Scrape email addresses from pdf