Nội dung từ PDF (cần cải tiến với OCR)