**Απαραίτητα link - steps :**

* download mongodb - <https://www.mongodb.com/try/download/community>
* download data - <https://www.aminer.org/citation>
* download mongodb compass - <https://www.mongodb.com/try/download/compass>
* in C:\Program Files\MongoDB\Server\4.4\bin download (extract) mongodb-database-tools-windows-x86\_64-100.2.0 (omngoimport command)
* cd C:\Program Files\MongoDB\Server\4.4\bin\mongodb-database-tools-windows-x86\_64-100.2.0\bin
* mongoimport --db MY\_DB --collection users --drop --jsonArray --batchSize 1 --file ./C:\Users\hp\Desktop\ΣΧΟΛΗ\ειδικό\dblp.v12.json
* download eclipse ide
* download and fix tomcat 9.0 server - <https://www.youtube.com/watch?v=tz1lrUcR0r4>
* Download rar file (Eidiko Thema.rar) and extract in eclipse-workspace (C:\Users\hp\eclipse-workspace)
* In eclipse ide ->import ->existing project into workspace ->choose file
* Finally index.html -> run as -> run on server

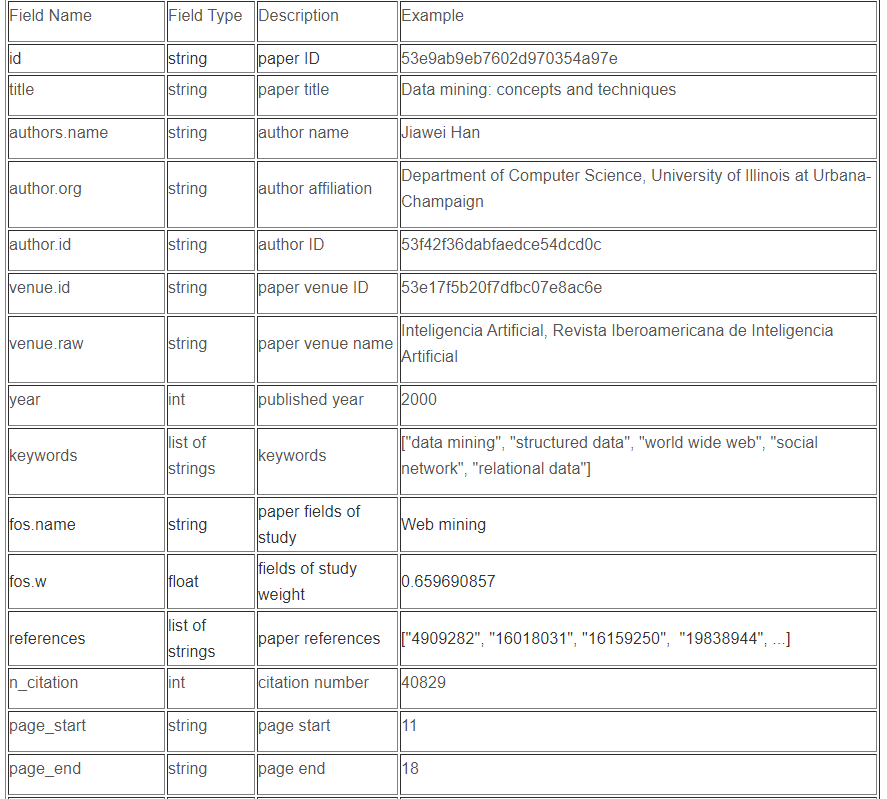
Link για database scaling : <https://learning.oreilly.com/library/view/scaling-mongodb/9781449303679/>

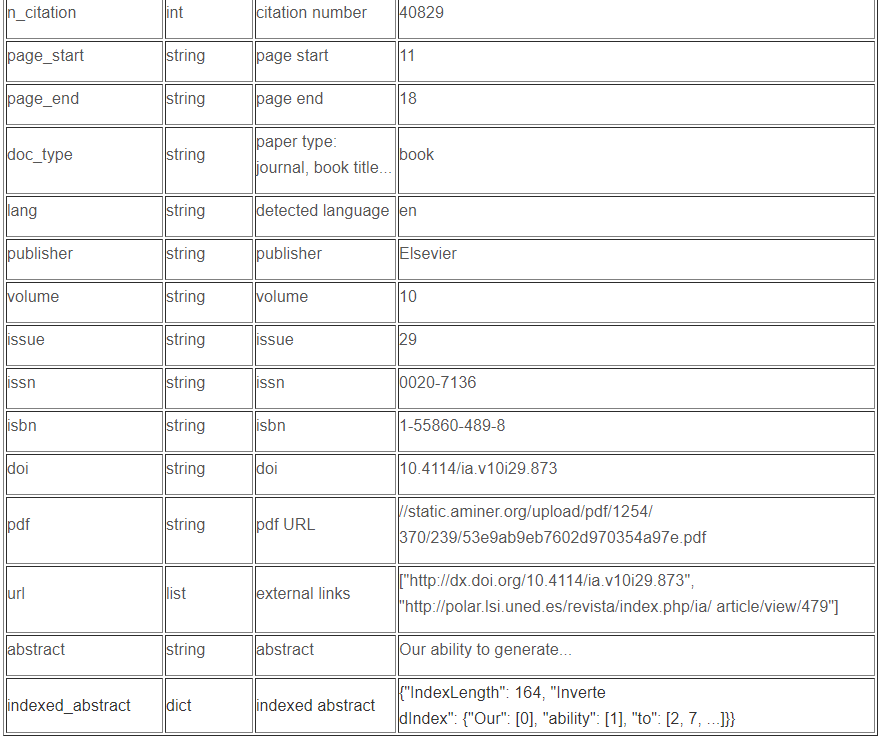
(Scaling MongoDB, by Kristina Chodorow, Publisher: O'Reilly Media, Inc., ISBN: 9781449303211)

**Περιγραφή Βάσης Δεδομένων που χρησιμοποιείται:**

The dataset is organized in a JSON file. You can load it as a whole in JSON format, or you can load it line by line (except the first and last line, each line is a paper dictionary).

The data schema is shown in Table 3 below. This version includes the following attributes: id, title, authors.name, authors.org, authors.id, venue.id, venue.raw, year, fos.name, fos.w, references, n\_citation, page\_start, page\_end, doc\_type, publisher, volume, issue, doi, indexed\_abstract. Id involved in V12 (paper id, author id, venue id) is represented as Long type (instead of string).

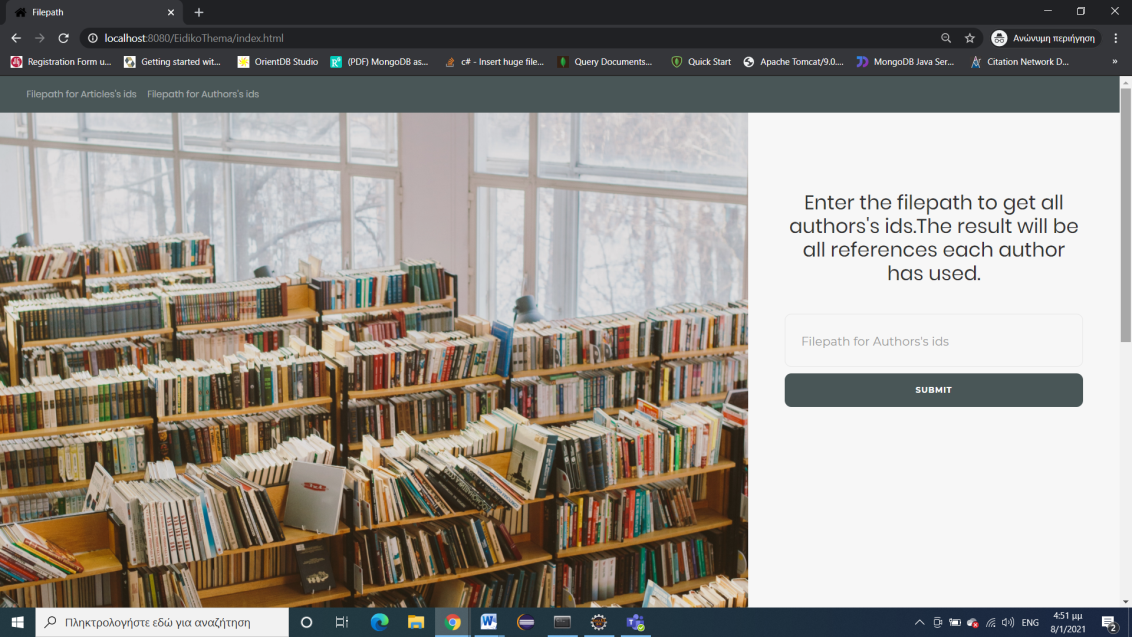




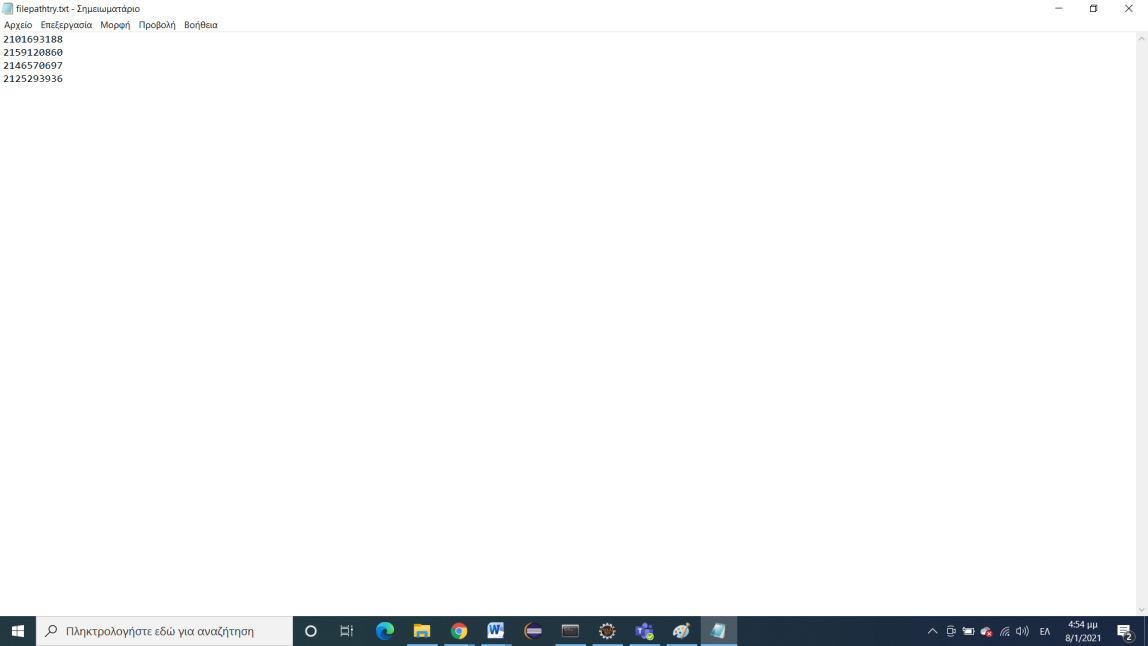
**Περιγραφή web application**

Στην εφαρμογή έχουν αναπτυχθεί 2 ερωτήσεις:

1. Ο χρήστης καλείται να δώσει το filepath ενός αρχείου που περιέχει ids συγγραφέων. Στην επόμενη σελίδα εμφανίζεται ένα block για κάθε συγγραφέα που δείχνει πόσα references έχει χρησιμοποιήσει συνολικά σε όλα τα άρθρα του και ποιά (το κάθε άρθρο εμφανίζεται μια μοναδική φορά). Το ίδιο αποτέλεσμα γράφεται και σε ένα αρχείο που δημιουργείται αυτόματα (queston1.txt).

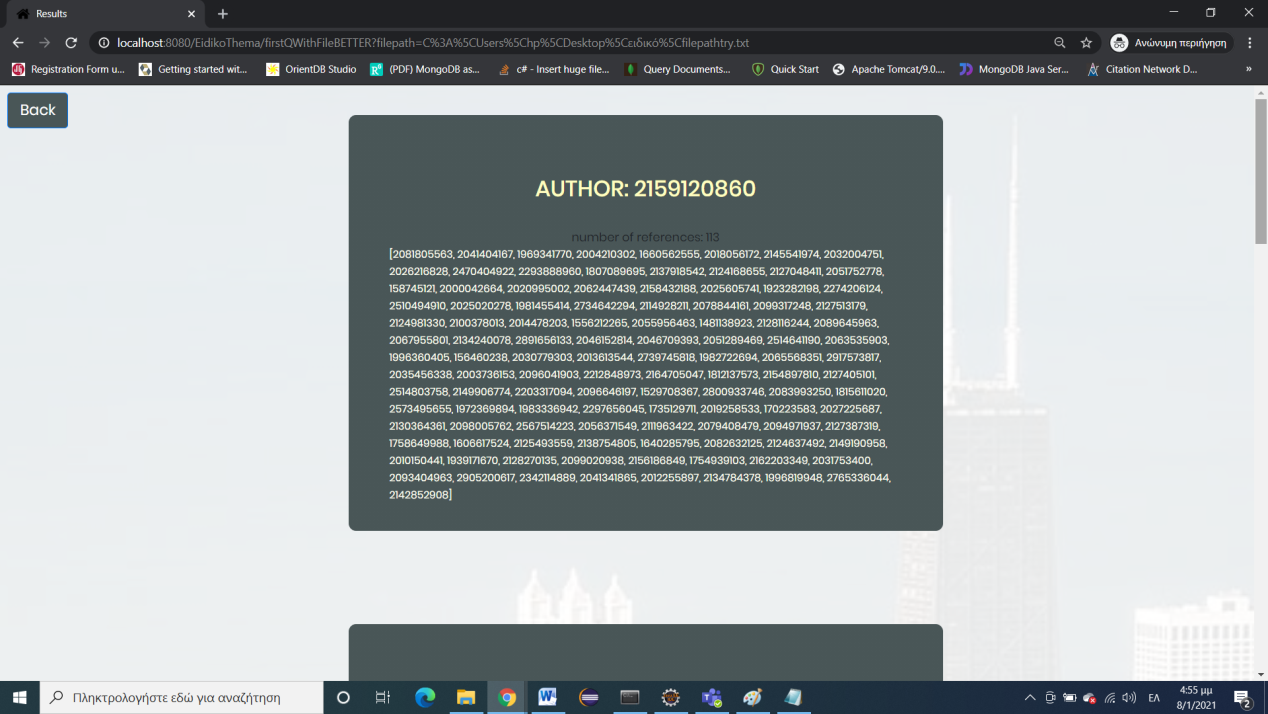


Για παράδειγμα βάζω ως input το παρακάτω αρχείο (filepathtry.txt):



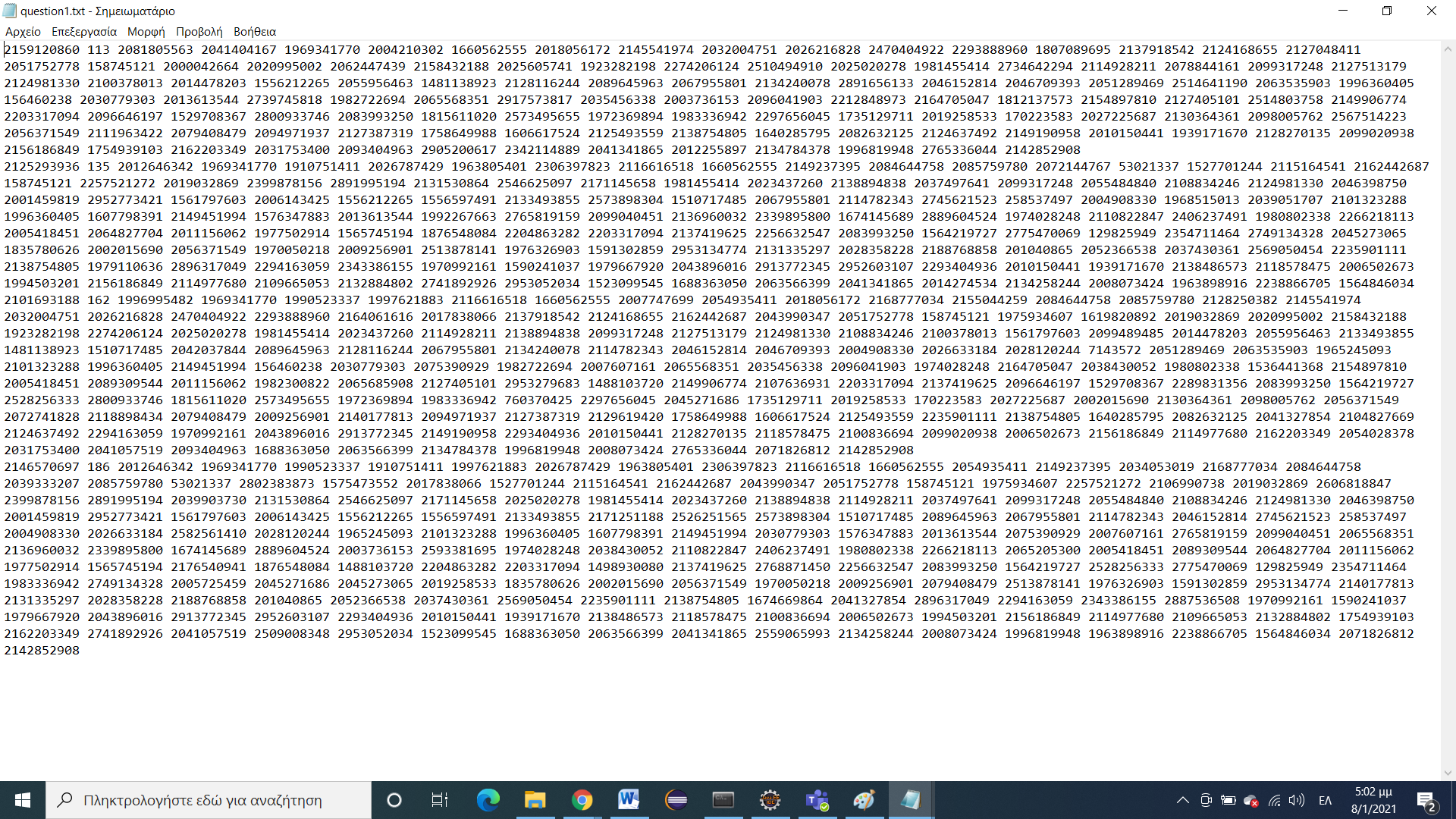
Στο οποίο έχω πάρει 4 τυχαία authors’ ids από τη mongodb βάση.

Ως έξοδος προκύπτει:

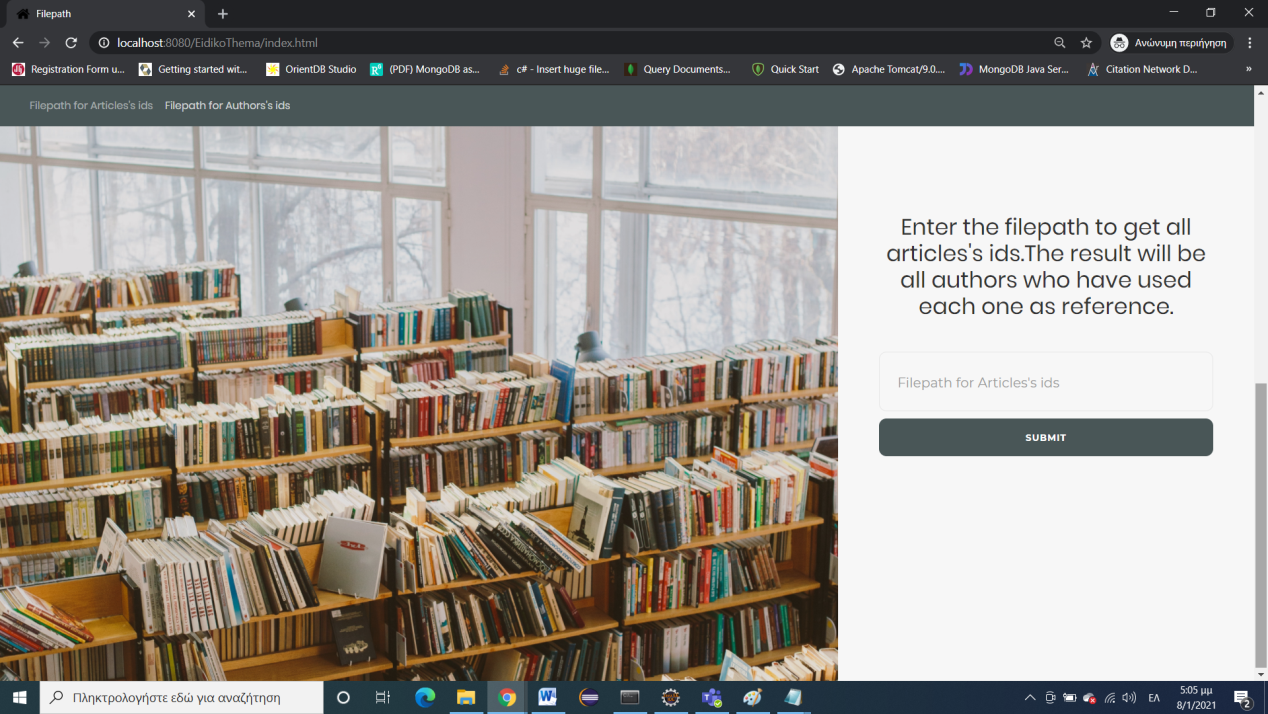


Όπως φαίνεται δημιουργείται ένα block για κάθε συγγραφέα που αναφέρεται πόσα references έχει χρησιμοποιήσει και ποια.

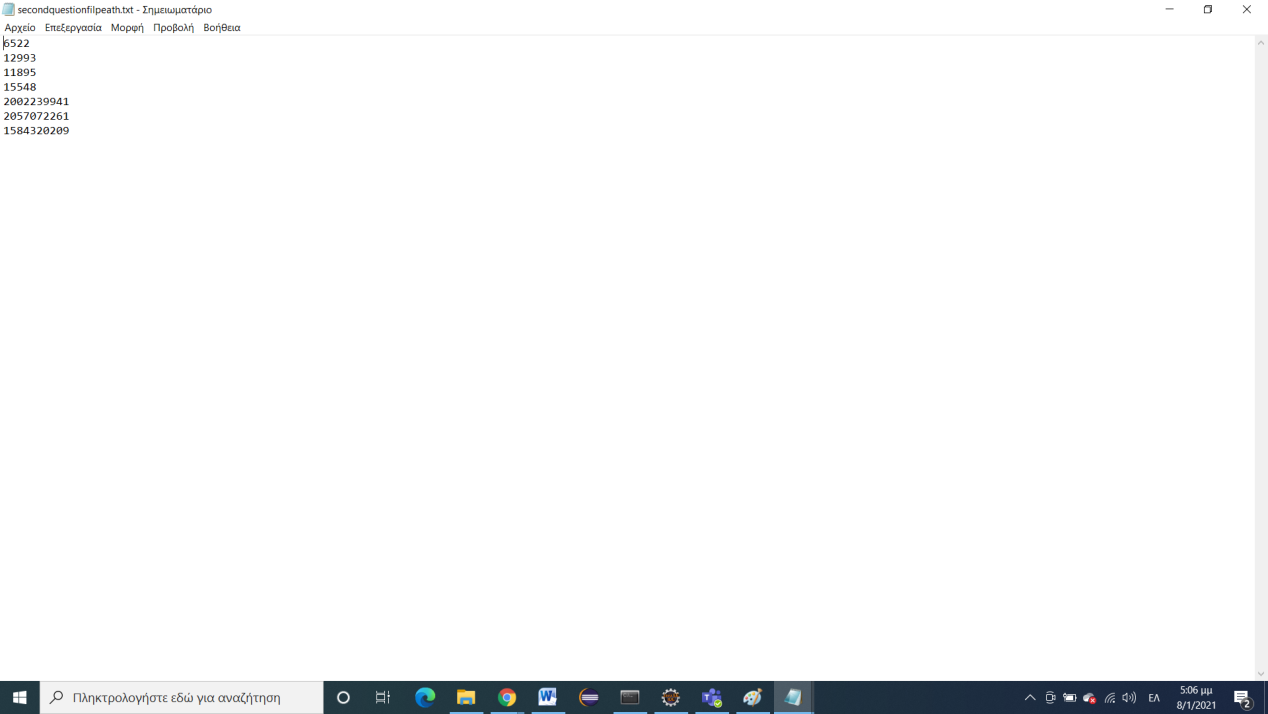
Ταυτόχρονα δημιουργείται και το αρχείο που προανέφερα (question1.txt) και έχει την παρακάτω μορφή:



2.Ο χρήστης καλείται να δώσει το filepath ενός αρχείου που περιέχει ids άρθρων. Στην επόμενη σελίδα εμφανίζεται ένα block για κάθε άρθρο που δείχνει πόσοι συγγραφείς το έχουν χρησιμοποιήσει ως reference και ποιοί (ο κάθε συγγραφέας εμφανίζεται μια μοναδική φορά). Το ίδιο αποτέλεσμα γράφεται και σε ένα αρχείο που δημιουργείται αυτόματα (queston2.txt).

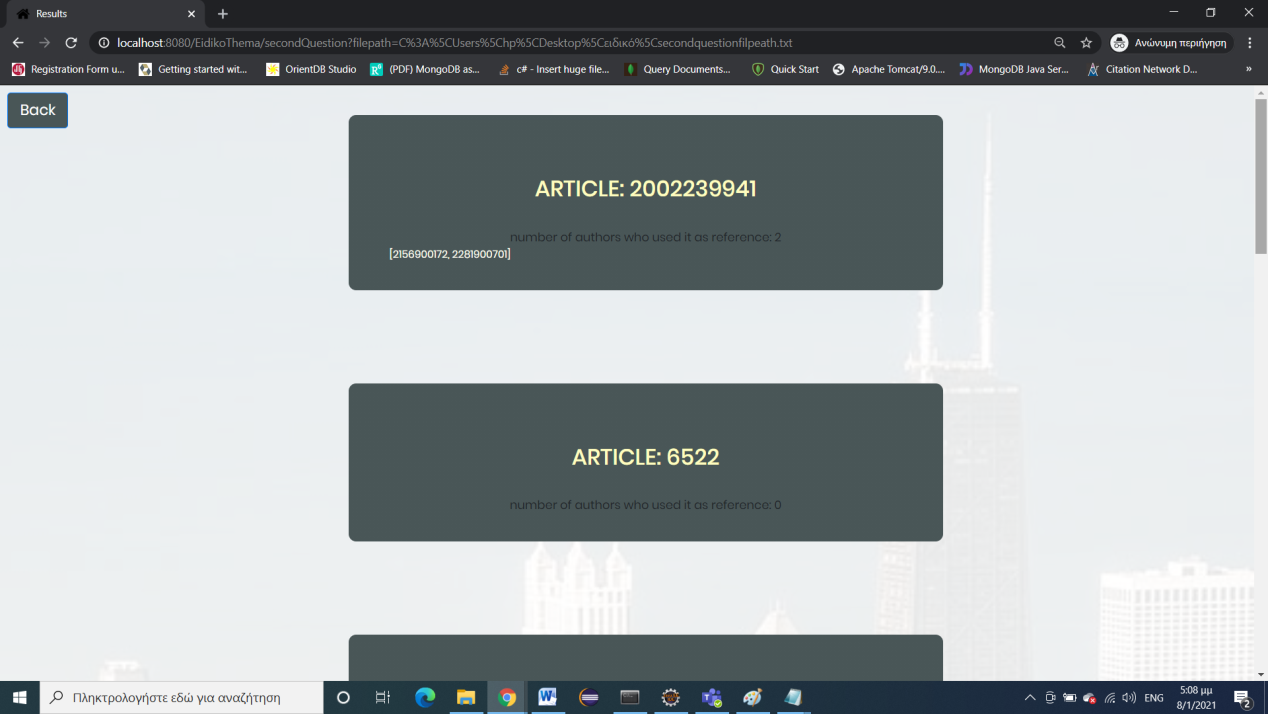


Για παράδειγμα βάζω ως input το παρακάτω αρχείο (secondquestionfilpeath.txt):



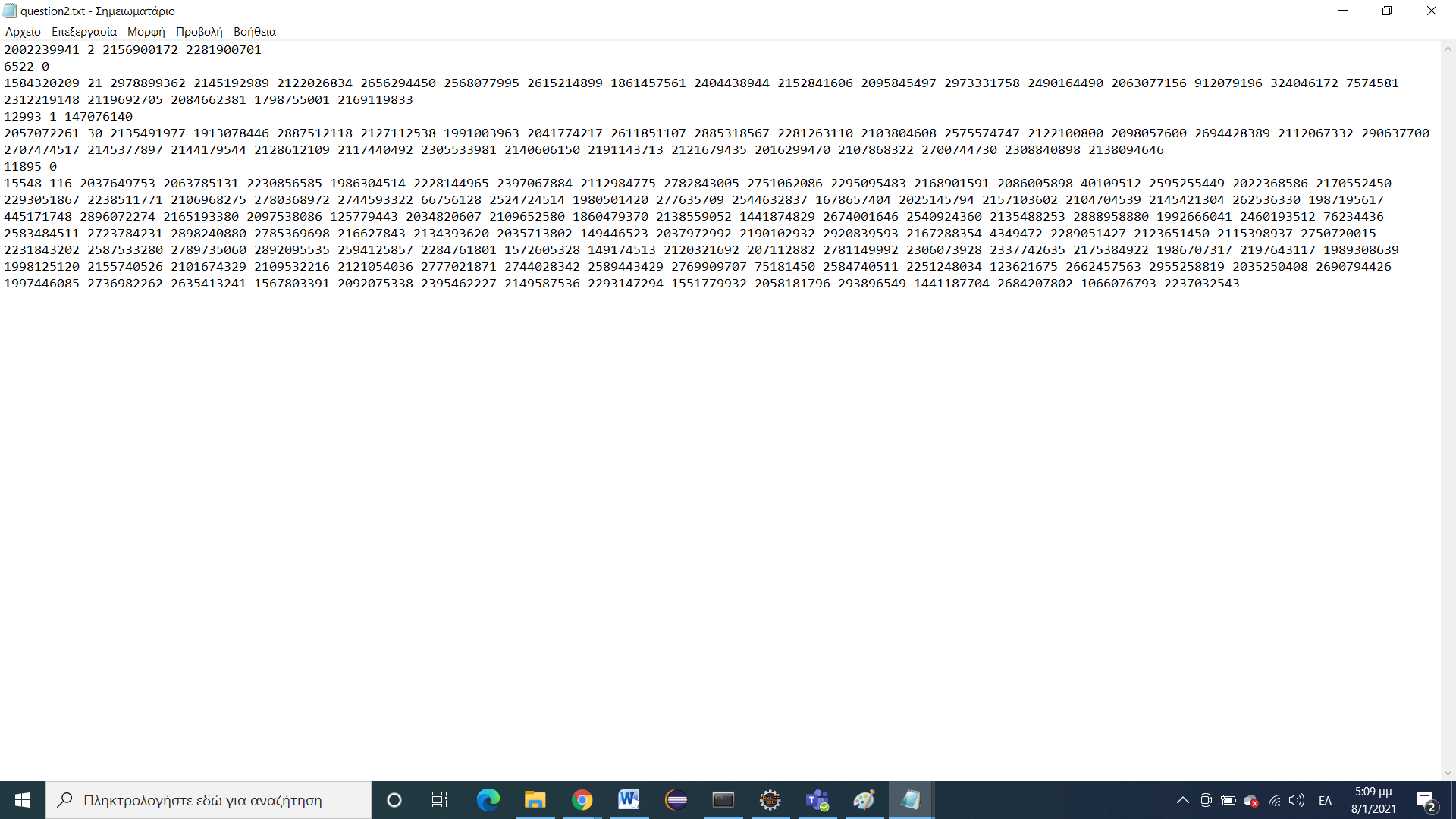
Στο οποίο έχω πάρει 7 τυχαία articles’ ids από τη mongodb βάση.

Ως έξοδος προκύπτει:



Όπως φαίνεται δημιουργείται ένα block για κάθε άρθρο που αναφέρεται πόσοι συγγραφείς το έχουν χρησιμοποιήσει ως references και ποια.

Ταυτόχρονα δημιουργείται και το αρχείο που προανέφερα (question2.txt) και έχει την παρακάτω μορφή:



**Python Υλοποίηση**

Την ίδια εφαρμογή υλοποίησα και σε γλώσσα python χωρίς να την αναπτύσσω ως web app.

Ο χρήστης καλείται να τρέξει το .ipynb αρχείο σε Jupyter notebook και να δώσει το filepath ακριβώς όπως και πριν και προκύπτουν πάλι τα ίδιας μορφής αρχεία.

1ο ερώτημα:



Output file: 1questionpython.txt

2ο ερώτημα:



Output file: 2questionpython.txt