



# Image to Text in Natural Language Processing

Joyce Woznica  
IST 664 (Tuesday 6:00 PM)

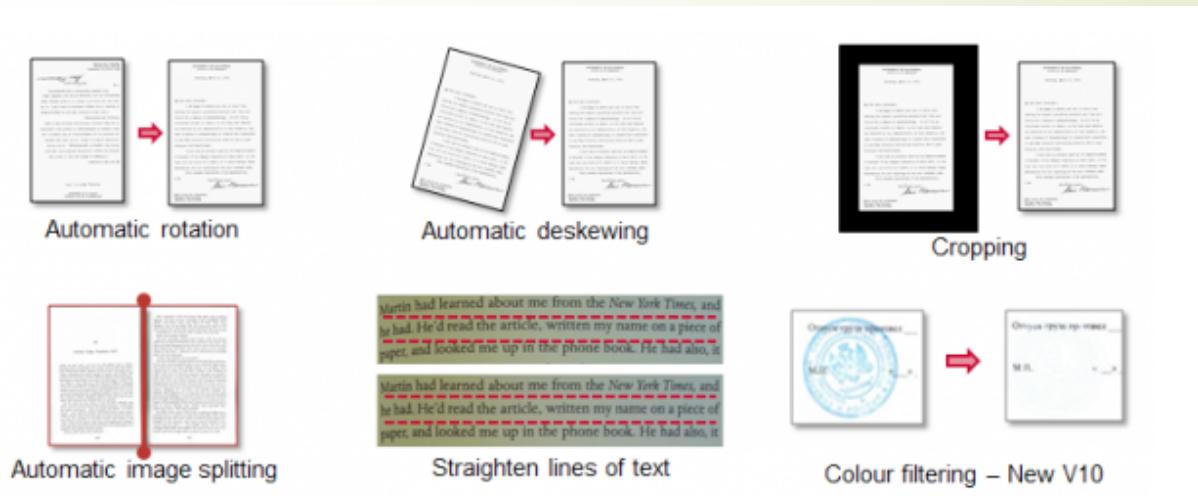


## What is Image to Text?

- ▶ The simplest way to describe Image to Text is to review Recognition:
  - ▶ Optical Character Recognition (OCR)
    - ▶ OCR converts pixels on a page into digital form that is then recognized by a computer and can be converted to individual letters, words, numbers, etc.
  - ▶ Optical Mark Recognition (OMR)
    - ▶ This converts marks in a particular location on a page (like filling in the black box for Yes/No) and records it as a value
  - ▶ Intelligent Character Recognition (ICR)
    - ▶ This is also the term used for recognizing handwriting which can be more complex because of how individual print or use cursive

# Processing Steps for Recognition

- ▶ Input
  - ▶ Read the input file(s)
    - ▶ On disk, from a scanner, from an Xray, from an image, etc.
- ▶ PreProcessing
  - ▶ Remove noise Despeckle (remove specks on the page)
  - ▶ Deskew: Straighten alignment of page
  - ▶ Cropping: eliminate extra white space
- ▶ Recognition
  - ▶ Look for words, letters, numbers, etc.
- ▶ Verification
  - ▶ Allow users to verify when there is a question like . . .
    - ▶ is it an “O” or a “0” or a “Q”

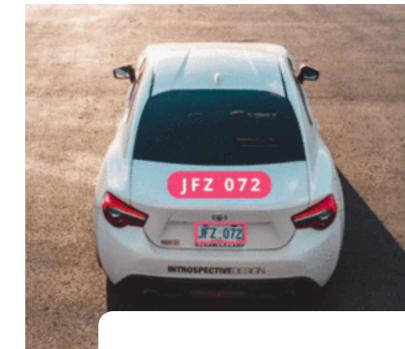


# Strong Engines for doing Recognition

- ▶ Historical Strong Players
  - ▶ ABBYY Finereader - <https://abbyy.technology/en:start>
  - ▶ Tesseract - <https://opensource.google/projects/tesseract>
  - ▶ OmniPage Ultimate - <https://www.kofax.com/Products/omnipage>
- ▶ New Players
  - ▶ Amazon Textract - <https://aws.amazon.com/textract/>
  - ▶ Adobe Acrobat - <https://acrobat.adobe.com/us/en/acrobat/how-to/ocr-software-convert-pdf-to-text.html>
  - ▶ Amazon Recognition - <https://aws.amazon.com/rekognition/>
- ▶ There are so many more - some free, some very costly

# Not just Paper Anymore

- ▶ Can be used in law enforcement to read plates
  - ▶ Already used at toll booths
- ▶ Can be used to provide intelligence to photographs and images
  - ▶ You can extract entities like company names or drug name
  - ▶ Can help to categorize content
  - ▶ Can be used to show relationships
    - ▶ White car to plate number

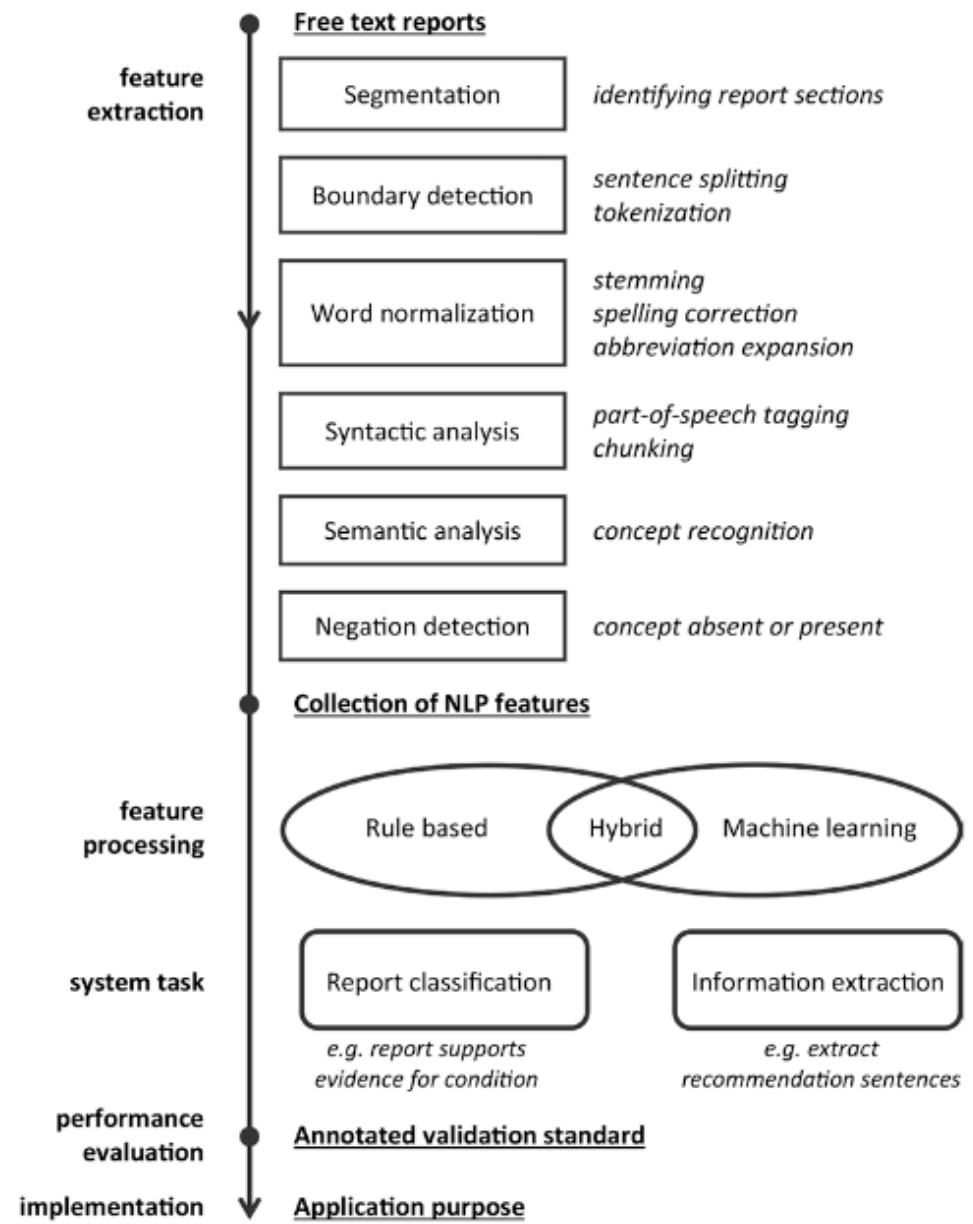


We live in times when any organisation or c



# But how does this apply to NLP?

- ▶ Image to Text platforms are getting smarter
  - ▶ First you apply Optical Character Recognition - then
    - ▶ You can now detect, understand and classify blocks of text to extract meaning
    - ▶ They have learning modules so they can be trained
    - ▶ Classification based on document structure, words in the documents (invoice, purchase order, etc) can be used to classify appropriate
    - ▶ Analysis example (Xray or Medical Test Reading)
      - ▶ Checking for information in thresholds (high counts/low counts, etc.)
        - ▶ This can be used to automatically notify a patient/doctor of an issue
        - ▶ Helps to ascertain next steps or validate treatment or more



# Use Case - Claims Processing

- ▶ Using Amazon Textract and Alfresco Content and Process Services to streamline triaging a claim:
  - ▶ <https://www.youtube.com/watch?v=s7iQG-jiLnw> - Overview
    - ▶ Using image to text to upload information in to a system
  - ▶ <https://www.youtube.com/watch?v=8gAwYv3dLvs> - Demonstration
    - ▶ In this demonstration:
      - ▶ Analyze Documents:
        - ▶ Classification is used to determine the type of document uploaded
      - ▶ Image to text is used to add claimed loss costs and verification

The Fresco Insurance claims processing interface consists of three main sections:

- Claim Details:** Shows Claim Number FNOL-2019-7-6792, Date of Loss 01.1.2019, Type Homeowners, Total Claim Amount \$10,000.00, Status Open, and Description My basement flooded.
- Customer Details:** Customer ID 54321, Name Sarah Smith.
- Actions:** Buttons for Accept Claim, Reject Claim, and Analyze Documents (highlighted with a green oval).

**Claim Files:** A list of five files with their sizes and modification times:

Name	Size	Modified
bookcase.jpg	272.37 KB	a few seconds ago
FloodVideo.mp4	68.09 MB	a minute ago
furniture.jpg	800.57 KB	a minute ago
pointing.jpg	1.82 MB	a minute ago
TV.png	1.79 MB	a minute ago

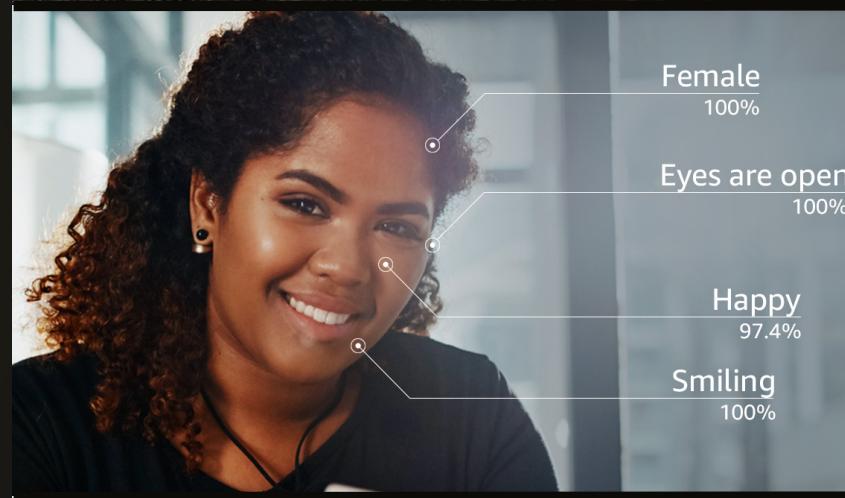
**Receipts:** A section showing four receipts with their detected amounts:

Receipt	Detected Amount
pointing.jpg	\$2338.95
TV.png	\$2483.99
bookcase.jpg	\$5137.00
furniture.jpg	\$1252.75

**Close-up of a Receipt:** A photograph of a crumpled receipt from RUTH WHITE GALLERY. The receipt details a purchase of a painting titled "Wild Spirit" for \$2,123.95, with a subtotal of \$2,123.95, sales tax of \$215.00, and a total of \$2,338.95. It also lists the address 42 East 57th Street, New York, NY 10022, and the recipient Sarah Smith at 789 Main Street, Keane, OH 45349, with phone number 555-1212.

# Where are we heading?

- ▶ Artificial Intelligence with Machine Learning
  - ▶ With companies like Amazon and their product Amazon Rekognition, we can now recognize:
    - ▶ People
    - ▶ Places
    - ▶ Objects
    - ▶ Sentiment
    - ▶ And more
- ▶ This type of technology is already being used by customers like those below



# References

- ▶ ABBYY Technology Portal:  
<https://abbyy.technology/en/features:ocr:processing-steps>
- ▶ Adobe: <https://acrobat.adobe.com/us/en/acrobat/how-to/ocr-software-convert-pdf-to-text.html>
- ▶ Alfresco: <http://www.alfresco.com>
- ▶ Amazon Rekognition:  
<https://pubs.rsna.org/doi/pdf/10.1148/radiol.16142770>
- ▶ Amazon Textract: <https://aws.amazon.com/textract/>
- ▶ Best OCR Engines: <https://beebom.com/best-ocr-software/>
- ▶ clarifai: <https://www.clarifai.com/blog/nlp>
- ▶ Deep Learning and OCR: <https://nanonets.com/blog/deep-learning-ocr/>
- ▶ Google Tessearct OCR:  
<https://opensource.google/projects/tesseract>
- ▶ Natural Language Processing in Radiology:  
<https://pubs.rsna.org/doi/pdf/10.1148/radiol.16142770>
- ▶ RSNA Radiology:  
<https://pubs.rsna.org/doi/10.1148/radiol.16142770>