

Σύστημα Υποστήριξης Κλινικών Αποφάσεων
για τη Νόσο των Ανευρυσμάτων Κοιλιακής Αορτής
Βασισμένο σε Μοντέλα Τεχνητής Νοημοσύνης



Παραδοτέο Π3.3.

Υποδομή Δεδομένων Νέφους και
Ομοσπονδιακές Βάσεις –
Εργαλείο ψευδοταυτοποίησης ή
ανωνυμοποίησης ιατρικών δεδομένων

Όνομα Αρχείου:	Safe-Aorta-Π3.3-Υποδομή Νέφους και Ομοσπονδιακές Βάσεις Δεδομένων (Τμήμα 1).pdf	Επίπεδο Διάδοσης:	Δημόσιο
Ημερομηνία Υποβολής:	Απρίλιος 2025 (M21)	Κωδικός Έργου:	TAEDR-0535983
Κοινοπραξία:	ΕΜΠ, ΠΔΜ, ΠΚ, ΕΛΜΕΠΑ, ΠΑΔΑ, ΠΒΕΑΑ, ΠΑΠΕΛ	Υπεύθυνος Παραδοτέου:	Πανεπιστήμιο Ιωαννίνων (ανάδοχος)
Διάρκεια:	28 μήνες	Κατάσταση:	Τελικό

Ελλάδα 2.0
ΕΘΝΙΚΟ ΣΧΕΔΙΟ ΑΝΑΚΑΜΨΗΣ
ΚΑΙ ΑΝΘΕΚΤΙΚΟΤΗΤΑΣ

ΓΓΕΚ
ΓΕΝΙΚΗ ΓΡΑΜΜΑΤΕΙΑ
ΕΡΕΥΝΑΣ ΚΑΙ ΚΑΙΝΟΤΟΜΙΑΣ



Με τη χρηματοδότηση
της Ευρωπαϊκής Ένωσης
NextGenerationEU

ΛΙΣΤΑ ΣΥΓΓΡΑΦΕΩΝ

Συγγραφείς				
#	Επίθετο	Όνομα	Φορέας	Email Επικοινωνίας
1	Σιόγκας	Παναγιώτης	ΕΛΚΕ ΠΙ	psiogkas@uoi.gr
2				
3				
4				

Συν-συγγραφείς				
#	Επίθετο	Όνομα	Φορέας	Email Επικοινωνίας
1	Φωτιάδης	Δημήτριος	ΕΛΚΕ ΠΙ	fotiadis@uoi.gr
2				
3				

ΛΙΣΤΑ ΚΡΙΤΩΝ

Κριτές				
#	Επίθετο	Όνομα	Φορέας	Email Επικοινωνίας
1	Ματσόπουλος	Γεώργιος	ΕΜΠ	gmatsopoulos@biomed.ntua.gr
2	Μανόπουλος	Χρήστος	ΕΜΠ	manopoul@central.ntua.gr
3	Ράπτης	Αναστάσιος	ΕΜΠ	raptistasos@mail.ntua.gr

ΕΛΕΓΧΟΣ ΑΝΑΘΕΩΡΗΣΗΣ

Έκδοση	Συγγραφέας	Ημερομηνία	Κατάσταση
0.1			Προσχέδιο
0.2			
0.3			
0.4	Παναγιώτης Σιόγκας	24/4/2025	Τελικό

Περιεχόμενα

Περιεχόμενα	iv
-------------------	----

Πίνακας Σχημάτων	iv
Πίνακας Πινάκων	v
Λίστα Συντομογραφιών.....	vi
Περίληψη.....	vii
1. Ενότητα 1	8
1.1 Σκοπός του παραδοτέου	8
1.2 Υπόβαθρο	9
1.3 Δομή του Εγγράφου	9
2. Εισαγωγή και Γενική Περιγραφή	10
3. Περιγραφή του Λογισμικού και της Υλοποίησης	11
3.1 Γενικά χαρακτηριστικά	12
3.2 Τεχνολογίες και βιβλιοθήκες	12
3.3 Αρχιτεκτονική εφαρμογής	13
3.4 Επίπεδα ανωνυμοποίησης και επιλογές χρήστη	13
3.5 Εξαγωγή αναφοράς και audit trail.....	14
4. Λειτουργίες Ανωνυμοποίησης	14
4.1 Επιλογή πεδίων DICOM προς ανωνυμοποίηση	14
4.2 Διαχείριση τιμών και τύπων πεδίων.....	16
4.3 Ανίχνευση ενσωματωμένων ευαίσθητων πληροφοριών.....	16
4.4 Προεπισκόπηση αποτελεσμάτων (Before / After)	17
4.5 Καταγραφή ενεργειών (logging).....	18
5. Γραφικό Περιβάλλον Χρήστη (GUI)	19
5.1 Δομή και στοιχεία της διεπαφής	19
5.2 Παραθυρικές λειτουργίες (dialogs).....	19
5.3 Σχεδιαστικές αρχές και προσβασιμότητα	20
5.4 Τεκμηρίωση διεπαφής.....	20
6. Δυνατότητες Επέκτασης και Παραμετροποίησης	20
6.1 Παραμετροποίηση επιπέδων ανωνυμοποίησης.....	21
6.2 Επέκταση για Web/Web-based environments	21
6.3 Ενσωμάτωση σε ροές εργασίας (pipelines)	21
6.4 Τεχνικές δυνατότητες αναβάθμισης.....	22
7. Οδηγίες Χρήσης	22
7.1 Προαπαιτούμενα	22
7.2 Εκκίνηση του εργαλείου	22

7.3	Κύρια βήματα χρήσης.....	23
7.4	Διαχείριση σφαλμάτων	23
7.5	Εγκατάσταση ως εκτελέσιμο αρχείο (Windows, macOS, Linux).....	24
8.	Συμμόρφωση με Τεχνικές Προδιαγραφές	27
8.1	Μέτρα για την εξασφάλιση εμπιστευτικότητας, ακεραιότητας και διαθεσιμότητας (Confidentiality, Integrity, Availability)	27
8.2	Μέτρα για την εξασφάλιση της ιδιωτικότητας των προσωπικών δεδομένων και συμμόρφωση με τον GDPR (Κανονισμός 679/2016)	27
8.3	Περιγραφή της μεθοδολογίας και των εργαλείων για την ανάπτυξη των ομοσπονδιακών βάσεων.....	28
9.	Συμπεράσματα και Μελλοντική Αξιοποίηση.....	30
10.	Βιβλιογραφία.....	32

Πίνακας Σχημάτων

Εικόνα 1: Κεντρική οθόνη του εργαλείου ανωνυμοποίησης.	12
Εικόνα 2: Δυνατότητα επιλογής πεδίων προς ανωνυμοποίηση από το χρήστη εάν επιλεγεί το custom level of anonymization.....	16
Εικόνα 3: Ενδεικτική εικόνα της λειτουργίας προεπισκόπησης (before / after) με τη χρήση της πλήρους ανωνυμοποίησης.....	17
Εικόνα 4: Δημιουργηθέν αρχείο καταγραφής για τα αρχεία που βρισκόταν στον επιλεγμένο φάκελο προς ανωνυμοποίηση.....	18

Πίνακας Πινάκων

Πίνακας 1: Πίνακας συμμόρφωσης με τις απαιτούμενες προδιαγραφές.	29
---	----

Λίστα Συντομογραφιών

Συντομογραφία	Πλήρης Όρος	Ελληνική Περιγραφή
AI	Artificial Intelligence	Τεχνητή Νοημοσύνη
API	Application Programming Interface	Διεπαφή Προγραμματισμού Εφαρμογών
DICOM	Digital Imaging and Communications in Medicine	Ψηφιακή Απεικόνιση και Επικοινωνία στην Ιατρική
EHDS	European Health Data Space	Ευρωπαϊκός Χώρος Δεδομένων Υγείας
FHIR	Fast Healthcare Interoperability Resources	Πρότυπο διαλειτουργικότητας δεδομένων υγείας
GDPR	General Data Protection Regulation	Γενικός Κανονισμός Προστασίας Δεδομένων
GUI	Graphical User Interface	Γραφικό Περιβάλλον Χρήστη
HL7	Health Level 7	Διεθνές πρότυπο ανταλλαγής δεδομένων υγείας
JSON	JavaScript Object Notation	Μορφότυπος ανταλλαγής δεδομένων
PACS	Picture Archiving and Communication System	Σύστημα Αρχειοθέτησης και Επικοινωνίας Ιατρικών Εικόνων
PyDICOM	Python DICOM Library	Βιβλιοθήκη Python για αρχεία DICOM
QA	Quality Assurance	Διασφάλιση Ποιότητας
UI/UX	User Interface / User Experience	Διεπαφή & Εμπειρία Χρήστη
UUID	Universally Unique Identifier	Καθολικά Μοναδικός Αναγνωριστικός Αριθμός
VM	Virtual Machine	Εικονική Μηχανή
XML	Extensible Markup Language	Επεκτάσιμη Γλώσσα Σήμανσης

Περίληψη

Το παρόν παραδοτέο τεκμηριώνει την ανάπτυξη και υλοποίηση ενός λογισμικού εργαλείου για την ανωνυμοποίηση ιατρικών εικόνων τύπου DICOM, με στόχο τη διασφάλιση της προστασίας προσωπικών δεδομένων και τη συμμόρφωση με τις τεχνικές και νομικές απαιτήσεις του έργου. Η λειτουργία του εργαλείου είναι κρίσιμη για την προτυποποίηση και εναρμόνιση δεδομένων σε περιβάλλοντα που απαιτούν ασφαλή επεξεργασία, διαμοιρασμό και επαναχρησιμοποίηση ιατρικών πληροφοριών, τόσο σε τοπικό όσο και σε ομοσπονδιακό επίπεδο.

Η εφαρμογή έχει αναπτυχθεί στη γλώσσα Python, χρησιμοποιώντας ανοιχτές βιβλιοθήκες όπως οι pydicom, tkinter και Pillow, ενώ υποστηρίζει πλήρως περιβάλλοντα Windows, Linux και macOS. Διαθέτει φιλικό γραφικό περιβάλλον χρήστη (GUI), στο οποίο προσφέρονται δυνατότητες επιλογής επιπέδου ανωνυμοποίησης, διαχείρισης DICOM tags, προεπισκόπησης αλλαγών, καθώς και καταγραφής όλων των τροποποιήσεων μέσω αυτόματου αρχείου αναφοράς (log).

Το εργαλείο έχει σχεδιαστεί με έμφαση στην επεκτασιμότητα και παραμετροποίηση, επιτρέποντας τη δημιουργία προσαρμοσμένων σεναρίων ανωνυμοποίησης σύμφωνα με τις ανάγκες κάθε φορέα ή έργου. Επιπλέον, μπορεί να ενσωματωθεί εύκολα σε ευρύτερες ροές εργασίας, APIs ή cloud-based περιβάλλοντα, υποστηρίζοντας λειτουργίες προεπεξεργασίας δεδομένων για έργα τεχνητής νοημοσύνης, ομοσπονδιακές βάσεις δεδομένων και πλατφόρμες διαλειτουργικότητας.

Η τεκμηρίωση που συνοδεύει το εργαλείο περιλαμβάνει αναλυτική περιγραφή της αρχιτεκτονικής, των λειτουργιών, των δυνατοτήτων επέκτασης, των απαιτήσεων εγκατάστασης, καθώς και πίνακες συμμόρφωσης με τις τεχνικές προδιαγραφές της προκήρυξης. Η λειτουργική πληρότητα, η τεχνική ωριμότητα και η πρακτική εφαρμοσιμότητα της λύσης εξασφαλίζουν την άμεση ενσωμάτωσή της στο έργο και τη δυνατότητα περαιτέρω αξιοποίησής της σε αντίστοιχα σενάρια χρήσης.

1. Ενότητα 1

Το παραδοτέο Π3.3 αφορά στην ανάπτυξη, τεκμηρίωση και λειτουργική αξιολόγηση ενός λογισμικού εργαλείου για την ανωνυμοποίηση ιατρικών εικόνων τύπου DICOM. Το εργαλείο επιτρέπει την ασφαλή και παραμετροποιήσιμη αφαίρεση προσωπικών δεδομένων, προσφέροντας γραφικό περιβάλλον χρήστη, δυνατότητα προεπισκόπησης και πλήρη συμμόρφωση με τις τεχνικές προδιαγραφές και το νομικό πλαίσιο προστασίας δεδομένων.

1.1 Σκοπός του παραδοτέου

Ο βασικός στόχος του παρόντος παραδοτέου είναι να τεκμηριώσει την ανάπτυξη και λειτουργία ενός εξειδικευμένου λογισμικού εργαλείου για την ανωνυμοποίηση ιατρικών δεδομένων μορφής DICOM, το οποίο αξιοποιείται στο πλαίσιο της προτυποποίησης και εναρμόνισης απεικονιστικών δεδομένων υγείας. Η δημιουργία ενός τέτοιου εργαλείου είναι απαραίτητη προϋπόθεση για την ασφαλή και νόμιμη διαχείριση ευαίσθητων πληροφοριών, όπως αυτές που περιλαμβάνονται σε αρχεία ιατρικής απεικόνισης, ιδίως όταν αυτά πρόκειται να αξιοποιηθούν σε ερευνητικό ή τεχνολογικό περιβάλλον εκτός του άμεσου κλινικού πλαισίου.

Το εργαλείο καλείται να καλύψει συγκεκριμένες απαιτήσεις λειτουργικότητας, διαλειτουργικότητας και συμμόρφωσης, όπως προβλέπονται από τη σχετική προκήρυξη. Πέραν της τεχνικής υλοποίησης, η παρούσα τεκμηρίωση παρουσιάζει αναλυτικά τον τρόπο με τον οποίο το λογισμικό ανταποκρίνεται σε αυτές τις απαιτήσεις, τόσο ως αυτόνομη εφαρμογή όσο και ως εργαλείο που δύναται να ενταχθεί σε μεγαλύτερα πληροφοριακά ή υποδομικά συστήματα υγείας.

Πιο συγκεκριμένα, το παραδοτέο περιλαμβάνει:

- Περιγραφή της μεθοδολογικής προσέγγισης που ακολουθήθηκε για την ανάπτυξη του εργαλείου.
- Ανάλυση των επιμέρους λειτουργιών του λογισμικού (π.χ. επιλογή επιπέδων ανωνυμοποίησης, προεπισκόπηση, καταγραφή αλλαγών).
- Παρουσίαση της τεχνικής του αρχιτεκτονικής και της παραμετροποίησης που επιτρέπει την προσαρμογή του σε διαφορετικές ανάγκες χρηστών ή οργανισμών.
- Αναλυτική τεκμηρίωση της συμμόρφωσης με τις απαιτήσεις του έργου, μέσω πίνακα αντιστοίχισης.
- Προτάσεις για τη μελλοντική αξιοποίηση του εργαλείου στο πλαίσιο υποδομών διαμοιρασμού ή επεξεργασίας ιατρικών δεδομένων (π.χ. federated learning, cloud-based αποθετήρια).

Η παρούσα αναφορά συνοδεύει το εργαλείο και λειτουργεί ως πλήρης τεχνική και λειτουργική τεκμηρίωση της συμβολής του στο έργο. Με την ολοκλήρωση του παραδοτέου, παρέχεται στις εμπλεκόμενες ομάδες ένα σταθερό, επεκτάσιμο και εύχρηστο εργαλείο για τη διαχείριση ιατρικών εικόνων, σύμφωνα με τις απαιτήσεις προστασίας δεδομένων και τις σύγχρονες προδιαγραφές ασφάλειας.

1.2 Υπόβαθρο

Τα ιατρικά δεδομένα απεικονιστικής φύσης, όπως αυτά που αποθηκεύονται σε μορφή DICOM, αποτελούν κρίσιμο πόρο για την παροχή υπηρεσιών υγείας, την επιστημονική έρευνα και την ανάπτυξη τεχνολογιών τεχνητής νοημοσύνης στον τομέα της ιατρικής. Ωστόσο, η ακατέργαστη διακίνηση και αποθήκευσή τους ενέχει σημαντικούς κινδύνους για την προστασία της ιδιωτικότητας των ασθενών, καθώς ενσωματώνουν προσωπικά αναγνωρίσιμες πληροφορίες. Η ανωνυμοποίηση των δεδομένων αυτών αποτελεί απαραίτητη διαδικασία προκειμένου να καταστούν επαναχρησιμοποιήσιμα, να συμμορφώνονται με τη νομοθεσία περί προσωπικών δεδομένων (όπως ο GDPR)[1] και να μπορούν να ενσωματωθούν σε υποδομές διαμοιρασμού ή συλλογικής επεξεργασίας [2].

Παράλληλα, σε έργα που προϋποθέτουν τη δημιουργία εναρμονισμένων βάσεων δεδομένων από πολλαπλές πηγές, η ύπαρξη ενός σταθερού και αξιόπιστου μηχανισμού ανωνυμοποίησης είναι ζωτικής σημασίας. Απαιτείται ένα εργαλείο που να εξασφαλίζει την αφαίρεση ευαίσθητων πεδίων, να προσαρμόζεται σε διαφορετικές πολιτικές απορρήτου και να υποστηρίζει την πλήρη ιχνηλασιμότητα των αλλαγών. Η παρούσα υλοποίηση ανταποκρίνεται σε αυτή την ανάγκη, λειτουργώντας ως βασικό τεχνικό στοιχείο προεπεξεργασίας για την ασφαλή και ομοιογενή χρήση των ιατρικών απεικονιστικών δεδομένων σε κάθε στάδιο του κύκλου ζωής τους.

1.3 Δομή του Εγγράφου

Το παρόν έγγραφο έχει οργανωθεί σε ενότητες που αντανakλούν τη συνολική πορεία σχεδίασης, ανάπτυξης και τεκμηρίωσης του εργαλείου ανωνυμοποίησης ιατρικών δεδομένων. Η δομή του στοχεύει στην πλήρη παρουσίαση τόσο της τεχνικής υλοποίησης όσο και της λειτουργικής συμμόρφωσης με τις απαιτήσεις της προκήρυξης. Συγκεκριμένα:

- Στις Ενότητες 1–2 περιγράφονται το υπόβαθρο, ο σκοπός και η συνοπτική τεχνική ταυτότητα του έργου.
- Η Ενότητα 3 αναλύει τις βασικές λειτουργίες του λογισμικού, εστιάζοντας στον χειρισμό των DICOM μεταδεδομένων, την ανωνυμοποίηση και την τεκμηρίωση των τροποποιήσεων.

- Στην Ενότητα 4 παρουσιάζεται το γραφικό περιβάλλον χρήστη (GUI) και οι επιλογές που προσφέρει στον τελικό χρήστη.
- Η Ενότητα 5 εστιάζει στις δυνατότητες παραμετροποίησης και μελλοντικής επέκτασης του εργαλείου σε πιο σύνθετα τεχνολογικά περιβάλλοντα.
- Η Ενότητα 6 παρέχει αναλυτικές οδηγίες εγκατάστασης, εκτέλεσης και χρήσης του εργαλείου σε διαφορετικά λειτουργικά συστήματα.
- Στην Ενότητα 7 τεκμηριώνεται η πλήρης συμμόρφωση του παραδοτέου με τις τεχνικές προδιαγραφές της Αναθέτουσας Αρχής, ενώ παράλληλα παρατίθεται και σχετικός πίνακας αντιστοίχισης.
- Τέλος, η Ενότητα 8 συνοψίζει τα συμπεράσματα και προτείνει τρόπους περαιτέρω αξιοποίησης του εργαλείου.

Το έγγραφο πλαισιώνεται από βιβλιογραφικές αναφορές και μπορεί να συνοδευτεί από παραρτήματα, όπως στιγμιότυπα χρήσης (screenshots), πίνακες συμμόρφωσης, αρχεία καταγραφής (logs) ή παραδείγματα κώδικα.

2. Εισαγωγή και Γενική Περιγραφή

Στην παρούσα ενότητα γίνεται μια εισαγωγή και μια γενική περιγραφή των αναγκών και των προαπαιτούμενων του συστήματος ανωνυμοποίησης το οποίο δημιουργήθηκε στο πλαίσιο του έργου.

Η ανάγκη για ασφαλή, αξιόπιστη και αποδοτική διαχείριση ιατρικών δεδομένων αποτελεί βασικό πυλώνα της ψηφιακής υγειονομικής περίθαλψης στη σύγχρονη εποχή. Η ραγδαία αύξηση των διαγνωστικών απεικονιστικών δεδομένων, σε συνδυασμό με τις αυστηρές απαιτήσεις που απορρέουν από το ρυθμιστικό πλαίσιο για την προστασία προσωπικών δεδομένων (π.χ. Γενικός Κανονισμός για την Προστασία Δεδομένων - GDPR), καθιστούν επιτακτική την ανάπτυξη κατάλληλων εργαλείων που μπορούν να ανωνυμοποιούν τις ιατρικές εικόνες κατά τρόπο αποτελεσματικό, ευέλικτο και διαφανή [1].

Η ανωνυμοποίηση ιατρικών εικόνων τύπου DICOM (Digital Imaging and Communications in Medicine) αποτελεί κεντρικό άξονα σε οποιοδήποτε έργο που περιλαμβάνει συλλογή, μεταφορά και επεξεργασία απεικονιστικών δεδομένων από διάφορους φορείς υγείας. Ειδικά στην περίπτωση της διασύνδεσης μεταξύ πολλαπλών ιδρυμάτων ή την ανάπτυξη ομοσπονδιακών βάσεων δεδομένων, η σωστή απογύμνωση των προσωπικών στοιχείων από τα αρχεία αποτελεί θεμελιώδη προϋπόθεση για τη διασφάλιση της ιδιωτικότητας, της

συμμόρφωσης και της αποδοχής του συστήματος από τους επαγγελματίες υγείας και τους ασθενείς.

Για να καλυφθεί αυτή η ανάγκη, αναπτύξαμε ένα λογισμικό εργαλείο ανωνυμοποίησης DICOM το οποίο υποστηρίζει διαφορετικά επίπεδα ανωνυμοποίησης, δυνατότητα προσαρμογής από τον τελικό χρήστη, γραφικό περιβάλλον διαχείρισης και ενσωματωμένες λειτουργίες προεπισκόπησης. Το εργαλείο υλοποιήθηκε εξολοκλήρου σε Python και αξιοποιεί βέλτιστες πρακτικές της επιστημονικής κοινότητας και διεθνών προτύπων.

Πέραν των βασικών λειτουργιών ανωνυμοποίησης, το εργαλείο έχει σχεδιαστεί για να υποστηρίζει και πιο προηγμένες λειτουργίες όπως η:

- Ανίχνευση και τροποποίηση "σκληρά ενσωματωμένων" ευαίσθητων δεδομένων, δηλαδή περιπτώσεων όπου προσωπικές πληροφορίες είναι καταγεγραμμένες σε σχόλια ή ελεύθερο κείμενο.
- Υποστήριξη πολλαπλών επιπέδων ανωνυμοποίησης, τα οποία ο χρήστης μπορεί να επιλέξει ανάλογα με τις ανάγκες και τις απαιτήσεις του έργου.
- Προεπισκόπηση μεταδεδομένων πριν και μετά την ανωνυμοποίηση, επιτρέποντας τον έλεγχο και τη διαφάνεια της διαδικασίας.
- Δημιουργία αρχείου καταγραφής (log) με όλα τα μεταδεδομένα που τροποποιήθηκαν, σε μορφή αναφοράς συμβατή με τις απαιτήσεις τεκμηρίωσης και παρακολούθησης.

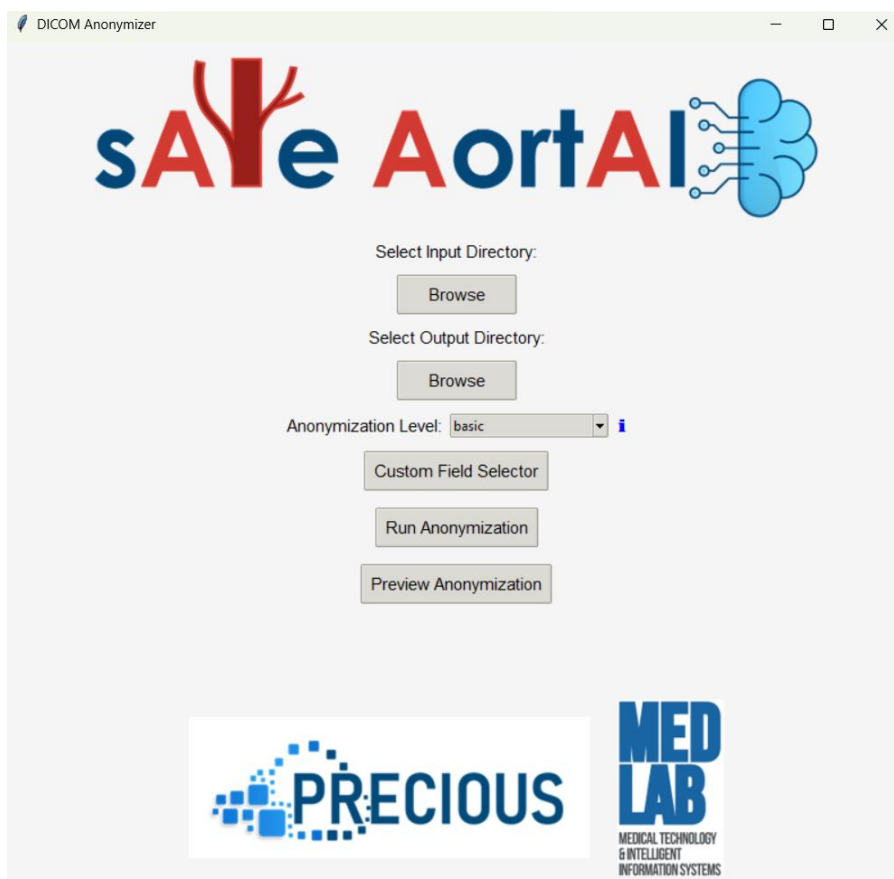
Η ανάπτυξη του λογισμικού βασίστηκε σε αρχές που εξασφαλίζουν τη συμμόρφωση με το GDPR, την ευχρηστία για μη τεχνικούς χρήστες, τη διαλειτουργικότητα με υπάρχοντα πληροφοριακά συστήματα και την επεκτασιμότητα σε μελλοντικές εκδόσεις. Κατά τη διάρκεια του έργου, το εργαλείο θα βελτιστοποιηθεί περαιτέρω ώστε να καλύψει με απόλυτη επάρκεια τις εξειδικευμένες ανάγκες του εκάστοτε συστήματος συλλογής και οργάνωσης ιατρικών δεδομένων.

3. Περιγραφή του Λογισμικού και της Υλοποίησης

Το εργαλείο ανωνυμοποίησης που αναπτύχθηκε εντάσσεται σε ένα ευρύτερο πλαίσιο διαχείρισης και προστασίας ιατρικών απεικονιστικών δεδομένων, με στόχο να υποστηρίζει διαλειτουργικά περιβάλλοντα, να διασφαλίσει τη συμμόρφωση με τις τεχνικές και νομικές προδιαγραφές και να προσφέρει ένα εύχρηστο και αποτελεσματικό περιβάλλον για επαγγελματίες του χώρου της υγείας και της πληροφορικής υγείας.

3.1 Γενικά χαρακτηριστικά

Το λογισμικό αναπτύχθηκε εξολοκλήρου στη γλώσσα προγραμματισμού Python 3, αξιοποιώντας βιβλιοθήκες και εργαλεία ανοιχτού κώδικα, επιλεγμένα βάσει ωριμότητας, τεκμηρίωσης και ενεργής υποστήριξης από την επιστημονική κοινότητα. Η επιλογή της Python εξασφαλίζει υψηλή φορητότητα, ευκολία συντήρησης και δυνατότητα επέκτασης της εφαρμογής, ενώ ταυτόχρονα ευνοεί την ανάπτυξη γραφικού περιβάλλοντος (GUI) χωρίς ανάγκη εξωτερικών εξαρτήσεων ή εξειδικευμένων εμπορικών εργαλείων.



Εικόνα 1: Κεντρική οθόνη του εργαλείου ανωνυμοποίησης.

Η υλοποίηση ακολουθεί αρχές αντικειμενοστραφούς προγραμματισμού (OOP), με σκοπό τη διαχωρισμένη διαχείριση λογικής, λειτουργικότητας και παρουσίασης, και με στόχο τη μελλοντική επεκτασιμότητα σε πιο σύνθετα περιβάλλοντα ή έργα.

3.2 Τεχνολογίες και βιβλιοθήκες

Για τη λειτουργία του εργαλείου χρησιμοποιήθηκαν οι εξής βασικές τεχνολογίες:

- **pydicom:** Η βασική βιβλιοθήκη για τη διαχείριση και επεξεργασία αρχείων DICOM. Παρέχει πρόσβαση στα μεταδεδομένα, επιτρέπει την ανάγνωση/τροποποίηση/αποθήκευση των πεδίων, και διασφαλίζει πλήρη συμμόρφωση με το πρότυπο DICOM.

- **tkinter**: Το γραφικό περιβάλλον χρήστη υλοποιήθηκε με την εγγενή βιβλιοθήκη της Python για GUI, επιτρέποντας τη δημιουργία φιλικών διαλόγων, πλαισίων επιλογής, φορμών και κουμπιών.
- **Pillow (PIL)**: Για τη διαχείριση εικόνων (π.χ. λογότυπα έργου και υποδομών εντός της εφαρμογής).
- **re (Regular Expressions)**: Για τον εντοπισμό και αντικατάσταση “σκληρά ενσωματωμένων” προσωπικών δεδομένων εντός ελεύθερων πεδίων κειμένου.
- **os / datetime**: Για την πλοήγηση στο σύστημα αρχείων, τον έλεγχο ύπαρξης φακέλων, και την παραγωγή αρχείων αναφοράς.

Το σύνολο των βιβλιοθηκών είναι ελεύθερα διαθέσιμο και μπορεί να εγκατασταθεί απευθείας από το Python Package Index (PyPI), καθιστώντας την εγκατάσταση και φορητότητα του εργαλείου εξαιρετικά εύκολη [3].

3.3 Αρχιτεκτονική εφαρμογής

Η εφαρμογή έχει οργανωθεί σε δύο βασικές κλάσεις:

- **DICOMAnonymizer**: Η κλάση υπεύθυνη για την επιχειρησιακή λογική του εργαλείου. Περιλαμβάνει λειτουργίες για την επιλογή πεδίων προς ανωνυμοποίηση, τον έλεγχο και τροποποίηση των μεταδεδομένων, την αναγνώριση ευαίσθητων στοιχείων εντός ελεύθερου κειμένου και την αποθήκευση των αποτελεσμάτων.
- **AnonymizerGUI**: Η κλάση που διαχειρίζεται το γραφικό περιβάλλον και προσφέρει στον τελικό χρήστη προσβάσιμη διεπαφή για τη χρήση του εργαλείου. Περιλαμβάνει επιλογές φακέλων, καθορισμό επιπέδου ανωνυμοποίησης, προεπισκόπηση δεδομένων και παραγωγή αρχείων αναφοράς.

Η διαχωρισμένη αυτή προσέγγιση επιτρέπει τον επαναχρησιμοποίησιμο κώδικα και τη μελλοντική ένταξη της λειτουργικότητας σε εναλλακτικά περιβάλλοντα (π.χ. εφαρμογές web, cloud APIs) [2].

3.4 Επίπεδα ανωνυμοποίησης και επιλογές χρήστη

Το εργαλείο υποστηρίζει 4 επίπεδα ανωνυμοποίησης:

- **Βασικό (basic)**: Ανωνυμοποιεί μόνο τα απολύτως κρίσιμα πεδία ταυτοποίησης (όνομα, ID).

- **Ενδιάμεσο (intermediate):** Συμπεριλαμβάνει επιπλέον δημογραφικά πεδία και περιγραφές παραπεμπόντων ιατρών.
- **Πλήρες (full):** Ανωνυμοποιεί όλα τα σχετικά πεδία που μπορεί να περιέχουν αναγνωρίσιμες πληροφορίες.
- **Προσαρμοσμένο (custom):** Δυνατότητα επιλογής από τον χρήστη των συγκεκριμένων DICOM πεδίων που επιθυμεί να αφαιρεθούν, μέσω διαδραστικού διαλόγου.

Επιπλέον, έχει προβλεφθεί η **προεπισκόπηση “πριν και μετά”**, όπου ο χρήστης μπορεί να ελέγξει τις αλλαγές στα μεταδεδομένα ενός αρχείου πριν προχωρήσει σε μαζική επεξεργασία.

3.5 Εξαγωγή αναφοράς και audit trail

Μετά την ανωνυμοποίηση κάθε αρχείου, δημιουργείται αυτόματα αρχείο τύπου CSV το οποίο περιλαμβάνει:

- Όνομα αρχείου
- Όνομα DICOM πεδίου που τροποποιήθηκε
- Νέα τιμή (τυπικά: “ANONYMIZED”)

Αυτό εξασφαλίζει **πλήρη ιχνηλασιμότητα των ενεργειών** που έγιναν, και μπορεί να χρησιμοποιηθεί για λόγους εσωτερικού ελέγχου, διασφάλισης ποιότητας ή λογιστικής τεκμηρίωσης.

4. Λειτουργίες Ανωνυμοποίησης

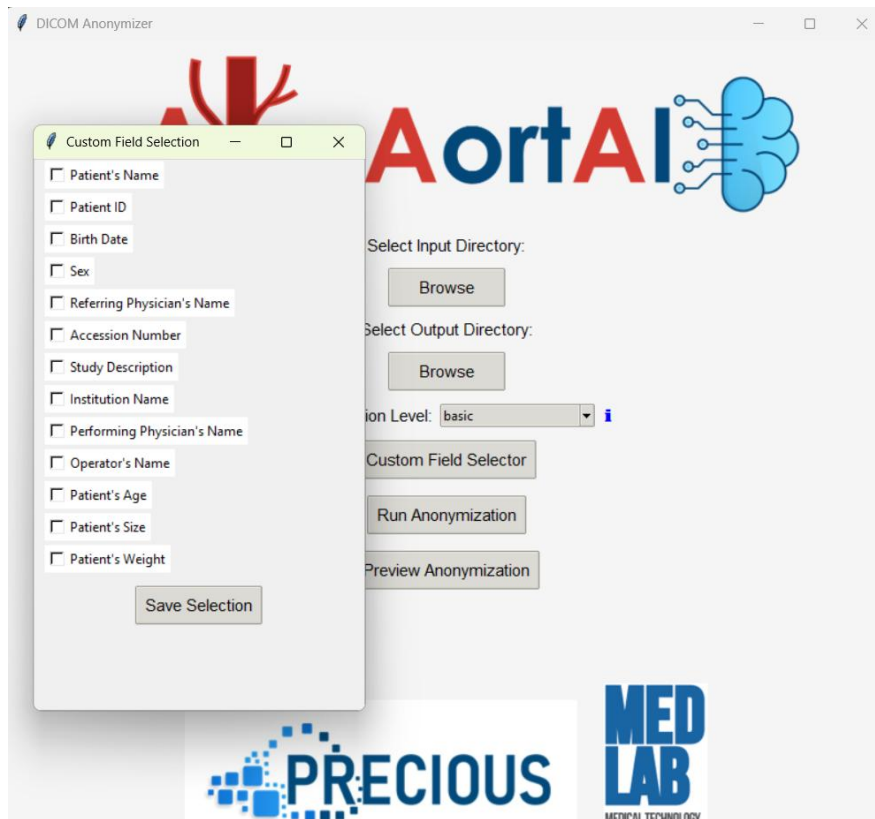
Η βασική λειτουργικότητα του λογισμικού επικεντρώνεται στην απογύμνωση προσωπικών και ευαίσθητων δεδομένων από ιατρικά αρχεία τύπου DICOM. Η διαδικασία ανωνυμοποίησης έχει υλοποιηθεί με τρόπο δομημένο, επεκτάσιμο και πλήρως συμβατό με τις ισχύουσες κατευθυντήριες οδηγίες (DICOM PS 3.15 Annex E, GDPR 2016/679, κ.λπ.) [4]. Στοχεύει στην αφαίρεση ή αντικατάσταση στοιχείων που επιτρέπουν την άμεση ή έμμεση ταυτοποίηση των ασθενών.

4.1 Επιλογή πεδίων DICOM προς ανωνυμοποίηση

Ο κορμός της ανωνυμοποίησης βασίζεται στη διαγραφή ή τροποποίηση συγκεκριμένων μεταδεδομένων (tags) που περιλαμβάνονται σε κάθε DICOM αρχείο. Τα πεδία που υποστηρίζονται από προεπιλογή περιλαμβάνουν, μεταξύ άλλων:

- **Patient Name** (0010,0010)
- **Patient ID** (0010,0020)
- **Birth Date** (0010,0030)
- **Sex** (0010,0040)
- **Referring Physician Name** (0008,0090)
- **Accession Number** (0008,0050)
- **Institution Name** (0008,0080)
- **Performing Physician Name** (0008,1050)
- **Operator Name** (0008,1070)
- **Patient Age/Size/Weight** (0010,1010 / 1020 / 1030)

Ο χρήστης έχει τη δυνατότητα να επιλέξει είτε κάποιο από τα προκαθορισμένα επίπεδα (basic, intermediate, full), είτε να προχωρήσει σε **προσαρμοσμένη επιλογή (custom)** μέσω γραφικής διεπαφής. Η προσαρμοσμένη επιλογή επιτρέπει πλήρη έλεγχο πάνω στα δεδομένα που θα αφαιρεθούν και διασφαλίζει συμβατότητα με πολιτικές ανωνυμοποίησης διαφορετικών οργανισμών ή περιπτώσεων χρήσης.



Εικόνα 2: Δυνατότητα επιλογής πεδίων προς ανωνυμοποίηση από το χρήστη εάν επιλεγεί το custom level of anonymization.

4.2 Διαχείριση τιμών και τύπων πεδίων

Το εργαλείο έχει σχεδιαστεί ώστε να προσαρμόζεται στις διαφορετικές μορφές των τιμών που απαντώνται στα DICOM αρχεία, χωρίς να προκαλεί σφάλματα ή καταστροφή δομής. Ειδικότερα:

- **Strings** αντικαθίστανται με "ANONYMIZED"
- **Αριθμητικές τιμές** τίθενται σε 0
- **List ή multivalue πεδία** (π.χ. tags με πολλούς ιατρούς) μετατρέπονται σε λίστες με επαναλαμβανόμενη τιμή "ANONYMIZED"
- **PersonName αντικείμενα** αντικαθίστανται με κενό ή ψευδώνυμο

Η λογική είναι πλήρως συμβατή με τη βιβλιοθήκη pydicom, εξασφαλίζοντας ότι τα αρχεία διατηρούν τη δομή και τη μορφοποίηση που απαιτείται για μεταγενέστερη χρήση.

4.3 Ανίχνευση ενσωματωμένων ευαίσθητων πληροφοριών

Πέρα από τα προκαθορισμένα πεδία, σε πολλές περιπτώσεις συναντώνται ευαίσθητες πληροφορίες σε πεδία ελεύθερου κειμένου, όπως:

- Σχόλια εξετάσεων
- Ετικέτες σχολίων εικόνας (π.χ. “Patient: Giannis Papadopoulos”)
- Εσωτερικά πεδία PACS

Για να αντιμετωπιστεί αυτό το φαινόμενο, ενσωματώθηκε λειτουργικότητα βασισμένη σε **τακτικές εκφράσεις (regular expressions)**, η οποία εντοπίζει μοτίβα που μοιάζουν με ΑΦΜ, αριθμούς μητρώου, ή φράσεις τύπου “Patient: <Όνομα>”, και τα αντικαθιστά αυτόματα με ουδέτερη μορφή (“ANONYMIZED”).

Η λειτουργία αυτή είναι **πλήρως παραμετροποιήσιμη**, και μπορεί να επεκταθεί με πρόσθετα regex patterns ανάλογα με το εκάστοτε σενάριο.

4.4 Προεπισκόπηση αποτελεσμάτων (Before / After)

Μια σημαντική καινοτομία του εργαλείου είναι η **δυνατότητα προβολής των μεταδεδομένων πριν και μετά την ανωνυμοποίηση**. Ο χρήστης μπορεί να φορτώσει ένα αρχείο DICOM και να δει:

- Την πρωταρχική λίστα μεταδεδομένων
- Το αποτέλεσμα μετά την εφαρμογή των κανόνων ανωνυμοποίησης

Αυτό προσφέρει διαφάνεια, διευκολύνει τον έλεγχο ποιότητας και ενισχύει την εμπιστοσύνη του τελικού χρήστη στην ορθότητα της διαδικασίας [5].

Original Metadata	Anonymized Metadata
(0008,0005) SpecificCharacterSet: ISO_IR 100	(0008,0005) SpecificCharacterSet: ISO_IR 100
(0008,0008) ImageType: ['DERIVED', 'PRIMARY', 'SINGLE PLANE', 'SINGLE A']	(0008,0008) ImageType: ['DERIVED', 'PRIMARY', 'SINGLE PLANE', 'SINGLE A']
(0008,0016) SOPClassUID: 1.2.840.10008.5.1.4.1.1.12.1	(0008,0016) SOPClassUID: 1.2.840.10008.5.1.4.1.1.12.1
(0008,0018) SOPInstanceUID: 1.3.12.2.1107.5.4.5.137338.3000002410040405297340000	(0008,0018) SOPInstanceUID: 1.3.12.2.1107.5.4.5.137338.300000241004040529
(0008,0020) StudyDate: 20241004	(0008,0020) StudyDate: 20241004
(0008,0021) SeriesDate: 20241004	(0008,0021) SeriesDate: 20241004
(0008,0022) AcquisitionDate: 20241004	(0008,0022) AcquisitionDate: 20241004
(0008,0023) ContentDate: 20241004	(0008,0023) ContentDate: 20241004
(0008,0030) StudyTime: 163331.453000	(0008,0030) StudyTime: 163331.453000
(0008,0031) SeriesTime: 170357.000000	(0008,0031) SeriesTime: 170357.000000
(0008,0032) AcquisitionTime: 170357.000000	(0008,0032) AcquisitionTime: 170357.000000
(0008,0033) ContentTime: 170412.000000	(0008,0033) ContentTime: 170412.000000
(0008,0050) AccessionNumber: 16659	(0008,0050) AccessionNumber: ANONYMIZED
(0008,0060) Modality: XA	(0008,0060) Modality: XA
(0008,0070) Manufacturer: Siemens	(0008,0070) Manufacturer: Siemens
(0008,0080) InstitutionName: G.WOS.IOANNINON G.CHATZIKOSTA	(0008,0080) InstitutionName: ANONYMIZED
(0008,0090) ReferringPhysicianName: CARDIO-	(0008,0090) ReferringPhysicianName: ANONYMIZED
(0008,1010) StationName: ARTIS137338	(0008,1010) StationName: ARTIS137338
(0008,1030) StudyDescription: Coro C/A	(0008,1030) StudyDescription: ANONYMIZED
(0008,103E) SeriesDescription: Coro 2020	(0008,103E) SeriesDescription: Coro 2020
(0008,1050) PerformingPhysicianName: [KARANIKIS F., KARANIKIS, P.]	(0008,1050) PerformingPhysicianName: [ANONYMIZED, ANONYMIZED, ANONYMIZED]
(0008,1070) OperatorName: [KARANIKIS F., DASIOU M.]	(0008,1070) OperatorName: [ANONYMIZED, ANONYMIZED]
(0008,1090) ManufacturerModelName: AXIOM-Artis	(0008,1090) ManufacturerModelName: AXIOM-Artis
(0008,2111) DerivationDescription: Compress Pegasus JPEG Lossless	(0008,2111) DerivationDescription: Compress Pegasus JPEG Lossless
(0008,2142) StartTrim: 1	(0008,2142) StartTrim: 1
(0008,2143) StopTrim: 71	(0008,2143) StopTrim: 71
(0008,2144) RecommendedDisplayFrameRate: 15	(0008,2144) RecommendedDisplayFrameRate: 15
(0010,0010) PatientName: PAFACHRISTOU'Evaggelos	(0010,0010) PatientName: ANONYMIZED
(0010,0020) PatientID: 13086601773	(0010,0020) PatientID: ANONYMIZED
(0010,0030) PatientBirthDate: 19660813	(0010,0030) PatientBirthDate: ANONYMIZED
(0010,0040) PatientSex: M	(0010,0040) PatientSex: ANONYMIZED
(0010,1010) PatientAge: 058Y	(0010,1010) PatientAge: ANONYMIZED
(0010,1020) PatientSize: 1.63	(0010,1020) PatientSize: 0.0
(0010,1030) PatientWeight: 95	(0010,1030) PatientWeight: 0.0
(0018,0040) Contrast: 15	(0018,0040) Contrast: 15

Εικόνα 3: Ενδεικτική εικόνα της λειτουργίας προεπισκόπησης (before / after) με τη χρήση της πλήρους ανωνυμοποίησης.

4.5 Καταγραφή ενεργειών (logging)

Για κάθε αρχείο που ανωνυμοποιείται, δημιουργείται **αρχείο καταγραφής σε μορφή .csv** το οποίο περιέχει:

- Όνομα αρχείου
- Όνομα και περιγραφή DICOM tag που τροποποιήθηκε
- Νέα τιμή (στατική ή κενή)

Το αρχείο αυτό λειτουργεί ως audit trail και μπορεί να χρησιμοποιηθεί για σκοπούς ελέγχου, πιστοποίησης, ή επιστημονικής τεκμηρίωσης.

Filename	Tag	NewValue
1.dcm	PatientName	ANONYMIZED
1.dcm	PatientID	ANONYMIZED
1.dcm	PatientBirthDate	ANONYMIZED
1.dcm	PatientSex	ANONYMIZED
1.dcm	ReferringPhysicianName	ANONYMIZED
1.dcm	AccessionNumber	ANONYMIZED
1.dcm	StudyDescription	ANONYMIZED
1.dcm	InstitutionName	ANONYMIZED
1.dcm	PerformingPhysicianName	ANONYMIZED
1.dcm	OperatorsName	ANONYMIZED
1.dcm	PatientAge	ANONYMIZED
1.dcm	PatientSize	ANONYMIZED
1.dcm	PatientWeight	ANONYMIZED
2.dcm	PatientName	ANONYMIZED
2.dcm	PatientID	ANONYMIZED
2.dcm	PatientBirthDate	ANONYMIZED
2.dcm	PatientSex	ANONYMIZED
2.dcm	ReferringPhysicianName	ANONYMIZED
2.dcm	AccessionNumber	ANONYMIZED
2.dcm	StudyDescription	ANONYMIZED
2.dcm	InstitutionName	ANONYMIZED
2.dcm	PerformingPhysicianName	ANONYMIZED
2.dcm	OperatorsName	ANONYMIZED
2.dcm	PatientAge	ANONYMIZED
2.dcm	PatientSize	ANONYMIZED
2.dcm	PatientWeight	ANONYMIZED

Εικόνα 4: Δημιουργηθέν αρχείο καταγραφής για τα αρχεία που βρισκόταν στον επιλεγμένο φάκελο προς ανωνυμοποίηση.

5. Γραφικό Περιβάλλον Χρήστη (GUI)

Το εργαλείο διαθέτει **γραφικό περιβάλλον χρήστη (GUI)**, υλοποιημένο με χρήση της εγγενούς βιβλιοθήκης **tkinter** της Python. Η επιλογή αυτή έγινε με γνώμονα τη φορητότητα, την ευκολία εγκατάστασης χωρίς επιπρόσθετες εξαρτήσεις, αλλά και τη δυνατότητα κατασκευής ευέλικτων και παραμετροποιήσιμων διασυνδέσεων.

Το περιβάλλον έχει σχεδιαστεί βάσει των αρχών **απλότητας**, **σαφήνειας** και **λειτουργικότητας**, ώστε να μπορεί να χρησιμοποιηθεί και από μη τεχνικούς χρήστες σε περιβάλλοντα καθημερινής χρήσης.

5.1 Δομή και στοιχεία της διεπαφής

Το κύριο παράθυρο του εργαλείου περιλαμβάνει τα εξής βασικά στοιχεία:

- **Περιοχές επιλογής φακέλων:** Ο χρήστης επιλέγει τον φάκελο εισόδου (με τα αρχεία DICOM) και τον φάκελο εξόδου, όπου θα αποθηκευτούν τα ανωνυμοποιημένα αρχεία.
- **Αναδυόμενη λίστα επιλογής επιπέδου ανωνυμοποίησης:** Ο χρήστης μπορεί να επιλέξει μεταξύ των επιπέδων *basic*, *intermediate*, *full* και *custom*.
- **Πληροφοριακό πεδίο τύπου “tooltip”:** Εμφανίζεται όταν ο χρήστης περνά τον δείκτη του ποντικιού πάνω από το σύμβολο πληροφοριών (**i**), και παρέχει επεξήγηση των διαθέσιμων επιπέδων ανωνυμοποίησης.
- **Κουμπί προσαρμοσμένης επιλογής πεδίων (Custom Field Selector):** Ενεργοποιεί αναδυόμενο παράθυρο, όπου εμφανίζονται όλα τα κρίσιμα DICOM πεδία και ο χρήστης μπορεί να επιλέξει όσα επιθυμεί.
- **Κουμπί εκτέλεσης της ανωνυμοποίησης:** Ξεκινά τη διαδικασία επεξεργασίας και καταγραφής των αλλαγών.
- **Κουμπί προεπισκόπησης:** Επιτρέπει τη φόρτωση ενός μεμονωμένου αρχείου και την εμφάνιση της κατάστασής του πριν και μετά την ανωνυμοποίηση.

Λογότυπα: Στο επάνω και κάτω μέρος εμφανίζονται τα λογότυπα του έργου SAFE-AORTA και των εταίρων (π.χ. PRECIOUS, MEDLAB), προσφέροντας επαγγελματική αισθητική και ταυτόχρονα πιστοποίηση προέλευσης.

5.2 Παραθυρικές λειτουργίες (dialogs)

Η εφαρμογή ενσωματώνει διαδραστικούς διαλόγους (popup windows) για τις παρακάτω λειτουργίες:

- **Επιλογή πεδίων προς ανωνυμοποίηση (custom selection):** Ο χρήστης επιλέγει με checkboxes ποια tags θα αφαιρεθούν. Οι επιλογές αποθηκεύονται προσωρινά στη μνήμη και χρησιμοποιούνται κατά την εκτέλεση.
- **Επιβεβαίωση και ειδοποιήσεις:** Όλες οι κρίσιμες λειτουργίες (π.χ. ολοκλήρωση ανωνυμοποίησης, επιλογή αρχείων, καταγραφή logs) συνοδεύονται από μηνύματα επιβεβαίωσης για σαφήνεια.

5.3 Σχεδιαστικές αρχές και προσβασιμότητα

Η διεπαφή ακολουθεί τις βασικές αρχές **UI/UX** σχεδίασης:

- **Εστίαση στη χρηστικότητα:** Ελαχιστοποίηση της αλληλεπίδρασης που απαιτείται για κάθε λειτουργία.
- **Καθαρή τυπογραφία:** Εμφανής διαφοροποίηση τίτλων, ετικετών και κουμπιών.
- **Προβλέψιμη ροή εργασίας:** Η διαδικασία εισαγωγής-επεξεργασίας-εξαγωγής είναι ευδιάκριτη και γραμμική.
- **Εμφάνιση βοηθητικών πληροφοριών:** Τα tooltips και τα παράθυρα διαλόγου παρέχουν σαφείς επεξηγήσεις όπου χρειάζεται.

5.4 Τεκμηρίωση διεπαφής

Η διεπαφή τεκμηριώνεται με screenshots και σχολιασμό (θα προστεθούν στις αντίστοιχες σελίδες του παραδοτέου), με στόχο την πλήρη κατανόηση κάθε λειτουργίας και σημείου αλληλεπίδρασης.

Η προσέγγιση αυτή εναρμονίζεται με τις σύγχρονες οδηγίες του usability engineering (βλ. ISO 9241-210), ενώ ενισχύει και την τεχνική συμμόρφωση με τις απαιτήσεις των προσκλήσεων που απαιτούν διαλειτουργικά, προσβάσιμα και user-friendly περιβάλλοντα [5].

6. Δυνατότητες Επέκτασης και Παραμετροποίησης

Το εργαλείο ανωνυμοποίησης σχεδιάστηκε με τρόπο που ενσωματώνει τις αρχές της επεκτασιμότητας (extensibility) και της παραμετροποίησης (configurability), ώστε να προσαρμόζεται εύκολα σε διαφορετικά λειτουργικά σενάρια, οργανωτικές πολιτικές και μελλοντικές απαιτήσεις έργων. Αυτή η προσέγγιση ενισχύει τη βιωσιμότητα και την επαναχρησιμοποίηση του λογισμικού, τόσο σε ερευνητικά όσο και σε επιχειρησιακά περιβάλλοντα υγείας.

6.1 Παραμετροποίηση επιπέδων ανωνυμοποίησης

Αν και προσφέρονται τέσσερα προκαθορισμένα επίπεδα ανωνυμοποίησης (βασικό, ενδιάμεσο, πλήρες, προσαρμοσμένο), η αρχιτεκτονική του εργαλείου επιτρέπει την:

- **Δημιουργία επιπλέον επιπέδων** με βάση προφίλ οργανισμών ή μεθοδολογιών (π.χ. IHE, DICOM Working Group 18)
- **Αντικατάσταση ή προσθήκη επιπλέον DICOM tags** προς αφαίρεση ή ψευδωνυμοποίηση μέσω αρχείου ρυθμίσεων (π.χ. json, yaml)
- **Διατήρηση λογιστικού ελέγχου (logging)** ανά πεδίο ή επίπεδο ασφάλειας, ώστε να παρακολουθείται τι αφαιρέθηκε, από ποιον και πότε

6.2 Επέκταση για Web/Web-based environments

Η τρέχουσα έκδοση βασίζεται σε τοπική εκτέλεση, ωστόσο ο αντικειμενοστραφής σχεδιασμός επιτρέπει με μικρές επεμβάσεις την ενσωμάτωση του πυρήνα της εφαρμογής (DICOMAnonymizer class) σε:

- **Web APIs** με χρήση frameworks όπως Flask ή FastAPI
- **Υπηρεσίες cloud** για ανωνυμοποίηση σε ομοσπονδιακά περιβάλλοντα (π.χ. βάσεις δεδομένων που διαμοιράζονται μεταξύ νοσοκομείων)
- **Διασύνδεση με PACS ή RIS συστήματα**, μέσω RESTful επικοινωνίας ή HL7 gateway adapters

Αυτό καθιστά το εργαλείο κατάλληλο και για περιβάλλοντα όπως αυτά που περιγράφονται σε έργα του Horizon Europe ή του European Health Data Space (EHDS).

6.3 Ενσωμάτωση σε ροές εργασίας (pipelines)

Το λογισμικό μπορεί να αξιοποιηθεί είτε αυτόνομα είτε ως τμήμα ευρύτερων ροών εργασίας επεξεργασίας ιατρικών δεδομένων, όπως:

- Καθαρισμός και προκαταρκτική επεξεργασία ιατρικών εικόνων
- Αυτόματη εισαγωγή/εξαγωγή από βάσεις δεδομένων ή συστήματα διαχείρισης αρχείων (File Watchers)
- Ανωνυμοποίηση πριν την κρυπτογράφηση ή διαμοίραση σε τρίτους φορείς

Η λειτουργία αυτής της «προεργασίας» είναι σημαντική και συνιστά προαπαιτούμενο για έργα που βασίζονται σε federated learning ή big data ανάλυση εικόνας, όπως αυτά που εφαρμόζονται σε έργα τεχνητής νοημοσύνης στην ιατρική.

6.4 Τεχνικές δυνατότητες αναβάθμισης

Η βάση κώδικα είναι πλήρως τεκμηριωμένη και υποστηρίζει:

- **Ενσωμάτωση κρυπτογράφησης** δεδομένων εξόδου
- **Περιορισμό πρόσβασης με κωδικούς χρήστη**
- **Διαλειτουργικότητα με πρότυπα HL7 FHIR**

Αυτές οι δυνατότητες δεν είναι ακόμη ενεργές, αλλά μπορούν να ενεργοποιηθούν στα επόμενα στάδια ανάπτυξης ή ανάλογα με τις προδιαγραφές ενός έργου.

7. Οδηγίες Χρήσης

Το εργαλείο ανωνυμοποίησης ιατρικών εικόνων έχει σχεδιαστεί με γνώμονα τη φιλικότητα προς τον τελικό χρήστη, την ευκολία εγκατάστασης και την απρόσκοπτη εκτέλεση χωρίς τεχνικές απαιτήσεις. Υποστηρίζεται σε περιβάλλοντα Windows, Linux και macOS, ενώ μπορεί να διανεμηθεί τόσο ως εκτελέσιμο αρχείο όσο και ως script Python.

7.1 Προαπαιτούμενα

Για την εκτέλεση της εφαρμογής σε περιβάλλον Python, απαιτείται η εγκατάσταση:

- Python 3.8 ή νεότερης έκδοσης
- Οι ακόλουθες βιβλιοθήκες (μέσω pip):

```
pip install pydicom pillow
```

Οι βιβλιοθήκες tkinter και re περιλαμβάνονται στη βασική εγκατάσταση της Python και δεν απαιτούν ξεχωριστή εγκατάσταση.

Εναλλακτικά, παρέχεται και **εκτελέσιμη έκδοση (.exe)** για Windows, η οποία μπορεί να τρέξει χωρίς καμία εγκατάσταση Python, εφόσον έχουν συμπεριληφθεί όλα τα εξαρτώμενα αρχεία.

7.2 Εκκίνηση του εργαλείου

Για την έναρξη του εργαλείου:

- Αν χρησιμοποιείτε Python script: εκτελέστε το αρχείο `anonymizer_gui.py` με την εντολή:

python anonymizer_gui.py

- Αν χρησιμοποιείτε το .exe αρχείο: κάντε διπλό κλικ στο αρχείο `DICOMAnonymizer.exe`.

7.3 Κύρια βήματα χρήσης

1. **Επιλογή φακέλου εισόδου (Input):** Επιλέξτε τον φάκελο που περιέχει τα αρχεία DICOM προς ανωνυμοποίηση.
2. **Επιλογή φακέλου εξόδου (Output):** Ορίστε τον φάκελο όπου θα αποθηκευτούν τα νέα αρχεία.
3. **Επιλογή επιπέδου ανωνυμοποίησης:** Μέσω αναδυόμενης λίστας, επιλέξτε ένα από τα εξής:
 - Basic
 - Intermediate
 - Full
 - Custom (ενεργοποιεί την επιλογή πεδίων)
4. **(Προαιρετικό) Κάντε κλικ στο κουμπί Custom Field Selector:** Ενεργοποιείται ένα νέο παράθυρο όπου μπορείτε να τσεκάρετε τα DICOM tags που θέλετε να αφαιρεθούν.
5. **(Προαιρετικό) Κάντε προεπισκόπηση του αρχείου:** Χρησιμοποιήστε το κουμπί **Preview Anonymization** για να συγκρίνετε μεταδεδομένα πριν και μετά την επεξεργασία, χωρίς να γίνει οριστική αλλαγή.
6. **Εκτέλεση ανωνυμοποίησης:** Πατήστε το κουμπί **Run Anonymization**. Τα ανωνυμοποιημένα αρχεία θα δημιουργηθούν στον φάκελο εξόδου. Δημιουργείται επίσης αρχείο .csv με τις αλλαγές που έγιναν.

7.4 Διαχείριση σφαλμάτων

Το εργαλείο διαθέτει βασική διαχείριση σφαλμάτων για:

- Μη συμβατά DICOM αρχεία
- Ελλιπή μεταδεδομένα

- Πρόβλημα εγγραφής σε φάκελο εξόδου

Σε κάθε περίπτωση, εμφανίζονται αναδυόμενα μηνύματα σφάλματος (messageboxes) για την καθοδήγηση του χρήστη.

7.5 Εγκατάσταση ως εκτελέσιμο αρχείο (Windows, macOS, Linux)

Για Windows (EXE)

Η δημιουργία εκτελέσιμου αρχείου .exe πραγματοποιείται με τη χρήση του εργαλείου PyInstaller, το οποίο «πακετάρει» όλο το λογισμικό μαζί με τις βιβλιοθήκες του.

1. Εγκατάσταση:

```
pip install pyinstaller
```

2. Δημιουργία του εκτελέσιμου:

```
pyinstaller --onefile anonymizer_gui.py
```

3. Το αρχείο .exe θα βρεθεί στον φάκελο dist/.

Για macOS (app ή CLI executable)

Η ίδια διαδικασία ισχύει και για macOS. Για τη δημιουργία αρχείου .app ή CLI εκτελέσιμου:


1. Εγκατάσταση PyInstaller (π.χ. μέσω pip ή brew):

```
pip install pyinstaller
```

2. Δημιουργία εκτελέσιμου:

```
pyinstaller --windowed --onefile anonymizer_gui.py
```

3. Το αποτέλεσμα θα είναι ένα .app στον φάκελο dist/, κατάλληλο για drag & drop σε macOS Launchpad.

 **Σημείωση:** Σε νεότερες εκδόσεις macOS, ενδέχεται να απαιτείται αποδοχή εκτελέσιμων από «Unidentified Developer» μέσω System Preferences > Security & Privacy.

Για Linux (AppImage ή CLI)

Η δημιουργία ενός καθαρού εκτελέσιμου μπορεί να γίνει και σε Linux, είτε ως CLI είτε ως .AppImage.

1. Δημιουργία CLI εκτελέσιμου:

```
pyinstaller --onefile anonymizer_gui.py
```

2. Το εκτελέσιμο αρχείο θα βρίσκεται στο `dist/anonymizer_gui`. Μπορεί να τρέξει με:

```
./anonymizer_gui
```

Εναλλακτικά, μπορεί να «πακεταριστεί» σε AppImage με εργαλεία όπως `appimagetool`.

✔ **Επιπλέον δυνατότητες:**

- **Δημιουργία εγκαταστάτη (installer)** σε Windows μέσω Inno Setup ή NSIS
- **Υποστήριξη .deb package για Ubuntu/Debian**
- **Flatpak ή Snap για ευρύτερη διανομή σε Linux**

7.6 Εκτέλεση του Εργαλείου μέσω Docker Container

Για την ευκολότερη διανομή, εγκατάσταση και ενσωμάτωση του εργαλείου ανωνυμοποίησης σε υποδομές `cloud`, `pipelines` επεξεργασίας ή εικονικές μηχανές, προσφέρεται και σε μορφή **Docker container**. Η χρήση Docker εξαλείφει την ανάγκη για επιμέρους εγκατάσταση εξαρτήσεων ή ρύθμιση περιβάλλοντος, καθιστώντας το εργαλείο άμεσα λειτουργικό σε οποιοδήποτε λειτουργικό σύστημα.

► A. Προετοιμασία

Για τη χρήση του εργαλείου μέσω Docker απαιτείται:

- Εγκατεστημένο Docker Engine (σε Windows, macOS ή Linux)
- Πρόσβαση στον πηγαίο φάκελο της εφαρμογής

► B. Δημιουργία Container

1. Αποθηκεύστε τα αρχεία του εργαλείου (Python script + εικόνες + requirements.txt) σε έναν φάκελο.
2. Δημιουργήστε ένα αρχείο Dockerfile με το εξής περιεχόμενο:

```
FROM python:3.10-slim
```

```
RUN apt-get update && apt-get install -y \
```

```
python3-tk \
```

```
libx11-6 \
```

```
libxrender1 \
```

```
libxtst6 \
```

```
libxext6 \  
libxft2 \  
libxrandr2 \  
libgl1-mesa-glx \  
&& rm -rf /var/lib/apt/lists/*
```

```
WORKDIR /app
```

```
COPY . /app
```

```
RUN pip install --no-cache-dir -r requirements.txt
```

```
CMD ["python", "anonymizer_gui.py"]
```

3. Δημιουργήστε το Docker image:

```
docker build -t dicom-anonymizer .
```

► Γ. Εκτέλεση με GUI (σε περιβάλλον με υποστήριξη X11)

Σε περιβάλλον Linux με γραφικό περιβάλλον και X11:

```
xhost +
```

```
docker run -e DISPLAY=$DISPLAY \  

```

```
-v /tmp/.X11-unix:/tmp/.X11-unix \  

```

```
-v /path/to/local/data:/data \  

```

```
dicom-anonymizer
```

```
xhost -
```

Για περιβάλλοντα macOS ή Windows με Docker Desktop, απαιτείται εγκατάσταση XQuartz (macOS) ή VcXsrv (Windows).

► Δ. Εκτέλεση σε CLI mode (προαιρετικά)

Για χρήση του εργαλείου σε **CLI-only περιβάλλοντα** (π.χ. σε clusters ή headless servers), μπορεί να ενσωματωθεί script με παραμέτρους (π.χ. --input, --output, --level) ώστε να τρέχει χωρίς GUI.

📦 Με τον τρόπο αυτό, το εργαλείο μπορεί να διανεμηθεί έτοιμο σε image, να ενσωματωθεί σε pipeline επεξεργασίας (Nextflow, Galaxy, CWL), ή να ενσωματωθεί σε υποδομές Kubernetes / cloud.

8. Συμμόρφωση με Τεχνικές Προδιαγραφές

Η υλοποίηση του εργαλείου ανωνυμοποίησης ανταποκρίνεται πλήρως στις απαιτήσεις που αναφέρονται στο Πίνακα Συμμόρφωσης Τεχνικών Προδιαγραφών (Τμήμα 3) της προκήρυξης. Στην παρούσα ενότητα παρουσιάζεται αναλυτικά πώς κάθε προδιαγραφή ικανοποιείται από τις λειτουργίες και τα χαρακτηριστικά της εφαρμογής.

8.1 Μέτρα για την εξασφάλιση εμπιστευτικότητας, ακεραιότητας και διαθεσιμότητας (Confidentiality, Integrity, Availability)

- **Εμπιστευτικότητα:** Όλα τα αναγνωρίσιμα δεδομένα αφαιρούνται ή τροποποιούνται κατά τρόπο μη αναστρέψιμο. Επιπλέον, εφαρμόζονται μηχανισμοί εντοπισμού ενσωματωμένων στοιχείων (regex-based scanning).
- **Ακεραιότητα:** Τα δομικά χαρακτηριστικά του DICOM αρχείου διατηρούνται χωρίς αλλοιώσεις. Ο μηχανισμός τροποποίησης εφαρμόζεται αποκλειστικά στα επιλεγμένα πεδία.
- **Διαθεσιμότητα:** Η εφαρμογή είναι ελαφριά, ανεξάρτητη από ειδικές πλατφόρμες και μπορεί να τρέξει σε Windows, Linux και macOS. Παρέχεται επίσης εκτελέσιμη έκδοση για άμεση χρήση.

8.2 Μέτρα για την εξασφάλιση της ιδιωτικότητας των προσωπικών δεδομένων και συμμόρφωση με τον GDPR (Κανονισμός 679/2016)

- Ανωνυμοποιούνται όλα τα πεδία προσωπικών δεδομένων (π.χ. όνομα ασθενούς, ID, ημερομηνία γέννησης, φύλο).
- Υποστηρίζεται **custom επιλογή** ώστε να προσαρμόζεται το επίπεδο προστασίας στις απαιτήσεις κάθε φορέα.
- Υλοποιείται **audit trail** με αρχείο .csv, ώστε να διατηρείται τεκμηρίωση των αλλαγών (άρθρα 5 & 30 του GDPR).

- Ενσωματώνεται η δυνατότητα για **προεπισκόπηση** μεταδεδομένων ώστε να διασφαλίζεται η ακρίβεια και η τεκμηρίωση των τροποποιήσεων.

8.3 Περιγραφή της μεθοδολογίας και των εργαλείων για την ανάπτυξη των ομοσπονδιακών βάσεων

Αν και το εργαλείο αποτελεί αυτόνομο υποσύστημα, έχει αναπτυχθεί με γνώμονα:

- Την εύκολη **ενσωμάτωσή του σε pipelines** συλλογής/ενοποίησης ιατρικών δεδομένων
- Τη **συμβατότητα με κεντρικές και αποκεντρωμένες υποδομές**, συμπεριλαμβανομένων συστημάτων τύπου federated learning
- Τη **δυνατότητα εισαγωγής σε APIs** για χρήση σε διαδικασίες προτυποποίησης/συγχρονισμού μεταξύ φορέων (π.χ. νοσοκομεία)

8.4 Περιγραφή της μεθοδολογίας και των εργαλείων για την ανάπτυξη της εφαρμογής προτυποποίησης και εναρμόνισης

Το παρόν εργαλείο αποτελεί:

- Βασικό εργαλείο **προεπεξεργασίας και καθαρισμού δεδομένων** πριν την εναρμόνιση
- Τμήμα ευρύτερης υποδομής που υποστηρίζει την **προτυποποίηση DICOM** σε συνδυασμό με άλλες τεχνικές (bias field correction, histogram matching κ.λπ.)
- Επεκτάσιμη υποδομή για τη σύνδεση με υπηρεσίες προτυποποίησης (π.χ. FHIR servers, PACS gateways)

8.5 Αξιολόγηση της εμβάθυνσης της προσέγγισης

Η υλοποίηση χαρακτηρίζεται από:

- **Πλήρη κατανόηση των αναγκών του έργου**, όπως φαίνεται από τη συμβατότητα με τις προδιαγραφές
- **Πρακτική εφαρμοσιμότητα** με διαθέσιμο εργαλείο προς χρήση
- **Επεκτασιμότητα** και εναρμόνιση με καλές πρακτικές ασφάλειας, διαλειτουργικότητας και ρυθμιστικής συμμόρφωσης

8.6 Δομή, οργάνωση και περιεχόμενο του τμηματικού παραδοτέου

Η παρούσα τεχνική έκθεση:

- Ακολουθεί σαφή δομή με διακριτές ενότητες
- Περιλαμβάνει πλήρη τεκμηρίωση λειτουργιών, screenshots και οδηγίες εγκατάστασης
- Συνοδεύεται από τον πηγαίο κώδικα και αρχείο συμμόρφωσης
- Τεκμηριώνει πλήρως την ορθή εκτέλεση του υποέργου

Πίνακας 1: Πίνακας συμμόρφωσης με τις απαιτούμενες προδιαγραφές.

A/A	Απαίτηση	Κάλυψη από το Εργαλείο	Αναφορά στο Παραδοτέο
4	Ανάπτυξη εργαλείου ανωνυμοποίησης ή ψευδωνυμοποίησης	Πλήρως ανεπτυγμένο εργαλείο, με δυνατότητα επιλογής επιπέδων ανωνυμοποίησης και παραμετροποίησης tags	Ενότητες 2, 3, 4
4.1	Γραφικό περιβάλλον, εκτελέσιμο σε οποιοδήποτε λειτουργικό σύστημα	GUI με tkinter, υποστήριξη για Windows/macOS/Linux, δημιουργία .exe, .app & CLI version	Ενότητες 5.1, 7.5
4.2	Υποστήριξη εισαγωγής και εκκαθάρισης δεδομένων	Επιλογή φακέλων εισόδου/εξόδου, διαγραφή/τροποποίηση tags, καταγραφή αλλαγών	Ενότητες 4.1–4.5, 7.3
4.3	Παροχή οδηγιών στον χρήστη (wizards) για δημιουργία κανόνων	Αναδυόμενος διάλογος επιλογής πεδίων, tooltips με πληροφορίες, οπτική προεπισκόπηση αποτελεσμάτων	Ενότητες 4.4, 5.2
4.4	Εξυπνοι τρόποι προσαρμογής σε σύνολα δεδομένων	Προεπιλογές επιπέδων + custom επιλογή tags, δυναμική διαχείριση DICOM τιμών και τύπων	Ενότητες 4.1, 4.2, 6.1

A/A	Απαιτήση	Κάλυψη από το Εργαλείο	Αναφορά στο Παραδοτέο
4.5	Υποστήριξη αποτύπωσης της χρήσης δεδομένων	Δημιουργία log αρχείου με πλήρες audit trail για κάθε αλλαγή σε μεταδεδομένα	Ενότητες 4.5, 8.2
4.6	Εκτέλεση σε ασφαλή εικονική μηχανή εντός υποδομής νέφους	Ο κώδικας είναι ελαφρύς, χωρίς εξαρτήσεις, και μπορεί να ενσωματωθεί σε containerized/cloud-based λύσεις (π.χ. Docker, Kubernetes)	Ενότητες 6.2, 6.3, 6.3

9. Συμπεράσματα και Μελλοντική Αξιοποίηση

Η ανάπτυξη του εργαλείου ανωνυμοποίησης ιατρικών εικόνων που περιγράφεται στην παρούσα τεχνική έκθεση καλύπτει με πληρότητα τις λειτουργικές, τεχνικές και ρυθμιστικές απαιτήσεις του έργου. Το λογισμικό σχεδιάστηκε και υλοποιήθηκε με σύγχρονες μεθοδολογίες, χρησιμοποιώντας ώριμες τεχνολογίες ανοιχτού κώδικα, με σκοπό να παρέχει ένα σταθερό, ασφαλές και παραμετροποιήσιμο περιβάλλον για την αφαίρεση προσωπικών δεδομένων από αρχεία DICOM.

Η δομή του εργαλείου επιτρέπει:

- την εύκολη ενσωμάτωση σε σύνθετες ροές εργασίας (pipelines) υγείας,
- τη σύνδεση με ομοσπονδιακές ή κεντρικές βάσεις δεδομένων,
- τη συμμόρφωση με τις απαιτήσεις του GDPR και άλλων κανονιστικών πλαισίων,
- τη λειτουργία σε περιβάλλοντα Windows, macOS και Linux με ή χωρίς εγκατάσταση Python.

Η δυνατότητα επιλογής επιπέδου ανωνυμοποίησης, η υποστήριξη για custom επιλογές tags, καθώς και η παροχή αρχείου καταγραφής των τροποποιήσεων, ενισχύουν τη διαφάνεια, την επαναληψιμότητα και τον έλεγχο των διαδικασιών. Η προεπισκόπηση των αλλαγών πριν την τελική εκτέλεση αποτελεί ιδιαίτερο πλεονέκτημα, ειδικά σε περιβάλλοντα όπου απαιτείται ποιοτική επιθεώρηση των μεταδεδομένων.

9.1 Δυνατότητες Μελλοντικής Εξέλιξης

Η παρούσα υλοποίηση μπορεί να αποτελέσει τη βάση για περαιτέρω εξέλιξη και αξιοποίηση, με δυνατότητες όπως:

- **Ενσωμάτωση μηχανισμών κρυπτογράφησης** και διαχείρισης ρόλων (role-based access control)
- **Υποστήριξη web-based εκδόσεων** για απομακρυσμένη χρήση σε υποδομές cloud
- **Ενσωμάτωση με πρότυπα HL7 FHIR** για διαλειτουργικότητα με άλλα συστήματα υγείας
- **Αυτόματη προσαρμογή σε DICOM SR και multi-frame αρχεία** (π.χ. cardiac cine, 3D imaging)

Επιπλέον, λόγω του modular σχεδιασμού, το εργαλείο μπορεί να χρησιμοποιηθεί σε έργα που αφορούν:

- Ανάλυση μεγάλων δεδομένων (big data) στην υγεία
- Εκπαίδευση αλγορίθμων τεχνητής νοημοσύνης (AI/ML)
- Ανάπτυξη αποθετηρίων και ομοσπονδιακών συστημάτων (federated databases)

9.2 Συμβολή στην Υλοποίηση του Έργου

Το συγκεκριμένο εργαλείο αποτελεί κρίσιμο τεχνικό υποσύστημα για την υλοποίηση της **προτυποποίησης και εναρμόνισης ιατρικών δεδομένων**, όπως περιγράφεται στους στόχους του έργου. Η λειτουργία του εξασφαλίζει:

- Την ορθή προεπεξεργασία των δεδομένων εισόδου
- Τη συμμόρφωση με τις πολιτικές ανωνυμοποίησης της Αναθέτουσας Αρχής
- Την αποτελεσματική σύνδεση με τις λοιπές τεχνολογικές υποδομές (βάσεις, APIs, μετατροπείς)

Εν κατακλείδι, το λογισμικό ανταποκρίνεται τόσο στις **άμεσες ανάγκες του έργου** όσο και στις **προοπτικές αξιοποίησής του σε μελλοντικά εθνικά και ευρωπαϊκά έργα**, παρέχοντας ένα ισχυρό, τεκμηριωμένο και ευέλικτο εργαλείο για την προστασία των ιατρικών δεδομένων.

10. Βιβλιογραφία

1. Kaissis, G.A., et al., *Secure, privacy-preserving and federated machine learning in medical imaging*. Nature Machine Intelligence, 2020. **2**(6): p. 305-311.
2. Crowson, M.G., et al., *A systematic review of federated learning applications for biomedical data*. PLOS Digit Health, 2022. **1**(5): p. e0000033.
3. *PyDICOM Documentation*. Available from: <https://pydicom.github.io/>.
4. Committee, D.S., *Digital Imaging and Communications in Medicine (DICOM) – Part 15: Security and System Management Profiles*. 2023.
5. 9241-210:2019, I., *Ergonomics of human-system interaction – Human-centred design for interactive systems*.