



Δημοκρίτειο Πανεπιστήμιο Θράκης

Τμήμα Ελληνικής Φιλολογίας

Εργαστήριο Παπυρολογίας &
Παλαιογραφίας

Ε.Τ.Ε.Π. Άγγελος Σισμάνης

asismani@helit.duth.gr



ΕΡΓΑΣΤΗΡΙΑΚΗ ΑΣΚΗΣΗ ΔΗ-13

Κωδικοποίηση κειμένου με βάση το πρότυπο TEI



Στοιχεία Φοιτητή

Ονοματεπώνυμο:

A.M.

Email:

Σημειώσεις

1. Δημητρούλια, Ξ., Τικτοπούλου, Α. (2015). Κωδικοποιήσεις κειμένου [Κεφάλαιο 6].

Αθήνα: Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις.

Διαθέσιμο στο: <https://repository.kallipos.gr/handle/11419/5833?&locale=en>

2. Περιγραφή των TEI στοιχείων είναι διαθέσιμη στην παρακάτω σελίδα

<https://tei-c.org/release/doc/tei-p5-doc/en/html/REF-ELEMENTS.html>

3. TEI P5 Customization and Encoding Guidelines for Manuscript Description διαθέσιμες στη σελίδα παρακάτω:

<https://msdesc.github.io/consolidated-tei-schema/msdesc.html>

ΑΣΚΗΣΗ DH-13.1

Αυτόματη κωδικοποίηση κειμένου με βάση το πρότυπο TEI

1. Αποθηκεύστε με τίτλο TEIXH σε φάκελο της επιφάνειας εργασίας του υπολογιστή σας το παρακάτω ποίημα του Κ. Π. Καβάφη ως έγγραφο απλού κειμένου χρησιμοποιώντας το Σημειωματάριο της Microsoft (μορφή .txt).

ΤΕΙΧΗ

Χωρίς περίσκεψιν, χωρίς λύπην, χωρίς αιδῶ
μεγάλα κ' υψηλά τριγύρω μου έκτισαν τείχη.

Και κάθομαι και απελπίζομαι τώρα εδώ.
Άλλο δεν σκέπτομαι: τον νουν μου τρώγει αυτή η τύχη·

διότι πράγματα πολλά έξω να κάμω είχαν.
Α όταν έκτιζαν τα τείχη πώς να μην προσέξω.

Αλλά δεν άκουσα ποτέ κρότον κτιστών ή ήχον.
Ανεπαισθήτως μ' έκλεισαν από τον κόσμον έξω.

2. Χρησιμοποιήστε το online εργαλείο **OxGarage** για να μετατρέψετε το παραπάνω txt αρχείο σε xml έγγραφο. Αποθηκεύστε τον ίδιο τίτλο στον ίδιο φάκελο στην επιφάνεια εργασίας.

<https://oxgarage.tei-c.org/>

TEI < Text Encoding Initiative >

OxGarage Conversion

File submitted for conversion. Please wait until the download dialog appears...

Select file to convert: ?

Browse... TEIXH.txt

+ Show advanced options ?

Upload images: ?

You can upload image files and .zip files containing images

You do not need to upload any images, as images are included in the document file.

Convert

Reset

Convert from: ?

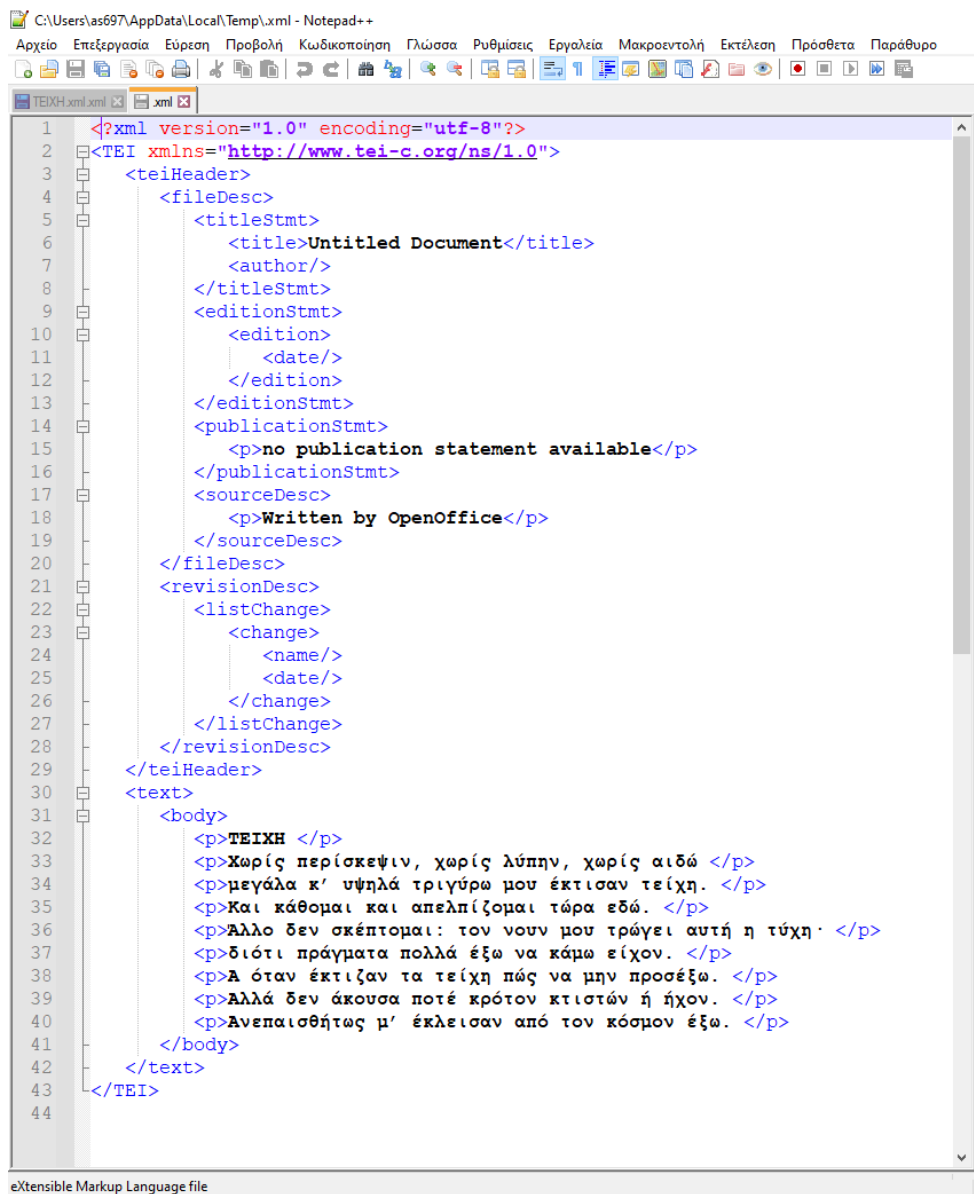
Documents

Plain Text (.txt)

Convert to: ?

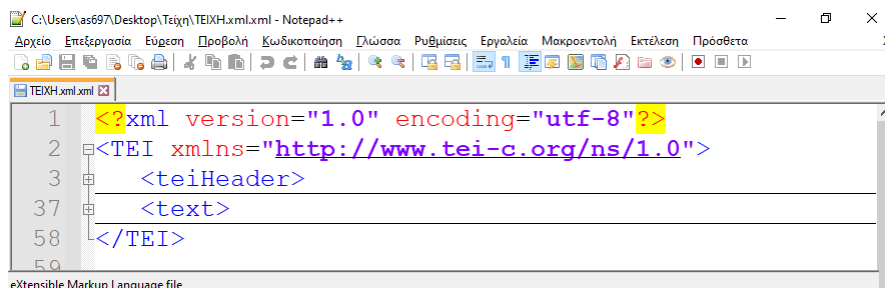
XML Document

3. Ανοίξτε το έγγραφο TEIΧΗ.xml με την εφαρμογή **Notepad++** και αποθηκεύστε (Ctrl+S) στον ίδιο φάκελο στην επιφάνεια εργασίας.



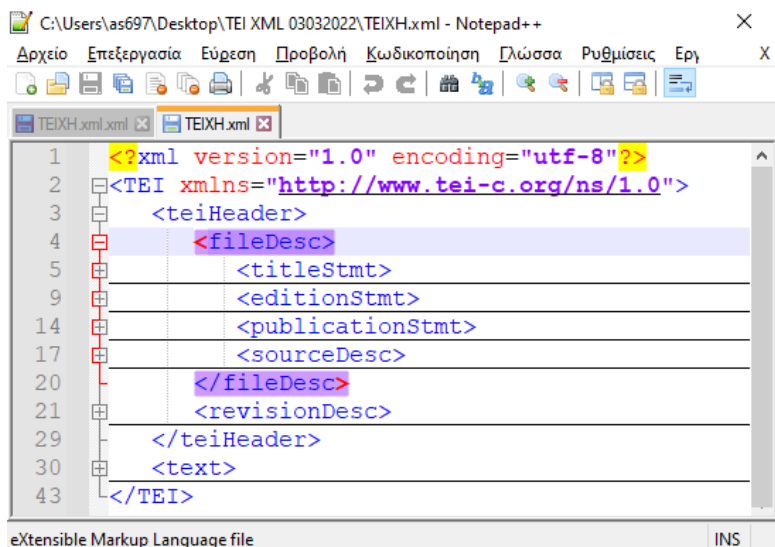
```
1 <?xml version="1.0" encoding="utf-8"?>
2 <TEI xmlns="http://www.tei-c.org/ns/1.0">
3   <teiHeader>
4     <fileDesc>
5       <titleStmt>
6         <title>Untitled Document</title>
7         <author/>
8       </titleStmt>
9       <editionStmt>
10        <edition>
11          <date/>
12        </edition>
13      </editionStmt>
14      <publicationStmt>
15        <p>no publication statement available</p>
16      </publicationStmt>
17      <sourceDesc>
18        <p>Written by OpenOffice</p>
19      </sourceDesc>
20    </fileDesc>
21    <revisionDesc>
22      <listChange>
23        <change>
24          <name/>
25          <date/>
26        </change>
27      </listChange>
28    </revisionDesc>
29  </teiHeader>
30  <text>
31    <body>
32      <p>ΤΕΙΧΗ </p>
33      <p>Χωρίς περίσκεψιν, χωρίς λύπην, χωρίς αιδώ </p>
34      <p>μεγάλα κ' υψηλά τριγύρω μου έκτισαν τείχη. </p>
35      <p>Και κάθομαι και απελπίζομαι τώρα εδώ. </p>
36      <p>Άλλο δεν σκέπτομαι: τον νουν μου τρώγει αυτή η τύχη. </p>
37      <p>Διότι πράγματα πολλά έξω να κάμω είχον. </p>
38      <p>Α όταν έκτιζαν τα τείχη πώς να μην προσέξω. </p>
39      <p>Αλλά δεν άκουσα ποτέ κρότον κτιστών ή ήχον. </p>
40      <p>Ανεπαισθήτως μ' έκλεισαν από τον κόσμο έξω. </p>
41    </body>
42  </text>
43 </TEI>
44
```

4. Εντοπίστε στο xml αρχείο το στοιχείο **<TEI>** και επιβεβαιώστε ότι περιλαμβάνει τα στοιχεία **<teiHeader>** και **<text>** για την αναπαράσταση των μεταδεδομένων και του κειμένου του ποιήματος αντίστοιχα.



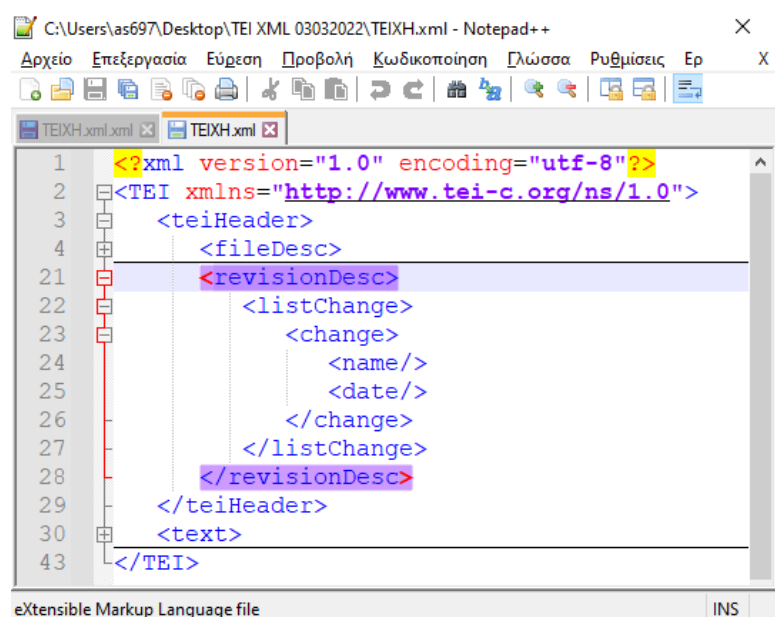
```
1 <?xml version="1.0" encoding="utf-8"?>
2 <TEI xmlns="http://www.tei-c.org/ns/1.0">
3   <teiHeader>
4     <fileDesc>
5       <title>Untitled Document</title>
6       <author/>
7     </title>
8     <edition>
9       <date/>
10    </edition>
11    <publication>
12      <p>no publication statement available</p>
13    </publication>
14    <sourceDesc>
15      <p>Written by OpenOffice</p>
16    </sourceDesc>
17  </fileDesc>
18  <revisionDesc>
19    <listChange>
20      <change>
21        <name/>
22        <date/>
23      </change>
24    </listChange>
25  </revisionDesc>
26 </teiHeader>
27 <text>
28   <body>
29     <p>ΤΕΙΧΗ </p>
30     <p>Χωρίς περίσκεψιν, χωρίς λύπην, χωρίς αιδώ </p>
31     <p>μεγάλα κ' υψηλά τριγύρω μου έκτισαν τείχη. </p>
32     <p>Και κάθομαι και απελπίζομαι τώρα εδώ. </p>
33     <p>Άλλο δεν σκέπτομαι: τον νουν μου τρώγει αυτή η τύχη. </p>
34     <p>Διότι πράγματα πολλά έξω να κάμω είχον. </p>
35     <p>Α όταν έκτιζαν τα τείχη πώς να μην προσέξω. </p>
36     <p>Αλλά δεν άκουσα ποτέ κρότον κτιστών ή ήχον. </p>
37     <p>Ανεπαισθήτως μ' έκλεισαν από τον κόσμο έξω. </p>
38   </body>
39 </text>
40 </TEI>
41
```

5. Εντοπίστε το στοιχείο [<teiHeader>](#) και επιβεβαιώστε ότι είναι η ρίζα του στοιχείου [<fileDesc>](#), το οποίο περιλαμβάνει τα παρακάτω στοιχεία [<titleStmt>](#), [<publicationStmt>](#), [<sourceDesc>](#).



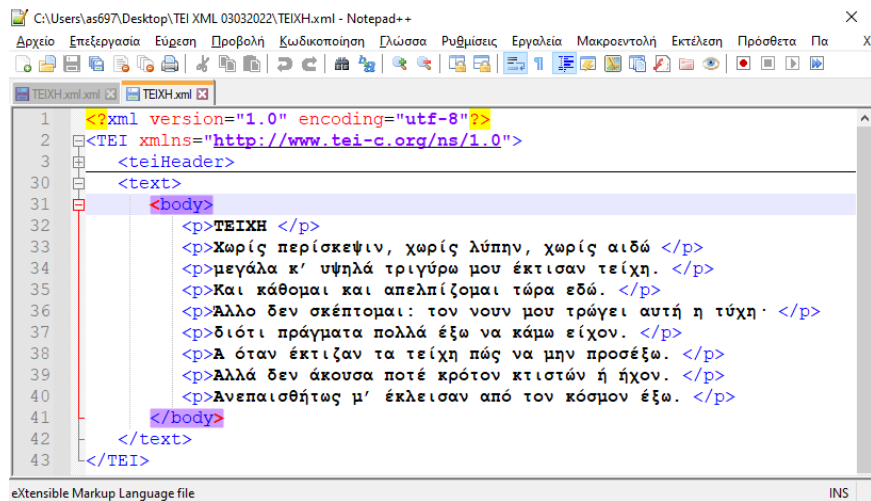
```
1  <?xml version="1.0" encoding="utf-8"?>
2  <TEI xmlns="http://www.tei-c.org/ns/1.0">
3    <teiHeader>
4      <fileDesc>
5        <titleStmt>
9        <editionStmt>
14       <publicationStmt>
17       <sourceDesc>
20     </fileDesc>
21     <revisionDesc>
29   </teiHeader>
30   <text>
43 </TEI>
```

6. Εντοπίστε το στοιχείο [<teiHeader>](#) και επιβεβαιώστε ότι είναι η ρίζα του στοιχείου [<revisionDesc>](#), το οποίο περιλαμβάνει το στοιχείο [<listChange>](#) το οποίο εσωκλείει στοιχείο [<change>](#) που περιλαμβάνει τα [<name>](#) και [<date>](#).



```
1  <?xml version="1.0" encoding="utf-8"?>
2  <TEI xmlns="http://www.tei-c.org/ns/1.0">
3    <teiHeader>
4      <fileDesc>
21     <revisionDesc>
22       <listChange>
23         <change>
24           <name/>
25           <date/>
26         </change>
27       </listChange>
28     </revisionDesc>
29   </teiHeader>
30   <text>
43 </TEI>
```

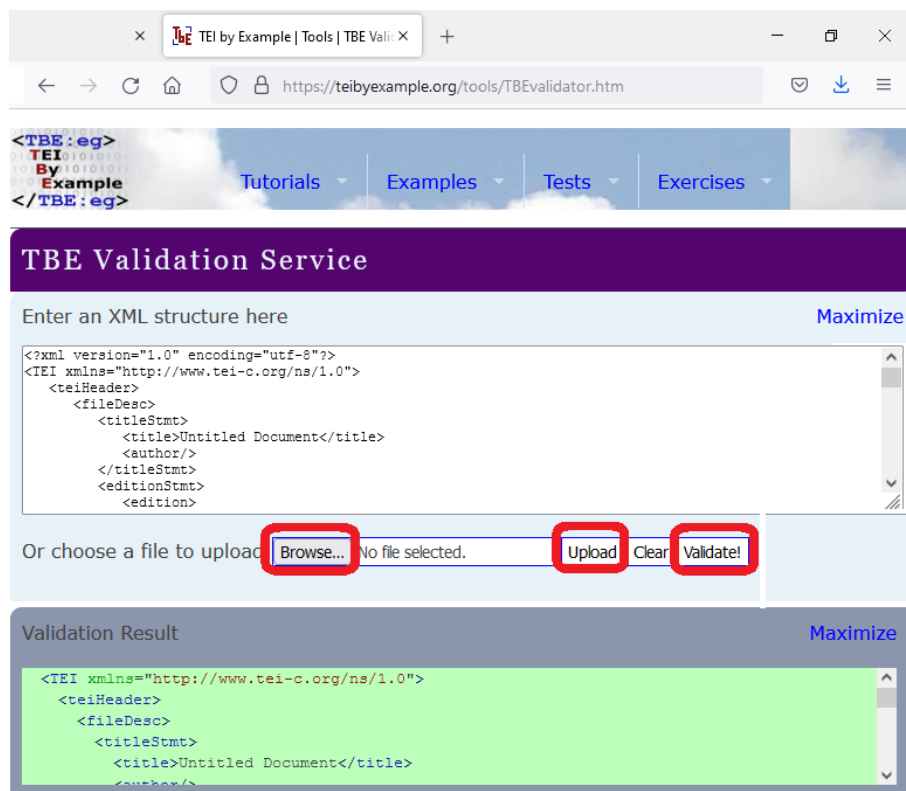
7. Εντοπίστε το στοιχείο [<text>](#) και επιβεβαιώστε ότι περιλαμβάνει το στοιχείο [<body>](#).



```
1 <?xml version="1.0" encoding="utf-8"?>
2 <TEI xmlns="http://www.tei-c.org/ns/1.0">
3   <teiHeader>
30  <text>
31    <body>
32      <p>ΤΕΙΧΗ </p>
33      <p>Χωρίς περίσκεψιν, χωρίς λύπην, χωρίς αιδώ </p>
34      <p>μεγάλα κ' υψηλά τριγύρω μου έκτισαν τείχη. </p>
35      <p>Και κάθομαι και απελπίζομαι τώρα εδώ. </p>
36      <p>Άλλο δεν σκέπτομαι: τον νουν μου τρώγει αυτή η τύχη. </p>
37      <p>διότι πράγματα πολλά έξω να κάμω είχον. </p>
38      <p>Α όταν έκτιζαν τα τείχη πώς να μην προσέξω. </p>
39      <p>Αλλά δεν άκουσα ποτέ κρότον κτιστών ή ήχον. </p>
40      <p>Ανεπαισθήτως μ' έκλεισαν από τον κόσμον έξω. </p>
41    </body>
42  </text>
43 </TEI>
```

8. Χρησιμοποιήστε το online εργαλείο **TBE Validator** για να ελέγξετε την εγκυρότητα του TEIXH.xml αρχείου σύμφωνα με το πρότυπο TEI.

<https://teibyexample.org/tools/TBEvalidator.htm>



TEI by Example | Tools | TBE Validator

https://teibyexample.org/tools/TBEvalidator.htm

Tutorials Examples Tests Exercises

TBE Validation Service

Enter an XML structure here [Maximize](#)

```
<?xml version="1.0" encoding="utf-8"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>Untitled Document</title>
        <author/>
      </titleStmt>
      <editionStmt>
        <edition>
```

Or choose a file to upload [Browse...](#) No file selected. [Upload](#) [Clear](#) [Validate!](#)

Validation Result [Maximize](#)

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>Untitled Document</title>
        <author/>
```

9. Τροποποιήστε το όνομα μιας ετικέτας, π.χ. αλλαγή της **<fileDesc>** σε **<fileDescription>**. Παρατηρήστε ότι στο πλαίσιο **"Validation result"** αυτόματα εμφανίζεται το παρακάτω μήνυμα:

The element type "fileDescription" must be terminated by the matching end-tag "</fileDescription>".

10. Διορθώστε (επαναφέρετε) το όνομα της ετικέτας, ώστε να μην εμφανίζεται το παραπάνω μήνυμα.

ΑΣΚΗΣΗ DH-13.2

Καταχώρηση μεταδεδομένων στο <teiHeader>

1. Εντοπίστε τη δήλωση τίτλου (στοιχείο [<titleStmt>](#)) και προχωρήστε στις απαραίτητες διορθώσεις / προσθήκες, ώστε να περιλαμβάνει τις παρακάτω πληροφορίες σχετικά με τον τίτλο και τον συγγραφέα και αποθηκεύστε.

```
<fileDesc>
  <titleStmt>
    <title>Τείχη</title>
    <author>Κ. Π. Καβάφης</author>
  </titleStmt>
```

2. Εντοπίστε την εκδοτική δήλωση (στοιχείο [<editionStmt>](#)) και προχωρήστε στις απαραίτητες διορθώσεις / προσθήκες ώστε, να περιλαμβάνει τις παρακάτω πληροφορίες σχετικά με την ημερομηνία έκδοσης του κειμένου και αποθηκεύστε.

```
<fileDesc>
  <titleStmt>
  <editionStmt>
    <edition>
      <date>1896,1897</date>
    </edition>
```

3. Εντοπίστε τη δήλωση δημοσίευσης (στοιχείο [<publicationStmt>](#)) και προχωρήστε στις απαραίτητες διορθώσεις / προσθήκες ώστε, να περιλαμβάνει τις παρακάτω πληροφορίες που αφορούν στη δημοσίευση και τη διανομή του ηλεκτρονικού κειμένου και αποθηκεύστε.

```
<fileDesc>
  <titleStmt>
  <editionStmt>
  <publicationStmt>
    <p>Από τα Ποιήματα 1897-1933, Ίκαρος 1984</p>
  </publicationStmt>
```

4. Εντοπίστε το στοιχείο [<sourceDesc>](#) και προχωρήστε στις απαραίτητες διορθώσεις / προσθήκες ώστε, να περιλαμβάνει τις παρακάτω βιβλιογραφικές πληροφορίες που αφορούν στο βιβλίο στο οποίο περιλαμβάνεται το συγκεκριμένο ποίημα και αποθηκεύστε.

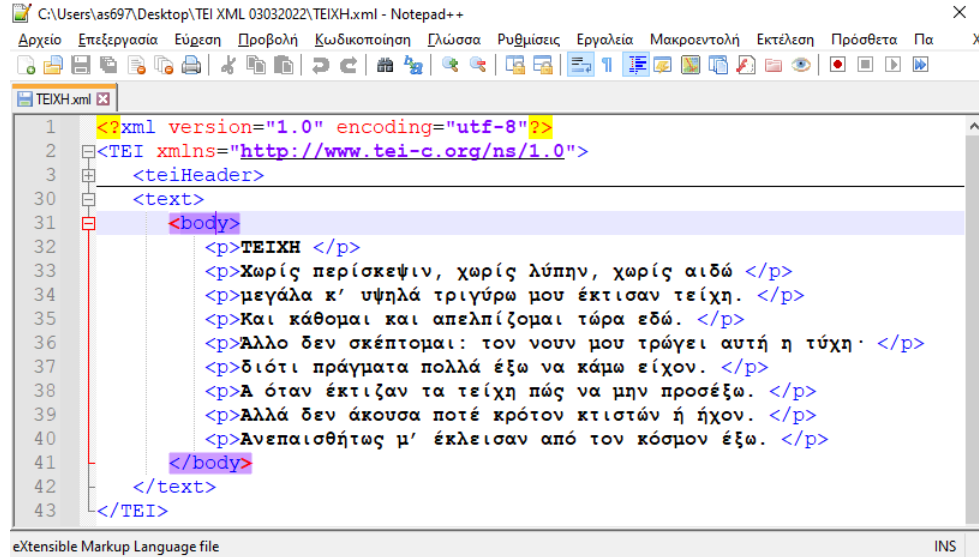
```
<fileDesc>
  <titleStmt>
  <editionStmt>
  <publicationStmt>
  <sourceDesc>
    <p>ISBN:9789607233141, Τόμος Α'</p>
  </sourceDesc>
```

5. Χρησιμοποιήστε το on line εργαλείο **TBE Validator** για να ελέγξετε την εγκυρότητα του xml αρχείου σύμφωνα με το πρότυπο TEI.

ΑΣΚΗΣΗ DH-13.3

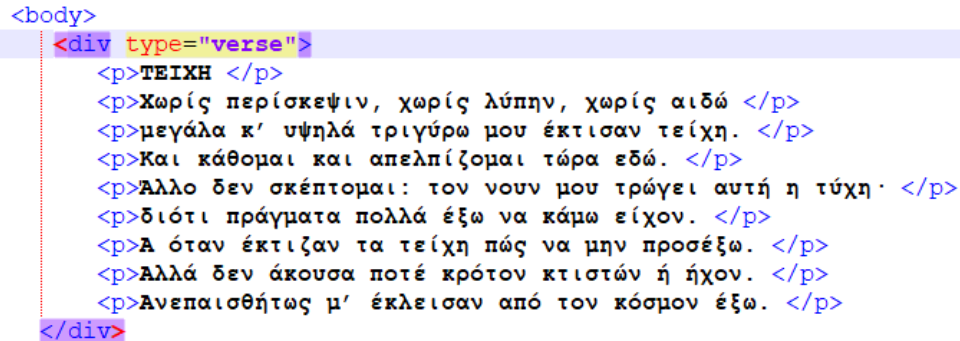
Καταχώρηση πληροφοριών στο στοιχείο <text>.

1. Εντοπίστε το στοιχείο `<body>` εντός του `<text>` και επιβεβαιώστε ότι περιλαμβάνει το κείμενο (ποίημα).



```
1 <?xml version="1.0" encoding="utf-8"?>
2 <TEI xmlns="http://www.tei-c.org/ns/1.0">
3   <teiHeader>
30  <text>
31    <body>
32      <p>ΤΕΙΧΗ </p>
33      <p>Χωρίς περίσκεψιν, χωρίς λύπην, χωρίς αιδώ </p>
34      <p>μεγάλα κ' υψηλά τριγύρω μου έκτισαν τείχη. </p>
35      <p>Και κάθομαι και απελπίζομαι τώρα εδώ. </p>
36      <p>Άλλο δεν σκέπτομαι: τον νουν μου τρώγει αυτή η τύχη. </p>
37      <p>διότι πράγματα πολλά έξω να κάμω είχον. </p>
38      <p>Α όταν έκτιζαν τα τείχη πώς να μην προσέξω. </p>
39      <p>Αλλά δεν άκουσα ποτέ κρότον κτιστών ή ήχον. </p>
40      <p>Ανεπαισθήτως μ' έκλεισαν από τον κόσμο έξω. </p>
41    </body>
42  </text>
43 </TEI>
```

2. Προχωρήστε στις απαραίτητες διορθώσεις / προσθήκες ώστε, να περιλαμβάνει / αναπαριστά πληροφορίες σχετικά με το είδος του κειμένου (χρησιμοποιήστε την ετικέτα `<div>`)



```
<body>
  <div type="verse">
    <p>ΤΕΙΧΗ </p>
    <p>Χωρίς περίσκεψιν, χωρίς λύπην, χωρίς αιδώ </p>
    <p>μεγάλα κ' υψηλά τριγύρω μου έκτισαν τείχη. </p>
    <p>Και κάθομαι και απελπίζομαι τώρα εδώ. </p>
    <p>Άλλο δεν σκέπτομαι: τον νουν μου τρώγει αυτή η τύχη. </p>
    <p>διότι πράγματα πολλά έξω να κάμω είχον. </p>
    <p>Α όταν έκτιζαν τα τείχη πώς να μην προσέξω. </p>
    <p>Αλλά δεν άκουσα ποτέ κρότον κτιστών ή ήχον. </p>
    <p>Ανεπαισθήτως μ' έκλεισαν από τον κόσμο έξω. </p>
  </div>
```

3. Προχωρήστε στις απαραίτητες διορθώσεις / προσθήκες ώστε, να περιλαμβάνει / αναπαριστά πληροφορίες σχετικά με:
 - Τον τίτλο (χρησιμοποιήστε την ετικέτα `<head>`)
 - Τους αριθμούς κάθε στροφής (χρησιμοποιήστε την ετικέτα `<lg>`)
 - Τους αριθμούς κάθε στίχου (χρησιμοποιήστε την ετικέτα `<l>`)

```

<div type="verse">
  <head>ΤΕΙΧΗ</head>
  <lg n="1">
    <l n="1">Χωρίς περίσκεψιν, χωρίς λύπην, χωρίς αιδῶ </l>
    <l n="2">μεγάλα κ' υψηλά τριγύρω μου ἐκτίσαν τείχη. </l>
  </lg>
  <lg n="2">
    <l n="3">Και κάθομαι και απελπίζομαι τώρα εδώ. </l>
    <l n="4">Ἄλλο δεν σκέπτομαι: τον νουν μου τρώγει αὐτὴ ἡ τύχη· </l>
  </lg>
  <lg n="3">
    <l n="5">διότι πράγματα πολλά ἐξω να κάμω εἶχον. </l>
    <l n="6">Ἄ όταν ἐκτίζαν τα τείχη πῶς να μην προσέξω. </l>
  </lg>
  <lg n="4">
    <l n="7">Ἀλλὰ δεν ἄκουσα ποτέ κρότον κτιστῶν ἢ ἤχον. </l>
    <l n="8">Ἀνεπαισθήτως μ' ἐκλείσαν ἀπὸ τον κόσμον ἐξω. </l>
  </lg>
</div>

```

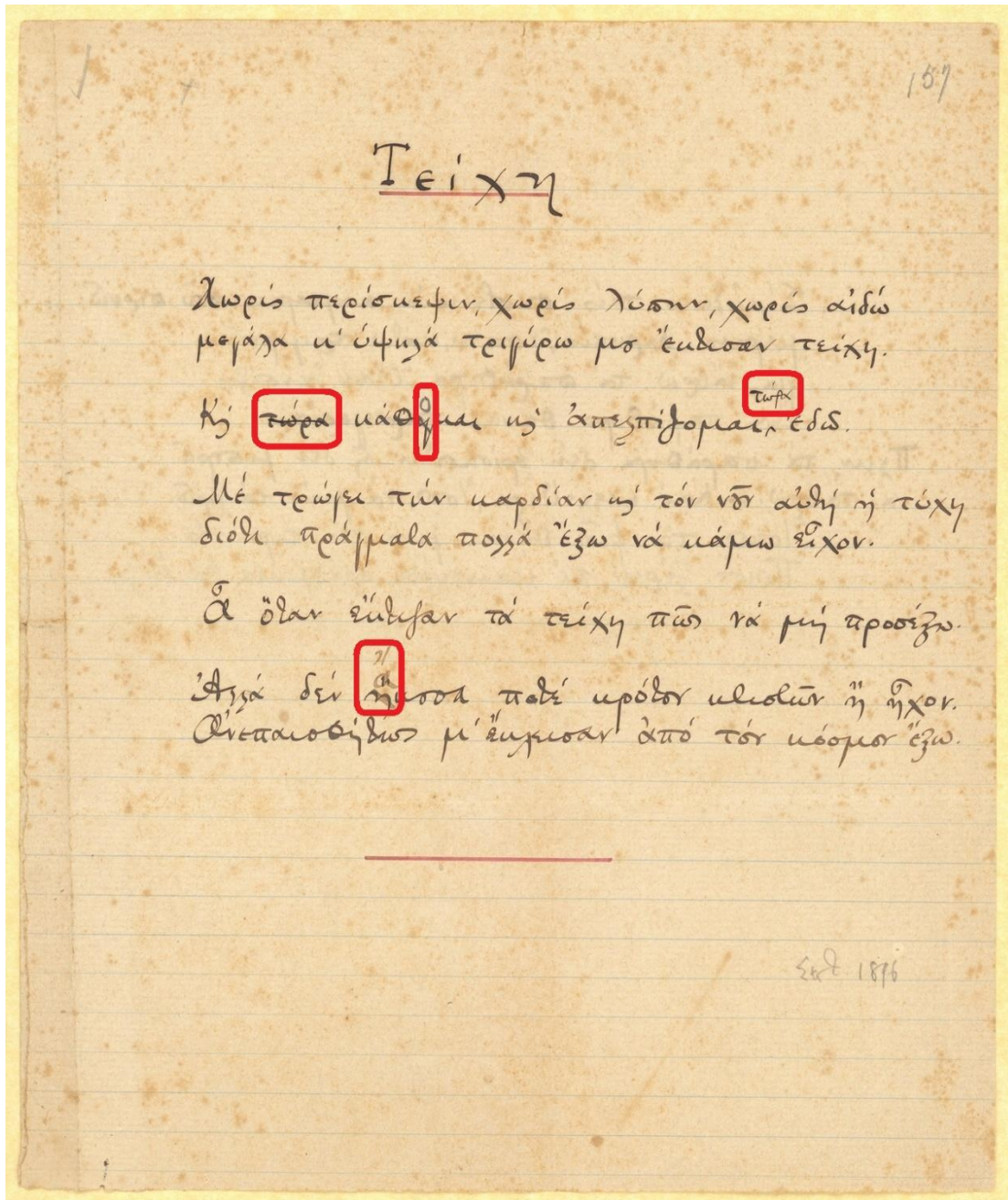
4. Αποθηκεύστε και χρησιμοποιήστε το online εργαλείο **TBE Validator** για να ελέγξετε την εγκυρότητα του xml αρχείου σύμφωνα με το πρότυπο TEI.

ΑΣΚΗΣΗ DH-13.4

Κωδικοποίηση διαγραφής, προσθήκης και διόρθωσης λαθών κατά την μεταγραφή κειμένου.

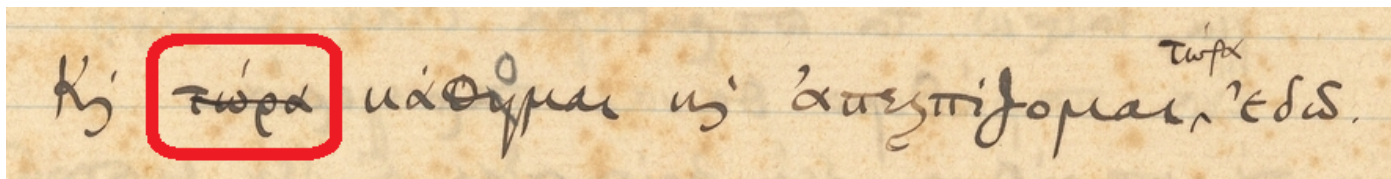
Στην παρακάτω διεύθυνση είναι διαθέσιμο ψηφιοποιημένο το χειρόγραφο ποίημα «Τείχη» του Κ. Π. Καβάφη.

<https://cavafy.onassis.org/object/cmyq-ybk8-desg/>



Παρατηρήστε τις παρακάτω διορθώσεις / επεμβάσεις του συγγραφέα, λαμβάνοντας υπόψη το τελικό (δημοσιευμένο) κείμενο του 3ου στίχου του ποιήματος «Τείχη»:

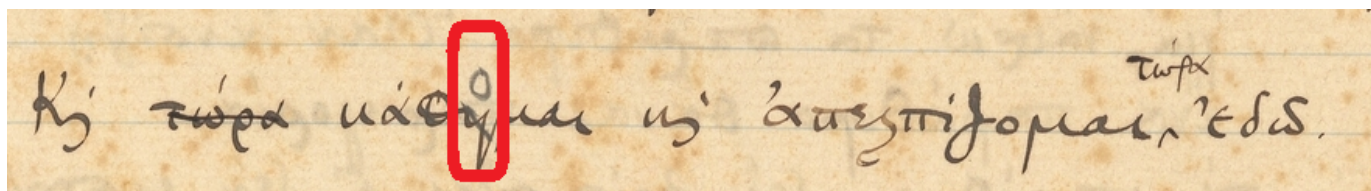
Στο χειρόγραφο κείμενο η 2^η λέξη του 3ου στίχου έχει διαγραφεί.



Σύμφωνα με το πρότυπο TEI, οι διαγραφές από τον γραφέα ή νεότερους σχολιαστές θα πρέπει να καταγράφονται χρησιμοποιώντας το στοιχείο [](#) το οποίο περιέχει ένα γράμμα, λέξη ή απόσπασμα που έχει διαγραφεί, έχει επισημανθεί ως διαγραμμένο ή με άλλο τρόπο υποδεικνύεται ως περιττό ή ψευδές στο κείμενο αντιγραφής από έναν συγγραφέα, γραφέα ή προηγούμενο σχολιαστή ή διορθωτή.

[](#)τώρα[](#)

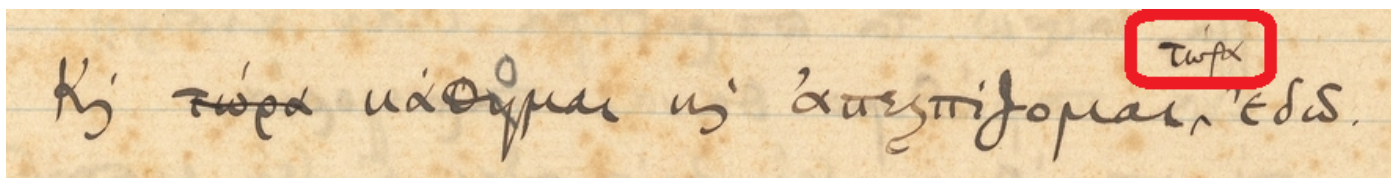
Στο χειρόγραφο κείμενο η 3^η λέξη του 3ου στίχου έχει λάθος.



Σύμφωνα με το πρότυπο TEI, οι επισημάνσεις λαθών στο χειρόγραφο κείμενο και οι υποδείξεις των σχετικών διορθώσεων θα πρέπει να καταγράφονται χρησιμοποιώντας τα στοιχεία (ετικέτες) [<sic>](#) και [<corr>](#) αντίστοιχα.

[<sic>](#)κάθημαι[</sic>](#) [<corr>](#)κάθομαι[</corr>](#)

Στο χειρόγραφο κείμενο έγινε προσθήκη πριν την τελευταία λέξη του 3ου στίχου.



Σύμφωνα με το πρότυπο TEI, το στοιχείο [<add>](#) χρησιμοποιείται για να περιέχει γράμματα, λέξεις ή φράσεις που έχουν εισαχθεί στο κείμενο πηγής από συγγραφέα, γραφέα ή προηγούμενο σχολιαστή ή διορθωτή. Σε περίπτωση που το πρωτότυπο δεν μπορεί να διαβαστεί λόγω φυσικής βλάβης ή λόγω προφανούς παράλειψης του συγγραφέα ή του γραφέα, τότε χρησιμοποιείται η ετικέτα [<supplied>](#) για το κείμενο που παρέχεται από τον μεταγραφέα ή τον επιμελητή.

<add>τώρα</add>

Προχωρήστε στην προσθήκη των παραπάνω πληροφοριών στο xml αρχείο, χρησιμοποιώντας την εφαρμογή Notepad++ και αποθηκεύστε.

```
<l n="3">Και  
    <del>τώρα</del>  
    <sic>κάθημαι</sic> <corr>κάθομαι</corr>  
    και απελπίζομαι  
    <add>τώρα</add>  
    εδώ. </l>
```

Χρησιμοποιήστε το online εργαλείο **TBE Validator** για να ελέγξετε την εγκυρότητα του TEIΧΗ.xml αρχείου σύμφωνα με το πρότυπο TEI.

<https://teibyexample.org/tools/TBEvalidator.htm>