Interns

TASK 1: Extract all the labeled entities in the html section and create a JSON dictionary (PRATYUSH)

For every document, extract id, extract html, extract labeled entities:

Example:

"id": "e4c5dfe7-351b-413c-b763-501505d4d636",

"html": "\n<p><person>ST. ALBANS</person> - A Milton man has been charged with assault for allegedly threatening another man at what was an otherwise peaceful demonstration in <location>Taylor Park</location> here on Tuesday.</p>\n<p>According to the <org>St. Albans Police Dept</org>.(SAPD), <person>Christopher Reynolds</person>, 22, threatened another man while resting his hand on a handgun at his hip. Reynolds was cited for aggravated assault and the handgun seized as evidence, SAPD reports.</p>\n\n<p>SAPD officers were called at approximately <chron>6:41 p.m.</chron> for a report of a protestor arguing with a passerby.</p>\n\n<p>The alleged victim told police he was confronted by Reynolds after saying "all lives matter" to the protestors. An argument followed with the victim crossing <location>Main Street</location> and entering the park.</p>\n\n<p>Video shows the victim entering the park and being confronted by Reynolds who said, "'do something or I certainly will' while approaching the victim with his hand on the grip..."

person: ST. ALBANS

org:St. Albans Police Dep

chron: 6:41 p.m.

location:Main Street

JSON dictionary: { person: [ ], location: [ ]......}

TASK 2: Extract all the entities under enriched_text.entities and create a JSON dictionary (JAYESH)

id: e4c5dfe7-351b-413c-b763-501505d4d636",

enriched_text.entities: extract all types {Person, organization, facility, location,....}

Ex: Facility under enriched_text

```
        "text": "Taylor Park",
                              "relevance": 0.441593,
                              "type": "Facility"
```

TASK 3: Run Spacy NLP libraries to extract named entities from text (SAYAN)

id: e4c5dfe7-351b-413c-b763-501505d4d636",

"text": "ST. ALBANS \\- A Milton man has been charged with assault for allegedly\nthreatening another man at what was an otherwise peaceful demonstration in\nTaylor Park here on Tuesday.\n\nAccording to the St. Albans Police Dept.(SAPD), Christopher Reynolds, 22,\nthreatened another man while resting his hand on a handgun at his hip.\nReynolds was cited for aggravated assault and the handgun seized as evidence,\nSAPD reports.\n\nSAPD officers were called at approximately 6:41 p.m. for a report of a\nprotestor arguing with a passerby.\n\nThe alleged victim told police he was confronted by Reynolds after saying \"all\nlives matter\" to the protestors. An argument followed with the victim crossing\nMain Street and entering the park.\n\nVideo shows the victim entering the park and being confronted by Reynolds who\nsaid, \"'do something or I certainly will' while approaching the victim with\nhis hand on the grip of his gun, which was being openly carried on his hip,\"\naccording to the statement from SAPD.\n\nReynolds was taken into custody, pro...",