

Συγγραφή Επιστημονικών Εργασιών

Απόστολος Μπουρνέτας
Τμήμα Μαθηματικών ΕΚΠΑ

Κύρια σημεία

- Οργάνωση Εργασίας
 - Αυτόματη αρίθμηση κεφαλαίων, υποκεφαλαίων κλπ
 - Εισαγωγή μαθηματικών εξισώσεων
 - Αρίθμηση εξισώσεων και αναφορές
- Οργάνωση Βιβλιογραφίας
 - Εργαλεία Βιβλιογραφίας
 - Mendeley, Zotero, Endnote
 - Εισαγωγή αναφορών και βιβλιογραφίας
- Επαναληψιμότητα/Τεκμηρίωση
 - Επαναληψιμότητα αποτελεσμάτων αναλύσεων
 - Τεκμηρίωση
 - R Markdown, Jupyter Notebooks, κλπ

Οργάνωση Εργασίας σε Κεφάλαια

- Όλα τα προγράμματα επεξεργασίας κειμένου έχουν επιλογές για αυτόματη εισαγωγή κεφαλαίων κλπ
- Η οργάνωση αυτή είναι πολύ χρήσιμο να γίνεται στην αρχή της συγγραφής του κειμένου
- Διευκολύνει αν αργότερα χρειάζεται ανακατάταξη/εισαγωγή νέων κεφαλαίων/υποκεφαλαίων κλπ
- Αυτόματη εισαγωγή κεφαλαίων/υποκεφαλαίων κλπ
 - Word: Heading1/Heading2,...
 - Multilevel list για αυτόματη αρίθμηση
 - Συνολική διαμόρφωση μέσω modify style
 - Αυτόματη ανανέωση σε μετακινήσεις κεφαλαίων κλπ

Μαθηματικά σύμβολα και εξισώσεις

- Οι μαθηματικές εκφράσεις (μεταβλητές, εξισώσεις κλπ) πρέπει να εισάγονται μέσω εργαλείων μαθηματικού συμβολισμού ανεξάρτητα από το μέγεθός τους.
- Ακόμα και η μεταβλητή πρέπει να εισάγεται μέσω μαθηματικής εξίσωσης και όχι ως μεταβλητή x ή x .
- Οι μαθηματικές εκφράσεις διακρίνονται σε εντός κειμένου (inline equations) και σε χωριστή γραμμή (displayed equations).
- Οι εκφράσεις εντός κειμένου ακολουθούν τη ροή και τη σύνταξη του κειμένου.
- Οι εκφράσεις σε χωριστή γραμμή αριθμούνται
 - Όλες ανεξαιρέτως, ή
 - Μόνο αυτές στις οποίες γίνεται αναφορά μέσα στο κείμενο.

Εργαλεία Εισαγωγής Μαθηματικών Εκφράσεων

- Στο Word χρησιμοποιείται ο Equation Editor
- Συνδυασμός πληκτρολόγησης και επιλογών από μενού.
- Όταν χρειάζεται αριθμηση
 - Η αριθμηση να γίνεται αυτόματα (insert caption...)
 - Η αναφορά σε μια εξίσωση μέσα στο κείμενο να γίνεται μέσω cross-reference.
 - Όταν γίνεται ανακατάταξη, εισαγωγή νέων εξισώσεων κλπ χρειάζεται ενημέρωση της αριθμησης (update fields).

Οργάνωση Βιβλιογραφίας

- Στη βιβλιογραφία στο τέλος της εργασίας περιλαμβάνονται **μόνο** εργασίες στις οποίες γίνεται αναφορά (citation) μέσα στο κείμενο.
- Τα περισσότερα περιοδικά ακολουθούν προτυποποιημένο τρόπο για εισαγωγή της βιβλιογραφίας και των citations.
- Οι περισσότεροι επεξεργαστές κειμένου έχουν τρόπους για οργανωμένη εισαγωγή της βιβλιογραφίας και των αναφορών (citations) μέσα στο κείμενο.
- Είναι χρήσιμο η βιβλιογραφία να εισάγεται με ένα συστηματικό τρόπο, για να διευκολύνεται η αυτόματη εισαγωγή των αναφορών με συνεπή τρόπο και χωρίς να γίνονται λάθη.

Προσωπικές Βιβλιογραφικές Βάσεις

- Οι γενικές βιβλιογραφικές βάσεις είναι πολύ χρήσιμες για την αναζήτηση εργασιών στη βιβλιογραφία μιας περιοχής.
- Όμως κάθε ερευνητής που ασχολείται συστηματικά σε μια περιοχή τυπικά συλλέγει ένα μεγάλο αριθμό εργασιών για μελέτη και πολλές από αυτές πρέπει να τις εισάγει ως αναφορές σε περισσότερες από μια δικές του εργασίες.
- Για τη διευκόλυνση αυτής της διαδικασίας υπάρχουν διάφορα εργαλεία που αυτοματοποιούν την οργάνωση και χρήση της προσωπικής βιβλιογραφικής βάσης ενός ερευνητή.
- Τα εργαλεία αυτά μπορεί να είναι
 - Ηλεκτρονικές πλατφόρμες
 - Αυτόνομες εφαρμογές
 - Συνδυασμός (το πιο συνηθισμένο).

Εφαρμογές βιβλιογραφικής βάσης δεδομένων

- Μια εφαρμογή προσωπικής βιβλιογραφικής βάσης δεδομένων συνήθως επιτρέπει
 - την αναζήτηση βιβλιογραφίας σε διάφορες γενικές βάσεις, ιστοσελίδες, πλατφόρμες κλπ
 - την εισαγωγή αναφορών από γενικές βάσεις
 - την οργάνωση της προσωπικής βάσης σε κατηγορίες
 - την αποθήκευση του πλήρους κειμένου της εργασίας (όταν αυτό είναι διαθέσιμο)
 - την εισαγωγή αναφορών από τη βιβλιογραφική βάση σε έναν επεξεργαστή κειμένου και την αντίστοιχη εισαγωγή στη βιβλιογραφία του κειμένου
- Για τη διασύνδεση μεταξύ εφαρμογής βιβλιογραφικής βάσης και επεξεργαστή κειμένου συνήθως χρειάζεται εγκατάσταση κατάλληλου εργαλείου στην εφαρμογή του επεξεργαστή.

Δημοφιλείς εφαρμογές βιβλιογραφίας

- Mendeley
 - Εργαλείο του εκδοτικού οίκου Elsevier
 - Συνδυασμός online πλατφόρμας και τοπικής εφαρμογής
 - Ελεύθερη εγγραφή/χρήση με περιορισμούς
 - Εγγραφή με email ΕΚΠΑ επιτρέπει περισσότερες δυνατότητες αποθήκευσης κλπ
 - <http://www.mendeley.com>
- EndNote
 - Εμπορική εφαρμογή
 - Πολύ διαδεδομένη
- Zotero
 - Ελεύθερη εφαρμογή και online πλατφόρμα
 - Γενικού σκοπού
 - <http://www.zotero.com>

Επαναληψιμότητα/Τεκμηρίωση

- Σε μια επιστημονική έρευνα είναι σημαντικό τα αποτελέσματα που δημοσιεύονται να μπορούν να ελεγχθούν/επαληθευτούν από άλλους ερευνητές.
- Όταν πρόκειται για καθαρά μαθηματική εργασία αυτό είναι εύκολο: ο αναγνώστης της εργασίας μπορεί να ακολουθήσει και να ελέγξει τα βήματα των αποδείξεων.
 - Πολλά περιοδικά ζητούν οι αποδείξεις να δίνονται είτε σε appendix στο τέλος της εργασίας είτε σε χωριστό online supplement που δεν συμπεριλαμβάνεται στο εκτυπωμένο κείμενο αλλά είναι διαθέσιμο ηλεκτρονικά.
- Όταν πρόκειται για εργασία με πειραματικά, στατιστικά ή προσομοιωμένα δεδομένα, είναι πιο περίπλοκο.

Reproducibility/Replicability

- Από τις Εθνικές Ακαδημίες Επιστημών και Μηχανικών των ΗΠΑ
<https://www.nationalacademies.org/news/2019/09/reproducibility-and-replicability-in-research>
- Reproducibility is defined as obtaining consistent results using the same data and code as the original study (synonymous with computational reproducibility).
- Replicability means obtaining consistent results across studies aimed at answering the same scientific question using new data or other new computational methods.

Reproducible Research Methodology

- Τα αρχεία και τα δεδομένα που χρησιμοποιούνται στην έρευνα/ανάλυση πρέπει να είναι οργανωμένα, με σαφείς ονομασίες και ικανή τεκμηρίωση
- Όλες οι αναλύσεις, έρευνες, επιμέρους εργασίες κλπ πρέπει να είναι πλήρως τεκμηριωμένες και ελέγξιμες, αποφεύγοντας παρεμβάσεις που δεν έχουν καταγραφεί.
- Η ροή της εργασίας πρέπει να είναι σαφώς τεκμηριωμένη και χωρισμένη σε επιμέρους υποεργασίες και βήματα που καταγράφονται.
- Διαδοχικές εκδόσεις (versions) μιας ανάλυσης ή μιας εργασίας πρέπει να διατηρούνται με σαφή τρόπο ώστε να μπορεί να ελεγχθεί η ιστορία του γενικού project.

Reproducible Research and Biostatistics

2.1 *Dimensions of reproducibility*

The Associate Editor for reproducibility (AER) will handle submissions of reproducible articles. Currently, the AER's involvement with a submission begins only when an article has been accepted for publication. The AER will consider three different criteria when evaluating the reproducibility of an article.

1. **Data:** The analytic data from which the principal results were derived are made available on the journal's Web site. The authors are responsible for ensuring that necessary permissions are obtained before the data are distributed.
2. **Code:** Any computer code, software, or other computer instructions that were used to compute published results are provided. For software that is widely available from central repositories (e.g. CRAN, Statlib), a reference to where they can be obtained will suffice.
3. **Reproducible:** An article is designated as reproducible if the AER succeeds in executing the code on the data provided and produces results matching those that the authors claim are reproducible. In reproducing these results, reasonable bounds for numerical tolerance will be considered.

Biostatistics (2009), **10**, 3, pp. 405–408

Τρόποι/Εργαλεία Τεκμηρίωσης (Documentation)

- Οι στατιστικές αναλύσεις πρέπει να είναι πλήρως τεκμηριωμένες
- Συστηματική περιγραφή των δεδομένων (ονόματα μεταβλητών, labels, ονόματα αρχείων κλπ).
- Αναλυτικά σχόλια στα προγράμματα που γράφονται για
 - διαχείριση των δεδομένων (δημιουργία νέων μεταβλητών, συμπλήρωση ελλειπουσών τιμών κλπ)
 - στατιστική ανάλυση
 - προγραμματισμό γενικότερα
- Το επίπεδο λεπτομέρειας στην τεκμηρίωση υπαγορεύεται από τον κανόνα: αν ακόμη και μετά από μεγάλο χρονικό διάστημα ζητηθεί από τον ερευνητή να εξηγήσει τη διαχείριση που έκανε στα δεδομένα, ή/και τη μεθοδολογία που ακολούθησε, θα πρέπει να μπορεί να το κάνει με βάση τα στοιχεία που έχει κρατήσει.

Εργαλεία Markdown για τεκμηρίωση

- Το Markdown είναι μια απλή γλώσσα επεξεργασίας κειμένου που δεν απαιτεί ειδικό επεξεργαστή και λειτουργεί με απλά αρχεία κειμένου.
- Αρκετές δυνατότητες μορφοποίησης του κειμένου
- Το κυριότερο πλεονέκτημα που αφορά την τεκμηρίωση είναι ότι
 - μπορεί να ενσωματωθεί ως εργαλείο σε εφαρμογές στατιστικών γλωσσών ή γενικότερων γλωσσών προγραμματισμού
 - συνδυάζει κείμενο με εντολές προγραμματισμού
 - επιτρέπει αυτόματη επανεκτέλεση των ενσωματωμένων προγραμμάτων και ενημέρωση των αποτελεσμάτων στο κείμενο.
- Για τη γλώσσα R υπάρχει το λογισμικό Rmarkdown (ενσωματωμένο στο Rstudio, αλλά μπορεί να χρησιμοποιηθεί και σε άλλα περιβάλλοντα εργασίας).