**Εφαρμογή αποθήκευσης και ανάκτησης επιστημονικών άρθρων**

Oμάδα 13

Βαγγέλης Χριστόφορος

Α.Μ:1053715  
 Τμήμα ΗΜΤΥ  
 Πανεπιστήμιο Πατρών  
 Πάτρα,Ελλάδα  
 crisvaggelis@gmail.com

Γκανάς Ιωάννης  
Α.Μ:1053577

Τμημα ΗΜΤΥ  
 Πανεπιστήμιο Πατρών  
 Πάτρα,Ελλάδα  
john13ganas@gmail.com

**Περίληψη**

Θέμα του πρότζεκτ που μας ανατέθηκε ήταν η δημιουργία μίας εφαρμογής, η οποία θα δίνει την δυνατότητα στον χρήστη να αναζητά άρθρα επιστημονικών περιοδικών που ανήκουν στην βάση δεδομένων μας και να τα αποθηκεύει. Η αναζήτηση αυτή μπορεί να γίνει αρχικά μέσω φιλτραρίσματος των άρθρων με κριτήρια όπως το όνομα του συγγραφέα, το επιστημονικό πεδίο ή τα επιστημονικά πεδία στα οποία γίνεται αναφορά κ.α.. Επιπλέον ο χρήστης καλείται να διαλέξει μεταξύ δυο τρόπων αναζήτησης: του άμεσου (ο χρήστης μετά την επιλογή του φίλτρου που θέλει να χρησιμοποιήσει πληκτρολογεί το όνομα που αντιστοιχεί στο πεδίο που αναζητεί και του παρουσιάζονται τα αντίστοιχα αποτελέσματα) και του έμμεσου (μετά την επιλογή του άρθρου παρουσιάζονται όλες οι διαθέσιμες επιλογές του επιλεγμένου πεδίου π.χ. όλοι οι συγγραφείς που περιλαμβάνονται στην βάση μας και πατώντας τον αντίστοιχο αριθμό του παρουσιάζονται τα αποτελέσματα).Παράλληλα με την εύρεση του άρθρου δίνονται διάφορες επιλογές στον χρήστη όπως το άνοιγμα του σε μορφή .pdf, η παρουσίαση αναλυτικών πληροφοριών για αυτό (όπως η γλώσσα, το τεύχος έκδοσης του κ.α.), το κατέβασμα-αποθήκευση του κ.α.. Δικός μας στόχος ήταν η δημιουργία μιας εύχρηστης εφαρμογής η οποία θα δίνει αρκετές δυνατότητες σε αυτόν που την χρησιμοποιεί.

**Λέξεις Κλειδιά**

Βάση δεδομένων, Οντότητες, Συσχετίσεις, Πρωτεύοντα κλειδιά, Ξένα κλειδιά, Πίνακες

1.**Μεθοδολογία**

Το πρώτο βήμα που κάναμε όταν μας δόθηκε η άσκηση ήταν να σκεφτούμε την χρήση της εφαρμογής που έπρεπε να φτιάξουμε και κατ’ επέκταση τον μικρόκοσμο του συγκεκριμένου προβλήματος.

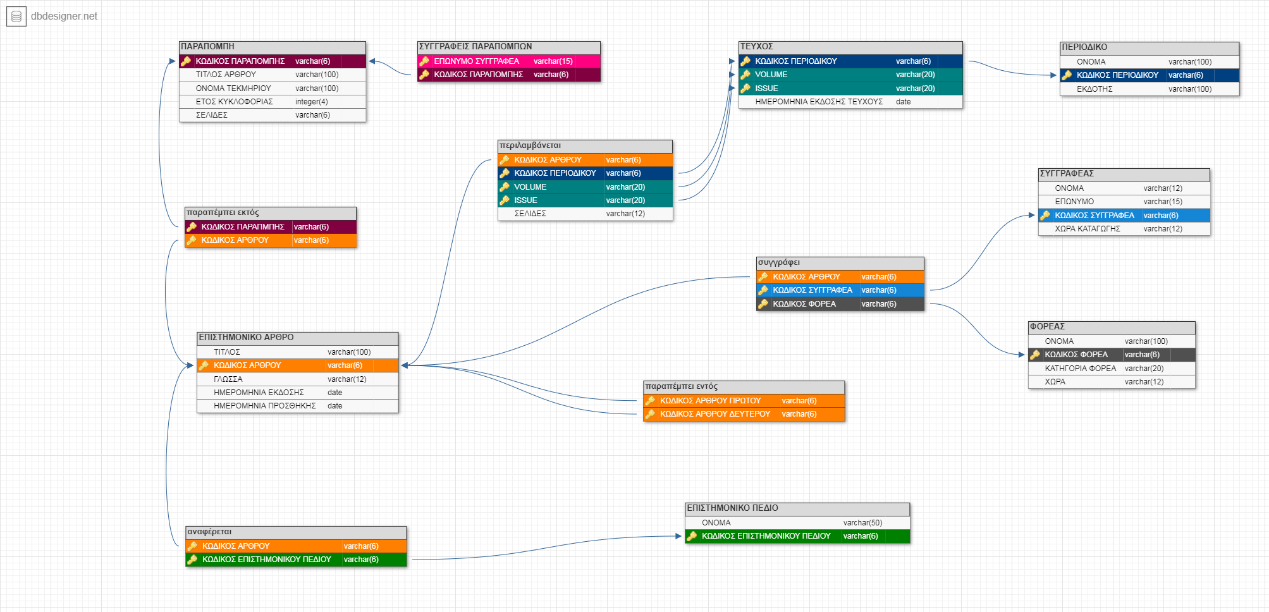
*Μικρόκοσμος:* Στην βάση μας θα αποθηκεύονται επιστημονικά άρθρα επιστημονικών περιοδικών με σκοπό την δημιουργία μιας ψηφιακής βιβλιοθήκης. Κάθε επιστημονικό άρθρο που θα χαρακτηρίζεται από βασικά στοιχεία όπως ο τίτλος, η γλώσσα, η ημερομηνία έκδοσης, θα αναφέρεται σε Επιστημονικά πεδία και θα εκδίδεται από συγκεκριμένο τεύχος ενός περιοδικού. Επίσης κάθε συγγραφέας θα συνεργάζεται με έναν φορέα(π.χ. ένα εκπαιδευτικό ίδρυμα) για την συγγραφή του. Τέλος κάθε άρθρο μπορεί να παραπέμπει σε άλλα άρθρα ή τεκμήρια που είτε θα υπάρχουν είτε όχι στην βάση έτσι ώστε να βοηθάει τον χρήστη στην αναζήτηση της πληροφορίας που τον ενδιαφέρει.

Το δεύτερο βήμα που κάναμε ήταν η μετάβαση από τον μικρόκοσμο στο λογικό μοντέλο ERD με την χρήση του ηλεκτρονικού εργαλείου **ERD Plus**. Μετά από διάφορα προβλήματα που αντιμετωπίσαμε και μετά από διάφορες αλλαγές στο μοντέλο μας καταλήξαμε στο ακόλουθο:

Εικόνα που περιέχει κείμενο, χάρτης

Περιγραφή που δημιουργήθηκε αυτόματα

Το τρίτο βήμα κάναμε ήταν η μετάβαση από το λογικό μοντέλο στο σχεσιακό μοντέλο (πίνακες) με την χρήση του ηλεκτρονικού εργαλείου **dbdesigner.** Έτσι προέκυψαν οι ακόλουθοι πίνακες:



Το τέταρτο βήμα μας ήταν η δημιουργία της βάσης μας με εντολές SQL δίνοντας ιδιαίτερη έμφαση στους τύπους δεδομένων και στους περιορισμούς αναφορικής ακεραιότητας. Αρχικά εξήγαμε τον κώδικα από το **dbdesigner** αλλά στην συνέχεια αντιμετωπίσαμε κάποια προβλήματα και έπρεπε να τον αλλάξουμε σε αρκετά σημεία. Εν τέλει δημιουργήσαμε την βάση μας με την εφαρμογή **Phpmyadmin** που χρησιμοποιήσαμε στο εργαστήριο.

Το πέμπτο μας βήμα ήταν η εισαγωγή των δεδομένων .

Τέλος, η υλοποίηση της εφαρμογής έγινε με την χρήση της γλώσσας Python καθώς και την εγκατάσταση κάποιων πρόσθετων βιβλιοθηκών της όπως η Pymysql (έτσι ώστε να συνδέσουμε την εφαρμογή μας με την βάση μας, η requests και η google\_drive\_downloader (για το κατέβασμα του άρθρου) και η webbrowser για την σύνδεση με τον φάκελο που περιλαμβάνει τα άρθρα και βρίσκεται στο Google Drive .

2. Αξιολόγηση

Το βασικό κριτήριο για την σχεδίαση της βάσης μας είναι η τοποθέτηση των πληροφοριών σε πίνακες βάση θέματος με σκοπό την μείωση των επιπρόσθετων δεδομένων καθώς και η εύκολη πρόσβαση τους. Ακόμα ένα σημαντικό κριτήριο αποτελεί η ακρίβεια και η ακεραιότητα των πληροφοριών μας καθώς και η αποφυγή αλληλοεπικάλυψης δεδομένων στην βάση μας.

Στην συνέχεια για την περαιτέρω αξιολόγηση της βάσης μας κάναμε ένα τεστ ταχύτητας της βάσης μας επιλέγοντας κάποια queries και βρίσκοντας τον χρόνο που χρειάζεται για να εκτελεστούν. Χρησιμοποιήσαμε την βιβλιοθήκη time και κάναμε χρήση του κώδικα:

Start=time()

Author\_search()

print('It took', time()-start, 'seconds.')



**Χρόνος εκτέλεσης του query 1**

****

**Χρόνος εκτέλεσης του query 2**

****

**Χρόνος εκτέλεσης του query 3**

Παρατηρούμε ότι ο χρόνος εκτέλεσης των queries είναι αρκετά μικρός (ο μέσος χρόνος εκτέλεσης ενός query είναι σχεδόν ίδιος όπως στις φωτογραφίες), το οποίο είναι απαραίτητο για μια βάση.

Ελέγξαμε επίσης τη σωστή λειτουργία και τον μικρό χρόνο εκτέλεσης των λειτουργιών στη βάσης μέσω εκτέλεση μίας εντολής insert (ή delete) σε πίνακα.

Ακόμα ελέγξαμε ότι δεν επιτρέπεται σε όλα τα child rows να εκτελεστούν οι εντολές UPDATE και DELETE, διότι προσθέσαμε περιορισμούς αναφορικής ακεραιότητας.

Τέλος ένα άλλο σημαντικό κριτήριο είναι η ασφάλεια της βάσης μας. Εάν ο χρήστης πληκτρολογήσει λανθασμένο password, username ή host ID η σύνδεση είναι ανεπιτυχής.

3. Δεδομένα

Η συλλογή δεδομένων των άρθρων και των αντίστοιχων πληροφοριών έγινε από τους ακόλουθους ιστότοπους:

<https://dl.acm.org/>

<https://www.openarchives.gr/aggregator-openarchives/portal/>

<http://www.astronomy.com/>.

Η ενσωμάτωση των δεδομένων αυτών στην ΒΔ μας έγινε με την χρήση της γλώσσας χειρισμού δεδομένων DML-SQL και συγκεκριμένα με την εντολή INSERT.

4. Ενέργειες υλοποίησης του project

Η υλοποίηση του λογικού και του σχεσιακού μοντέλου καθώς και η κατασκευή-εισαγωγή δεδομένων στην βάση έγινε κατόπιν συνεργασίας και των δύο μελών της ομάδας. Στην εφαρμογή συνεργάστηκαν στα περισσότερα κομμάτια κώδικα και τα δύο μέλη της ομάδας δίνοντας έμφαση σε συγκεκριμένα κομμάτια:

Σύνδεση με την βάση, συναρτήσεις άμεσης αναζήτησης, δημιουργία gui (tkinter) από τον Χριστόφορο Βαγγέλη

Συναρτήσεις έμμεσης αναζήτησης, βασικό μενού, πληροφορίες, παραπομπές, άνοιγμα και κατέβασμα άρθρων από τον Γκανά Γιάννη.

5. Χρονοδιάγραμμα

Τις πρώτες τρεις εβδομάδες ασχοληθήκαμε με το διάγραμμα ERD.

Την επόμενη εβδομάδα ασχοληθήκαμε με το Σχεσιακό Διάγραμμα.

Τις επόμενες δύο εβδομάδες ασχοληθήκαμε με την δημιουργία της βάσης (SQL) και το γέμισμα αυτής με δεδομένα.

Τις επόμενες δύο εβδομάδες ασχοληθήκαμε με την δημιουργία της εφαρμογής.

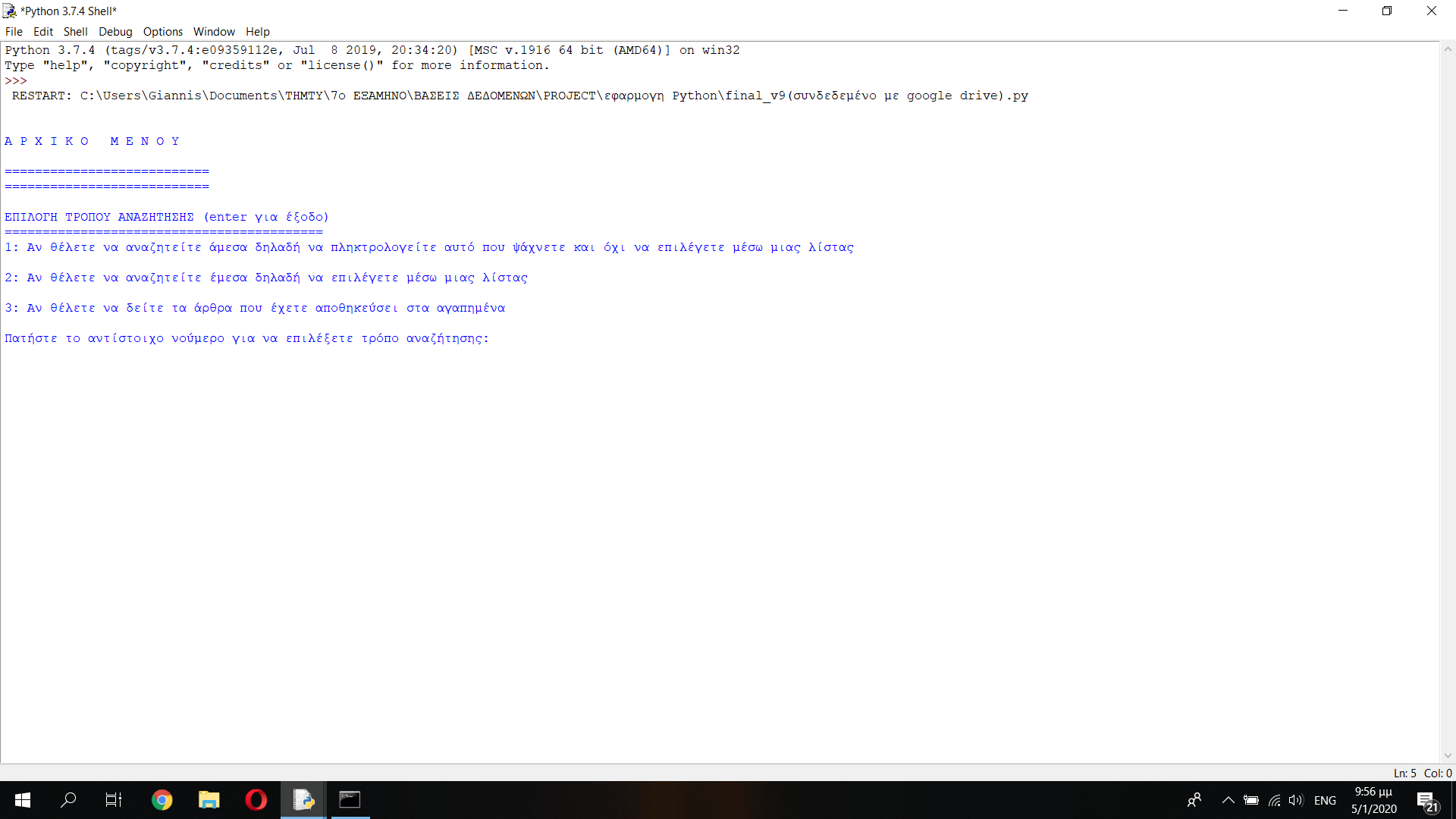
Τέλος κάναμε την έκθεση.

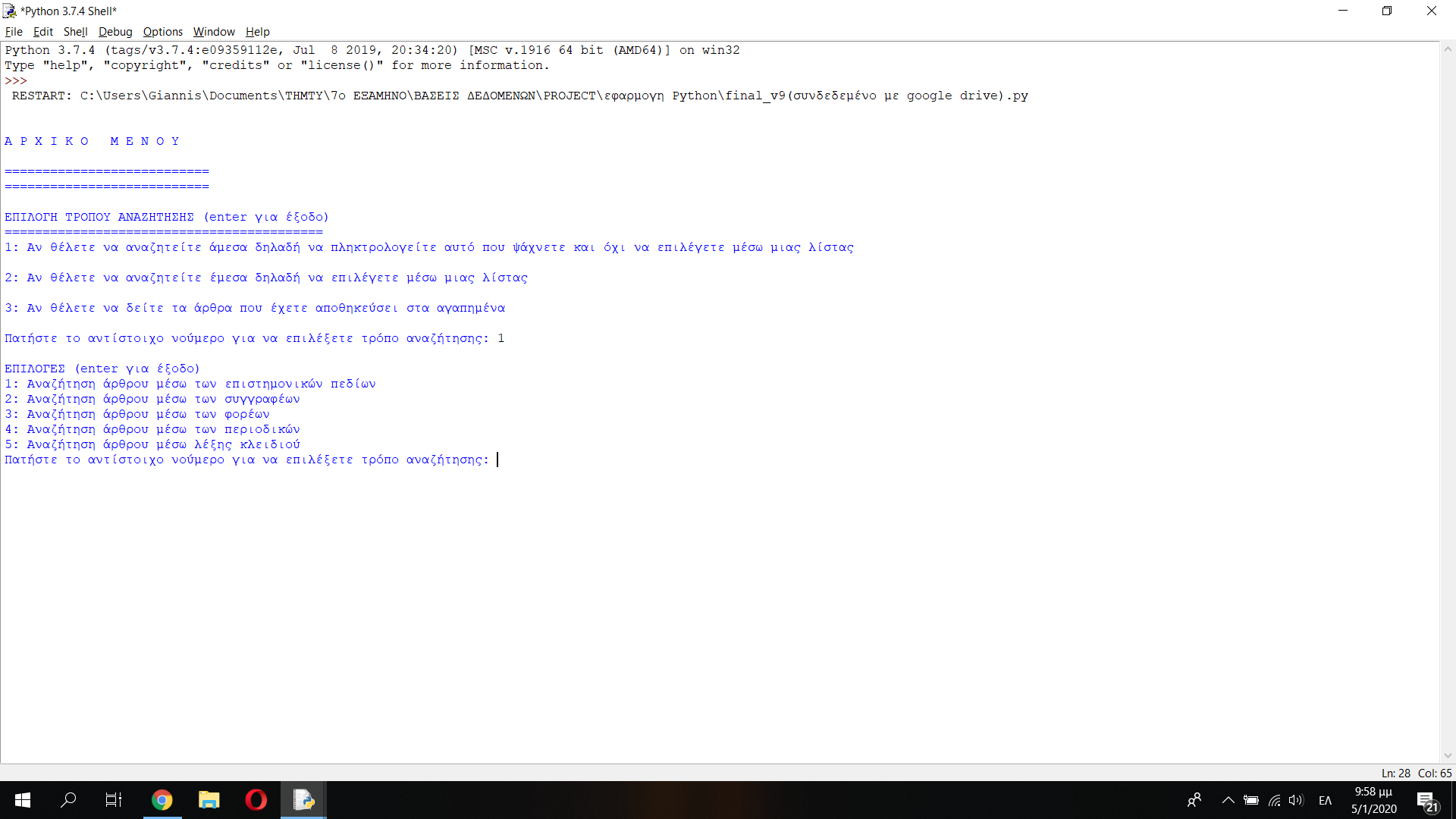
6. Oδηγίες εγκατάστασης & παραδείγματα χρήσης του προγράμματος7

1. Κατεβάστε και αποσυμπιέστε το αρχείο scientific\_articles. Μέσα σε αυτό το αρχείο βρίσκονται δύο αρχεία Python ένα κενός φάκελος που θα μπαίνουν τα άρθρα που κατεβάζουμε και ένα κενό txt στο οποίο θα μπαίνουν οι κωδικοί των αγαπημένων.
2. Ανοίξτε την γραμμή εντολών των windows και εκτελέστε τις παρακάτω εντολές: pip install pymysql , pip install requests και pip install google\_drive\_downloader .
3. Το ένα από τα δύο προγράμματα (tkinter) είναι σε γραφικό περιβάλλον και το άλλο (final) τρέχει μέσω τερματικού. Είναι ξεχωριστα!

Παρακάτω παρουσιάζονται κάποια παραδείγματα χρήσης του προγράμματος στο τερματικό της Python:

Στις παρακάτω εικόνες φαίνονται κατά σειρά το αρχικό μενού και οι επιλογές της άμεσης αναζήτησης πατώντας το 1.



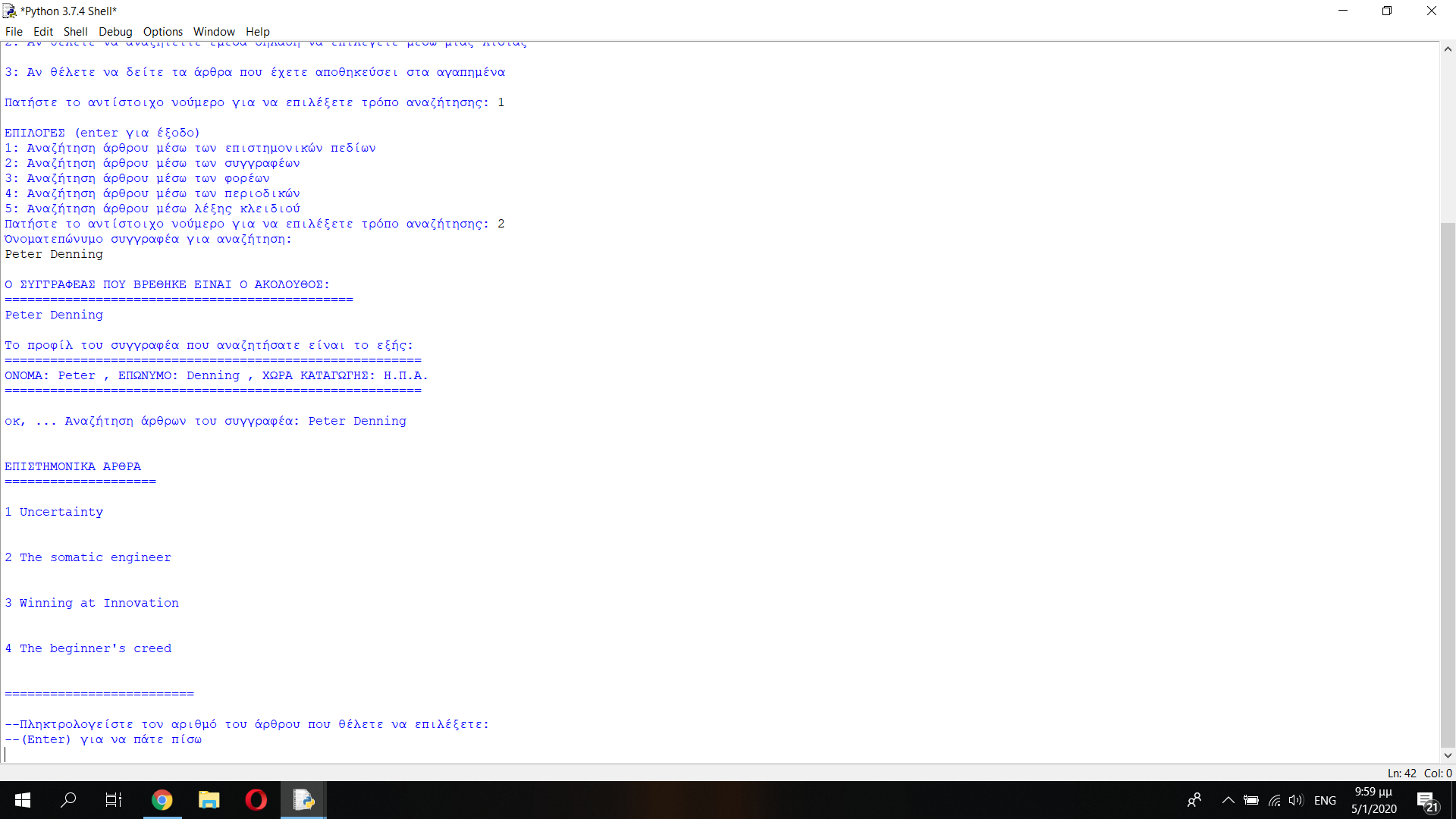


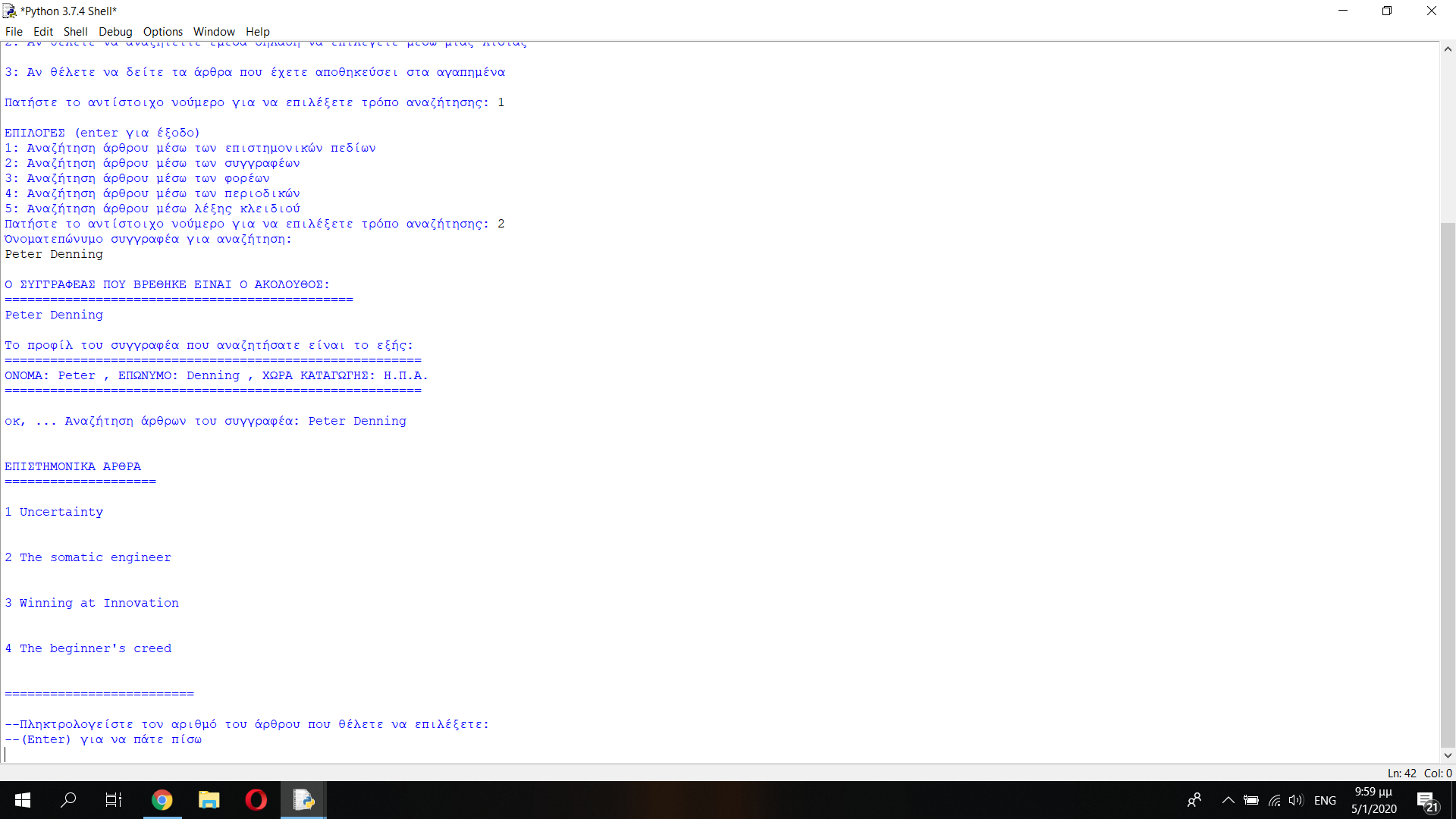
Στις παρακάτω εικόνες φαίνεται η πληκτρολόγηση του ονόματος

ή ενός τμήματος από το όνομα και η παρουσίαση των άρθρων του

συγκεκριμένου συγγραφέα, καθώς και οι προσωπικές

πληροφορίες του συγγραφέα.





Πατώντας το 1 επιλέγουμε το πρώτο άρθρο με τίτλο Uncertainty.

Και οδηγούμαστε στο μενού του άρθρου. Στην συνέχεια

πατώντας το 1 το προσθέτουμε στα αγαπημένα πράγμα που

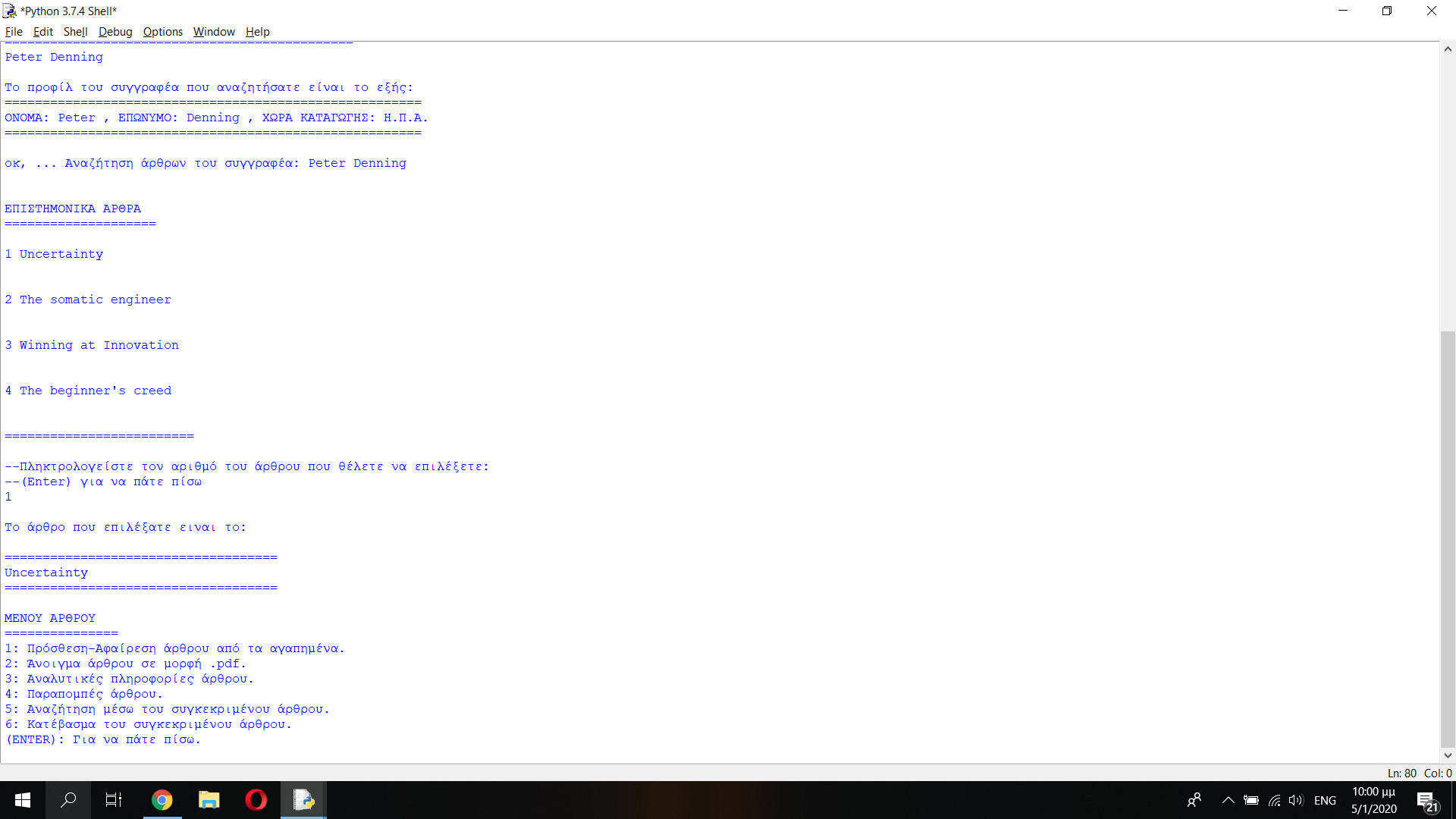
φαίνεται στην τρίτη φωτογραφία όταν από το αρχικό μενού

πατώντας το 3 πάμε στην λίστα με τα αγαπημένα. Πατώντας πάλι

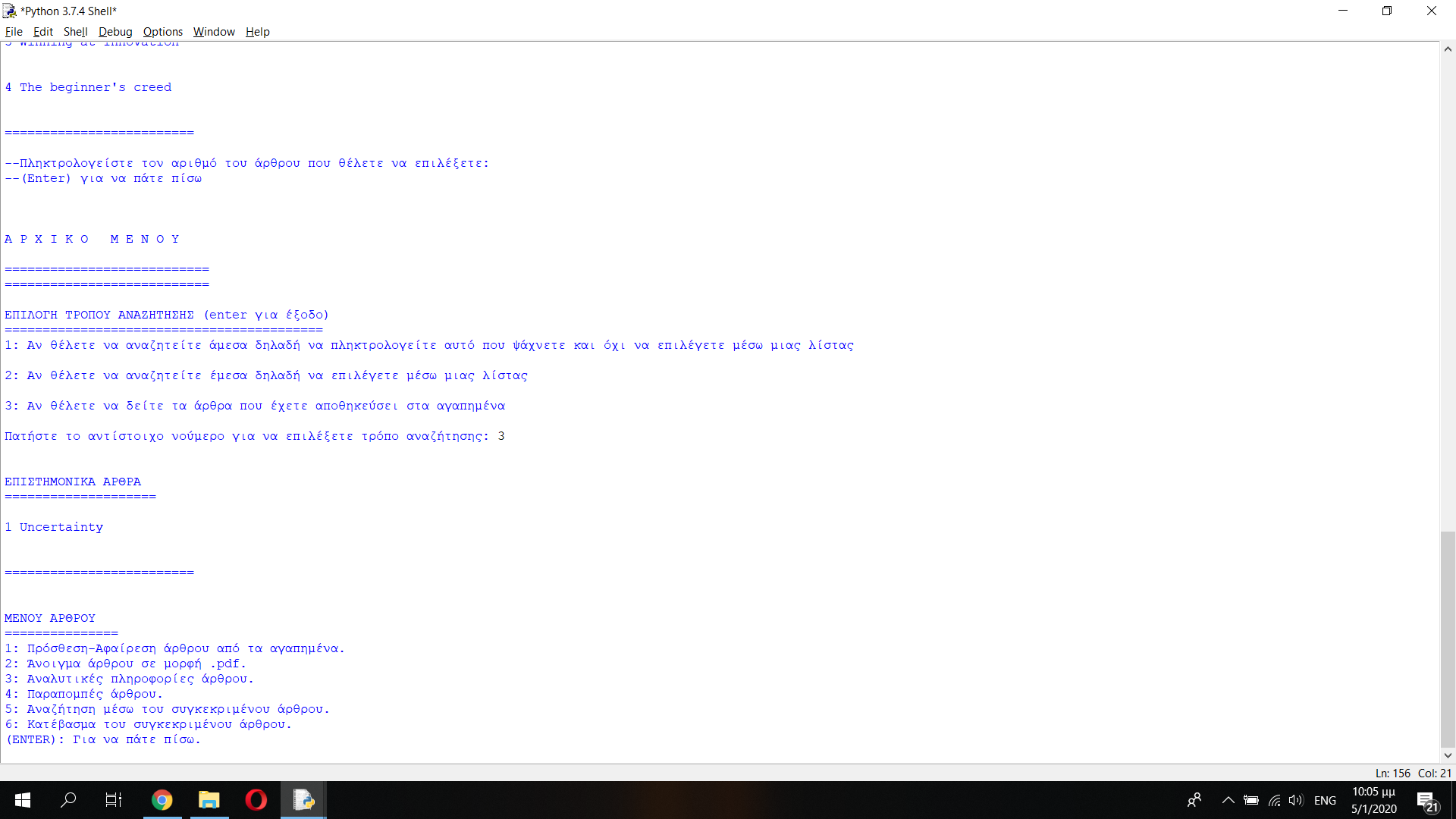
1 στο συγκεκριμένο άρθρο βλέπουμε ότι μπορούμε να το

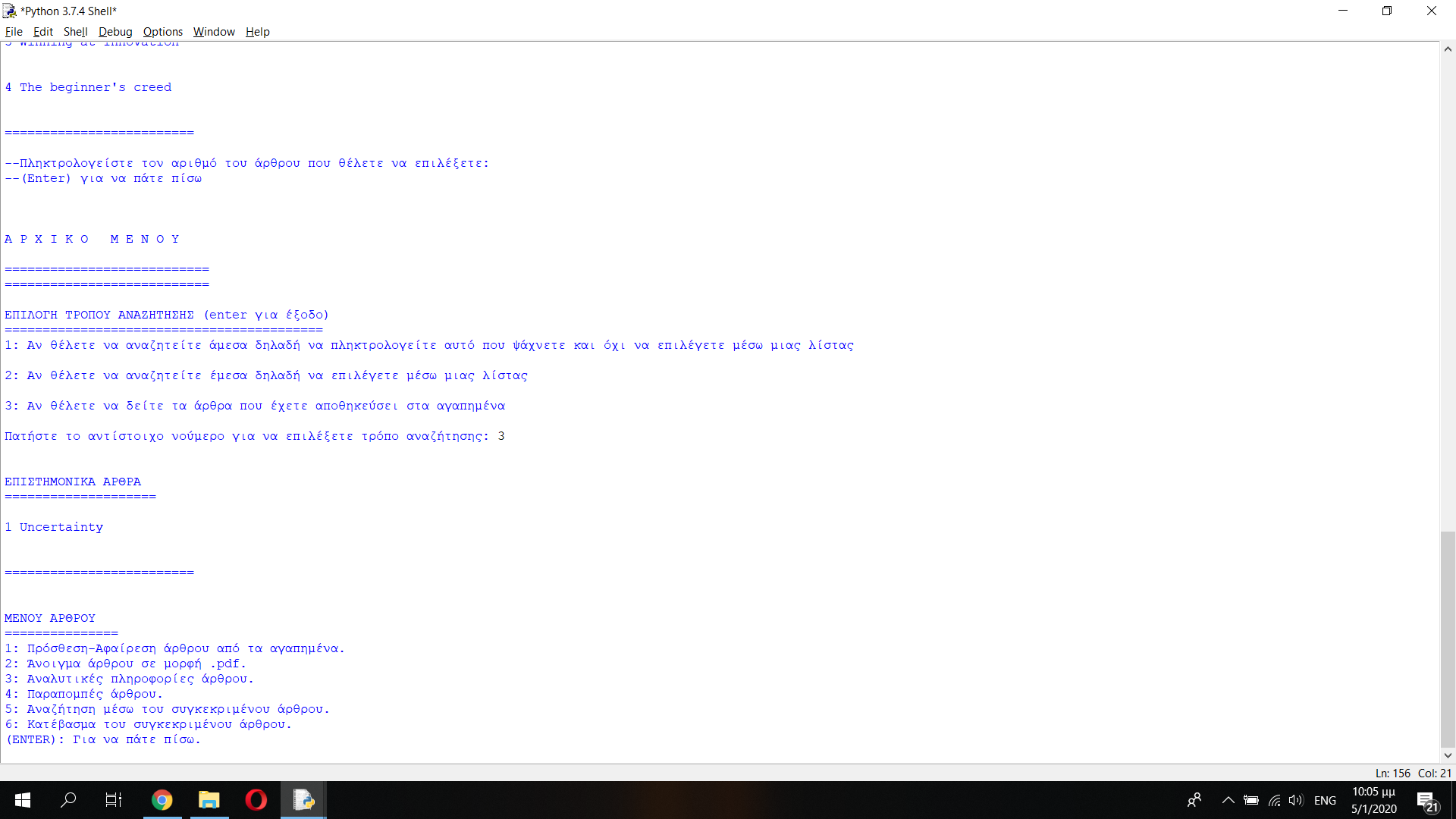
διαγράψουμε από τα αγαπημένα πράγμα που εν τέλει φαινεται

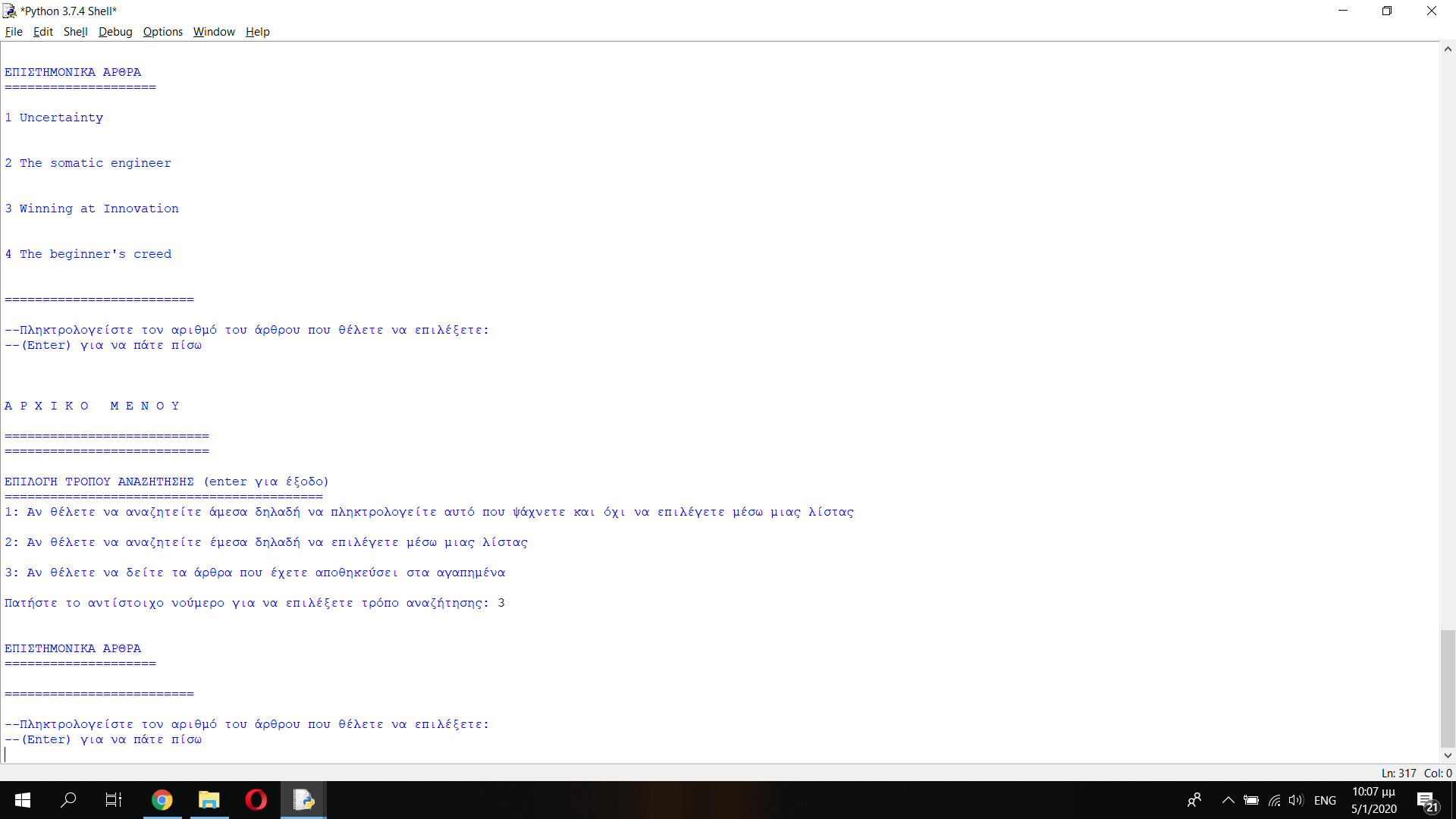
στην τελευταία εικόνα.



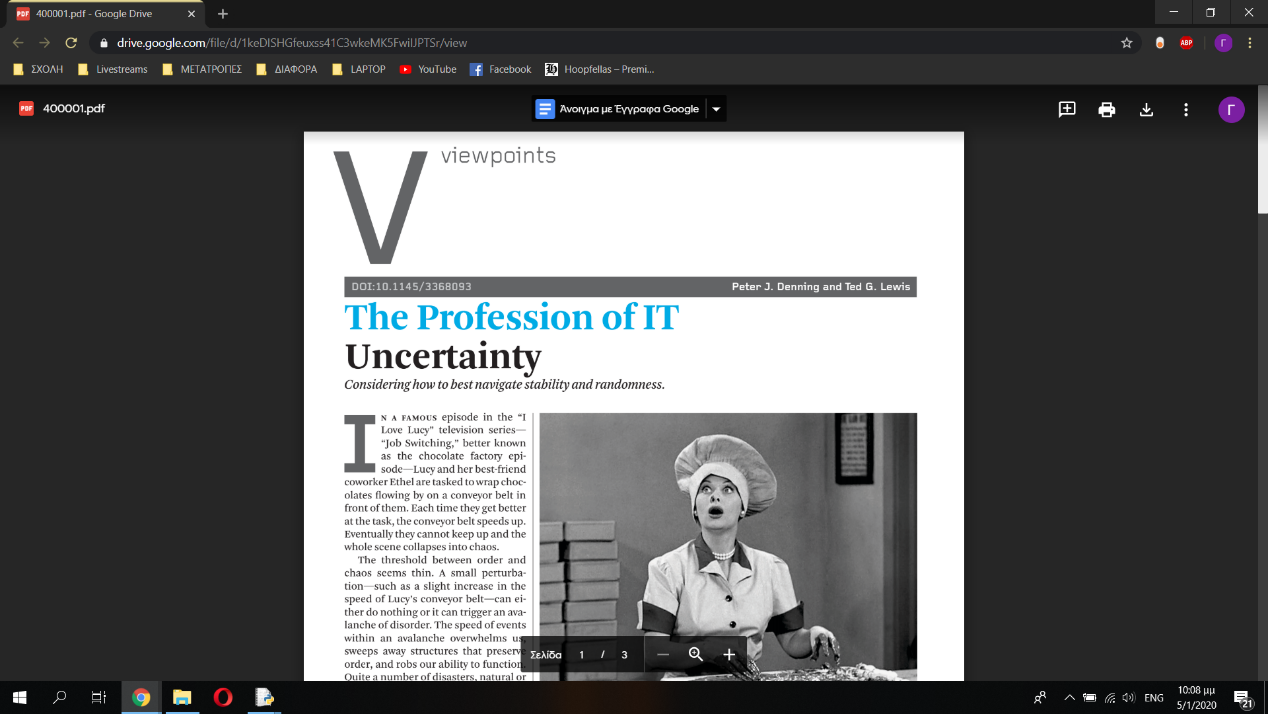






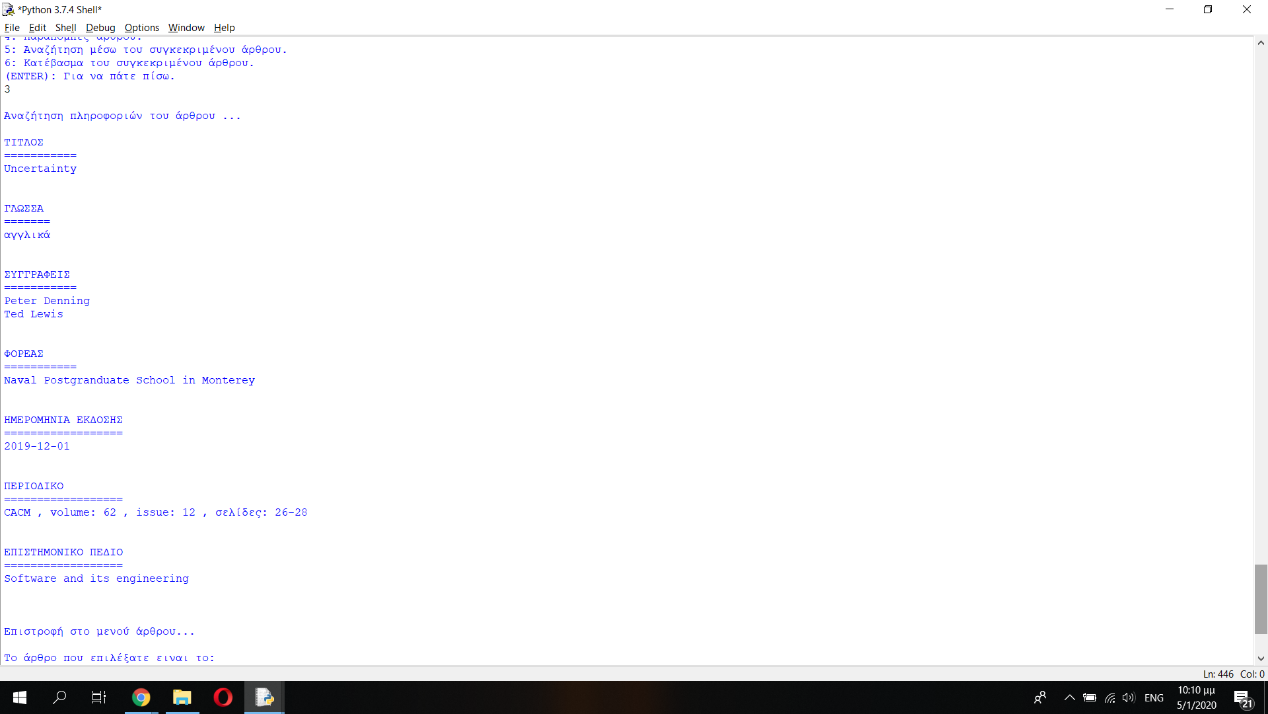


Αν στο μενού άρθρου πατήσουμε το 2 θα ανοίξουμε το άρθρο.



Στην συνέχεια αν πατήσουμε 3 θα βγουν αναλυτικές πληροφορίες

του άρθρου.

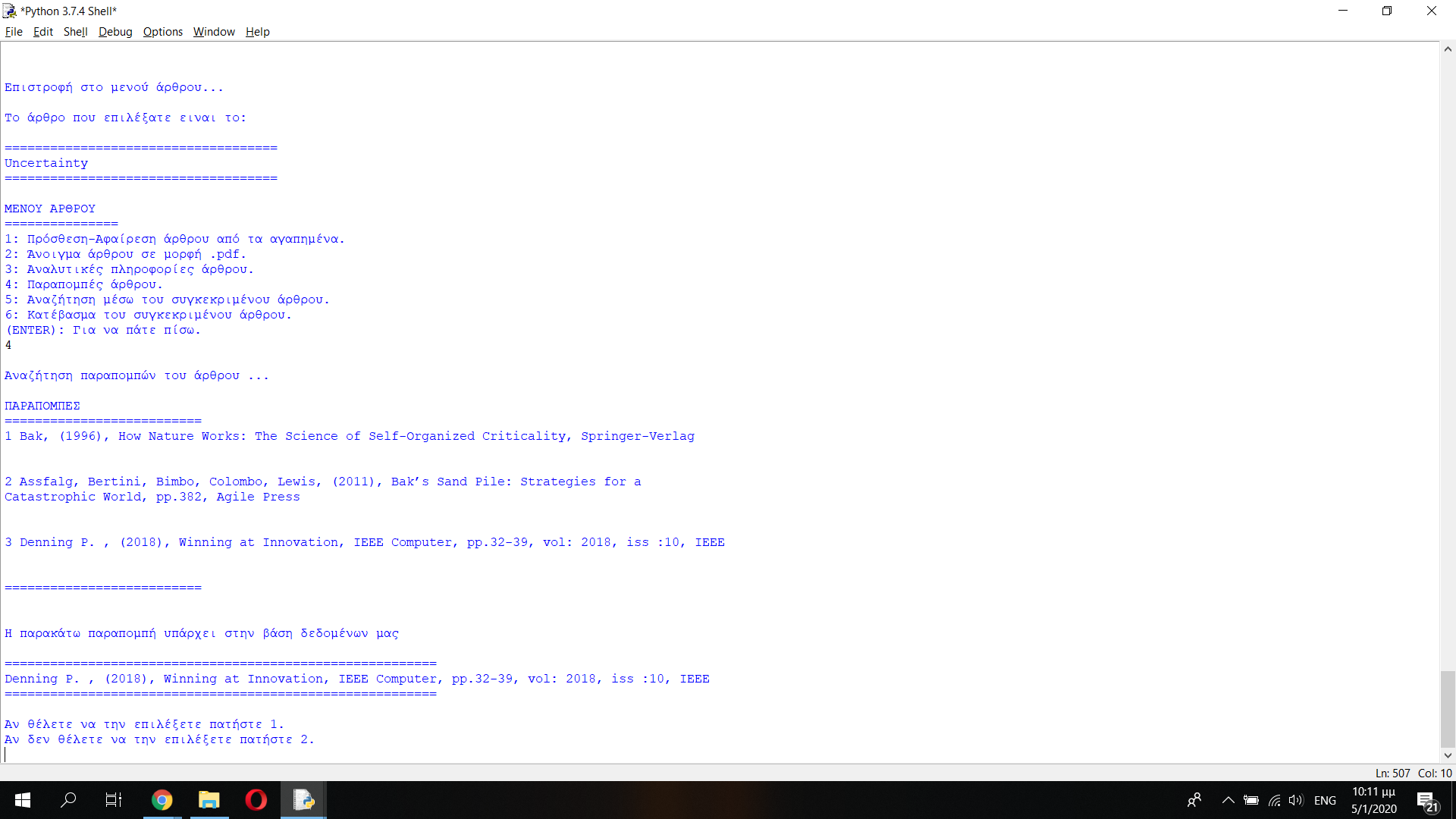


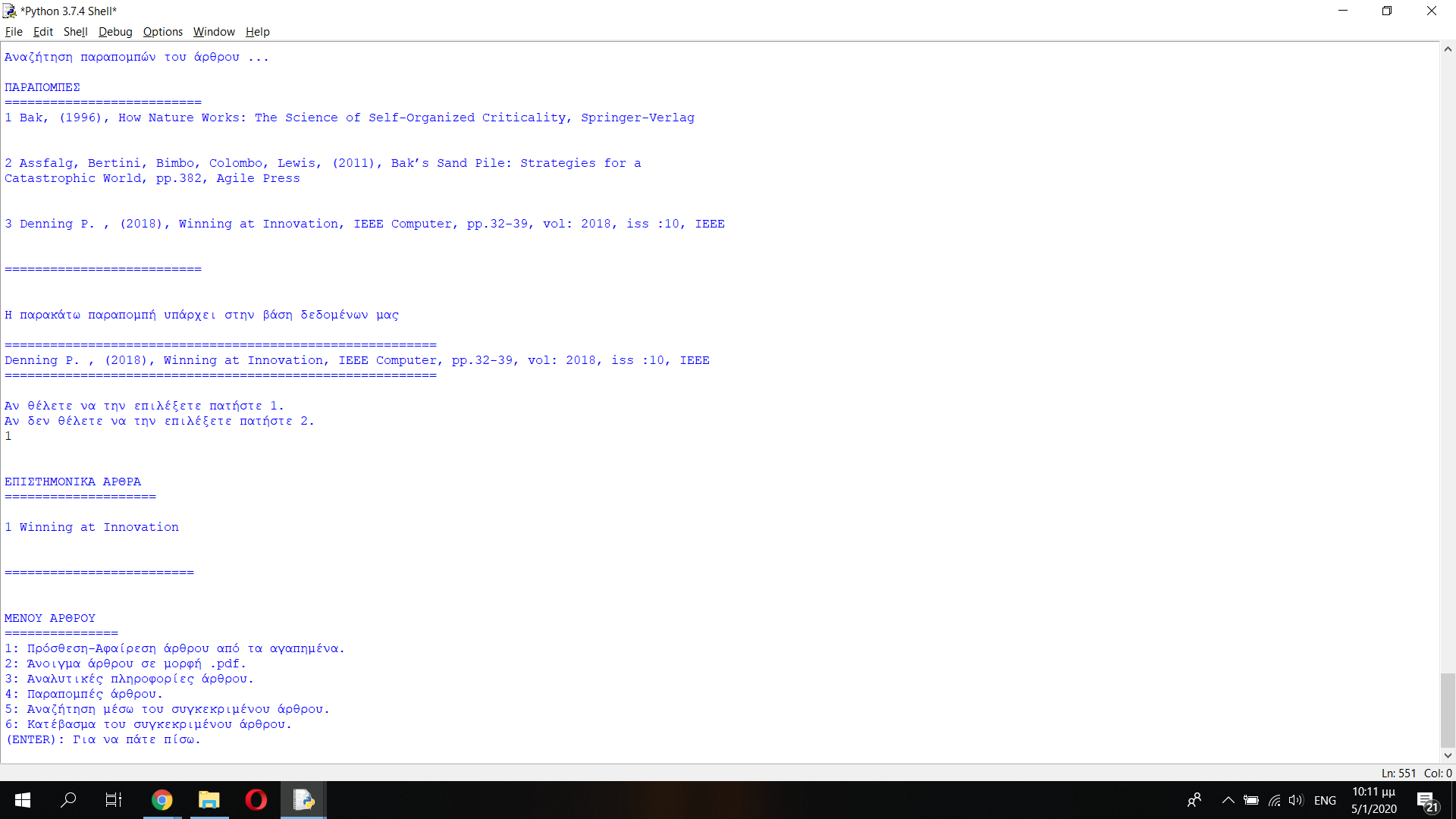
Πατώντας το 3 θα βγουν οι παραπομπές. Στην συγκεκριμένη

περίπτωση μία από τις παραπομπές υπάρχει στην βάση μας. Έτσι

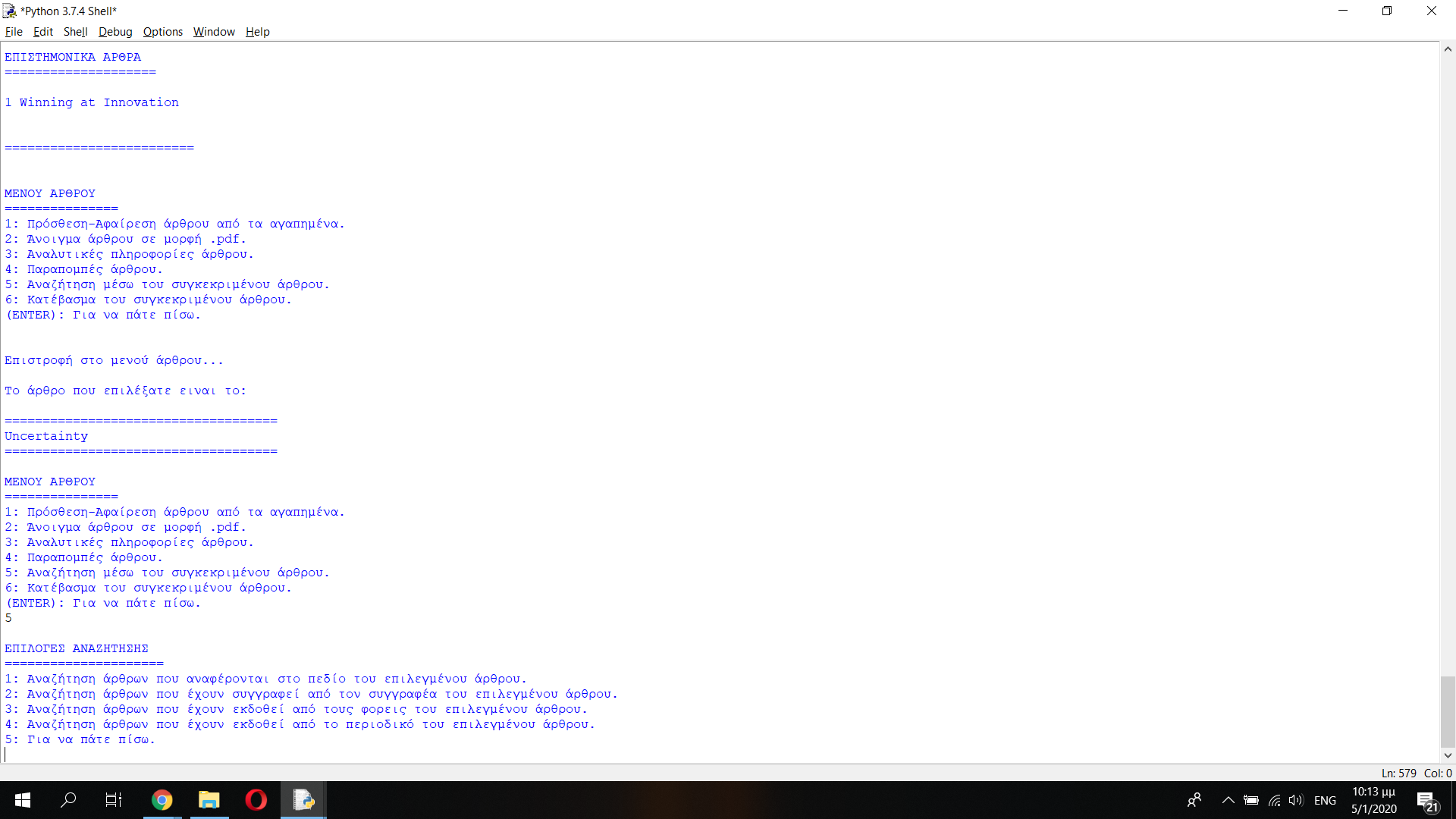
μας δίνεται η δυνατότητα να την ανοίξουμε αν θέλουμε πατώντας

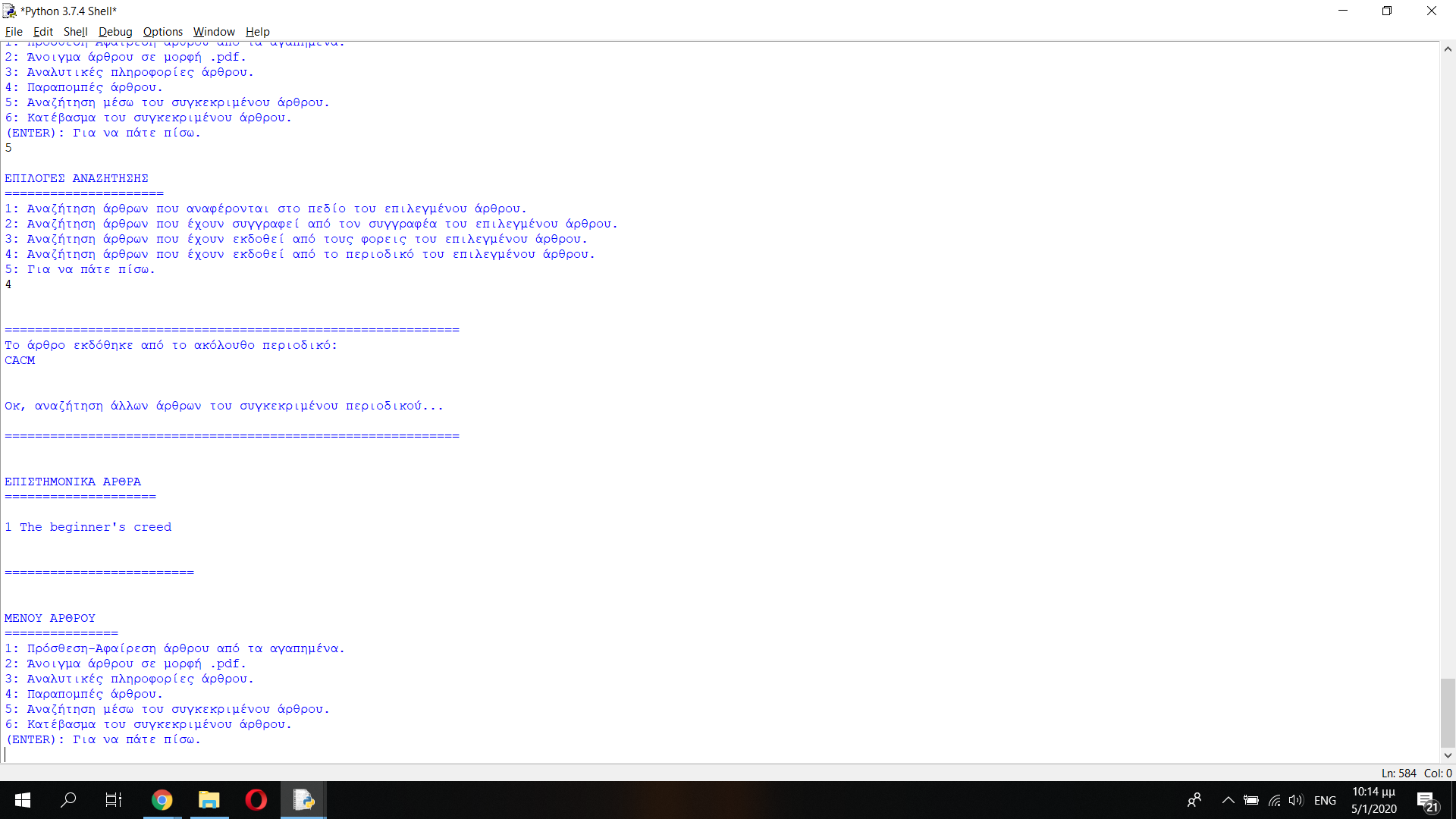
το 1.





Πατώντας το 5 μπορούμε να ξεκινήσουμε μια σύνθετη αναζήτηση ψάχνοντας άρθρα του ίδιο περιοδικού.

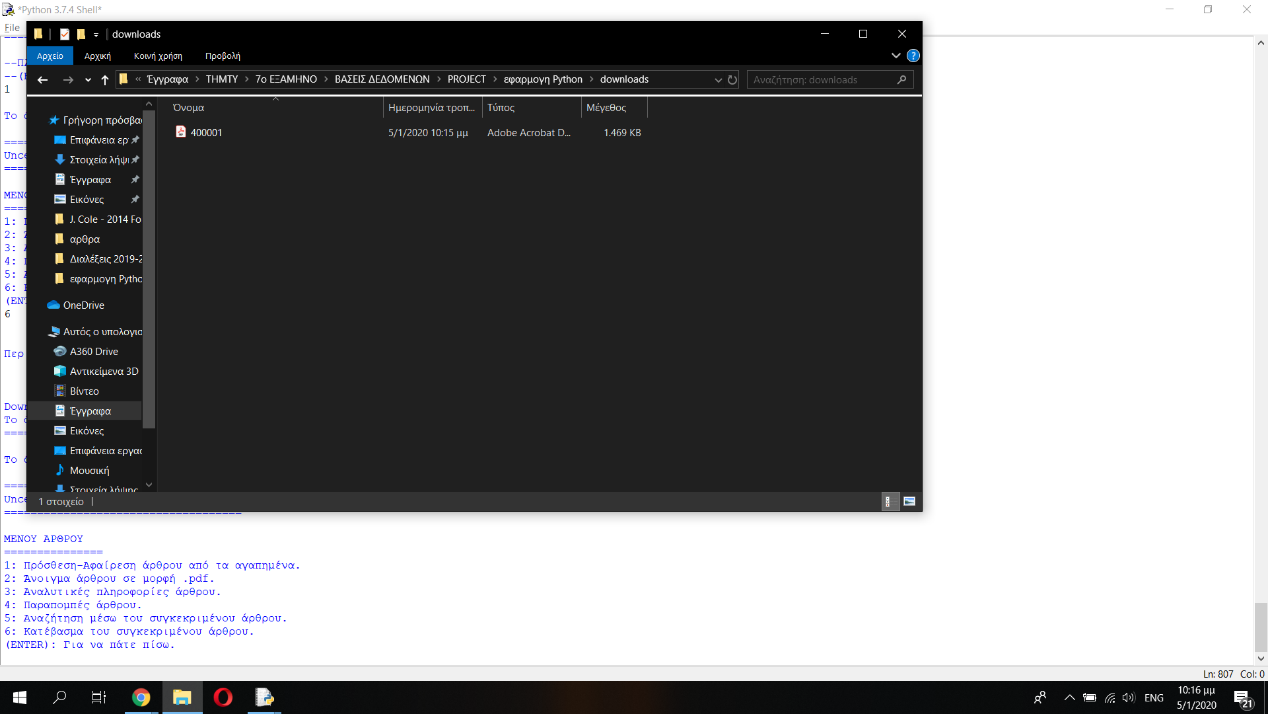




Πατώντας το 6 κατεβάζουμε το άρθρο όπως φαίνεται στις

παρακάτω εικόνες.



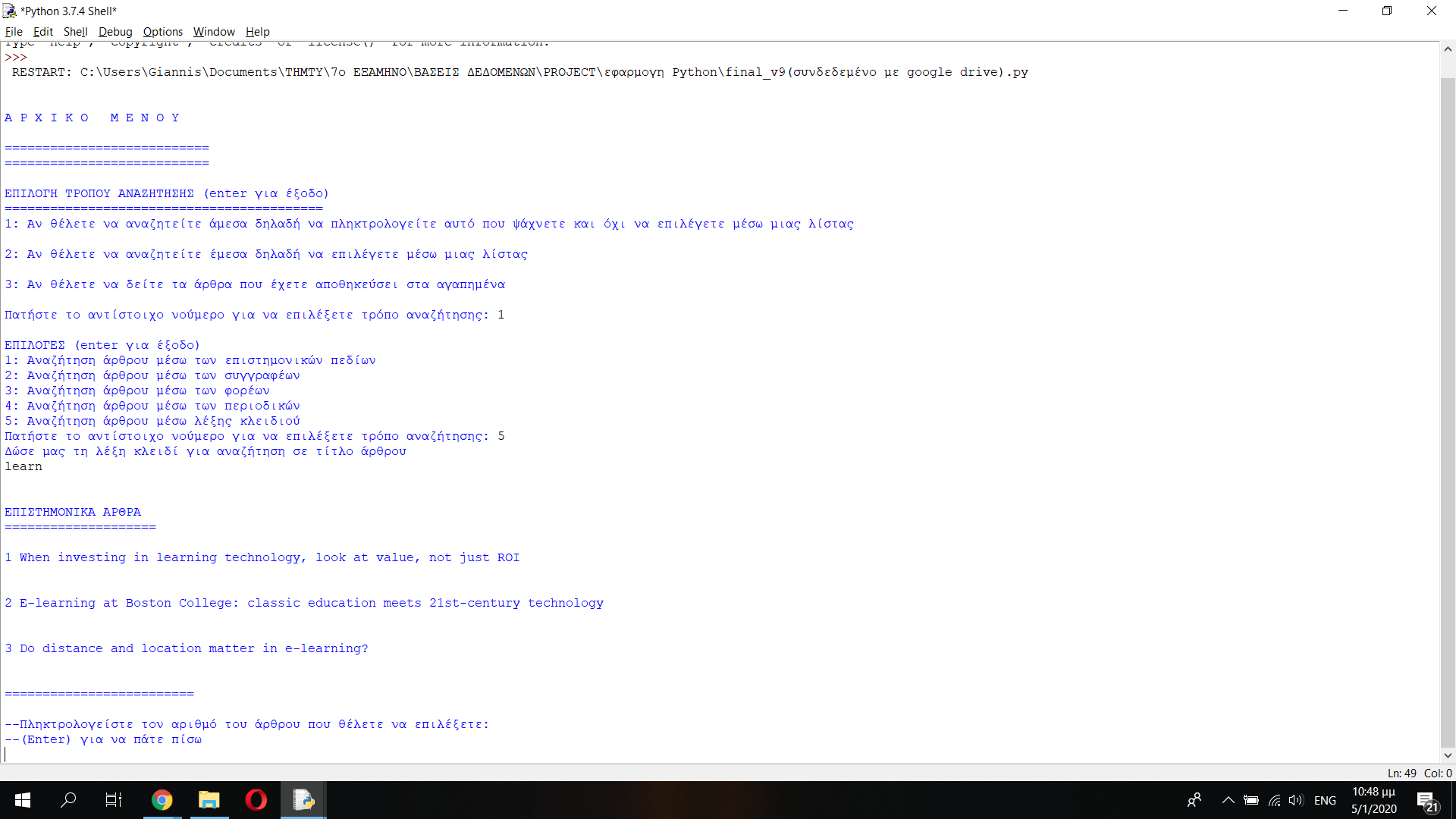


Ξεχωριστή αναφορά πρέπει να γίνει στις λέξεις κλειδιά της

άμεσης αναζήτησης όπου πληκτρολογείς μια λέξη ή ένα τμήμα

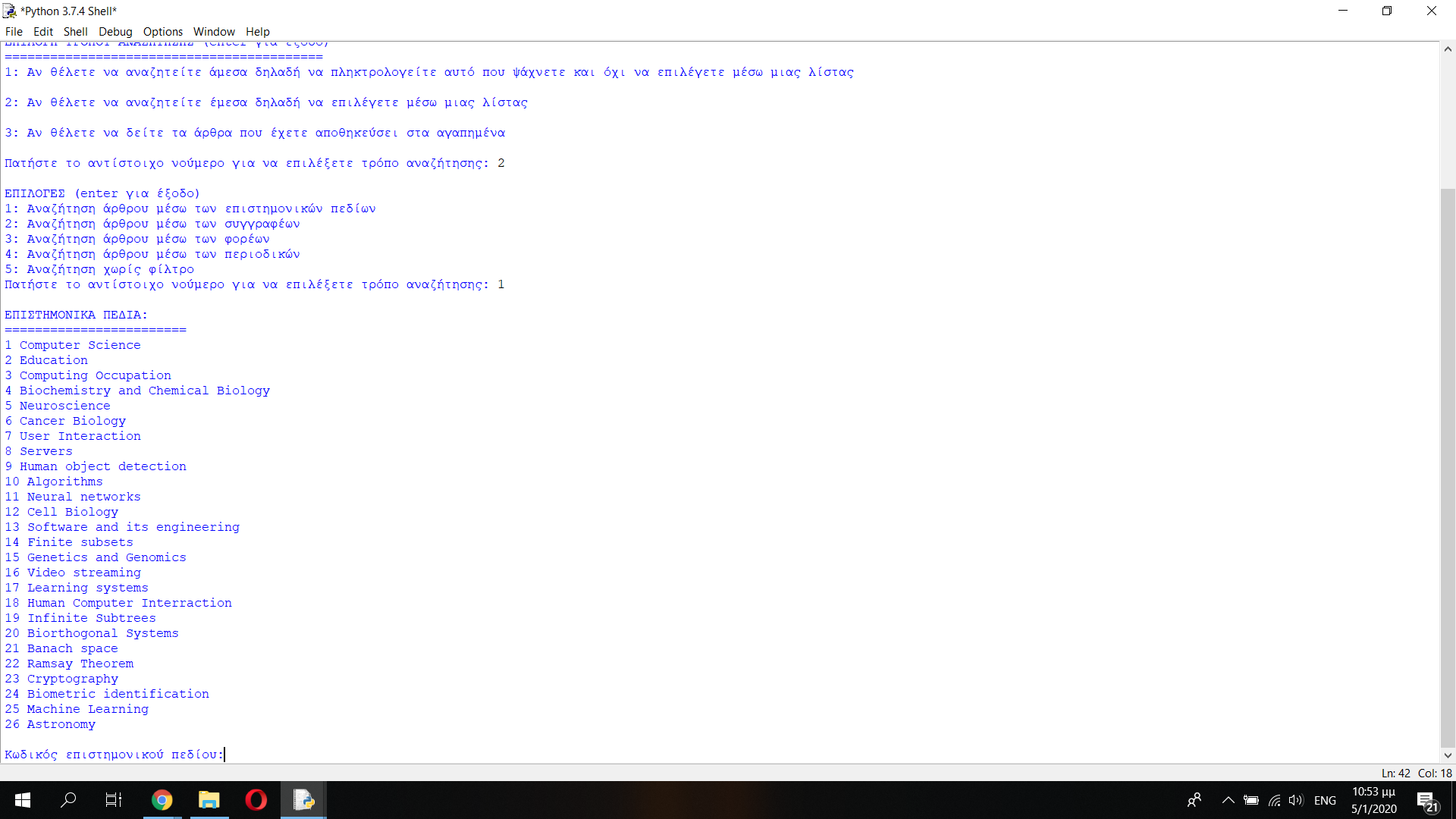
λέξης και η εφαρμογή βρίσκει άρθρα που την περιέχουν στον

τίτλο τους.



Για την έμμεση αναζήτηση ισχύουν όλα τα παραπάνω απλώς στην

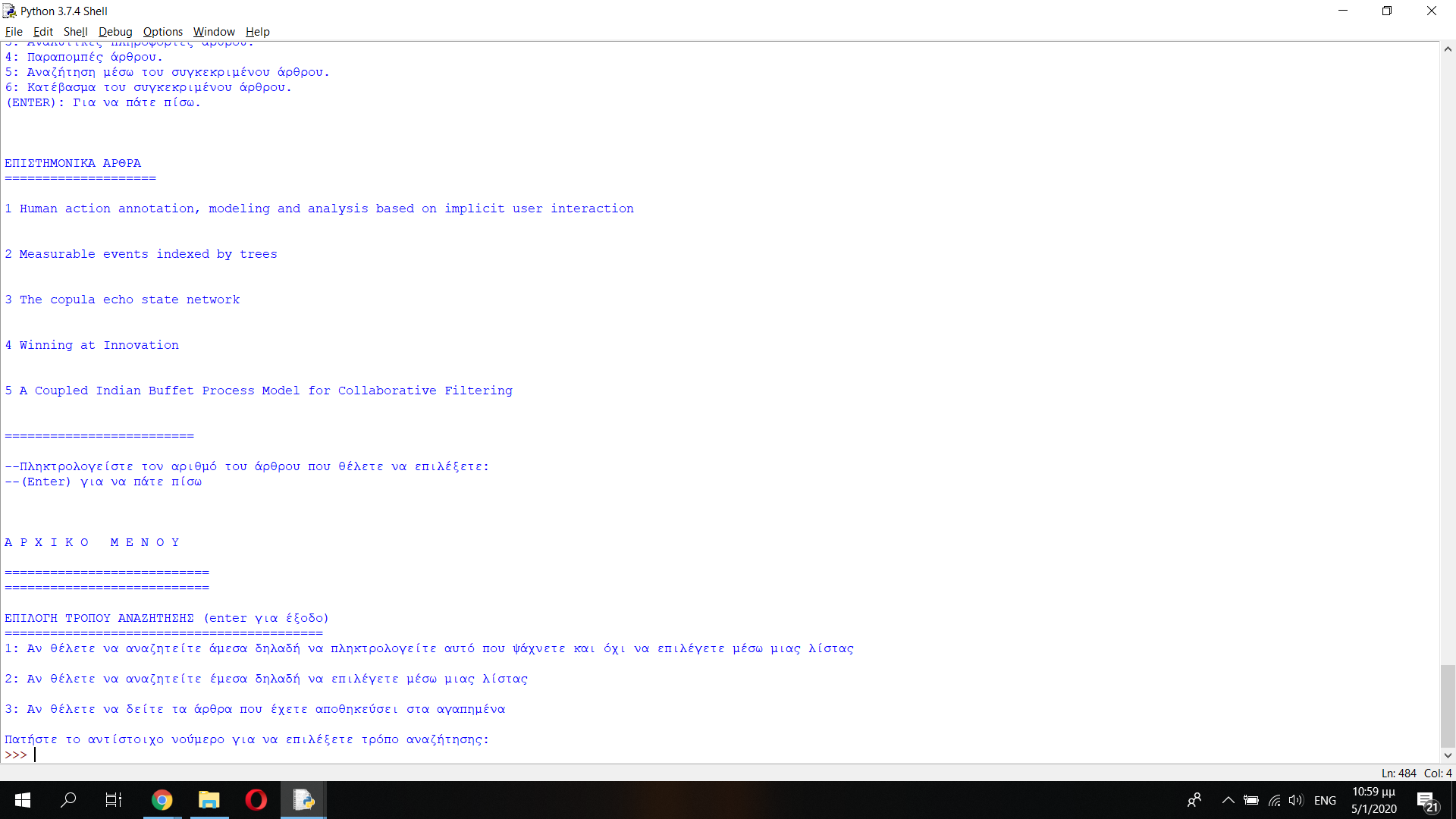
αρχική αναζήτηση παρουσιάζεται το παρακάτω.





Πατώντας το Enter στο αρχικό μενού η εφαρμογή κλείνει, ενώ

πατώντας Enter οπουδήποτε αλλού πάει πίσω.



ΔΗΜΙΟΥΡΓΙΑ ΓΡΑΦΙΚΟΥ ΠΕΡΙΒΑΛΛΟΝΤΟΣ ΜΕΣΩ tkinter:

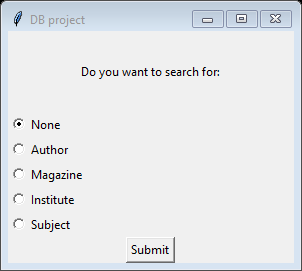
Αρχικά δημιουργήσαμε ένα παράθυρο εισαγωγής στο

οποίο ο χρήστης επιλέγει το φίλτρο αναζήτησης του

και πατάει το κουμπί submit έτσι ώστε να συνεχίσει

στο πρόγραμμα.(Εάν επιλεγεί η ετικέτα None τότε το

πρόγραμμα τερματίζεται!)

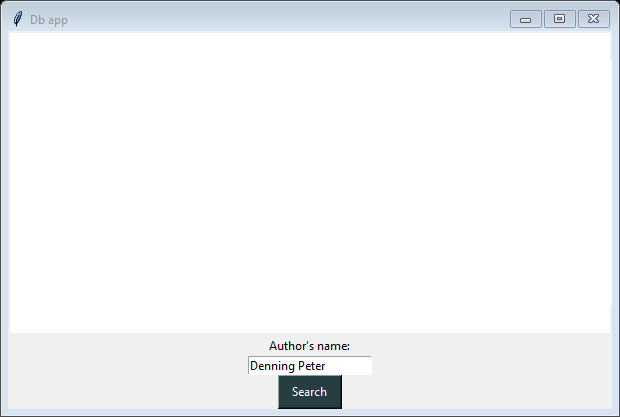


Στη συνέχεια o χρήστης (έστω ότι έχει επιλεγεί το φίλτρο author)

κάνει αναζήτηση ενός αρθρογράφου πληκτρολογώντας το

επώνυμο και το όνομα του στην είσοδο όπως φαίνεται

παρακάτω.



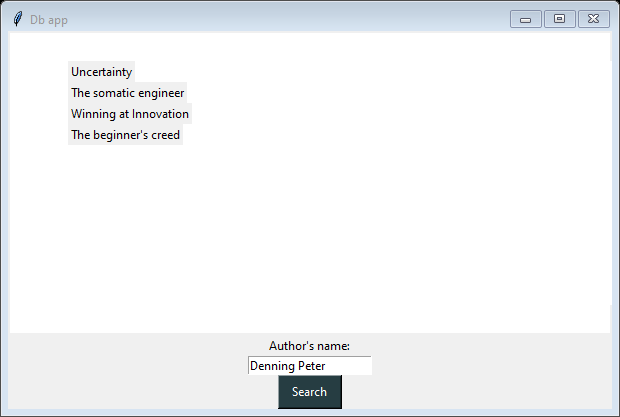
Έπειτα πατώντας το κουμπί search γίνεται αναζήτηση στη βάση

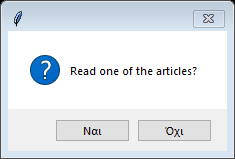
των άρθρων που έχουν σχέση με τον συγκεκριμένο αρθρογράφο.

Στη συνέχεια γίνεται εμφάνιση στην οθόνη ενός μηνύματος που

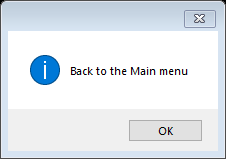
δίνει τη δυνατότητα στο χρήστη να διαλέξει για ανάγνωση κάποιο

άρθρο εφόσον επιλέξει το κουμπί ναι .





Εάν ο χρήστης επιλέξει το κουμπί όχι τότε εμφανίζεται μήνυμα στην οθόνη μήνυμα που δείχνει ότι γίνεται επιστροφή στο αρχικό μενού αναζήτησης.



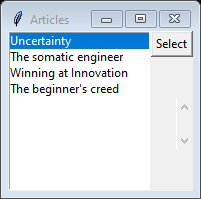
Εάν ο χρήστης επιλέξει το κουμπί ναι για ανάγνωση των άρθρων

τότε εμφανίζεται ένα listbox στο οποίο μπορεί να επιλέξει το

άρθρο που θέλει. Το άρθρο που επιλέγει ανοίγεται στο

googledrive και δίνεται και η πιθανότητα αποθήκευσης του

άρθρου σε φάκελο μέσω πάλι ενός μηνύματος στην οθόνη.



Εικόνα που περιέχει στιγμιότυπο οθόνης

Περιγραφή που δημιουργήθηκε αυτόματα

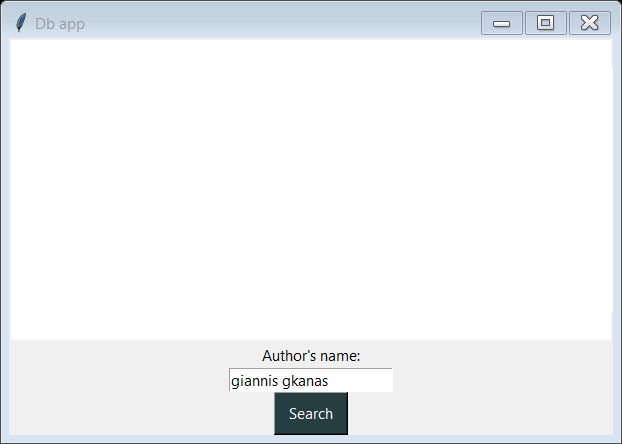
Έπειτα το πρόγραμμα τερματίζει.

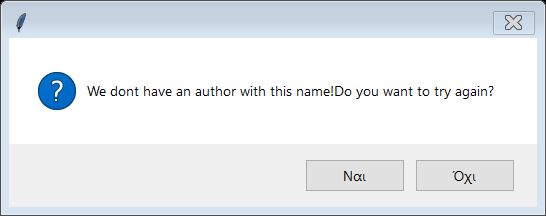
ΠΑΡΑΤΗΡΗΣΗ:

Εφόσον γίνει αναζήτηση ενός ονόματος που δεν υπάρχει στη

βάση εμφανίζεται μήνυμα και ο χρήστης μπορεί να κάνει

αναζήτηση εκ΄ νέου .





**REFERENCES**

[1] <https://pymysql.readthedocs.io/en/latest/user/examples.html>

[2] <https://www.w3schools.com/python/python_mysql_getstarted.asp>

[3] https://www.tutorialspoint.com/mysql/index.htm

[4] https://www.youtube.com/

[5] <https://stackoverflow.com/questions/56038193/will-connection-cursor-as-cursor-create-new-connection>