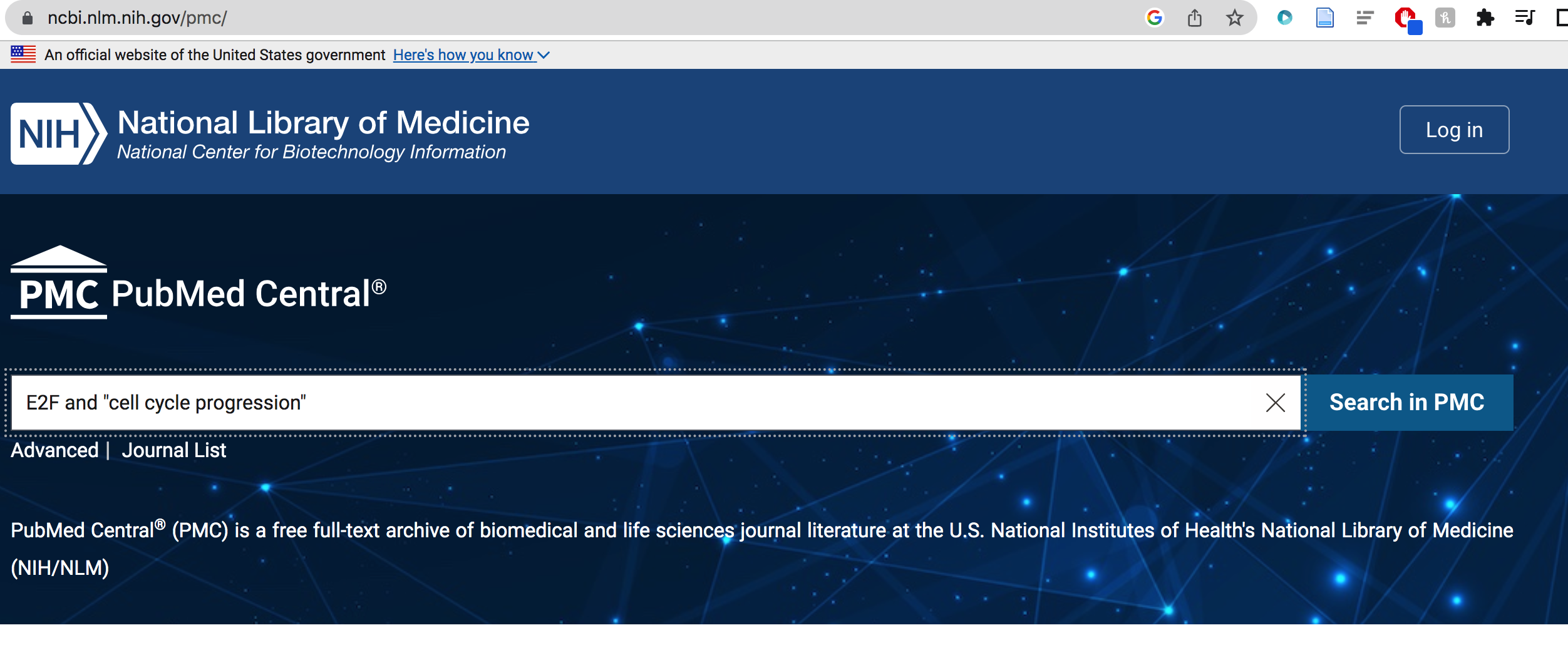
Retrieving statements from the INDRA DB:

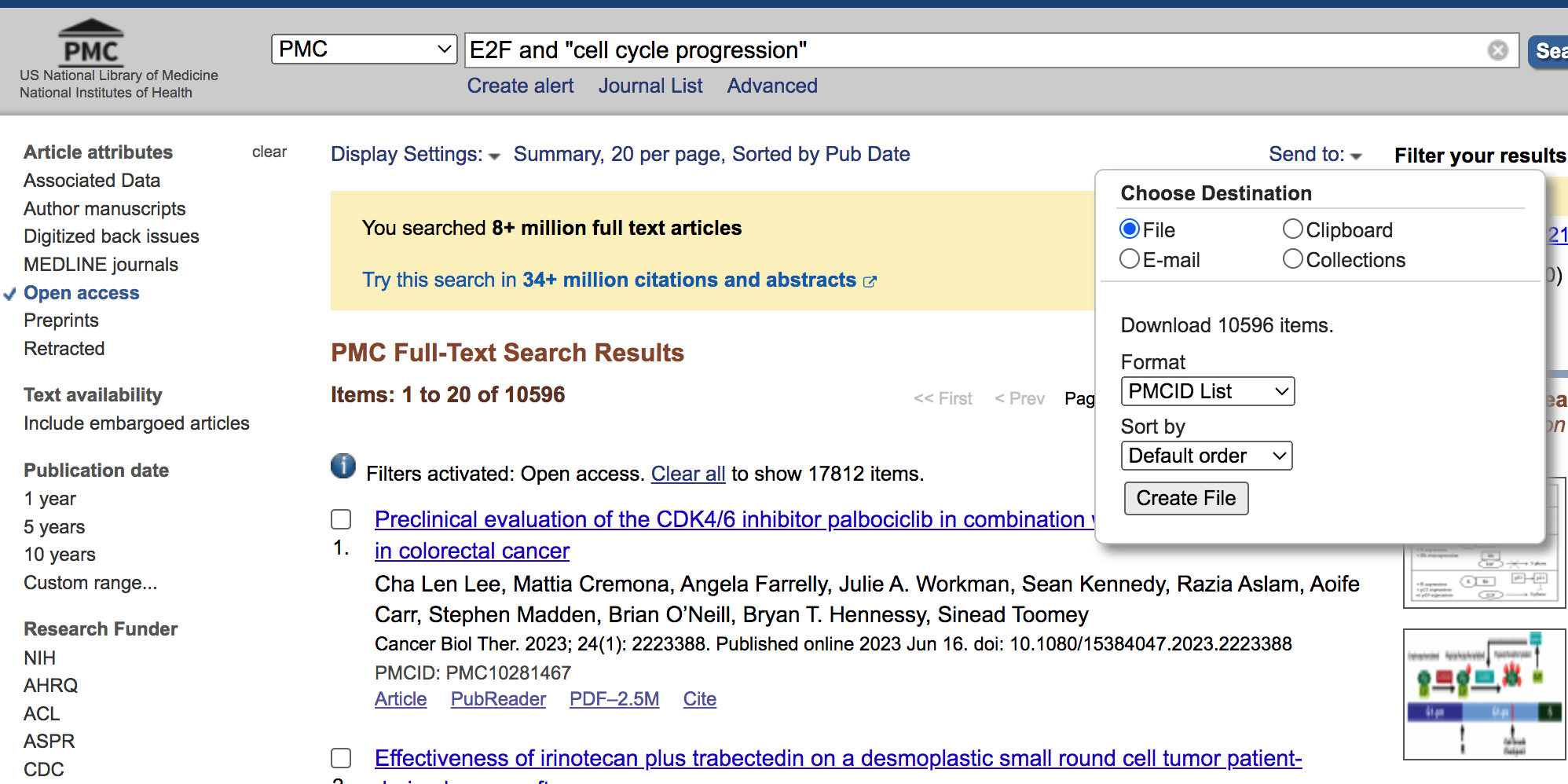
1. Write one or more literature queries for use with PMC
   1. Combine terms with AND, OR, or NOT
   2. Use quotation marks for phrases containing whitespace
   3. Example:

E2F AND “cell cycle progression”

1. Retrieve literature corpus from [PMC](https://www.ncbi.nlm.nih.gov/pmc/) by entering query into search box



1. Save results as a list of PMCIDs by first filtering to “Open access”, then selecting “File” as the destination and “PMCID List” as the format.



1. Use script to search the INDRA database for all statements from literature corpus
   1. [Install INDRA](https://indra.readthedocs.io/en/latest/installation.html#installing-releases-with-pip) with pip or by cloning the git repository.
   2. Install any other dependencies (you can skip Pyjnius, there are several workarounds needed for Mac users)
   3. Set up config file (Mac users):
      1. Open the config file using the following command:



* + 1. Set the URL key, the API key can remain blank:

Text

Description automatically generated with medium confidence

* 1. Run script and input one argument – the name of the CSV file of PMCIDs

Python run\_indra\_pmcids\_biorecipe.py -i [CSV file of PMCIDs] -o [BioRECIPE]

* 1. The script will translate all formatted statements into BioRECIPE interaction lists, and also print to screen the ids of all papers that had no statements.
     1. Potential pitfall – the INDRA database may not contain reading or database results from new papers. For the papers that are missing from the INDRA database, it may be possible to process them separately through a reader like REACH, Eidos, etc.