

Πολιτική απόθεσης δεδομένων στο Αποθετήριο Ερευνητικών Δεδομένων του ΣΕΑΒ HARDMIN

εκδ. 1

Hellenic Academic
Research Data Management
Initiative



Hellenic Academic Libraries Link

HEALLINK

Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών

Πίνακας περιεχομένων

Εισαγωγή	3
Αναγκαιότητα χρήσης του HARDMIN	3
Απόθεση δεδομένων	4
Γενική πολιτική	4
Ειδική πολιτική: χειρισμός ευαίσθητων δεδομένων	5
Δικαιώματα πρόσβασης και ιδιοκτησία	5
Περίοδος χρονικού αποκλεισμού	6
Αναγνώριση πόρων	6
Μόνιμα αναγνωριστικά ψηφιακών πόρων	6
Μόνιμα αναγνωριστικά φυσικών προσώπων	7
Αδειοδότηση πόρων	7
Ανοικτές Άδειες συνόλων δεδομένων	7
Περιορισμένες άδειες	8
Υποχρεώσεις των ερευνητών	8
Ιδιότητες συνόλων δεδομένων	9
Έκδοση	9
Μεταδεδομένα	9
Τύποι και μορφές	9
Απόσυρση συνόλου δεδομένων	10

- Απόφαση διαβούλευσης, Γενική Συνέλευση ΣΕΑΒ, 16 Ιουνίου 2021
- Απόφαση έγκρισης, ΔΣ ΣΕΑΒ, XX XX XX
- Απόφαση έγκρισης, Γενική Συνέλευση ΣΕΑΒ, XX XX XX

Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών, 2022, CC-BY

Έκδοση 1

Εισαγωγή

Τα ερευνητικά δεδομένα είναι πολύτιμα για την πρόοδο της Επιστήμης και δη της Ανοικτής που αποτελεί πλέον το κύριο παράδειγμα και υποστηρίζει ο ΣΕΑΒ με τις πολιτικές και πρακτικές του. Λόγω της αξίας τους πρέπει να διατηρούνται συστηματικά και συντεταγμένα, ούτως ώστε να μπορούν να αξιοποιούνται μακροπρόθεσμα, μέσω της πρόσβασης και της δυναμικής επαναχρησιμοποίησής τους στο μέλλον. Η παρούσα πολιτική αφορά τα δεδομένα (αλλιώς γνωστά ως «ψηφιακά δεδομένα») που αποτίθενται και αποθηκεύονται ηλεκτρονικά στο Αποθετήριο Ερευνητικών Δεδομένων του ΣΕΑΒ (HARDMIN - Hellenic Academic Research Data Management Initiative).

Αναγκαιότητα χρήσης του HARDMIN

Καθώς όλα τα ψηφιακά δεδομένα εξαρτώνται από το λογισμικό που χρησιμοποιείται κάθε φορά, ο κίνδυνος να αχρηστευθούν ή να χαθούν είναι πολύ πιθανός, εφόσον κατά κανόνα το λογισμικό εξελίσσεται ή/και καθίσταται παρωχημένο. Η αποθήκευση των δεδομένων και η διατήρησή τους πρέπει να εξασφαλίζεται με τη χρήση μηχανισμών backup για την προστασία από ψηφιακές ή φυσικές καταστροφές. Οι δύο παραπάνω συνθήκες διασφαλίζονται με τη χρήση του HARDMIN και της κεντροποιημένης υποδομής του ΣΕΑΒ.

Σκοπός της λειτουργίας του κεντρικού αυτού αποθετηρίου είναι η συγκέντρωση του συνόλου των ερευνητικών δεδομένων που δημιουργούνται από Έλληνες ερευνητές και ακαδημαϊκούς των ελληνικών Πανεπιστημίων και η διάθεση τους με όσο το δυνατόν πιο ανοικτό και ασφαλή τρόπο.

Το αποθετήριο έρχεται να καλύψει την καίρια ανάγκη ασφαλούς φύλαξης και δημοσίευσης ερευνητικών δεδομένων της ελληνικής επιστημονικής κοινότητας. Απώτερος στόχος είναι η αύξηση της διαφάνειας στην έρευνα, η δυνατότητα επαναχρησιμοποίησης των ερευνητικών δεδομένων από τους ενδιαφερόμενους ερευνητές ανά τον κόσμο, η επιτάχυνση του ψηφιακού μετασχηματισμού του ερευνητικού πεδίου στη χώρα μας και η υιοθέτηση ανταγωνιστικών πρακτικών στον στίβο των ερευνητικών προτάσεων και της επιστημονικής επικοινωνίας.

Η παρούσα πολιτική περιγράφει συνοπτικά τα ελάχιστα πρότυπα που πρέπει να τηρούνται για τη διασφάλιση της ποιότητας των δεδομένων, συγκεκριμενοποιεί τη μορφή των υποστηριζόμενων αρχείων, ορίζει τις δυνατότητες αδειοδότησης και όλες τις διαδικασίες που πρέπει να ακολουθούνται κατά την απόθεση ερευνητικών δεδομένων στο HARDMIN.

Η πρόσβαση και η χρήση του Αποθετηρίου Ερευνητικών Δεδομένων HARDMIN για απόθεση επιτρέπεται από όλους/ες τους ερευνητές/ριες που έχουν ενεργή σχέση απασχόλησης σε ίδρυμα μέλος του ΣΕΑΒ. Σε περιπτώσεις, η πρόσβαση και η χρήση του HARDMIN για απόθεση μπορεί να επιτραπεί και σε μη ενεργούς/ες ερευνητές/ριες, σύμφωνα πάντα όμως με τις ρυθμίσεις που επιβάλλονται από τον ΣΕΑΒ και σε άλλες υπηρεσίες, όπως για παράδειγμα η συμμετοχή σε προγράμματα Ανοικτής Πρόσβασης. Η πρόσβαση γίνεται αποκλειστικά με τον ιδρυματικό λογαριασμό που κατέχει ενεργά κάποιος/α ερευνητής/ρία, όπως αυτός υποστηρίζεται από τα ομόσπονδα συστήματα εξουσιοδότησης στα οποία συμμετέχει ο Σύνδεσμος.

Για την αποτελεσματική χρήση του Αποθετηρίου Ερευνητικών Δεδομένων HARDMIN από τους ερευνητές/τριες των ακαδημαϊκών ιδρυμάτων της χώρας προτείνεται η σύσταση Ομάδας Υποστήριξης Απόθεσης Ερευνητικών Δεδομένων (ΟΥΑΕΔ) σε κάθε ακαδημαϊκό/ερευνητικό ίδρυμα που συμμετέχει στον ΣΕΑΒ. Η ΟΥΑΕΔ μπορεί να στελεχώνεται σύμφωνα με τις δυνατότητες

κάθε ιδρύματος, αλλά συνιστάται να συμμετέχουν τουλάχιστον ένας βιβλιοθηκονόμος της κεντρικής βιβλιοθήκης και ένας πληροφορικός. Άλλα μέλη της μπορούν να προέρχονται από το προσωπικό του Ειδικού Λογαριασμού Κονδυλίων Έρευνας ή τη Νομική Υπηρεσία.

Απόθεση δεδομένων

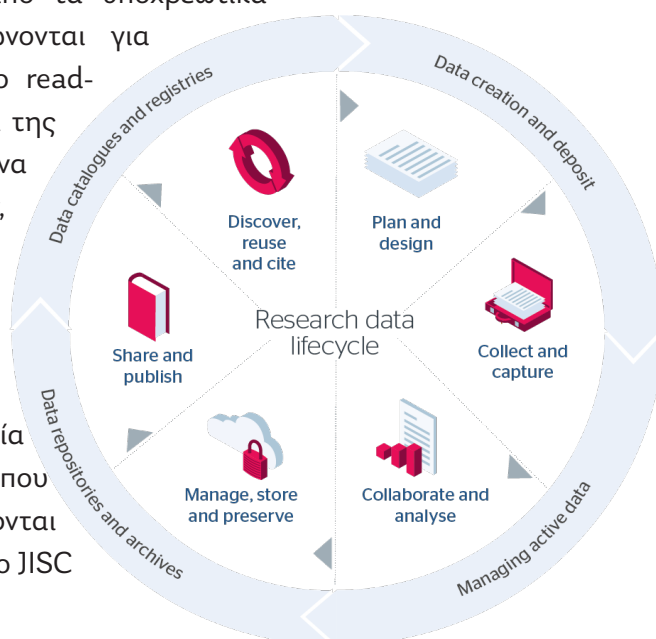
Γενική πολιτική

Όλοι οι ερευνητές που ενδιαφέρονται να αποθηκεύσουν δεδομένα στο HARDMIN συστήνεται να συνεργαστούν με την ΟΥΑΕΔ του ιδρύματός τους, προκειμένου να αποσαφηνίσουν λεπτομέρειες για τη διαδικασία απόθεσης δεδομένων, τη μορφή των μεταδεδομένων, την προετοιμασία των αρχείων, την εκπόνηση ενός Σχεδίου Διαχείρισης Ερευνητικών Δεδομένων (Research Data Management Plan), κλπ.

Για να διατηρηθούν και να προσπελαστούν τα σύνολα δεδομένων στο HARDMIN πρέπει να τηρούνται οι ακόλουθες προϋποθέσεις:

- Το μέγεθος κάθε μεμονωμένου αρχείου που σχετίζεται με ένα σύνολο δεδομένων οφείλει να βρίσκεται εντός των ορίων της τρέχουσας πολιτικής απόθεσης δεδομένων στο HARDMIN. Οι ερευνητές, ειδικά εκείνοι που διαθέτουν μεγέθη δεδομένων που υπερβαίνουν τα όρια της αποθήκευσης μέσω κοινών φυλλομετρητών (browsers), θα πρέπει να συμβουλευονται την ΟΥΑΕΔ του ιδρύματος τους για να προσδιορίζουν εναλλακτικούς τρόπους απόθεσης δεδομένων, όπως το FTP upload.
- Εξ' ορισμού, η εγγραφή του κάθε συνόλου δεδομένων στο HARDMIN είναι δημόσια προσβάσιμη μέσω του Διαδικτύου. Η πρόσβαση στα αρχεία του συνόλου δεδομένων διαφοροποιείται ανάλογα με το επίπεδο άδειας που έχει ορίσει ο ερευνητής (ανοικτή/Creative Commons ή περιορισμένη/restricted) – βλ. παράγραφο «Αδειοδότηση πόρων».
- Οι ερευνητές/ριες θα πρέπει να καταβάλλουν εύλογη προσπάθεια χρήσης των προτεινόμενων μορφών και τύπων αρχείων, ακολουθώντας τις γενικές συστάσεις της ενότητας «Τύποι και μορφές συνόλων δεδομένων».
- Συνιστάται τα σύνολα δεδομένων, πέρα από τα υποχρεωτικά μεταδεδομένα που πρέπει να συμπληρώνονται για την απόθεση, να συνοδεύονται από αρχείο readme, για τη διευκόλυνση της κατανόησης και της επαναχρησιμοποίησης τους, αλλά και από ένα Σχέδιο Διαχείρισης Ερευνητικών Δεδομένων, όπως υπαγορεύουν οι καλές πρακτικές, αλλά και οι απαιτήσεις πολλών χρηματοδοτών.

Το σύνολο της διαδικασίας, από τα βήματα που αφορούν αποκλειστικά στους/ις ερευνητές/ριες και τα βήματα που βασίζονται στη συνεργασία με την ΟΥΑΕΔ, όπως επίσης και οι διαδικασίες που διασφαλίζονται από το ίδιο το HARDMIN, παρουσιάζονται σχηματικά στη διπλανή εικόνα που προέρχεται από το JISC RDM Toolkit (2018).



Τα ιδρύματα-μέλη του ΣΕΑΒ μπορούν να προσθέτουν όρους σε δικές τους εκδοχές της Πολιτικής Απόθεσης Δεδομένων, οι οποίοι όμως δεν υπερисχύουν αυτών που προβλέπει η παρούσα πολιτική.

Ειδική πολιτική: χειρισμός ευαίσθητων δεδομένων

Η Πολιτική Απόθεσης Δεδομένων που περιγράφεται στο παρόν έγγραφο ισχύει για όλα τα παράγωγα έρευνας. Δύναται, εντούτοις, για την προστασία ευαίσθητων προσωπικών δεδομένων ή για τη συμμόρφωση με όρους που επιβάλλονται από τρίτους που δίνουν άδεια χρήσης των δεδομένων τους (third party restricted datasets), να επιβάλλονται από τους/ις ερευνητή/ριες εύλογοι περιορισμοί στη διαθεσιμότητα των δεδομένων. Η ΟΥΑΕΔ εντός του κάθε ιδρύματος μπορεί να παρέχει γενικές συμβουλές σχετικά με την απόθεση συνόλων δεδομένων που περιέχουν ευαίσθητες πληροφορίες. Για παράδειγμα, η ΟΥΑΕΔ θα μπορούσε να προσφέρει συγκεκριμένες πληροφορίες για την ανωνυμοποίηση δεδομένων και τη σύνδεση ευαίσθητων δεδομένων με δημοσιεύσεις, εάν και εφόσον αυτό κρίνεται απαραίτητο από τον ερευνητή. Με τον όρο «ευαίσθητα» δεδομένα εννοούνται, ενδεικτικά και μη περιοριστικά, τα:

- δεδομένα προσωπικού χαρακτήρα που αποκαλύπτουν φυλετική ή εθνοτική καταγωγή, πολιτικά φρονήματα, θρησκευτικές ή φιλοσοφικές πεποιθήσεις·
- δεδομένα συμμετοχής σε συνδικαλιστική οργάνωση·
- γενετικά δεδομένα, βιομετρικά δεδομένα που υποβάλλονται σε επεξεργασία αποκλειστικά για την ταυτοποίηση ενός ατόμου·
- δεδομένα σχετικά με την υγεία·
- δεδομένα σχετικά με τη σεξουαλική ζωή ή τον σεξουαλικό προσανατολισμό ενός ατόμου.

Τα ευαίσθητα προσωπικά δεδομένα υπόκεινται σε συγκεκριμένες προϋποθέσεις επεξεργασίας και η διαδικασία «ευαίσθητου» χειρισμού τους θα πρέπει να υφίσταται καθ' όλη τη διάρκεια της ερευνητικής διαδικασίας, δηλαδή κατά τη διάρκεια της συλλογής, επεξεργασίας και αποθήκευσης των δεδομένων. Η διαδικασία αυτή αφορά τους/ις ερευνητές/ριες και τις εγκρίσεις που λαμβάνουν από την Επιτροπή Ηθικής και Δεοντολογίας του ΕΛΚΕ κάθε ιδρύματος και η ΟΥΑΕΔ λειτουργεί υποστηρικτικά στη συγκεκριμένη διαδικασία. Για παράδειγμα, για τη διαδικασία της ανωνυμοποίησης δεδομένων, οι ερευνητές μπορούν να χρησιμοποιούν προτεινόμενα από την ΟΥΑΕΔ εργαλεία, όπως το AMNESIA του OpenAIRE.

Δικαιώματα πρόσβασης και ιδιοκτησία

Οι ερευνητές που αποθέτουν ερευνητικά δεδομένα στο HARDMIN πρέπει να είναι κάτοχοι της διανοητικής ιδιοκτησίας των συγκεκριμένων συνόλων δεδομένων και των δικαιωμάτων που εκπορεύονται από αυτή ή να έχουν τη συγκατάθεση του νόμιμου δικαιούχου για να το πράξουν.

Όλοι οι συνεισφέροντες στα δεδομένα και η ιδιότητά τους θα πρέπει να περιγράφονται λεπτομερώς κατά τη διαδικασία της απόθεσης, μέσω των αντίστοιχων πεδίων μεταδεδομένων. Τα πρόσωπα αυτά θα καταγράφονται ως «Κάτοχοι δεδομένων». Οποιαδήποτε επικοινωνία κριθεί σκόπιμη θα γίνεται με τους συγκεκριμένους κατόχους.

Ο περιορισμός της πρόσβασης δύναται να υλοποιηθεί μέσω του μηχανισμού «Περιορισμένης

πρόσβασης (restricted access)» του HARDMIN, με βάση τον οποίο κάποιος μπορεί να ζητήσει μέσα από το σύστημα από τους κατόχους των δεδομένων να έχει πρόσβαση σε ένα σύνολο δεδομένων. Συνιστάται να ακολουθείται αυτός ο τρόπος και να μην ισχύει κάποιος κωδικός πρόσβασης στα ίδια τα δεδομένα, π.χ. ένα password protected zip file.

Σε περίπτωση που οι κάτοχοι των δεδομένων επιλέξουν να βρίσκονται σε καθεστώς περιορισμένης πρόσβασης και να απαιτείται έγκριση για την αποθήκευση και τη χρήση τους από τρίτους, τότε απαιτείται έγκριση μόνο ενός από τους κατόχους δεδομένων για να παρασχεθεί η άδεια πρόσβασης στα σύνολα δεδομένων.

Τα ερευνητικά δεδομένα δεν θα πρέπει να παραβιάζουν την ισχύουσα νομοθεσία ούτε να παραβιάζουν τα πνευματικά δικαιώματα δικαιούχων ή τρίτων. Για δεδομένα τρίτων, τα οποία αποτελούν μέρος του ευρύτερου προς απόθεση συνόλου δεδομένων, θα πρέπει να διασφαλίζεται η άδεια επανάχρησης και όλα τα τυχόν νόμιμα δικαιώματα αυτών.

Όλοι οι νόμιμοι περιορισμοί στην πρόσβαση (restricted access) ή την περαιτέρω χρήση των ερευνητικών δεδομένων, συμπεριλαμβανομένης της περιόδου χρονικού αποκλεισμού (embargo period), πρέπει να προσδιορίζονται με σαφήνεια και να καταγράφονται στα αντίστοιχα πεδία που σχετίζονται με τον περιορισμό πρόσβασης (π.χ. όροι περιορισμένης πρόσβασης που πρέπει να πληρούνται για να αποκτήσει κάποιος πρόσβαση στα δεδομένα).

Σε περίπτωση που τα ερευνητικά δεδομένα εντάσσονται σε κάποιο χρηματοδοτούμενο πρόγραμμα, θα πρέπει να τηρούνται όλοι οι περιορισμοί που συνοδεύουν το πρόγραμμα αυτό (π.χ. χρονικοί περιορισμοί, περιορισμοί πρόσβασης, κ.ά.)

Περίοδος χρονικού αποκλεισμού

Οι ερευνητές μπορούν να καταθέτουν σύνολα δεδομένων υποκείμενα σε μια περίοδο χρονικού αποκλεισμού (γνωστή ως εμπάργκο, embargo period) και να προσδιορίζουν την ημερομηνία άρσης του. Το HARDMIN, αυτοματοποιημένα, θα περιορίζει την πρόσβαση στα δεδομένα μέχρι το τέλος της περιόδου αποκλεισμού και με την παρέλευση του οριζόμενου χρονικού διαστήματος, το περιεχόμενο θα είναι αυτόματα διαθέσιμο στο κοινό. Παρόλα αυτά, η πρόσβαση στα μεταδεδομένα της εγγραφής θα είναι κανονικά προσβάσιμη ακόμα και πριν τη λήξη της περιόδου εμπάργκο.

Αναγνώριση πόρων

Στο HARDMIN προσφέρεται εγγενώς η δυνατότητα ανάθεσης μοναδικών αναγνωριστικών ψηφιακών πόρων και φυσικών προσώπων, όπως αυτές οι υπηρεσίες υλοποιούνται στο πλαίσιο των ισχυουσών συμφωνιών με έγκριτους παρόχους. Στην παρούσα φάση, έχει κριθεί σκόπιμη η συνεργασία με παρόχους προνομιακών υπηρεσιών (premium) για αναγνωριστικά τύπου DOI και ORCID.

Μόνιμα αναγνωριστικά ψηφιακών πόρων

Σκοπός της δημιουργίας των μόνιμων αναγνωριστικών (persistent identifiers) είναι ο μονοσήμαντος προσδιορισμός κάθε παραγώγου έρευνας (ερευνητικά δεδομένα, δημοσιεύσεις

κ.ά.). Η υπηρεσία παραγωγής και ανάθεσης DOI (Digital Object Identifier) σε κάθε ψηφιακό πόρο παρέχεται αυτοματοποιημένα κατά τη διαδικασία της απόθεσης δεδομένων ενός συνόλου ερευνητικών δεδομένων στο HARDMIN. Αυτό σημαίνει πως, μετά τη συμπλήρωση των απαραίτητων μεταδεδομένων και το πέρας της επιτυχημένης αποθήκευσης του συνόλου δεδομένων, αποδίδεται από το σύστημα ένα μοναδικό αναγνωριστικό τύπου DOI.

Μόνιμα αναγνωριστικά φυσικών προσώπων

Στο Αποθετήριο Ερευνητικών Δεδομένων HARDMIN έχει ολοκληρωθεί η ενσωμάτωση των προνομιακών υπηρεσιών του μη κερδοσκοπικού οργανισμού ORCID (Open Researcher and Contributor Identifier) και οι ερευνητές έχουν την δυνατότητα, κατά τη διαδικασία της απόθεσης ενός συνόλου δεδομένων, να διασυνδέουν τον λογαριασμό τους στο HARDMIN με το ORCID iD τους. Το ORCID iD μπορεί να συγκεντρώσει πολλαπλά αναγνωριστικά που δίνονται από διαφορετικές υπηρεσίες (SCOPUS ID, Researcher ID, αναγνωριστικά από διαφορετικούς εκδότες/παρόχους) και καθίσταται πλέον υποχρεωτικό από πάρα πολλούς εκδότες/παρόχους για την κατάθεση του ερευνητικού τους έργου προς κρίση και δημοσίευση.

Αδειοδότηση πόρων

Τα σύνολα δεδομένων μπορεί να είναι είτε ανοικτά προς διάθεση είτε να υφίστανται περιορισμούς διάθεσης και χρήσης, οι οποίοι ορίζονται από τον εκάστοτε ερευνητή.

Ανοικτές Άδειες συνόλων δεδομένων

Με τον όρο Ανοικτές Άδειες αναφέρεται το σύνολο των αδειών χρήσης περιεχομένου, λογισμικού ή δεδομένων οι οποίες επιτρέπουν τη χρήση του αδειοδοτικού περιεχομένου με τους ελάχιστους δυνατούς περιορισμούς. Ένας από τους πλέον διαδεδομένους τρόπους εφαρμογής των Ανοικτών Αδειών είναι οι Creative Commons, ως ένα βασικό εργαλείο «για την εξισορρόπηση των δικαιωμάτων των τωρινών και μελλοντικών δημιουργών, και ως συμβατικό αντίβαρο στη διαρκή επέκταση στο χρόνο των δικαιωμάτων πνευματικής ιδιοκτησίας και το διαρκή περιορισμό των εξαιρέσεων και περιορισμών του δικαιώματος πνευματικής ιδιοκτησίας (Lessig, 2005).» όπως αναφέρεται από την υπηρεσία HEAL Legal του ΣΕΑΒ. Στην απόφαση για την επιλογή της κατάλληλης άδειας χρήσης λαμβάνονται υπόψη τρία βασικά ερωτήματα: (α) Απαιτείται η αναφορά του δημιουργού; (β) Επιτρέπεται η εμπορική χρήση; (γ) Επιτρέπεται η τροποποίηση του έργου; Εάν ναι, ο χρήστης θα πρέπει να χρησιμοποιήσει την ίδια άδεια χρήσης ή μπορεί να επιλέξει άλλη; Ο συνδυασμός των απαντήσεων σε αυτές τις ερωτήσεις αυτών των όρων βοηθά στο να δημιουργηθούν οι έξι τακτικά χρησιμοποιούμενες άδειες, συν μια που διαθέτει το περιεχόμενο ως Κοινό Κτήμα (Public Domain), και τις οποίες μπορούν οι καταθέτες να εφαρμόζουν στα σύνολα δεδομένων τους.

- CC-BY 4.0 - Creative Commons Attribution 4.0 International
- CC-BY-SA 4.0 - Creative Commons Attribution-ShareAlike 4.0 International
- CC-BY-NC 4.0 - Creative Commons Attribution-NonCommercial 4.0 International
- CC-BY-NC-SA 4.0 - Creative Commons Attribution-NonCommercial-ShareAlike 4.0

International

- CC BY-ND 4.0 - Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International
- CC BY-NC-ND 4.0 - Creative Commons Attribution-NonCommercial-NoDerivs 4.0 Generic
- CC0 1.0 Universal (CC0 1.0) Creative Commons Public Domain Dedication

Περιορισμένες άδειες

Οι ερευνητές μπορούν να καταθέτουν σύνολα δεδομένων, τα οποία θα συνοδεύονται από μια περιορισμένη άδεια και θα είναι ανοικτά για διανομή μόνο κατόπιν έγκρισης, εφόσον πληρούνται ορισμένες προϋποθέσεις που οι ίδιοι θέτουν. Αυτά τα αρχεία δεν θα είναι διαθέσιμα στο κοινό και η κοινή χρήση θα είναι δυνατή μόνο με την έγκριση του ερευνητή που κατέθεσε το σύνολο δεδομένων. Για τη διαδικασία έγκρισης, ο ερευνητής θα λαμβάνει προσωποποιημένο e-mail που θα τον ειδοποιεί πως κάποιος χρήστης αιτήθηκε πρόσβαση στο συγκεκριμένο σύνολο δεδομένων. Θα δίνεται η δυνατότητα αιτιολογημένης έγκρισης ή απόρριψης του αιτήματος ή επιπλέον ρητού περιορισμού. Εννοείται πως, και σε αυτή την περίπτωση όμως, τα μεταδεδομένα που περιγράφουν το σύνολο δεδομένων θα είναι δημόσια ορατά και προσβάσιμα.

Υποχρεώσεις των ερευνητών

Στο μέτρο του δυνατού, οι ερευνητές θα πρέπει να διασφαλίζουν ότι τα κατατεθειμένα σύνολα δεδομένων είναι υψηλής ποιότητας, δεν περιέχουν σφάλματα και είναι καλά τεκμηριωμένα. Το περιεχόμενο και η ποιότητα των δεδομένων είναι εξ ολοκλήρου ευθύνη του ερευνητή/ιδιοκτήτη των συνόλων δεδομένων και, πριν από την απόθεσή τους, θα πρέπει να πληρούνται τουλάχιστον τα ακόλουθα κριτήρια:

- Τα σύνολα δεδομένων να είναι πλήρως ελεγμένα για ιούς.
- Μη κρυπτογραφημένα.
- Χωρίς προστασία με κωδικό πρόσβαση.
- Σε μη συμπιεσμένη μορφή, όπου είναι δυνατόν.
- Να έχουν περάσει την διαδικασία ανωνυμοποίησης, όπως αυτή περιγράφεται στην ενότητα «Ειδική πολιτική: χειρισμός ευαίσθητων δεδομένων».

Η απόθεση δεδομένων έρευνας στο HARDMIN δεν αποκλείει την απόθεση σε οποιοδήποτε άλλο αποθετήριο δεδομένων. Σε ορισμένες περιπτώσεις, η απόθεση σε ένα εξωτερικό αποθετήριο επιβάλλεται λόγω των περιορισμών που θέτουν τρίτα μέρη, όπως για παράδειγμα ένας εκδότης μιας κυρίως δημοσίευσης που συνδέεται με τα σύνολα δεδομένων.

Εάν υπάρχει απόθεση του συνόλου δεδομένων σε εξωτερικό αποθετήριο, ο ερευνητής ενθαρρύνεται να δημιουργήσει εγγραφή και στο HARDMIN. Η εγγραφή θα συνοδεύεται με επαρκή τεκμηρίωση ως προς τα μεταδεδομένα και με σύνδεσμο προς το σύνολο δεδομένων (μέσω της συμπλήρωσης του αντίστοιχου πεδίου για DOIs τρίτων πηγών στα μεταδεδομένα), προκειμένου να διασφαλιστεί ότι τα δεδομένα παραμένουν ανιχνεύσιμα και σε εθνικό επίπεδο.

Ιδιότητες συνόλων δεδομένων

Ως γενική αρχή, θα πρέπει να καταβάλλεται κάθε προσπάθεια διασφάλισης των αρχών του FAIRness για τα ερευνητικά δεδομένα που αποτίθενται στο HARDMIN. Ως FAIR θεωρούνται τα δεδομένα που εκπληρώνουν τις εξής συνθήκες: Ανεύρεσιμα (Findable), Προσβάσιμα (Accessible), Διαλειτουργικά (Interoperable) και Επαναχρησιμοποιήσιμα (Reusable). Για κάποιες από τις παραπάνω ιδιότητες έχει ληφθεί μέριμνα που αναφέρεται σε προηγούμενες ενότητες, π.χ. Αναγνώριση πόρων και Δικαιώματα πρόσβασης και ιδιοκτησίας.

Έκδοση

Το σύνολο ή τα σύνολα δεδομένων, τα οποία θα κατατίθενται στο HARDMIN, πρέπει να αποτελούν το πρωτότυπο, πλήρες και τελικό αντίγραφο των ερευνητικών δεδομένων ή μια έκδοση (version) ενός συνόλου δεδομένων, το οποίο χρησιμοποιήθηκε για την έρευνα σε κάποια δεδομένη χρονική στιγμή. Η ενεργοποίηση της λειτουργίας του versioning του HARDMIN βοηθά προς την κατεύθυνση αυτή.

Όπου είναι απαραίτητο, μπορεί να κατατεθεί μια ενημερωμένη έκδοση ενός συνόλου δεδομένων. Σε τέτοιες περιπτώσεις η προηγούμενη έκδοση θα αφαιρείται από τη δημόσια προβολή, αλλά δεν θα διαγράφεται, και οι σύνδεσμοι θα παραμένουν ενεργοί, ούτως ώστε να τηρούνται οι όροι του “FAIRness” των δεδομένων.

Μεταδεδομένα

Τα επιμέρους στοιχεία των δεδομένων (π.χ. μεταδεδομένα, μορφή, τύπος, κτλ) θα αξιολογούνται για την καταλληλότητά τους για απόθεση και επαναχρησιμοποίηση από την ΟΥΑΕΔ του ιδρύματος. Η αξιολόγηση θα αποτελείται από μία σειρά ελέγχων που έχουν να κάνουν κυρίως με τα κριτήρια που προαναφέρθηκαν στην ενότητα «Υποχρεώσεις των ερευνητών» και δεν θα αφορούν την αξιοπιστία του περιεχομένου του συνόλου δεδομένων.

Όπου τα δεδομένα δεν συμμορφώνονται με τις ελάχιστες απαιτήσεις για τη διαδικασία της απόθεσης δεδομένων, η ΟΥΑΕΔ θα συνεργάζεται με τον ερευνητή, προκειμένου να διαμορφώσει ένα πλάνο αντιμετώπισης των σχετικών ζητημάτων, ούτως ώστε να καταστεί εφικτή η μεταφόρτωση του συνόλου δεδομένων.

Τύποι και μορφές

Λόγω της μεγάλης παραγωγής και ετερογένειας των ερευνητικών δεδομένων από τα ακαδημαϊκά ιδρύματα από τα οποία τα δεδομένα προέρχονται, είναι σαφές ότι δυσχεραίνεται η εκτέλεση ολοκληρωμένων εργασιών ψηφιακής επιμέλειας (data curation) σε όλους τους τύπους και μορφές δεδομένων. Παρόλα αυτά, προτείνεται να υπάρχει σχετική διαδικασία ψηφιακής επιμέλειας σε επίπεδο ιδρύματος, από την ΟΥΑΕΔ, η οποία θα εγγυάται συνέπεια και συνοχή.

Επί της αρχής, δεν υπάρχει περιορισμός στους τύπους δεδομένων που μπορούν να αποθετηθούν στο HARDMIN. Εφόσον όλοι οι τύποι των δεδομένων θα μπορούν να κατατεθούν στο HARDMIN, αυτό θα γίνεται με βάση τις γενικές συστάσεις από την ΟΥΑΕΔ. Ανάλογα

όμως με τον τύπο των δεδομένων σε συγκεκριμένα ερευνητικά πεδία, δύναται να υπάρχουν συγκεκριμένες κατευθυντήριες οδηγίες σχετικά με την ψηφιακή τους επιμέλεια. Τα σύνολα δεδομένων θα αξιολογούνται για απόθεση στο HARDMIN με βάση τα ακόλουθα κριτήρια, όπου και θα παρέχεται το ανάλογο επίπεδο υποστήριξης:

- Προτιμώμενες μορφές: Μορφές που υιοθετούνται εκτενώς και υποστηρίζονται από ένα ευρύ φάσμα πλατφορμών λογισμικού. Αυτές οι μορφές ευνοούν τη μακροχρόνια συντήρηση και επαναχρησιμοποίηση των δεδομένων και θα λαμβάνουν εκτενή υποστήριξη επιμέλειας από την Ομάδα Υποστήριξης Απόθεσης Ερευνητικών Δεδομένων. Επί παραδείγματι, είναι προτιμητέα μια μορφή csv έναντι μιας μορφής xls/xlsx.
- Αποδεκτές μορφές: Πρόκειται για μορφές που συνήθως δημιουργούνται από εμπορικά διαθέσιμο λογισμικό (π.χ. matlab, GIS data κτλ). Ωστόσο, οι μορφές αυτές υιοθετούνται ευρέως για πρακτικούς λόγους που έχουν να κάνουν με το εκάστοτε ερευνητικό πεδίο και είναι αδύνατο να μην χρησιμοποιηθούν. Αυτές θα λαμβάνουν περιορισμένη υποστήριξη επιμέλειας από την Ομάδα Υποστήριξης απόθεσης δεδομένων Ερευνητικών Δεδομένων.
- Προβληματικές μορφές: Πρόκειται για μορφές που συνήθως δημιουργούνται από εμπορικά διαθέσιμο λογισμικό, μη αναγνωρισμένες ή μη βιώσιμες στο διηνεκές, έχουν κακή τεκμηρίωση και δεν υιοθετούνται ευρέως. Αυτές οι μορφές δε θα λαμβάνουν υποστήριξη επιμέλειας που να εγγυάται τη μελλοντική χρηστικότητα.

Περαιτέρω, τα ονόματα των αρχείων, η έκδοση του αρχείου και η δομή των δεδομένων και των φακέλων θα αξιολογούνται για θέματα που έχουν να κάνουν με την πρόσβαση και τη διατήρησή τους. Αντίγραφα (backups) των συνόλων δεδομένων θα τηρούνται σε επίπεδο εξοπλισμού (infrastructure level) μέσω της κεντροποιημένης υποδομής του ΣΕΑΒ.

Όπου τα δεδομένα δεν συμμορφώνονται με τις ελάχιστες απαιτήσεις για τη διαδικασία της απόθεσης, η ΟΥΑΕΔ θα συνεργάζεται με τον ερευνητή, προκειμένου να διαμορφώσει ένα πλάνο αντιμετώπισης των σχετικών ζητημάτων, ούτως ώστε να καταστεί εφικτή η αναμενόμενη απόθεση του συνόλου δεδομένων.

Απόσυρση συνόλου δεδομένων

Δεν υπάρχει ορισμένος χρόνος ζωής των δεδομένων και η καταθεσή τους στο HARDMIN γίνεται επ' αόριστον. Ακόμη και αν υπάρχουν συμβατικές υποχρεώσεις (π.χ. προς κάποιον χρηματοδότη έρευνας) για την τήρηση δεδομένων για ένα ορισμένο χρονικό διάστημα, η αξία των δεδομένων δεν χάνεται απαραίτητα μετά το πέρας αυτού, αλλά μπορούν να είναι χρήσιμα για κάθε ερευνητή. Τυχόν βελτιώσεις σε ένα σύνολο δεδομένων εξυπηρετείται από τις διαδικασίες χαρακτηρισμού έκδοσης που έχει περιγραφεί σε προηγούμενη ενότητα.

Η ΟΥΑΕΔ κάθε ιδρύματος μπορεί να αποσύρει ένα σύνολο δεδομένων όταν συντρέχει κάποια από τις παρακάτω προϋποθέσεις:

- Αποδεδειγμένη παραβίαση πνευματικών δικαιωμάτων τρίτου προσώπου και λογοκλοπή
- Παραβίαση του Γενικού Κανονισμού Προστασίας Δεδομένων
- Ασφάλεια (εθνική ασφάλεια, ασφάλεια δικτύων, ασφάλεια ιδρυματικών υποδομών κλπ.)

Το Αποθετήριο HARDMIN διατηρεί το δικαίωμα τήρησης των μεταδεδομένων ενός συνόλου

δεδομένων, εάν απαιτείται από σοβαρούς λόγους, όπως για παράδειγμα ανάγκη διαρκούς αναφοράς.

