Testing extraction of PDF document