

ΤΙΤΛΟΣ ΕΡΕΥΝΗΤΙΚΗΣ ΠΡΟΤΑΣΗΣ:

Η εξαίρεση εξόρυξης κειμένων και δεδομένων στην πνευματική ιδιοκτησία και τις ακαδημαϊκές βιβλιοθήκες

A. Αναλυτική περιγραφή ερευνητικής πρότασης

A.1. Τα μεγάλα δεδομένα (Big Data) και η σημασία τους

Ένας απίστευτα μεγάλος όγκος πληροφοριών δημιουργείται ανά πάσα στιγμή από δισεκατομμύρια χρήστες σε όλη την υφήλιο. Περισσότερα δεδομένα έχουν δημιουργηθεί από τις μηχανές και τον άνθρωπο τα τελευταία δύο έτη απ' ό,τι σε ολόκληρη την ιστορία της ανθρωπότητας. Παραπάνω από 2,4 εκατομμύρια επιστημονικά άρθρα δημοσιεύονται κάθε χρόνο. Είναι πλέον, σχεδόν, αδύνατο να διαχειριστούμε αυτή τη θάλασσα των γνώσεων και των πληροφοριών και να επωφεληθούμε από αυτήν. Στην προσπάθειά μας να αξιοποιήσουμε τον πακτωλό γνώσεων, πληροφοριών και δεδομένων είναι απαραίτητη η υιοθέτηση νέων μεθόδων και η ανάπτυξη και παραμετροποίηση αντίστοιχων εργαλείων διαχείρισης της γνώσης. Ένα από τα βασικά εργαλεία σε αυτήν την προσπάθεια αναδεικνύονται οι μέθοδοι εξόρυξης κειμένων και δεδομένων (text and data mining methods), οι οποίες εντοπίζουν πληροφορίες μέσα σε αυτήν την απίστευτη μάζα των αδόμητων δεδομένων. Στη διαδικασία εξόρυξης κειμένων και δεδομένων εφαρμόζονται αλγοριθμικά συστήματα και προηγμένα εργαλεία λογισμικού που όχι μόνο συγκεντρώνουν και αναλύουν τεράστιες ποσότητες ψηφιακών ή ψηφιοποιημένων πληροφοριών αλλά επιπλέον δημιουργούν συσχετισμούς, οι οποίοι δεν ήταν προφανείς. Η εξόρυξη κειμένων και δεδομένων έχει μία τεράστια δυνατότητα να δημιουργεί, αποτελεσματικά και δραστικά, αξία από τα μαζικά δεδομένα (big data) και από τον διαρκώς αυξανόμενο όγκο των κειμένων.

Στην εποχή της Τέταρτης Βιομηχανικής Επανάστασης, ο κινητήριος μοχλός της οικονομίας φαίνεται να είναι αυτά τα μαζικά δεδομένα. Η διαχείριση των μαζικών δεδομένων προσδίδει προστιθέμενη αξία στην πληροφορία και τίθεται στο επίκεντρο της μελλοντικής οικονομίας της γνώσης. Η ανάγκη αξιοποίησης των μαζικών δεδομένων οδηγεί σε καινοτόμες τεχνολογικές εφαρμογές, στην ανάπτυξη νέων εργαλείων και νέων δεξιοτήτων, ενώ αναμένεται να μεταμορφώσει τις βιομηχανίες παροχής υπηρεσιών, δημιουργώντας ένα ευρύ πεδίο καινοτόμων πληροφοριακών προϊόντων και υπηρεσιών, αυξάνοντας την παραγωγικότητα όλων των τομέων της οικονομίας μέσω της βελτίωσης της έρευνας.

A.2. Η εξόρυξη κειμένων και δεδομένων ως τεχνική

Ως εξόρυξη κειμένων και δεδομένων εννοείται κάθε αυτοματοποιημένη αναλυτική τεχνική που αποσκοπεί στην ανάλυση κειμένων και δεδομένων σε ψηφιακή μορφή με στόχο την παραγωγή πληροφοριών, όπως προτύπων, τάσεων και συσχετισμών. Η εξόρυξη κειμένων και δεδομένων γίνεται σε σώματα κειμένων ή ψηφιακών/ψηφιοποιημένων δεδομένων ηλεκτρονικής μορφής.

Η εξόρυξη μπορεί να μας δώσει πληροφορίες πολλών 'βαθμίδων', εμπλέκοντας συσχετισμούς που μπορεί να επέλθουν μεταξύ οποιουδήποτε συνδυασμού όρων, δεδομένων ή εκφράσεων μέσα στο σώμα κειμένων που υπόκεινται σε εξόρυξη. Η εξόρυξη αφορά 'μετα-κειμενικά' θέματα, όπως π.χ. στην ανίχνευση της ιστορίας της επιστήμης, εντοπίζεται και υπολογίζεται πόσες φορές εμφανίζεται η φράση 'γρίπη των πουλερικών', στο περιεχόμενο άρθρων. Άλλοτε όμως η εξόρυξη αφορά ήδη διατυπωμένες πληροφορίες. Έτσι, ένα σώμα ερευνητικής χημικής βιβλιογραφίας μπορεί να υποστεί

εξόρυξη με ειδικό λογισμικό, ώστε να παράγει μια νέα βάση δεδομένων για τις μοριακές δομές, πολύτιμη πηγή για τους μοριακούς χημικούς, τους κρυσταλλογράφους και άλλους ερευνητές. Το πρόγραμμα Robots Reading Vogue του Πανεπιστημίου του Yale χρησιμοποιεί το αρχείο του περιοδικού Vogue με την πολυθεματική πλατφόρμα ProQuest για να μας δώσει ιδέες για τη μεταβαλλόμενη χρήση των εικόνων και της γλώσσας στη μόδα του 20ου αιώνα. Η ανάλυση της γραμματικής δομής εκατομμυρίων προτάσεων επιτρέπει να αποκτηθούν γνώσεις μέσα από τον τρόπο γραφής των συγγραφέων. Η αμερικανική βάση νομικών δεδομένων LEXIS NEXIS επιτρέπει στους χρήστες της τη φόρτωση και αποθήκευση 500 νομικών άρθρων ταυτόχρονα κάθε φορά σε ένα αρχείο αποθήκευσης, καθιστώντας ευχερή την εξόρυξη κειμένων και δεδομένων. Η βάση επιστημονικών δεδομένων JSTOR επιτρέπει τη μεταφόρτωση μεταδεδομένων, συχνότητα λέξεων, ετεροαναφορών και λέξεων - κλειδιών σε μέχρι 1.000 έγγραφα, ενώ εάν ένας ερευνητής επιθυμεί την εξόρυξη σε περισσότερα από 1.000, μπορεί να το κάνει κατόπιν συνεννόησης με την JSTOR.

Πολλά προγράμματα λογισμικού εξυπηρετούν χρήσεις εξόρυξης κειμένων και δεδομένων. Ενδεικτικά αναφέρουμε το CrossRef Text and Data Mining Tool, το Europeana Search API, το HathiTrust Research Center Analytics (που υποστηρίζει τη δωρεάν υπολογιστική ανάλυση 16 εκατομμυρίων έργων από το 1700 μέχρι σήμερα σε 400 γλώσσες που φιλοξενούνται στην HathiTrust Digital Library για να διευκολύνει μη κερδοσκοπικούς και εκπαιδευτικούς σκοπούς), το JSTOR Text Analyzer, το PLOS API, το Content Mine και το PubMed Central Text Mining Tools. Πολλά από αυτά προσφέρονται για χρήση στο διαδίκτυο μέσω μεγάλων ψηφιακών βιβλιοθηκών ξένων Πανεπιστημίων (όπως το UCLA, το Boston College, το University of Southern California κλπ), οι οποίες συχνά προσφέρουν για τους χρήστες τους στο διαδίκτυο έναν ηλεκτρονικό κατάλογο που περιέχει τις βάσεις δεδομένων που είναι διαθέσιμες για εξόρυξη, με ειδική αναφορά όρων χρήσης ανά βάση. Μεγάλοι εκδοτικοί οίκοι, όπως ο Wiley και η Elsevier, επίσης, παρέχουν τη δυνατότητα εξόρυξης (π.χ. η Wiley Online Library).

Η δυνατότητα χρήσης των τεχνολογιών εξόρυξης εξαρτάται συχνά από την ιδιότητα του χρήστη, την ανάγκη χορήγησης ειδικής άδειας, τη σύναψη σύμβασης ή και την ειδική διαπραγμάτευση των όρων μέσω μια βιβλιοθήκης (όπως η βιβλιοθήκη ETH της Ζυρίχης). Συνήθως αναφέρεται, όταν η εξόρυξη γίνεται μέσω ακαδημαϊκής βιβλιοθήκης, πως η χρήση της εξόρυξης θα πρέπει να αφορά σε μη εμπορικό, εκπαιδευτικό ή ερευνητικό σκοπό και πως η συστηματική μεταφόρτωση, διανομή περιεχομένου σε μη αδειοδοτημένους χρήστες ή η αέναη διατήρηση ουσιαστικών τμημάτων των βάσεων απαγορεύεται (βλ. ενδεικτικά οδηγίες της βιβλιοθήκης του Πανεπιστημίου του Yale).

Πολλές σημαντικές πρωτοβουλίες και συνέργειες σε διεθνές και ευρωπαϊκό επίπεδο στοχεύουν στην όσο το δυνατό ευρύτερη ανάπτυξη και χρήση αυτής της τεχνολογίας: α) το πρόγραμμα EUDATA2020 απευθύνεται σε διαχειριστές μεγάλων δεδομένων ώστε να ανεβάζουν τα δεδομένα τους στο διαδίκτυο προς χρήση με ανοικτές πλατφόρμες που χτίζει το ίδιο το πρόγραμμα, β) το OpenMintTeD στοχεύει, επίσης, να χτίσει μια διαδικτυακή υποδομή που θα διευκολύνει την εξόρυξη δεδομένων, καθώς τα καθιστά ευκολότερα στην εξόρυξη και δίνει δυνατότητες ανεύρεσής τους και διαλειτουργικότητας, γ) το Future TDM, επιδιώκει τη βελτίωση της εφαρμογής της τεχνικής της εξόρυξης στην Ευρώπη για την ερευνητική κοινότητα, και δ) το LEARN (Leaders Activating Research Networks) επιδιώκει τη μαζική ευαισθητοποίηση για την αποτελεσματικότερη διαχείριση δεδομένων για την έρευνα.

A.3. Η σχέση της εξόρυξης κειμένων και δεδομένων με την πνευματική ιδιοκτησία

Σε αυτήν την εποχή των μεγάλων δεδομένων, όπου η κοινωνία είναι βυθισμένη σε έναν ωκεανό νέων πληροφοριών το δίκαιο της πνευματικής ιδιοκτησίας αντιμετωπίζει πρωτόγνωρες προκλήσεις. Οι ερευνητές δεν έχουν άλλη επιλογή από το να στραφούν στους αλγόριθμους και σε άλλα εργαλεία για να διαχειριστούν αυτόν τον όγκο των πληροφοριών. Ωστόσο, εάν τα δεδομένα που χρησιμοποιούνται για την ανάλυση απολαμβάνουν προστασίας από το δίκαιο της πνευματικής ιδιοκτησίας, τότε γεννώνται περίπλοκα νομικά ζητήματα.

Παρόλο που θέσφατο της πνευματικής ιδιοκτησίας είναι ότι τα δεδομένα δεν προστατεύονται, συχνά το υλικό που είναι αντικείμενο της ανάλυσης κατά την εξόρυξη αποτελεί προστατευόμενο έργο, ενώ περισσότερο από συχνά υφίστανται δικαιώματα είτε πνευματικής ιδιοκτησίας είτε *sui generis* επί των βάσεων δεδομένων, στις οποίες περιλαμβάνονται τα πρωτογενή (μη προστατευόμενα καθαυτά) δεδομένα.

Σε κάθε περίπτωση στην έρευνα αντικείμενο της εξόρυξης είναι κυρίως επιστημονικά άρθρα και συγγράμματα, ενώ δεδομένης της χαμηλής στάθμης της πρωτοτυπίας που απαιτείται πλέον - σύμφωνα και με τη νομολογία του ΔΕΕ (απόφαση C- 5/08 Infopaq)- για την προστασία έργων από το δίκαιο της πνευματικής ιδιοκτησίας, είναι βέβαιο ότι συχνά θα προκύπτουν ζητήματα πνευματικής ιδιοκτησίας στα διάφορα στάδια της εξόρυξης κειμένων και δεδομένων.

Τα στάδια της εξόρυξης κειμένων και δεδομένων περιλαμβάνουν α) εντοπισμό του υλικού που θα αναλυθεί, όπως έργα ή δεδομένα συλλεγμένα ατομικά ή οργανωμένα σε μία προϋφιστάμενη βάση δεδομένων, β) αντιγραφή σημαντικών ποσοτήτων του υλικού, που περιλαμβάνει μετατροπή τους σε μηχαναγνώσιμη μορφή και ενδεχομένως μεταφόρτωσή τους σε μία πλατφόρμα ανάλογα με τη συγκεκριμένη τεχνική εξόρυξης που χρησιμοποιείται, γ) εξαγωγή των δεδομένων και επανασυνδυασμός τους για να γίνει αναγνώριση συγκεκριμένων μοτίβων.

Προκύπτει ότι κατά τη διάρκεια των διάφορων σταδίων της εξόρυξης είναι τεχνικά απαραίτητες ορισμένες πράξεις, οι οποίες, αν πραγματοποιούνται χωρίς την άδεια των δικαιούχων, μπορούν να παραβιάζουν εξουσίες πνευματικής ιδιοκτησίας.

Συνεπώς, για να υπάρχει πρόσβαση και σύννομη αξιοποίηση της εξόρυξης εγγράφων και δεδομένων σε επίπεδο πνευματικής ιδιοκτησίας θα πρέπει να πληρούνται οι προϋποθέσεις μίας εξαίρεσης. Σε ενωσιακό επίπεδο θα μπορούσε ενδεχομένως να εφαρμοστούν οι εξαιρέσεις της προσωρινής αναπαραγωγής, της έρευνας, της ιδιωτικής αναπαραγωγής, της νόμιμης χρήσης της βάσης δεδομένων και η εξαγωγή επουσιωδών μερών μιας βάσης δεδομένων.

Εντούτοις, η επίκληση των εξαιρέσεων αυτών δεν είναι ιδιαίτερα επιβλητική για τους ερευνητές, καθώς η πλειονότητα αυτών είναι προαιρετικές και συνεπώς υφίσταται ένα μωσαϊκό εξαιρέσεων σε επίπεδο ΕΕ, δημιουργώντας επιπλέον προβλήματα σε διασυνοριακό επίπεδο, ενώ είναι αβέβαιη η εφαρμογή τους, καθώς οι επικαλούμενες εξαιρέσεις δεν θεσπίστηκαν για τον συγκεκριμένο σκοπό.

A.4. Η εξαίρεση για την εξόρυξη κειμένων και δεδομένων ως νομικός κανόνας

Προκύπτει ότι το υφιστάμενο νομικό πλαίσιο στην ΕΕ – αλλά και στην Ελλάδα- παραμένει εν πολλοίς ασαφές. Οι ανεξάντλητες τεχνολογικές δυνατότητες εξόρυξης κειμένων και δεδομένων παραμένουν απλά δυνατότητες άνευ ουσίας, όταν το δίκαιο της πνευματικής ιδιοκτησίας εμποδίζει τη χρήση τους. Αυτό επισημαίνει και η Διακήρυξη της Χάγης για την Πρόσβαση στη Γνώση στην Ψηφιακή Εποχή του 2014.

Παρόλα αυτά υφίστανται ήδη στην ΕΕ ορισμένες εθνικές νομοθεσίες, οι οποίες αναγνωρίζοντας τη σημασία της εξόρυξης στην έρευνα και στην καινοτομία, προβλέπουν σχετικές εξαιρέσεις για την εξόρυξη δεδομένων, εκμεταλλευόμενες το υφιστάμενο ενωσιακό πλαίσιο των εξαιρέσεων.

A.4.1. Η εξαίρεση στα εθνικά δίκαια ευρωπαϊκών χωρών

Από τις ευρωπαϊκές χώρες, το Ηνωμένο Βασίλειο και η Γαλλία έχουν περισσότερο ασχοληθεί με την ενσωμάτωση της εξαίρεσης για την εξόρυξη κειμένων και δεδομένων. Η εξαίρεση πάντως έχει ενταχθεί και στα εθνικά δίκαια της Γερμανίας και της Εσθονίας, ενώ έχει συζητηθεί και στην Ιρλανδία.

A.4.1.α. Ηνωμένο Βασίλειο

Στο Ηνωμένο Βασίλειο η εξαίρεση για την εξόρυξη κειμένου και δεδομένων θεσμοθετήθηκε το 2014 με το άρθρο 29Α. Η εξαίρεση προβλέπει τη νόμιμη αντιγραφή κειμένου και δεδομένων από άτομο που ήδη έχει νόμιμη πρόσβαση στα έργα, εάν η υπολογιστική ανάλυση πραγματοποιείται αποκλειστικά για έρευνα χωρίς εμπορικό σκοπό και με επαρκή αναγνώριση της πηγής του έργου (εκτός εάν αυτό παρουσιάζει δυσχέρεια). Αντίγραφα έργων που δημιουργούνται με χρήση της εξαίρεσης δεν επιτρέπεται να μεταβιβαστούν σε τρίτο, ούτε να χρησιμοποιηθούν για άλλους σκοπούς. Συμβατικές ρήτρες, όμως, που περιορίζουν την εξαίρεση σε συμβάσεις πνευματικής ιδιοκτησίας είναι άκυρες.

Η εξαίρεση δεν μπορεί να ακυρωθεί με σύμβαση. ενώ οι εκδότες δικαιούνται να περιορίζουν την πρόσβαση στα έργα με τεχνολογικά μέτρα προστασίας. Όμως το Γραφείο Διανοητικής Ιδιοκτησίας του ΗΒ έχει υποστηρίξει ότι τα μέτρα αυτά πρέπει να είναι εύλογα και να μην αποκλείουν ή να περιορίζουν αδόκητα τη δυνατότητα των ερευνητών να επωφεληθούν της εξαίρεσης. Η εξαίρεση δεν εφαρμόζεται έναντι του κατασκευαστή μη πρωτότυπης βάσης δεδομένων, που διατηρεί τα δικαιώματα αφαίρεσης και επαναχρησιμοποίησης ουσιώδους μέρους της βάσης.

A.4.1.β. Γαλλία

Μετά από πολυετείς αντιπαραθέσεις το 2016 ο Κώδικας Πνευματικής Ιδιοκτησίας τροποποιήθηκε και υιοθετήθηκε με το άρθρο L. 122-5, 10 η εξαίρεση για την εξόρυξη κειμένων και δεδομένων. Η πολύ περιορισμένη αυτή εξαίρεση καλύπτει τα αντίγραφα και ψηφιακές αναπαραγωγές από νόμιμη πηγή για τον σκοπό της εξόρυξης κειμένων και δεδομένων που περιλαμβάνονται ή σχετίζονται με επιστημονικές δημοσιεύσεις για σκοπό δημόσιας έρευνας, αποκλείοντας κάθε εμπορικό σκοπό. Για να ενεργοποιηθεί η εφαρμογή της εξαίρεσης αυτής πρέπει να εκδοθεί ειδικό διάταγμα, το οποίο μέχρι σήμερα όμως δεν έχει πραγματοποιηθεί, καθώς αναμένεται η υιοθέτηση της Πρότασης Οδηγίας για την ενιαία ψηφιακή αγορά. Έτσι σήμερα η εξαίρεση παραμένει ανενεργή.

Η εξαίρεση αφορά και σε βάσεις δεδομένων που προστατεύονται με το ειδικής φύσης δικαίωμα του κατασκευαστή για μη εμπορικό σκοπό σε σχέση με επιστημονικές δημοσιεύσεις. Κάθε αντίγραφο που προκύπτει από την εφαρμογή τη εξαίρεσης καταστρέφεται, αφού εκπληρωθεί ο μη εμπορικός σκοπός της χρήσης του.

A.4.2. Ελληνικό δίκαιο

Η εξόρυξη κειμένων και δεδομένων στην Ελλάδα θα μπορούσε καταρχήν να εμπίπτει σε μία από τις εξαιρέσεις που ήδη έχουν θεσμοθετηθεί σε ευρωπαϊκό επίπεδο με την Οδηγία 2001/29, όπως η εξαίρεση για τον τεχνολογικό περιορισμό για τις προσωρινές πράξεις αναπαραγωγής. Στη θεωρία υποστηρίζεται επίσης ότι η εξαίρεση για την επιστημονική έρευνα της Οδηγίας (α. 5.3.α) θα μπορούσε από μόνη της να καλύπτει την εξόρυξη. Ο Έλληνας νομοθέτης, όμως, δεν ενσωμάτωσε στην εθνική νομοθεσία την προαιρετική αυτή εξαίρεση για την έρευνα, όπως αντίθετα το έπραξε ως εξαίρεση από το ειδικής φύσης δικαίωμα του κατασκευαστή βάσης δεδομένων.

Σήμερα, δεν υπάρχει θεσμοθετημένη εξαίρεση για την εξόρυξη κειμένων και δεδομένων στην Ελλάδα.

A.4.3 Ενωσιακό Δίκαιο

Δεδομένης της ανασφάλειας σε σχέση με την εξόρυξη κειμένων και δεδομένων και των εξαιρέσεων από το δίκαιο της πνευματικής ιδιοκτησίας και με γνώμονα την ανάπτυξη της έρευνας και της καινοτομίας σε ενωσιακό επίπεδο, κρίθηκε αναγκαία η καθιέρωση μιας υποχρεωτικής εξαίρεσης από το περιουσιακό δικαίωμα του δημιουργού με την Πρόταση Οδηγίας για τα δικαιώματα πνευματικής ιδιοκτησίας στην ψηφιακή ενιαία αγορά. Ένα σαφές νομοθετικό πλαίσιο θα επιτρέψει την αποφυγή της περίπλοκης εκκαθάρισης των πνευματικών δικαιωμάτων και θα μειώσει σημαντικά το επενδυτικό ρίσκο, με αποτέλεσμα την ενίσχυση της καινοτομίας.

Η υποχρεωτική νέα εξαίρεση (άρθρο 3 της Πρότασης Οδηγίας) προβλέπει ότι ερευνητικοί οργανισμοί (αλλά και ιδρύματα πολιτιστικής κληρονομιάς), στους οποίους περιλαμβάνονται οι βιβλιοθήκες των πανεπιστημίων αλλά και οι προσιτές στο κοινό βιβλιοθήκες μπορούν να αναπαράγουν προστατευόμενα έργα και αντικείμενα προστασίας και να εξάγουν πληροφορίες χρησιμοποιώντας τις τεχνολογίες εξόρυξης κειμένων και δεδομένων, υπό τον όρο ότι οι ερευνητές έχουν νόμιμη πρόσβαση στο περιεχόμενο επί του οποίου υλοποιείται η ανάλυση. Οι ερευνητές θα πρέπει να αποθηκεύουν τα αντίγραφα σε ασφαλές ηλεκτρονικό περιβάλλον και να τα διατηρούν όχι περισσότερο από όσο είναι αναγκαίο για να επιτύχουν τους σκοπούς της επιστημονικής έρευνας.

Η προτεινόμενη εξαίρεση θα αφορά εξαίρεση στα δικαιώματα της αναπαραγωγής, του δημιουργού βάσης δεδομένων, του κατασκευαστή μη πρωτότυπης βάσης δεδομένων και στο δικαίωμα εκδοτών τύπου σχετικά με τις ψηφιακές χρήσεις των δημοσιεύσεών τους.

Η προτεινόμενη εξαίρεση έχει γίνει δεκτή με ενθουσιασμό από τη θεωρία, ωστόσο το πεδίο εφαρμογής της εξαίρεσης έχει υποστεί κριτική ότι είναι περιορισμένο. Ειδικότερα, αποκλείονται από το πεδίο εφαρμογής οι κερδοσκοπικοί οργανισμοί/εταιρείες. Παρότι γίνεται ειδική μνεία στις συνέργειες δημόσιου και ιδιωτικού τομέα, αποκλείονται οι ιδιωτικοί οργανισμοί, με αποτέλεσμα ανεξάρτητοι επιστήμονες να μην μπορούν να επωφεληθούν της εξαίρεσης. Ο 'σκοπός της επιστημονικής έρευνας' θα πρέπει να διασαφηνισθεί για να μην προκληθεί δικαϊκή ανασφάλεια. Τέλος, περιοριστικά στο πεδίο εφαρμογής της εξαίρεσης λειτουργεί και η προϋπόθεση της νόμιμης πρόσβασης για να τεθεί σε εφαρμογή η εξαίρεση αυτή.

Η διάρκεια αποθήκευσης σε ασφαλές ηλεκτρονικό περιβάλλον για λόγους επαλήθευσης αποτελεί μείζον θέμα. Σύμφωνα με την Πρόταση Οδηγίας, οι δικαιούχοι επιτρέπεται να εφαρμόζουν μέτρα με σκοπό την εξασφάλιση της ασφάλειας και της ακεραιότητας των δικτύων και των βάσεων

δεδομένων όπου φιλοξενούνται τα έργα ή άλλο υλικό. Τα μέτρα αυτά δεν πρέπει να υπερβαίνουν τα αναγκαία όρια για την επίτευξη του στόχου αυτού. Ωστόσο, κρίσιμο στοιχείο για τους ερευνητές αποτελεί η επαλήθευση των αποτελεσμάτων και η φύλαξη των δεδομένων προκειμένου να γίνει επιβεβαίωση των αποτελεσμάτων. Πώς θα εξισορροπηθούν τα αντιτιθέμενα συμφέροντα εν προκειμένω; Μεγάλη πρακτική σημασία εμφανίζει η σχέση της εξαίρεσης με έργα, προστατευόμενα αντικείμενα και βάσεις δεδομένων που προστατεύονται με τεχνολογικά μέτρα προστασίας και ο τρόπος εξεύρεσης νομοθετικής λύσης για την πρακτική και 'ειρηνική' συνύπαρξή τους. Ιδιαίτερα θετικό όμως το γεγονός ότι δεν επιτρέπεται ο συμβατικός αποκλεισμός της εξαίρεσης.

Επειδή τη στιγμή της σύνταξης της παρούσας ερευνητικής πρότασης, έχουν προσωρινά διακοπεί οι τριμερείς διαπραγματεύσεις επί της Πρότασης Οδηγίας για την πνευματική ιδιοκτησία στην ψηφιακή ενιαία αγορά και δεν έχει λάβει την οριστική της μορφή, είναι βέβαιο ότι θα προκύψουν νέα ερωτήματα κατά την υιοθέτηση του τελικού κειμένου της εξαίρεσης. Ενδεικτικό της καίριας σημασίας που αποδίδεται στην εν λόγω εξαίρεση είναι ότι πλέον προστέθηκε και μία δεύτερη γενικότερη εξαίρεση κατόπιν αιτήματος κρατών μελών, για την οποία όμως μέχρι τη διακοπή των διαπραγματεύσεων της εν λόγω Πρότασης Οδηγίας εριζόταν κατά πόσο θα είναι προαιρετική ή υποχρεωτική. Άγνωστο παραμένει όχι μόνο η τελική μορφή αλλά ακόμα και το μέλλον της ίδιας της Πρότασης Οδηγίας.

Σε κάθε περίπτωση παρόλο που σκοπός της νέας πρότασης Οδηγίας είναι να απαλειφθούν οι νομικές ασάφειες ως προς τη χρήση της εξόρυξης κειμένων και δεδομένων, η ίδια θα δημιουργήσει νέες ασάφειες ως προς τους όρους εφαρμογής της. Η παρούσα μελέτη έχει σκοπό να εντοπίσει τις ασάφειες αυτές και να συμβάλει στην αποσαφήνισή τους.

A.4.4. Η εξαίρεση στις ΗΠΑ και Ιαπωνία

Εκτός ΕΕ, στο αμερικανικό δίκαιο, οι εξαιρέσεις από το δικαίωμα αναπαραγωγής στην πνευματική ιδιοκτησία, συγκεντρώνονται σε μια, αυτή της δίκαιης χρήσης, ενώ τα Δικαστήρια ερμηνεύουν τη γενική αυτή ρήτρα ανά περίπτωση, με βάση βέβαια κανόνες που έχουν διαμορφωθεί. Με την απόφαση του Εφετείου στην υπόθεση *Authors' Guild v. Google* του 2015 διευκρινίσθηκε ότι η μαζική ψηφιοποίηση έργων ουσιαστικά για σκοπούς εξόρυξης κειμένων και δεδομένων από τη Google δεν παραβιάζε τα δικαιώματα συγγραφέων και εκδοτών, ενώ η απόφαση του Εφετείου κρίθηκε από το Ανώτατο Δικαστήριο των ΗΠΑ πως δεν μπορεί να τεθεί σε νομική κρίση από το ίδιο, επικυρώνοντας σε τελικό βαθμό την ισχύ της. Συνεπώς, στις ΗΠΑ αναγνωρίζεται μια ευρεία εξαίρεση εξόρυξης κειμένων και δεδομένων, η οποία τυγχάνει εκτεταμένης και πολύπλευρης εφαρμογής, ενώ συνεισφέρει στην ανάπτυξη καινοτομίας, επιχειρηματικότητας και ανάπτυξης της ανοικτής επιστήμης.

Το 2009, η Ιαπωνία εισήγαγε μια εξαίρεση για κάθε πληροφοριακή ανάλυση χωρίς καμία διάκριση, και επομένως συμπεριέλαβε την εξόρυξη για εμπορικούς σκοπούς στον ιαπωνικό Νόμο για την Πνευματική Ιδιοκτησία.

A.5 Οι ακαδημαϊκές βιβλιοθήκες και η εξαίρεση εξόρυξης κειμένων και δεδομένων

Ο ρόλος των βιβλιοθηκών γενικά αλλά και ειδικότερα των ακαδημαϊκών βιβλιοθηκών στην ανάπτυξη και τη διευκόλυνση της εξόρυξης κειμένων και δεδομένων είναι κομβικός. Η μεγαλύτερη

πλειονότητα των ερευνητών που επιθυμούν να προβούν σε εξόρυξη κειμένων και δεδομένων θα πρέπει να το πραγματοποιούν μέσω μιας ακαδημαϊκής βιβλιοθήκης, η οποία όχι μόνο διαθέτει την πρόσβαση σε βάσεις δεδομένων, έργα, ψηφιακές βιβλιοθήκες, περιοδικά κ.λπ. αλλά θα είναι ουσιαστικά και ο επωφελούμενος της εν λόγω εξαίρεσης. Οι βιβλιοθήκες θα μπορούν να συμβάλλονται για την αδειοδότηση πρόσβασης των ερευνητών στα έργα και τα δεδομένα για σκοπούς εξόρυξης κειμένων και δεδομένων και θα πρέπει να διαπραγματεύονται με τους εκδότες και τους δικαιούχους. Οι βιβλιοθήκες συχνά καλούνται να μεσολαβήσουν όταν οι ερευνητές αντιμετωπίζουν μια αυτόματη διακοπή της εξόρυξης στην οποία προβαίνουν, όπως όταν ειδοποιούνται από τον υπολογιστή τους ότι έληξε το όριο για φόρτωση έργων ή δεδομένων για σκοπό εξόρυξης, ότι διακόπτεται η πρόσβασή τους λόγω μιας 'ασυνήθους συμπεριφοράς' τους στη χρήση της βάσης ή όταν τους εμποδίζουν τεχνολογικά μέτρα προστασίας του υλικού. Εκτός όμως από τον συμβουλευτικό τους ρόλο και τη συνδρομή τους στην υποστήριξη των ερευνητών ως προς το πλαίσιο και το πεδίο εφαρμογής της εξαίρεσης της εξόρυξης κειμένων και δεδομένων οι βιβλιοθήκες θα κληθούν να εφαρμόσουν οι ίδιες την εξαίρεση για την εξόρυξη. Θα πρέπει για τον λόγο αυτό να κατέχουν εις βάθος το νομικό πλαίσιο της εξαίρεσης και θα πρέπει να επενδύουν σε τεχνικές εξόρυξης προσανατολισμένες για υποβοηθητικές υπηρεσίες προς όφελος των ερευνητών αλλά και των ίδιων των βιβλιοθηκών. Το προσωπικό των βιβλιοθηκών θα πρέπει να εξοπλιστεί με τις απαραίτητες δεξιότητες για να υποστηρίξει τις τεχνικές εξόρυξης κειμένων και δεδομένων, δημιουργώντας ερευνητικές ομάδες. Μπορούν επίσης να ενσωματώσουν ρήτρες αναφορικά με τις τεχνικές εξόρυξης κειμένων στις άδειες χρήσης, οι οποίες να αναφέρουν την εξαίρεση και να καθορίζουν τους σχετικούς όρους πρόσβασης. Σε ευρύτερο πλαίσιο μπορούν οι βιβλιοθήκες να αναπτύξουν μελέτες περιπτώσεων και καλές πρακτικές, παρουσιάζοντας τις δυνατότητες των τεχνικών εξόρυξης δεδομένων και κειμένων οικοδομώντας έναν γόνιμο και εποικοδομητικό διάλογο με τους εκδότες.

Οι παρεμβάσεις των βιβλιοθηκών και των οργανώσεών τους υπέρ της εξαίρεσης της εξόρυξης κειμένων και δεδομένων είναι πολλές. Η IFLA ήδη από το 2013 έχει εκδώσει Δήλωση για την εξαίρεση εξόρυξης κειμένων και δεδομένων, υποστηρίζοντας την ανάγκη να νομοθετηθεί επίσημα, εξασφαλίζοντας ελεύθερη πρόσβαση στην πληροφορία προς το εκπαιδευτικό και πολιτισμικό συμφέρον της ολότητας, τονίζοντας ότι η εξόρυξη είναι απαραίτητο εργαλείο για την πρόοδο της μάθησης και για νέες μορφές δημιουργίας. Παρόμοιες δηλώσεις και θέσεις έχουμε από την LIBER, την Europeana, την EBLIDA, την ALLEA (All European Academies) κ.λπ.. Στις επιστολές ζητείται η άμεση διεύρυνση της εξαίρεσης της Πρότασης Οδηγίας.

A.6. Ερευνητικά ερωτήματα

Με βάση τα παραπάνω η έρευνα θα επιχειρήσει να απαντήσει σε μια σειρά ερωτημάτων για την εξαίρεση εξόρυξης κειμένων και δεδομένων. Ενδεικτικά αναφέρονται:

i. Ποιες χώρες εντός και εκτός ΕΕ σήμερα έχουν υιοθετήσει την εξαίρεση για την εξόρυξη κειμένου και δεδομένων, ποιες είναι οι συνέπειές τους στην καινοτομία και την έρευνα και με ποια ειδικότερα χαρακτηριστικά σε σχέση με:

α. τα είδη των δικαιούχων

β. τον εμπορικό ή μη σκοπό της έρευνας

γ. τη δυνατότητα ή όχι διατήρησης αντιγράφου των εξορυχθέντων δεδομένων

δ. την ακυρότητα ή όχι αντίθετων με την εξαίρεση συμβατικών όρων

ε. την ισχύ ή όχι των τεχνολογικών μέτρων προστασίας του υλικού προς εξόρυξη και πιθανές διαδικασίες απενεργοποίησης των μέτρων αυτών με βάση την εξαίρεση.

ii. Σε σχέση με την Πρόταση Οδηγίας, ποιες είναι οι τροποποιήσεις που θα έπρεπε να γίνουν ώστε η εξαίρεση για την εξόρυξη κειμένων και δεδομένων να μπορεί να επιτρέψει την αληθινή πρόσβαση στην έρευνα σε συμφωνία με την Διακήρυξη της Χάγης για την πρόσβαση στη γνώση στην ψηφιακή εποχή;

iii. Σε ποιές συμβάσεις έχουν προβεί ακαδημαϊκές βιβλιοθήκες στο εξωτερικό και στην Ελλάδα με τους εκδότες, ώστε να επιτρέπεται μέσω αυτών η εξόρυξη κειμένων και δεδομένων; Ποιες αλλαγές στις συμβατικές αυτές ρυθμίσεις θα πρέπει να ζητηθούν από τους εκδότες/παρόχους πληροφορίας προς τη διασφάλιση πρόσβασης;

iv. Όταν η πρόσβαση σε κείμενα/δεδομένα προς εξόρυξη γίνεται στις ακαδημαϊκές βιβλιοθήκες όχι με βάση ειδική σύμβαση με τον εκδότη, με ποιους άλλους τρόπους διασφαλίζεται η πρόσβαση στην πληροφορία προς εξόρυξη;

v. Πως μπορεί να διασφαλισθεί η δυνατότητα αποθήκευσης και διατήρησης των δεδομένων από τον χρήστη που προέβη σε εξόρυξη αλλά και ο διαμοιρασμός τους με άλλους ερευνητές;

vi. Εφόσον προκύπτει ότι τα κείμενα και δεδομένα προς εξόρυξη φυλάσσονται με πληροφορικά συστήματα διαφορετικά μεταξύ τους, κάτι που αποδεδειγμένα δυσχεραίνει την εφαρμογή της εξαίρεσης σε μεγάλη κλίμακα, με ποιον νομικό τρόπο μπορεί να υποβοηθηθεί η ομοιογένεια των συστημάτων;

vii. Όταν ο Γενικός Κανονισμός Προστασίας Δεδομένων εμποδίζει τη χρήση της εξαίρεσης για την εξόρυξη, σε ποιες τροποποιήσεις θα πρέπει να προβούμε ώστε να μην εμποδίζεται η λειτουργία της εξαίρεσης, χωρίς να θίγονται δικαιώματα υποκειμένων των δεδομένων;

viii. Όταν η εξαίρεση εξόρυξης συγκρούεται με το δικαίωμα δημιουργού ή κατασκευαστή βάσης δεδομένων, με ποιο νομικό τρόπο θα πρέπει να επιλυθεί αυτή η σύγκρουση αυτή ώστε να διασφαλισθεί η νόμιμη άσκηση της εξαίρεσης;

ΕΔ4: Σύνταξη και διακίνηση ερωτηματολογίων και διεξαγωγή συνεντεύξεων με εμπειρογνώμονες στην εξόρυξη δεδομένων (π.χ. νομικούς και βιβλιοθηκονόμους).

ΕΔ5: Ανάλυση δεδομένων. Τα δεδομένα (βιβλιογραφικά και στατιστικά) που έχουν ανακτηθεί σε όλο το διάστημα της έρευνας (συστηματική ερμηνεία, βιβλιογραφική και συμμετοχική έρευνα) θα καταχωρηθούν, αποτυπωθούν, επικυρωθούν με σκοπό την εξαγωγή ασφαλών συμπερασμάτων.

ΕΔ6: Συγγραφή αποτελεσμάτων, ερμηνεία και σχολιασμός ευρημάτων έρευνας, δημοσίευση άρθρου, ανακοίνωση σε συνέδριο.