumbnail pane) and the OCR will recognise the Text.



(If you save it at this stage you are still just saving the .tif file)

3) Copy and Paste Text

- While still in **Microsoft Office Document Imaging**.
- Select the text and use **Copy**.
- Go to **Word** and **Paste** the text into a new document.

You will probably need to proof-read the text to make sure all the characters have been translated properly.
- Save your document.

¹ OCR (Optical Character Recognition) translates/converts images of text, such as scanned documents, into text characters.