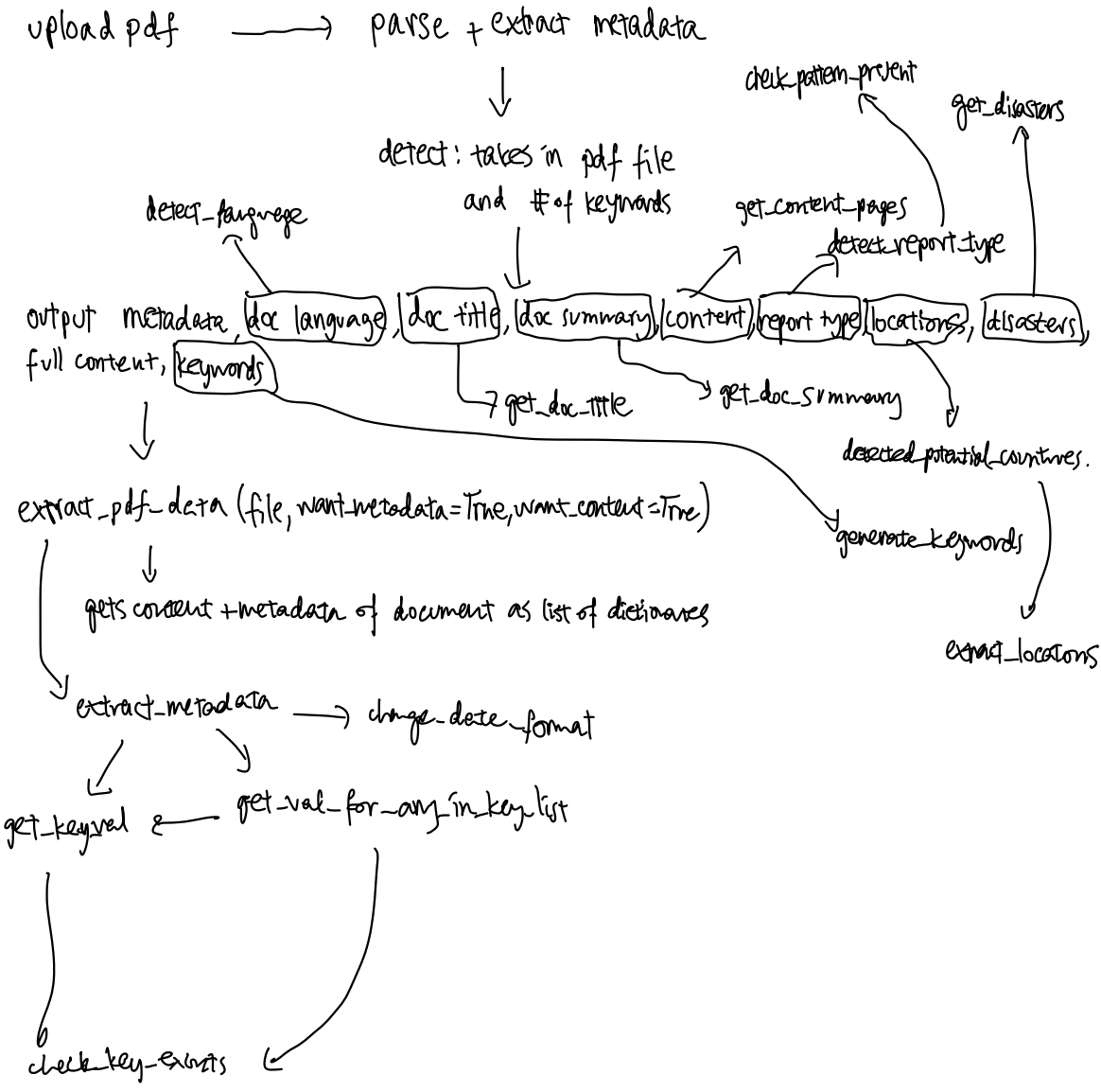
upload pdf
parse + extract metadata
detect: takes in pdf file and # of keywords
check_pattern_present
get_disasters
detect_language
get_content_pages
detect_report_type
output metadata, doc language, doc title, doc summary, content, report type, locations, disasters,
full content, keywords
get_doc_title
get_doc_summary
detected_potential_countries.
extract_pdf_data (file, want_metadata=True, want_content=True)
generate_keywords
gets content + metadata of document as list of dictionaries
extract_locations
extract_metadata
change_date_format
get_keyval
get_val_for_any_in_key_list
check_key_exists