

Homework 2--Capturing an image by scanning and converting it to text using OCR

Directions

- 1) In the lab, run the ReadIRIS OCR program on two different scanned documents:
 - one will be a page you scanned in class on Jan 30
 - the other will be the document in our Handouts folder: **Four ads scanned.jpg**
- 2) For each one, follow our class handout on how to run ReadIRIS to convert jpegs to text.
- 3) Write up a discussion of issues that came up in obtaining editable text for the two documents.

Your write-up essay should include these aspects:

- What language is your text in?
- Do you have any diacritical marks or special characters in your text?
- What errors did you have to correct?
- How would you categorize the kinds of recognition errors?
- What errors did not get noticed by the program?
- Did page layout or graphics affect the success of your scan?
- What other troublesome issues did you encounter?
- If you had 50 pages of such texts to input, would scanning or retyping it be more useful? Why?

Final product

Save the following five pages to turn in:

- One electronic copy of your original scanned document
- A copy of the page after the OCR has been done to it, and
- 3 pages of discussion (typed, double spaced)

Due date

Upload your response materials as **one single pdf file**.

It's due in Blackboard by the start of class on **Wednesday, Feb. 13**