Here some guidance on the PubAnnotation format:

<http://www.pubannotation.org/docs/annotation-format/>

For each text section you need to make a separate json file with the annotations. So, one file for the title, one for the abstract and one for each section in the body\_text. The file name is constructed like this:

jjpi5gjm-0-title, jjpi5gjm-1-abstract, jjpi5gjm-2-body\_text, jjpi5gjm-3-body\_text, …

You can make an annotation in PubAnnotation and then compare it to your output <http://www.pubannotation.org/docs/create-annotation/>

For an example click on the example on this page and then TextAE to look at it in the viewer.

And here is an example of the file with id jjpi5gjm (first in your set of 100) which I have annotated with the MONDO annotations. Do not remove \n or \u0026 and similar expressions when extracting the text from the original json file. Also, note how they are counted. To illustrate this, I kept title and abstract together in this example.

{"cord\_uid":"jjpi5gjm","sourcedb":"PMC","sourceid":"PMC3516577","divid":1,"text":" Cathepsin B \u0026 L are not required for ebola virus replication.\nEbola virus (EBOV), family Filoviridae, emerged in 1976 on the African continent. Since then it caused several outbreaks of viral hemorrhagic fever in humans with case fatality rates up to 90% and remains a serious Public Health concern and biothreat pathogen. The most pathogenic and best-studied species is Zaire ebolavirus (ZEBOV). EBOV encodes one viral surface glycoprotein (GP), which is essential for replication, a determinant of pathogenicity and an important immunogen. GP mediates viral entry through interaction with cellular surface molecules, which results in the uptake of virus particles via macropinocytosis. Later in this pathway endosomal acidification activates the cysteine proteases Cathepsin B and L (CatB, CatL), which have been shown to cleave ZEBOV-GP leading to subsequent exposure of the putative receptor-binding and fusion domain and productive infection. We studied the effect of CatB and CatL on in vitro and in vivo replication of EBOV. Similar to previous findings, our results show an effect of CatB, but not CatL, on ZEBOV entry into cultured cells. Interestingly, cell entry by other EBOV species (Bundibugyo, C\u00f4te d'Ivoire, Reston and Sudan ebolavirus) was independent of CatB or CatL as was EBOV replication in general. To investigate whether CatB and CatL have a role in vivo during infection, we utilized the mouse model for ZEBOV. Wild-type (control), catB(-/-) and catL(-/-) mice were equally susceptible to lethal challenge with mouse-adapted ZEBOV with no difference in virus replication and time to death. In conclusion, our results show that CatB and CatL activity is not required for EBOV replication. Furthermore, EBOV glycoprotein cleavage seems to be mediated by an array of proteases making targeted therapeutic approaches difficult.","project":"cdlai\_CORD-19","denotations":[{"id":"PD-MONDO\_T1","span":{"begin":37,"end":42},"obj":"http://purl.obolibrary.org/obo/MONDO\_0005737"},{"id":"PD-MONDO\_T2","span":{"begin":62,"end":67},"obj":"http://purl.obolibrary.org/obo/MONDO\_0005737"},{"id":"PD-MONDO\_T3","span":{"begin":186,"end":209},"obj":"http://purl.obolibrary.org/obo/MONDO\_0018087"},{"id":"PD-MONDO\_T4","span":{"begin":937,"end":946},"obj":"http://purl.obolibrary.org/obo/MONDO\_0005550"},{"id":"PD-MONDO\_T5","span":{"begin":1385,"end":1394},"obj":"http://purl.obolibrary.org/obo/MONDO\_0005550"}]}