



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Σχεδιασμός και ανάπτυξη πρωτότυπου
ομοσπονδιακού δικτύου (federated network)
για χρήση σε τράπεζα βιολογικών δεδομένων
(biobank), σύμφωνα με τα ευρωπαϊκά πρότυπα
διαχείρισης δεδομένων.**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

ΜΑΡΔΙΚΗ ΚΩΝΣΤΑΝΤΙΝΟΥ

Επιβλέπων: Γ. Ματσόπουλος
Καθηγητής Ε.Μ.Π.

Συνεπιβλέπουσα: Ο. Πετροπούλου
ΕΔΙΠ Α'

Ιούνιος 2025



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Σχεδιασμός και ανάπτυξη πρωτότυπου
ομοσπονδιακού δικτύου (federated network) για
χρήση σε τράπεζα βιολογικών δεδομένων (biobank),
σύμφωνα με τα ευρωπαϊκά πρότυπα διαχείρισης
δεδομένων.**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

ΜΑΡΔΙΚΗ ΚΩΝΣΤΑΝΤΙΝΟΥ

Επιβλέπων: Γ. Ματσόπουλος
Καθηγητής Ε.Μ.Π.

Συνεπιβλέπουσα: Ο. Πετροπούλου
ΕΔΙΠ Α'

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 4^η Ιουνίου 2025.

.....		
Γ. Ματσόπουλος	Α. Παναγόπουλος	Π. Τσανάκας
Καθηγητής Ε.Μ.Π.	Καθηγητής Ε.Μ.Π.	Καθηγητής Ε.Μ.Π.

Ιούνιος 2025



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

.....

Κωνσταντίνος Μαρδίκης, 2025.

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © – All rights reserved. Με την επιφύλαξη παντός δικαιώματος.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται στον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Οι σύγχρονες βιοτράπεζες αντιμετωπίζουν την πρόκληση της διαχείρισης εκθετικά αυξανόμενων, ετερογενών βιολογικών δεδομένων, ενώ οφείλουν ταυτόχρονα να τηρούν αυστηρά νομοθετικά πλαίσια και να εξασφαλίζουν διαλειτουργικότητα μεταξύ ανεξάρτητων φορέων. Η παρούσα διπλωματική εργασία μελετά, σχεδιάζει και υλοποιεί ένα ομοσπονδιακό δικτυακό μοντέλο που στοχεύει σε ασφαλή, ευέλικτη και FAIR-συμβατή ανταλλαγή δεδομένων μεταξύ βιοτραπεζών. Στο θεωρητικό μέρος παρουσιάζονται αναλυτικά οι αρχιτεκτονικές κεντριοποιημένων, αποκεντρωμένων, κατανεμημένων και ομοσπονδιακών δικτύων, οι τύποι βιοτραπεζών και οι κατηγορίες δεδομένων που μπορούν να αποθηκεύσουν. Χαρτογραφείται το ευρωπαϊκό νομοθετικό πλαίσιο και πραγματοποιείται βιβλιογραφική ανασκόπηση υπαρχόντων διεθνών πρωτοβουλιών. Στο πρακτικό μέρος προτείνεται και υλοποιείται αρχιτεκτονική δύο επιπέδων: τοπικοί κόμβοι, εγκατεστημένοι σε φορείς, και κεντρικός κόμβος για την αναζήτηση με metadata. Οι τοπικοί κόμβοι φιλοξενούν διεπαφές ιατρικού προσωπικού, διαχείριση συγκαταθέσεων δοτών/ασθενών και υπηρεσία ερωτημάτων, ενώ ο κεντρικός κόμβος παρέχει ομοιόμορφη αναζήτηση χωρίς μεταφορά ευαίσθητων δεδομένων. Η πλατφόρμα υλοποιείται με React, FastAPI, MongoDB, Keycloak, προσφέροντας εύκολη εγκατάσταση και επεκτασιμότητα.

Λέξεις Κλειδιά

ομοσπονδιακή αρχιτεκτονική, GDPR, βιοτράπεζα, metadata, React, FastAPI, MongoDB, Keycloak

Abstract

Modern biobanks face the challenge of managing exponentially growing, heterogeneous biological data, while at the same time having to comply with strict legislative frameworks and ensure interoperability between independent institutions. This thesis studies, designs and implements a federal network model that aims to achieve secure, flexible, and FAIR-compliant data exchange between biobanks. In the theoretical part, centralized architectures are presented in detail, decentralized, distributed and federated networks, the types of biobanks and the categories of data they can store. The European legislative framework is mapped and a literature review is carried out existing international initiatives. In the practical part, a two-tier architecture is proposed and implemented: local nodes, installed in institutions, and a central node for searching with metadata. The local nodes host medical staff interfaces, donor/patient consent management and query service, while the central node provides uniform search without transferring sensitive data. The platform is implemented with React, FastAPI, MongoDB, Keycloak, offering easy deployment and scalability.

Keywords

federated architecture, biobanking, metadata, React, FastAPI, MongoDB, Keycloak

Ευχαριστίες

Θα ήθελα, καταρχάς, να ευχαριστήσω θερμά τον καθηγητή κ. Ματσόπουλο για την επίβλεψη της παρούσας διπλωματικής εργασίας, καθώς και για την ευκαιρία που μου έδωσε να την εκπονήσω. Ιδιαίτερες ευχαριστίες οφείλω στην κ. Πετροπούλου για την εξαιρετική συνεργασία που είχαμε. Δεν θα μπορούσα να παραλείψω τον κ. Σταύρο Πιτόγλου και την κ. Θέλμα Ανδρούτσου, για την πολύτιμη βοήθεια και καθοδήγησή τους καθ' όλη τη διάρκεια της συγγραφής της εργασίας. Τέλος, θα ήθελα να εκφράσω την ευγνωμοσύνη μου στην οικογένειά μου και στους φίλους μου για τη συνεχή ηθική στήριξη που μου προσέφεραν όλα αυτά τα χρόνια.

Ιούνιος 2025

Κωνσταντίνος Μαρδίκης

Περιεχόμενα

Περίληψη	5
Abstract	7
Ευχαριστίες	9
Εισαγωγή	17
1 Θεωρητικό υπόβαθρο	18
1.1 Δίκτυα	18
1.1.1 Κεντριοποιημένο (Centralized) Δίκτυο	18
1.1.2 Αποκεντρωμένο (Decentralized) Δίκτυο	19
1.1.3 Κατανεμημένο (Distributed) Δίκτυο	20
1.1.4 Ομοσπονδιακό (Federated) Δίκτυο	21
1.1.5 Σύγκριση Αρχιτεκτονικών Δικτύων	22
1.2 Τράπεζα βιολογικών δεδομένων (biobank)	23
1.2.1 Τύποι τράπεζας βιολογικών δεδομένων	24
1.2.2 Τύποι δεδομένων σε τράπεζα βιολογικών δεδομένων	24
1.3 Ευρωπαϊκή πολιτική προστασίας και διαχείρισης προσωπικών δεδομένων	26
1.3.1 (General Data Protection Regulation - GDPR)	27
1.3.2 Το πρότυπο ISO 27001 και η εφαρμογή του στις βιοτράπεζες	28
1.3.3 Ασφάλεια Πληροφοριών Υγείας (ISO 27799)	29
1.3.4 Αρχές FAIR (Findable, Accessible, Interoperable, Reusable)	29
1.3.5 Πλαίσιο Διαχείρισης Δεδομένων ELIXIR	30
1.3.6 Πρότυπο για Βιοτράπεζες (ISO 20387:2018)	30
1.4 Βιβλιογραφική ανασκόπηση εφαρμογής δικτύου σε τράπεζα βιολογικών δεδομένων	31
1.4.1 UK Biobank	31
1.4.2 All of Us Research Program	32
1.4.3 EuroBioBank	32
1.4.4 AminoChain (Decentralized Biobank Platform)	32
1.4.5 Δίκτυο RD-Connect για Σπάνιες Νόσους	32
1.4.6 BioSHaRE-EU - DataSHIELD	33
1.4.7 BBMRI-ERIC	33
1.4.8 Federated European Genome-Phenome Archive (FEGA)	34
1.4.9 GA4GH Beacon Networks	34

1.4.10CINECA	34
1.4.11 Συγκεντρωτικός πίνακας δικτύων σε βιοτράπεζα	35
2 Πρακτικό Μέρος	37
2.1 Αρχιτεκτονική Δικτύου	37
2.1.1 Τοπικοί Κόμβοι	38
2.1.2 Κεντρικός Κόμβος	38
2.1.3 Μηχανισμός Σύνδεσης Τοπικού και Κεντρικού Κόμβου	38
2.2 Περιγραφή Αρχιτεκτονικής Τοπικού Κόμβου	40
2.2.1 Διεπαφή Διασύνδεσης Ιατρικού Προσωπικού	40
2.2.2 Διεπαφή Διαχείρισης Συγκρατάσεων (Δότης/Ασθενής)	41
2.2.3 Queries Service	42
2.3 Περιγραφή Αρχιτεκτονικής Κεντρικού Κόμβου	43
2.3.1 Διεπαφή Αναζήτησης	43
2.4 Authenticator Servers	44
2.5 Τεχνολογίες Ανάπτυξης Δικτύου	45
2.5.1 Frontend-React	46
2.5.2 Backend-FastAPI	47
2.5.3 Database-MongoDB	47
2.5.4 Reverse Proxy-Nginx	47
2.5.5 Docker	48
2.5.6 Keycloak	48
2.6 Υλοποίηση της αρχιτεκτονικής	49
Παράρτημα: Πηγαίος Κώδικας Διπλωματικής	65
Βιβλιογραφία	67

Κατάλογος Σχημάτων

1.1	Κεντριοποιημένο δίκτυο [1]	19
1.2	Αποκεντρωμένο δίκτυο [1]	20
1.3	Καταμεμημένο δίκτυο [1]	21
1.4	Ομοσπονδιακό δίκτυο	21
2.1	Περιγραφή της γενικής αρχιτεκτονικής μαζί με τις τεχνολογίες	37
2.2	Αρχιτεκτονική ενός τοπικού κόμβου με τον κεντρικό κόμβο	39
2.3	Περιγραφή της αρχιτεκτονικής του τοπικού κόμβου	40
2.4	Λειτουργίες διεπαφής δασύνδεσης ιατρικού προσωπικού	40
2.5	Λειτουργίες διεπαφής διαχείρισης συγκαταθέσεων (δότης/ασθενής)	42
2.6	Περιγραφή της αρχιτεκτονικής του κεντρικού κόμβου	43
2.7	Λειτουργίες διεπαφής αναζήτησης	44
2.8	Software stack του δικτύου	46

Κατάλογος Εικόνων

1.1	Σύγκριση των Δικτύων	23
1.2	UK Biobank: Περιγραφή δεδομένων [2]	26
2.1	Docker containers	50
2.2	διαχείριση των realms	50
2.3	Σύνδεση στην εφαρμογή του ιατρικού προσωπικού ως admin	51
2.4	Οθόνη Upload του ιατρικού προσωπικού	51
2.5	Οθόνη Upload του ιατρικού προσωπικού με συμπληρωμένα τα πεδία	52
2.6	Οθόνη View Data του ιατρικού προσωπικού	53
2.7	Οθόνη Manage Users του ιατρικού προσωπικού	53
2.8	Οθόνη Manage Users του ιατρικού προσωπικού με συμπληρωμένη την φόρμα δημιουργίας νέου χρήστη	54
2.9	Οθόνη Manage Users του ιατρικού προσωπικού με τις λειτουργίες διαχείρισης	55
2.10	Πρώτη σύνδεση στην εφαρμογή ιατρικού προσωπικού ως νέος χρήστης	55
2.11	Πρώτη σύνδεση στην εφαρμογή ιατρικού προσωπικού ως νέος χρήστης με αλλαγή κωδικού	56
2.12	Πρώτη σύνδεση στην εφαρμογή ιατρικού προσωπικού ως νέος χρήστης με ενημέρωση προσωπικών στοιχείων	56
2.13	Πρώτη σύνδεση στην εφαρμογή ιατρικού προσωπικού ως νέος χρήστης	57
2.14	Τρόπος εμφάνισης των logs του χρήστη	57
2.15	Εφαρμογή για την διαχείριση συγκατάθεσης ασθενή/δότη	58
2.16	Οθόνη Queries χωρίς την προσθήκη τοπικών κόμβων	59
2.17	Οθόνη Nodes με την εμφάνιση των Metadata χωρίς την προσθήκη τοπικών κόμβων	59
2.18	Οθόνη Nodes με την εμφάνιση διαχείρισης των τοπικών κόμβων χωρίς την προσθήκη τους	60
2.19	Οθόνη Nodes με την εμφάνιση διαχείρισης των τοπικών κόμβων με συμπληρωμένη την φόρμα προσθήκη τους	61
2.20	Οθόνη Nodes με την εμφάνιση διαχείρισης των τοπικών κόμβων με δύο κόμβους	61
2.21	Οθόνη Nodes με την εμφάνιση των Metadata των δύο τοπικών κόμβων	62
2.22	Οθόνη Queries με την προσθήκη τοπικών κόμβων	63

Κατάλογος Πινάκων

1.2	Πίνακας τύπων τράπεζας βιολογικών δεδομένων	24
1.3	Συγκεντρωτικός πίνακας δικτύων σε βιοτράπεζα	35

Εισαγωγή

Η ταχεία ανάπτυξη των βιοϊατρικών επιστημών, καθώς και η συνεχής εξέλιξη στον τομέα της τεχνολογίας, έχουν οδηγήσει σε μια εντυπωσιακή αύξηση των δεδομένων που συλλέγονται και αποθηκεύονται σε σύγχρονες τράπεζες βιολογικών δεδομένων, γνωστές και ως βιοτράπεζες (biobanks). Οι βιοτράπεζες αποτελούν θεμελιώδεις πόρους για τη σύγχρονη ιατρική και την επιστημονική έρευνα, παρέχοντας τη δυνατότητα για εις βάθος ανάλυση βιολογικών δειγμάτων και τη σύνδεσή τους με κλινικά, γονιδιακά και άλλα δεδομένα που μπορεί να συσχετίζονται με την υγεία των ασθενών.

Παρά τη σημασία τους, οι σύγχρονες βιοτράπεζες αντιμετωπίζουν σημαντικές προκλήσεις, όπως η διαχείριση εκθετικά αυξανόμενων και ετερογενών δεδομένων, οι αυστηροί περιορισμοί για τη διατήρηση της ιδιωτικότητας και της ασφάλειας, και η αναγκαιότητα συμμόρφωσης με πολύπλοκα ευρωπαϊκά πρότυπα όπως ο Γενικός Κανονισμός Προστασίας Δεδομένων (GDPR) και οι αρχές FAIR (Findable, Accessible, Interoperable, Reusable).

Η παρούσα διπλωματική εργασία επικεντρώνεται στον σχεδιασμό και την υλοποίηση ενός πρωτοτύπου ομοσπονδιακού δικτύου, για χρήση σε βιοτράπεζα. Κύριοι στόχοι του δικτύου είναι η ασφαλής, ευέλικτη και διαλειτουργική διαχείριση και ανταλλαγή βιοϊατρικών δεδομένων, χωρίς να παραβιάζονται οι αρχές προστασίας προσωπικών δεδομένων και ιδιωτικότητας που ορίζονται από την ευρωπαϊκή νομοθεσία. Η εργασία αναπτύσσεται σε δύο διακριτά μέρη: το θεωρητικό, όπου παρουσιάζονται οι αρχές λειτουργίας και οι αρχιτεκτονικές επιλογές για δίκτυα και βιοτράπεζες, και το πρακτικό μέρος, στο οποίο αναλύεται λεπτομερώς η αρχιτεκτονική και η υλοποίηση του προτεινόμενου συστήματος με χρήση σύγχρονων τεχνολογιών όπως React, FastAPI, MongoDB και Keycloak. Στόχος είναι να παρέχεται μια εύκολα διαχειρίσιμη, ασφαλής και επεκτάσιμη λύση που να ανταποκρίνεται πλήρως στις ανάγκες των σύγχρονων βιοτραπεζών και να αποτελεί υπόδειγμα για περαιτέρω εξελίξεις στον τομέα αυτό.

Κεφάλαιο **1**

Θεωρητικό υπόβαθρο

Το κεφάλαιο αυτό θέτει το θεωρητικό υπόβαθρο της εργασίας, παρουσιάζοντας τις βασικές προσεγγίσεις για την οργάνωση δικτύων και τη διαχείριση των βιοϊατρικών δεδομένων. Σε αυτή την ενότητα, αναλύονται συνοπτικά οι διαφορετικές δομές δικτύων, κεντρικοποιημένα, αποκεντρωμένα, κατανεμημένα και ομοσπονδιακά δίκτυα. Επισημαίνονται οι βασικές διαφορές μεταξύ αυτών των προσεγγίσεων, τονίζοντας τα πλεονεκτήματα και τις προκλήσεις που συνοδεύουν κάθε μία.

Στη συνέχεια, το κεφάλαιο στρέφει την προσοχή του στις βιοτράπεζες, παρουσιάζοντας συνοπτικά τη λειτουργία τους και την ποικιλία των δεδομένων που συλλέγουν, όπως πληροφορίες για την υγεία, τα γενετικά στοιχεία και τα δημογραφικά στοιχεία. Η περιγραφή γίνεται με τρόπο που αναδεικνύει την εκτεταμένη ποικιλομορφία δεδομένων που συναντάει κάποιος σε μια βιοτράπεζα.

Επιπρόσθετα, το κεφάλαιο εξετάζει το νομικό και τεχνικό πλαίσιο που διέπει την επεξεργασία και προστασία των δεδομένων, δίνοντας έμφαση στους ευρωπαϊκούς κανονισμούς και τα διεθνή πρότυπα που εξασφαλίζουν την ασφάλεια, τη διαφάνεια και τη διαλειτουργικότητα των πληροφοριών. Τέλος, συνοψίζονται παραδείγματα εφαρμογής δικτύωσης σε βιοτράπεζες.

1.1 Δίκτυα

1.1.1 Κεντρικοποιημένο (Centralized) Δίκτυο

Το κεντρικοποιημένο σύστημα αναφέρεται σε μια αρχιτεκτονική όπου η λήψη αποφάσεων, η διαχείριση των πόρων και ο έλεγχος πραγματοποιούνται από έναν κεντρικό κόμβο ή οντότητα. Στα συστήματα αυτά, όλες οι διεργασίες συντονίζονται κεντρικά, προσφέροντας συνολική οπτική και έλεγχο στο σύστημα, με στόχο τη βέλτιστη απόδοση και τη μεγιστοποίηση της αποδοτικότητας.[3, 4, 5]

Το Σχήμα 1.1 απεικονίζει το κεντρικοποιημένο δίκτυο με βάση το πως περιγράφεται στην βιβλιογραφία.[1, 6]

Παραδείγματα εφαρμογής δικτύου

Τραπεζικός Τομέας: Οι τράπεζες χρησιμοποιούν κεντρικά συστήματα για τη διαχείριση λογαριασμών και συναλλαγών, εξασφαλίζοντας ενοποιημένο έλεγχο και ασφάλεια [7].



Σχήμα 1.1: Κεντροποιημένο δίκτυο [1]

Επιχειρήσεις Λιανικής Πώλησης: Χρησιμοποιούν κεντρικά συστήματα POS για την παρακολούθηση αποθεμάτων και πωλήσεων, επιτρέποντας κεντρικό έλεγχο των δεδομένων [8].

1.1.2 Αποκεντρωμένο (Decentralized) Δίκτυο

Ένα αποκεντρωμένο σύστημα είναι ένα σύστημα χωρίς κεντρική αρχή που να ελέγχει την συνολική λειτουργία. Οι κόμβοι του συστήματος λειτουργούν αυτόνομα, λαμβάνοντας αποφάσεις σε τοπικό επίπεδο και συνεργάζονται μέσω πρωτοκόλλων για την επίτευξη κοινών στόχων.[4, 5, 9]

Το Σχήμα 1.2 απεικονίζει το αποκεντρωμένο δίκτυο με βάση το πως περιγράφεται στην βιβλιογραφία.[1, 6]

Παραδείγματα εφαρμογής δικτύου

Διανομείς Ενέργειας: Δίκτυα ηλεκτρικής ενέργειας με αποκεντρωμένη παραγωγή από ανανεώσιμες πηγές, επιτρέποντας την ευελιξία και την ανθεκτικότητα του συστήματος [10].

Διαχείριση Εφοδιαστικής Αλυσίδας: Επιτρέπει σε διάφορους φορείς να διαχειρίζονται τα δικά τους τμήματα ανεξάρτητα, βελτιώνοντας την αποδοτικότητα και την ανταπόκριση στις αλλαγές της αγοράς [11].



Σχήμα 1.2: Αποκεντρωμένο δίκτυο [1]

1.1.3 Κατανεμημένο (Distributed) Δίκτυο

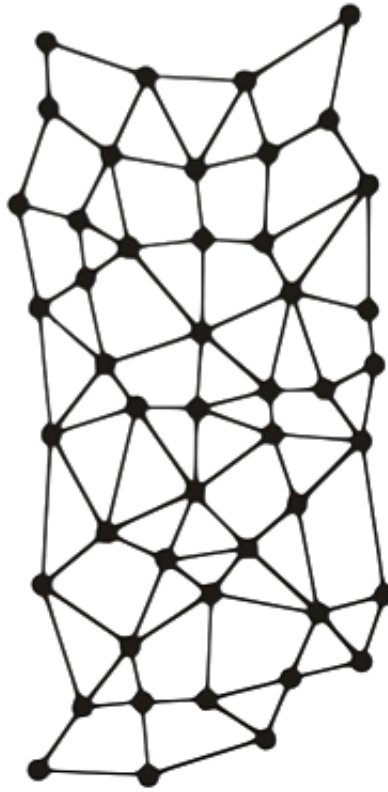
Το κατανεμημένο σύστημα αποτελείται από πολλαπλές αυτόνομες συσκευές ή κόμβους που επικοινωνούν μεταξύ τους. Η συνεργασία αυτών των κόμβων αποτελεί σημαντικό παράγοντα για την επίτευξη του κοινού σκοπού. Αυτοί οι κόμβοι είναι κατανεμημένοι σε διαφορετικές φυσικές τοποθεσίες, αλλά λειτουργούν ως μια ενιαία οντότητα. Η διαχείριση γίνεται από μια κεντρική αρχή (πρωτόκολλο) εξασφαλίζοντας την εύρυθμη λειτουργία.[5, 9, 12, 13]

Το Σχήμα 1.3 απεικονίζει το κατανεμημένο δίκτυο με βάση το πως περιγράφεται στην βιβλιογραφία.[1, 6]

Παραδείγματα εφαρμογής δικτύου

Υπηρεσίες Cloud Computing: Πλατφόρμες όπως το Amazon Web Services και το Google Cloud χρησιμοποιούν κατανεμημένα δίκτυα για την παροχή υπηρεσιών, εξασφαλίζοντας υψηλή επεκτασιμότητα και ανθεκτικότητα. [14]

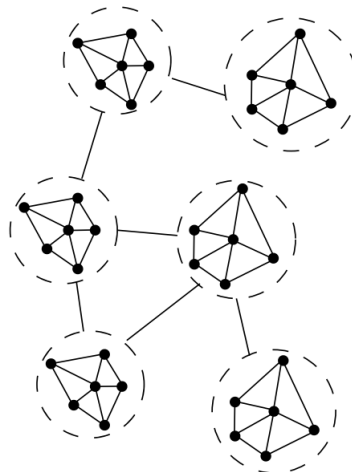
Δίκτυα Παράδοσης Περιεχομένου (CDNs): Διανέμουν περιεχόμενο μέσω πολλαπλών διακομιστών παγκοσμίως, βελτιώνοντας την ταχύτητα και την αξιοπιστία της παράδοσης περιεχομένου. [15]



Σχήμα 1.3: Κατανεμημένο δίκτυο [1]

1.1.4 Ομοσπονδιακό (Federated) Δίκτυο

Ένα ομοσπονδιακό σύστημα αποτελείται από πολλαπλά αυτόνομα υποσυστήματα ή κόμβους, τα οποία συνεργάζονται για την επίτευξη του κοινού σκοπού διατηρώντας την αυτονομία τους. Κάθε υποσύστημα εξασφαλίζει τον πλήρη έλεγχο των δεδομένων, των πολιτικών και των λειτουργιών του. Η επικοινωνία των υποσυστημάτων βασίζεται σε προκαθορισμένα πρωτόκολλα ή διεπαφές (APIs). Η ανταλλαγή των δεδομένων περιορίζεται με σκοπό την παραμονή τους τοπικά στους κόμβους, εξασφαλίζοντας ιδιωτικότητα και ασφάλεια.[12, 13]



Σχήμα 1.4: Ομοσπονδιακό δίκτυο

Το Σχήμα 1.4 απεικονίζει το ομοσπονδιακό δίκτυο που είναι ο συνδυασμός του αποκεντρωμένου και κατανεμημένου δικτύου.

Παραδείγματα εφαρμογής δικτύου

Υγειονομικός Τομέας: Νοσοκομεία διατηρούν τα δικά τους δεδομένα αλλά συνεργάζονται για την ανταλλαγή πληροφοριών ασθενών, διασφαλίζοντας την ιδιωτικότητα και την ασφάλεια των δεδομένων. [16]

Ακαδημαϊκά Ιδρύματα: Πανεπιστήμια διατηρούν τα δικά τους συστήματα αλλά συνεργάζονται σε ερευνητικά έργα, επιτρέποντας την ανταλλαγή γνώσης και πόρων χωρίς να θυσιάζεται η αυτονομία τους. [17]

1.1.5 Σύγκριση Αρχιτεκτονικών Δικτύων

Τα **Centralized Networks** χαρακτηρίζονται από έναν κεντρικό κόμβο που λαμβάνει όλες τις αποφάσεις (επεξεργασίας και διαχείρισης των δεδομένων), με τους υπόλοιπους κόμβους να εξαρτώνται πλήρως από αυτόν. Αυτή η αρχιτεκτονική προσφέρει απλότητα στη δομή, στη διαχείριση, στη συντήρηση και στην ανάπτυξη του δικτύου, δεδομένου ότι υπάρχει μόνο ένας κεντρικός κόμβος ελέγχου. Ωστόσο, ο κεντρικός κόμβος αποτελεί και το κύριο αδύναμο σημείο (single point of failure), καθώς οποιαδήποτε δυσλειτουργία του μπορεί να προκαλέσει ολική κατάρρευση του δικτύου, ενώ η συγκέντρωση του ελέγχου σε ένα σημείο περιορίζει την επεκτασιμότητα του δικτύου, δηλαδή τη δυνατότητα επέκτασης του συστήματος μέσω της προσθήκης νέων κόμβων. [3, 4, 5]

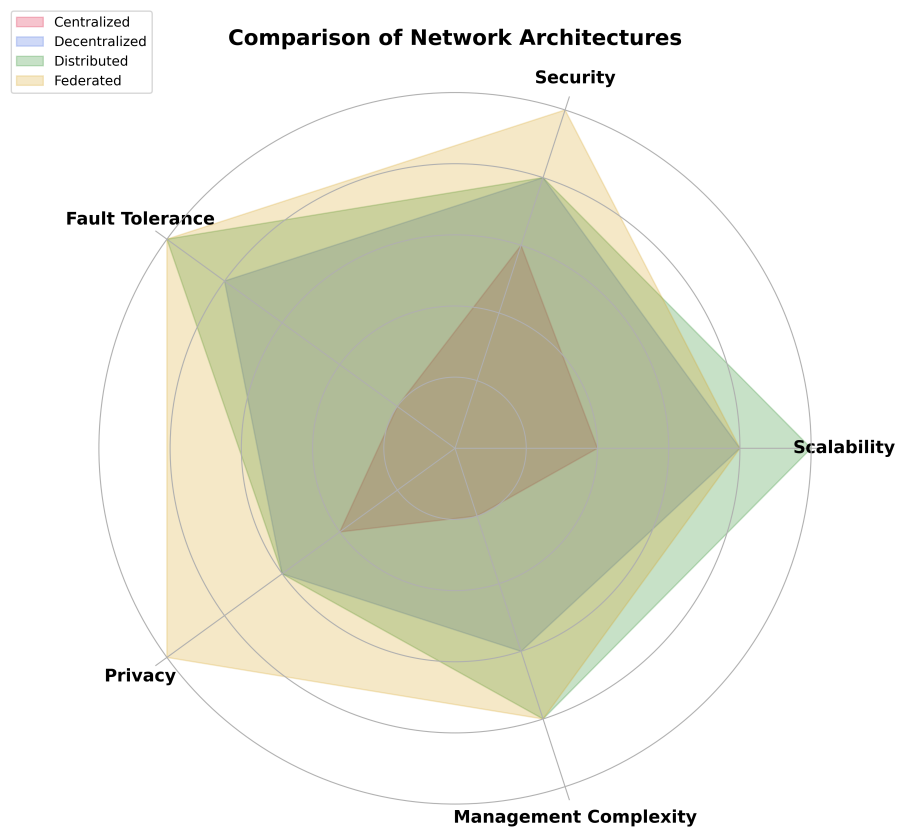
Αντίθετα, τα **Decentralized Networks** δεν στηρίζονται σε έναν ενιαίο, κεντρικό κόμβο ελέγχου, αλλά επιτρέπουν σε κάθε κόμβο ή ομάδα κόμβων να λαμβάνει αυτόνομα αποφάσεις. Αυτή η αρχιτεκτονική, χάρη στην απουσία μοναδικού σημείου αποτυχίας (δηλαδή, στη μείωση της πιθανότητας κατάρρευσης του συστήματος λόγω της αποκέντρωσης των λειτουργιών), προσφέρει αυξημένη ανθεκτικότητα και ενισχύει την επεκτασιμότητα του δικτύου. Ωστόσο, η έλλειψη κεντρικού ελέγχου οδηγεί σε προκλήσεις στον συντονισμό, τη διαχείριση και στην εμφάνιση του φαινομένου των *βρόχων πακέτων (networks looping)* (μια ανεπιθύμητη κατάσταση κατά την οποία τα πακέτα δεδομένων κυκλοφορούν συνεχώς μέσα στο δίκτυο λόγω περιπτώσεων συνδέσεων, όπως η σύνδεση ενός ενεργού και ενός κατεστραμένου κόμβου), απαιτώντας την υιοθέτηση πολύπλοκων αλγορίθμων και πρωτοκόλλων για την εξασφάλιση ομαλής λειτουργίας. [9, 4, 5]

Στα **Distributed Networks**, οι κόμβοι συνεργάζονται ισότιμα για την επίτευξη των στόχων του συστήματος μέσω πλήρους διασύνδεσης. Αυτή η αρχιτεκτονική χαρακτηρίζεται από υψηλή επεκτασιμότητα και αυξημένη ανθεκτικότητα, δεδομένου ότι η κατανομή των πόρων και ο έλεγχος γίνονται ομοιόμορφα μεταξύ των κόμβων. Ωστόσο, η πλήρης διασύνδεση οδηγεί σε αυξημένη πολυπλοκότητα στη διαχείριση και το συντονισμό, καθώς απαιτείται λεπτομερής παρακολούθηση και έλεγχος για την ομαλή λειτουργία του συστήματος. [5, 12]

Τέλος, τα **Federated Networks** προσφέρουν μια υβριδική προσέγγιση, όπου κάθε κόμβος ή υποσύστημα διατηρεί την αυτονομία του, ενώ παράλληλα επιτρέπεται η συνεργασία μεταξύ τους μέσω προκαθορισμένων διεπαφών και πρωτοκόλλων. Σε αυτό το μοντέλο, οι κόμβοι

ανταλλάσσουν μόνο τα αποτελέσματα, ενώ τα δεδομένα παραμένουν τοπικά, συμβάλλοντας στην προστασία της ιδιωτικότητας και της ασφάλειας. Παρότι αυτό το μοντέλο συνδυάζει τα πλεονεκτήματα της επεκτασιμότητας και της ανθεκτικότητας, απαιτεί πολύ προσεκτικό σχεδιασμό για την αποτελεσματική διασύνδεση και λειτουργία των κόμβων. [12, 13]

Συνολικά, ενώ τα **Centralized Systems** προσφέρουν απλότητα και εύκολο κεντρικό έλεγχο, τα **Decentralized** και **Distributed Networks** προσφέρουν ανθεκτικότητα και δυνατότητες επεκτασιμότητας, παρόλο που συνοδεύονται από προκλήσεις στον συντονισμό και τη διαχείριση. Τα **Federated Networks** επιτυγχάνουν έναν συμβιβασμό μεταξύ της διατήρησης της αυτονομίας και της ομαδικής συνεργασίας, καθιστώντας τα ιδανικά για εφαρμογές όπου απαιτείται υψηλό επίπεδο ασφάλειας, ιδιωτικότητας και αποτελεσματικής ανταλλαγής δεδομένων.



Εικόνα 1.1: Σύγκριση των Δικτύων

1.2 Τράπεζα βιολογικών δεδομένων (biobank)

Η τράπεζα βιολογικών δεδομένων (**biobank**) αρχικά ήταν η συστηματική συλλογή βιολογικών δειγμάτων (όπως το αίμα ή κυτταρικά δείγματα) και σχετικών πληροφοριών υγείας. Με την πάροδο του χρόνου, ο ορισμός επεκτάθηκε ώστε να περιλαμβάνει όχι μόνο βιολογικά δεδομένα αλλά και συνοδευτικά δεδομένα όπως το ιατρικό ιστορικό, ο τρόπος ζωής και περι-

βαλλοντικοί παράγοντες. Αυτό επιτρέπει μια πιο ολοκληρωμένη φαινοτυποποίηση, δηλαδή την ακριβή αξιολόγηση, την κατανόηση και τον συσχετισμό των κλινικών, των γενετικών και των περιβαλλοντικών στοιχείων που επηρεάζουν την υγεία των ασθενών. [18]

1.2.1 Τύποι τράπεζας βιολογικών δεδομένων

Biobank Type	Περιγραφή
Disease-specific	Συλλογή δειγμάτων από ασθενείς με συγκεκριμένες παθήσεις.
Population-based	Συλλογή δειγμάτων από γενικό πληθυσμό για την αξιολόγηση γενετικών και περιβαλλοντικών παραγόντων.
Genetic (DNA/RNA)	Εστίαση στη συλλογή και αποθήκευση γενετικού υλικού.
Project-driven	Δημιουργημένα για συγκεκριμένα ερευνητικά έργα.
Commercial	Διεξάγονται από ιδιωτικούς φορείς με κερδοσκοπικό χαρακτήρα.
Virtual	Διαχείριση και πρόσβαση σε δεδομένα από απομακρυσμένες τοποθεσίες μέσω διαδικτύου.

Πίνακας 1.2: Πίνακας τύπων τράπεζας βιολογικών δεδομένων

Ο πίνακας 1.2 απεικονίζει τους διάφορους τύπους τράπεζας βιολογικών δεδομένων και την περιγραφή τους, όπως αναφέρεται στην βιβλιογραφία. [19]

1.2.2 Τύποι δεδομένων σε τράπεζα βιολογικών δεδομένων

Κλινικά δεδομένα

Τα κλινικά δεδομένα περιλαμβάνουν πληροφορίες από το ιατρικό ιστορικό του ασθενούς, όπως διαγνώσεις, θεραπείες, αποτελέσματα εξετάσεων και την παρακολούθηση της πορείας της υγείας του. Αυτά τα δεδομένα βοηθούν τη συσχέτιση βιολογικών δειγμάτων με κλινικές εκβάσεις, βελτιώνοντας την κατανόηση των παραγόντων που επηρεάζουν μια νόσο και την ανταπόκριση στις θεραπείες.[20]

Δημογραφικά δεδομένα

Τα δημογραφικά δεδομένα αφορούν τα βασικά χαρακτηριστικά των δοτών/ασθενών, όπως η ηλικία, το φύλο, η εθνικότητα/φυλή, η οικογενειακή κατάσταση, η κοινωνικοοικονομική κατάσταση και ο τόπος διαμονής. Τα δημογραφικά στοιχεία είναι κρίσιμα για την ανάλυση αποτελεσμάτων, καθώς επιτρέπουν στους ερευνητές να εξετάζουν διαφορές μεταξύ πληθυσμιακών ομάδων και να προσαρμόζουν τις μελέτες πάνω σε συγκεκριμένους παράγοντες.[20]

Δεδομένα τρόπου ζωής

Τα δεδομένα τρόπου ζωής αφορούν τις καθημερινές συνήθειες και συμπεριφορές των συμμετεχόντων που επηρεάζουν την υγεία. Συλλέγονται συνήθως μέσω ερωτηματολογίων ή συνεντεύξεων κατά την ένταξη στη βιοτράπεζα ή σε μεταγενέστερες αξιολογήσεις. Σε αυτά

περιλαμβάνονται η διατροφή, η σωματική άσκηση, οι καπνιστικές συνήθειες, η κατανάλωση αλκοόλ και άλλες σχετικές συμπεριφορές.[21]

Περιβαλλοντικά δεδομένα

Τα περιβαλλοντικά δεδομένα καταγράφουν την έκθεση ενός ατόμου σε εξωτερικούς περιβαλλοντικούς παράγοντες καθώς και πληροφορίες για το περιβάλλον διαβίωσής του. Παραδείγματα τέτοιων δεδομένων είναι τα επίπεδα ατμοσφαιρικής ρύπανσης στην περιοχή κατοικίας, η έκθεση σε θόρυβο, οι κλιματολογικές συνθήκες, ή η παρουσία ρύπων σε νερό και έδαφος.[22]

Γονιδιακά και μοριακά δεδομένα

Οι βιοτράπεζες αποτελούν πολύτιμες πηγές γενετικών και άλλων μοριακών δεδομένων από τα αποθηκευμένα δείγματα. Συχνά διενεργούνται αναλύσεις ενΔΝΑ σε μεγάλη κλίμακα, παρέχοντας δεδομένα γονιδιωματικής (π.χ. αλληλούχιση γονιδιώματος) και "omics" (δηλαδή, δεδομένα που σχετίζονται με επιστημονικούς τομείς που μελετούν σύνολα βιολογικών μορίων, όπως το γενετικό υλικό (genomics), τις πρωτεΐνες (proteomics), τους μεταβολίτες (metabolomics) και άλλα, με στόχο την κατανόηση των συστημάτων και διαδικασιών σε ολόκληρο τον οργανισμό[23]). [21, 2, 24]

Δεδομένα απεικονιστικών τεχνικών

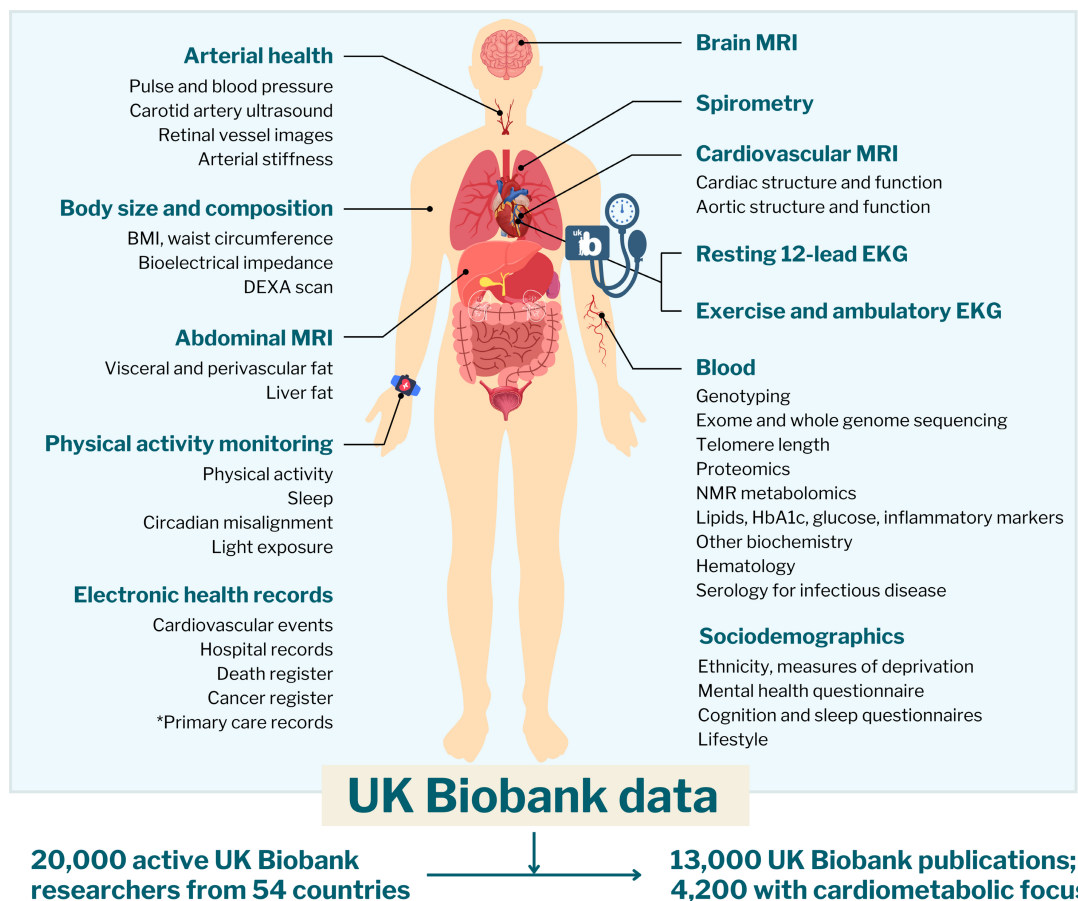
Η ιατρική απεικόνιση αποτελεί πλέον μέρος αρκετών μεγάλων βιοτραπεζών, προσθέτοντας πλούτο πληροφοριών για τη μορφολογία και λειτουργία οργάνων. Οι συμμετέχοντες μπορεί να υποβάλλονται σε εξετάσεις όπως MRI, CT, PET, υπερηχογραφήματα ή άλλες μεθόδους, και οι ψηφιακές εικόνες αποθηκεύονται στη βιοτράπεζα για ερευνητική χρήση.[21, 24]

Δεδομένα από φορετές συσκευές

Η εξάπλωση των φορητών συσκευών υγείας (wearables) έχει επιτρέψει τη συλλογή δεδομένων φυσιολογίας σε πραγματικό χρόνο από τους συμμετέχοντες, εκτός του κλινικού περιβάλλοντος. Με αυτόν τον τρόπο γίνεται καταγραφή με συνεχόμενο ή περιοδικό τρόπο μετρήσεις όπως ο καρδιακός ρυθμός, η αρτηριακή πίεση, η φυσική δραστηριότητα, ο ύπνος και άλλα χαρακτηριστικά της υγείας. [21, 2, 24]

Ψηφιακά δεδομένα παθολογίας

Οι βιοτράπεζες πλέον εντάσσουν και ψηφιοποιημένες εικόνες ιστών και άλλα παθολογοανατομικά δεδομένα στα αρχεία τους. Αυτά τα δεδομένα επιτρέπουν τη δυνατότητα στους ερευνητές να αναλύουν μορφολογικά χαρακτηριστικά των ιστών με τη βοήθεια υπολογιστικών εργαλείων και να τα συσχετίζουν με άλλα δεδομένα του ασθενούς.[2]



Εικόνα 1.2: UK Biobank: Περιγραφή δεδομένων [2]

Η εικόνα 1.2 περιγράφει εν συντομία τα δεδομένα της UK Biobank που περιέχει το μεγαλύτερο σύνολο των δεδομένων που περιγράφηκαν προηγουμένως.

1.3 Ευρωπαϊκή πολιτική προστασίας και διαχείρισης προσωπικών δεδομένων

Η διαχείριση προσωπικών δεδομένων στις βιοτράπεζες καθορίζεται από το ευρωπαϊκό νομικό πλαίσιο του GDPR, το οποίο θέτει θεμελιώδεις κανόνες για την προστασία και επεξεργασία προσωπικών και ευαίσθητων πληροφοριών. Τα διεθνή πρότυπα ασφάλειας και ποιότητας (ISO 27001, ISO 27799, ISO 20387) παρέχουν τις τεχνικές και οργανωτικές απαιτήσεις για την εφαρμογή αυτών των κανόνων στην πράξη, διασφαλίζοντας έτσι την εμπιστευτικότητα, την ακεραιότητα και τη διαθεσιμότητα των δεδομένων. Οι αρχές FAIR προωθούν την ευρεσιμότητα και τη διαλειτουργικότητα, ενισχύοντας την αξιοποίηση και επαναχρησιμοποίηση των δεδομένων για ερευνητικούς σκοπούς. Τέλος, υποδομές όπως το ELIXIR παρέχουν τα απαραίτητα εργαλεία και την τεχνογνωσία, ώστε όλα αυτά να εφαρμοστούν με τρόπο συνεπή, αποτελεσματικό και συμβατό με τις απαιτήσεις της νομοθεσίας και των προτύπων. Έτσι, διαμορφώνεται ένα ολοκληρωμένο και ασφαλές οικοσύστημα για τη διαχείριση των βιοϊατρικών δεδομένων, το οποίο προάγει την έρευνα και συμβάλλει στην ενίσχυση της δημόσιας υγείας.

1.3.1 (General Data Protection Regulation - GDPR)

Ο General Data Protection Regulation (GDPR), που καθορίζεται από τον Κανονισμό (ΕΕ) 2016/679, αποτελεί το πανευρωπαϊκό νομικό πλαίσιο για την προστασία προσωπικών δεδομένων και εφαρμόζεται και σε βιοτράπεζες που διαχειρίζονται προσωπικές και ευαίσθητες πληροφορίες δοτών.[25]

Για να είναι νόμιμη η επεξεργασία αυτών των δεδομένων, απαιτείται να πληρούνται συγκεκριμένες προϋποθέσεις που προβλέπονται από κανονισμούς όπως (GDPR), μεταξύ των οποίων η ρητή συγκατάθεση του δότη ή, κατ' εξαίρεση, η επεξεργασία για ερευνητικούς σκοπούς βάσει του άρθρου 9(2)(j) του GDPR. Επιπλέον, κατά την επεξεργασία πρέπει να τηρούνται κάποιες θεμελιώδεις αρχές όπως:

- **Νομιμότητα και διαφάνεια:** Η επεξεργασία των δεδομένων πρέπει να γίνεται νόμιμα με πλήρη ενημέρωση του ασθενή δότη.
- **Σκοπός συλλογής των δεδομένων:** Τα δεδομένα πρέπει να συλλέγονται για σαφώς καθορισμένους, συγκεκριμένους και νόμιμους σκοπούς.
- **Συλλογή των απαραίτητων δεδομένων:** Να συλλέγονται μόνο οι απολύτως απαραίτητες πληροφορίες που είναι απαραίτητες για τον επιδιωκόμενο σκοπό.
- **Ακρίβεια των δεδομένων:** Τα δεδομένα να είναι ακριβή και, όπου χρειάζεται, να ενημερώνονται διαρκώς.
- **Περιορισμός της περιόδου αποθήκευσης:** Να διατηρούνται μόνο για το διάστημα που απαιτείται για την εκπλήρωση των σκοπών της επεξεργασίας.
- **Ακεραιότητα και εμπιστευτικότητα:** Εφαρμογή κατάλληλων τεχνικών και οργανωτικών μέτρων για την προστασία από μη εξουσιοδοτημένη πρόσβαση, απώλεια ή καταστροφή των δεδομένων.

Ιδιαίτερη προσοχή δίνεται στη διαδικασία συγκατάθεσης των δοτών. Οι βιοτράπεζες οφείλουν να παρέχουν σαφή ενημέρωση για το πώς θα αξιοποιηθούν τα δείγματα και τα προσωπικά δεδομένα (σύμφωνα με τα άρθρα 12-13 του GDPR) και να διασφαλίζουν ότι η συγκατάθεση πραγματοποιείται με πλήρη επίγνωση. Η συγκατάθεση για την επεξεργασία προσωπικών δεδομένων πρέπει να διαχωρίζεται σαφώς από τη συναίνεση για τη δωρεά του βιολογικού δείγματος.[25]

Επιπλέον, οι υπάρχουσες υποδομές πρέπει να παρέχουν στους δότες τη δυνατότητα να ασκούν τα δικαιώματά τους (όπως ανάκληση της συγκατάθεσης, πρόσβαση, διόρθωση ή διαγραφή των δεδομένων τους), εφαρμόζοντας παράλληλα αυστηρά μέτρα ασφαλείας και προστασίας της ιδιωτικότητας (για παράδειγμα κρυπτογράφηση).[25]

Ο GDPR απαιτεί επίσης την τεκμηρίωση της συμμόρφωσης μέσω συγκεκριμένων πολιτικών και αρχείων επεξεργασίας, ενώ συστήνει την ανάθεση του ρόλου του (Data Protection Officer (DPO)) σε οργανισμούς που επεξεργάζονται ευαίσθητα δεδομένα μεγάλης κλίμακας, όπως συμβαίνει στις βιοτράπεζες. Τέλος, σε περιπτώσεις εκτεταμένης επεξεργασίας γενετικών ή υγειονομικών δεδομένων, είναι απαραίτητη η διενέργεια Εκτίμησης Αντικτύπου σχετικά με

την Προστασία Δεδομένων (Data Protection Impact Assessment (DPIA)) πριν την έναρξη της επεξεργασίας, ώστε να εντοπιστούν και να περιοριστούν πιθανοί κίνδυνοι για την ιδιωτικότητα των δοτών.[25]

Συνολικά, ο GDPR θέτει ένα αυστηρό και ολοκληρωμένο πλαίσιο που διασφαλίζει τα δικαιώματα των ατόμων και επιβάλλει στις βιοτράπεζες την εφαρμογή ισχυρών μέτρων ασφαλείας, διαφάνειας και ηθικής συμμόρφωσης.

1.3.2 Το πρότυπο ISO 27001 και η εφαρμογή του στις βιοτράπεζες

Το πρότυπο **ISO/IEC 27001** αποτελεί το διεθνώς αναγνωρισμένο πρότυπο για την ανάπτυξη, εφαρμογή και πιστοποίηση ενός Συστήματος Διαχείρισης Ασφάλειας Πληροφοριών (Information Security Management System - ISMS). Το πρότυπο παρέχει στις βιοτράπεζες ένα ολοκληρωμένο πλαίσιο για την προστασία της εμπιστευτικότητας, της ακεραιότητας και της διαθεσιμότητας των πληροφοριών που διαχειρίζονται.

Το ISO 27001 βασίζεται σε προσέγγιση εκτίμησης κινδύνου (risk-based approach), σύμφωνα με το οποίο κάθε οργανισμός οφείλει να αξιολογεί συστηματικά τους κινδύνους για την ασφάλεια των δεδομένων του και να λαμβάνει τα απαραίτητα μέτρα προστασίας. Στις βασικές απαιτήσεις του προτύπου περιλαμβάνονται [26, 27]:

- **Εκτίμηση και διαχείριση κινδύνων:** Συστηματική αναγνώριση και περιορισμός των κινδύνων που απειλούν την εμπιστευτικότητα, ακεραιότητα και διαθεσιμότητα των δεδομένων.
- **Διαχείριση πρόσβασης (Access Control):** Πρόσβαση μόνο από εξουσιοδοτημένα άτομα στα ευαίσθητα δεδομένα.
- **Οργάνωση και ευθύνες:** Καθορισμός συγκεκριμένων ρόλων, αρμοδιοτήτων και καθηκόντων του προσωπικού σχετικά με την ασφάλεια πληροφοριών.
- **Κρυπτογράφηση και ψευδωνυμοποίηση:** Εφαρμογή μέτρων για την προστασία των δεδομένων κατά την αποθήκευση και μεταφορά.
- **Προστασία πληροφοριακών περιουσιακών στοιχείων:** Διαχείριση και προστασία των δεδομένων και υποδομών που τα φιλοξενούν.
- **Έλεγχος και παρακολούθηση:** Τακτική αξιολόγηση της αποτελεσματικότητας των μέτρων ασφαλείας και συνεχής βελτίωση.

Η υιοθέτηση και πιστοποίηση κατά ISO 27001 διασφαλίζει ότι μια βιοτράπεζα τηρεί υψηλά επίπεδα προστασίας δεδομένων και συμμορφώνεται αποτελεσματικά με απαιτήσεις κανονισμών όπως ο GDPR. Σημαντικοί βιοϊατρικοί οργανισμοί, όπως η UK Biobank, έχουν ήδη υιοθετήσει το συγκεκριμένο πρότυπο, ενισχύοντας έτσι την εμπιστοσύνη των δοτών και της επιστημονικής κοινότητας.[28]

1.3.3 Ασφάλεια Πληροφοριών Υγείας (ISO 27799)

Το πρότυπο **ISO 27799:2016** αποτελεί εξειδικευμένο πρότυπο ασφάλειας πληροφοριών που επεκτείνει το πρότυπο ISO 27001/27002 ειδικά για τον τομέα της υγείας. Ο πλήρης τίτλος του είναι «Πληροφορική υγείας-Διαχείριση ασφάλειας πληροφορίας στην υγεία με χρήση του ISO/IEC 27002». Το ISO 27799 παρέχει οδηγίες για την εφαρμογή των ελέγχων ασφαλείας του ISO 27002 σε περιβάλλοντα υγειονομικών υπηρεσιών, λαμβάνοντας υπόψη τις ειδικές προκλήσεις και απειλές που αντιμετωπίζουν τα δεδομένα υγείας.

Το ISO 27799 καθοδηγεί φορείς υγείας, όπως νοσοκομεία, βιοτράπεζες και ερευνητικά κέντρα, στην υιοθέτηση βέλτιστων πρακτικών που διασφαλίζουν την εμπιστευτικότητα, ακεραιότητα, διαθεσιμότητα και ανιχνευσιμότητα των πληροφοριών υγείας. Συγκεκριμένα, επισημαίνει απειλές όπως η μη εξουσιοδοτημένη πρόσβαση από εσωτερικούς χρήστες, η πλαστοπροσωπία και οι παραβιάσεις συστημάτων ιατρικών δεδομένων, προτείνοντας εξειδικευμένους ελέγχους για τέτοια σενάρια. Μεταξύ αυτών περιλαμβάνονται η αυστηρή διαχείριση δικαιωμάτων πρόσβασης στα ιατρικά πληροφοριακά συστήματα, ειδικές διαδικασίες χειρισμού ευαίσθητων δεδομένων ασθενών, μηχανισμοί καταγραφής και ελέγχου προσβάσεων.

Παρόλο που το ISO 27799 δεν αποτελεί ανεξάρτητο πιστοποιήσιμο πρότυπο και εφαρμόζεται συμπληρωματικά μαζί με το ISO 27001, η συμμόρφωση με τις οδηγίες του παρέχει ένα πλαίσιο ασφάλειας προσαρμοσμένο στις ιδιαίτερες ανάγκες των δεδομένων υγείας. Έτσι, το ISO 27799 συμβάλλει στην κάλυψη των απαιτήσεων του GDPR και άλλων κανονιστικών πλαισίων μέσω τεχνικών και οργανωτικών μέτρων ειδικά σχεδιασμένων για την προστασία ευαίσθητων βιοϊατρικών πληροφοριών.[29, 30]

1.3.4 Αρχές FAIR (Findable, Accessible, Interoperable, Reusable)

Οι αρχές **FAIR (Findable, Accessible, Interoperable, Reusable)** αποτελούν ένα διεθνώς αναγνωρισμένο σύνολο κατευθυντηρίων οδηγιών για τη διαχείριση και τον διαμοιρασμό επιστημονικών δεδομένων, με στόχο τη μέγιστη αξιοποίησή τους μέσω της επαναχρησιμοποίησης. Οι τέσσερις αρχές ορίζουν ότι τα δεδομένα πρέπει να είναι [31]:

- **Findable (Ευρέσιμα):** Για να διασφαλιστεί η εύκολη εντοπισσιμότητα των δεδομένων, πρέπει να διαθέτουν επαρκή μεταδεδομένα, να φέρουν μια μοναδική και διαρκή ταυτότητα και να έχουν καταχωρηθεί ή ευρετηριαστεί σε έναν εύκολα αναζητήσιμο κατάλογο.
- **Accessible (Προσβάσιμα):** Τα δεδομένα και τα μεταδεδομένα θα πρέπει να είναι προσβάσιμα τόσο σε ανθρώπους όσο και σε μηχανές και να αποθηκεύονται σε ένα αξιόπιστο και ασφαλές αποθετήριο.
- **Interoperable (Διαλειτουργικά):** Η χρήση κοινής δομής στα δεδομένα είναι απαραίτητη, ενώ τα μεταδεδομένα πρέπει να περιγράφονται με επίσημες και αναγνωρισμένες ορολογίες, ώστε να διευκολύνεται η ανταλλαγή πληροφοριών.
- **Reusable (Επαναχρησιμοποιήσιμα):** Για να υποστηριχθεί η μελλοντική αξιοποίηση, τα δεδομένα και οι συλλογές τους θα πρέπει να συνοδεύονται από σαφείς άδειες χρήσης

και πληροφορίες για την προέλευσή τους, τηρώντας τα πρότυπα που ορίζει η αντίστοιχη κοινότητα.

Στο πλαίσιο των βιοτραπεζών, οι αρχές FAIR είναι ιδιαίτερα κρίσιμες. Η τήρησή τους σημαίνει ότι τα δεδομένα από τα δείγματα και οι σχετικές πληροφορίες (π.χ. κλινικά δεδομένα δοτών, γενετικές αναλύσεις) μπορούν να είναι διαθέσιμα και αξιοποιήσιμα από πολλούς ερευνητές για διαφορετικές μελέτες, ενισχύοντας έτσι τη συνεργασία και την επιστημονική απόδοση. Για παράδειγμα, μια βιοτράπεζα που ακολουθεί τις αρχές FAIR, καταχωρεί τα δείγματά της σε ευρετήρια ώστε να είναι ευρέσιμα, παρέχει μεταδεδομένα σε τυποποιημένες μορφές (όπως το (MIABIS) Minimum Information About Biobank Data), επιτρέπει την πρόσβαση μέσω διαλειτουργικών πλατφορμών και ορίζει σαφείς όρους χρήσης.[32]

1.3.5 Πλαίσιο Διαχείρισης Δεδομένων ELIXIR

Το **ELIXIR** μια πανευρωπαϊκή υποδομή που υποστηρίζει τη διαχείριση δεδομένων βιοεπιστημών και προωθεί τις αρχές FAIR. Στόχος του είναι να παρέχει πρακτικές οδηγίες σε ερευνητές και βιοτράπεζες, ώστε να εφαρμόζουν αποτελεσματικά τα διεθνή πρότυπα διαχείρισης δεδομένων.

Το **ELIXIR Data Management Framework** βοηθά στην υλοποίηση των αρχών FAIR, παρέχοντας το Research Data Management Kit (RDMkit) [33], έναν διαδικτυακό οδηγό που καθοδηγεί τους χρήστες ανάλογα με τον κύκλο ζωής των δεδομένων και τους ρόλους τους. Επιπλέον, το ELIXIR παρέχει πλαίσια για τη συλλογή, τεκμηρίωση, αποθήκευση και κοινοποίηση των δεδομένων, προάγοντας τη διαλειτουργικότητα και την επαναχρησιμοποίηση.

Σε ευρωπαϊκό επίπεδο, έχει δημιουργηθεί ένα δίκτυο Συντονιστών Διαχείρισης Δεδομένων (ELIXIR Data Management Coordinators Network), που διευκολύνει την ανταλλαγή γνώσεων και την εναρμόνιση των πρακτικών διαχείρισης δεδομένων μεταξύ των χωρών [34]. Τέλος, το ELIXIR διασυνδέεται με μεγάλους βιοϊατρικούς αποθηκευτικούς χώρους (π.χ. Local EGA), διευκολύνοντας έτσι την κατάθεση, διαμοιρασμό και επαναχρησιμοποίηση δεδομένων, σύμφωνα με τον κανονισμό GDPR και διεθνή πρότυπα.

1.3.6 Πρότυπο για Βιοτράπεζες (ISO 20387:2018)

Το πρότυπο ISO 20387 αποτελεί το πρώτο διεθνές πρότυπο που καθορίζει γενικές απαιτήσεις για την ικανότητα, την αμεροληψία και τη συνεπή λειτουργία των βιοτραπεζών. Το πρότυπο αυτό λειτουργεί ως διεθνώς αναγνωρισμένο σημείο αναφοράς για την οργάνωση και τη διαχείριση ποιότητας των βιοτραπεζών και παρέχει τη δυνατότητα διαπίστευσης από ανεξάρτητους φορείς πιστοποίησης της συμμόρφωσης τους [25].

Οι απαιτήσεις του προτύπου ομαδοποιούνται σε πέντε βασικούς τομείς:

1. **Γενικές απαιτήσεις (General requirements):** Απαιτείται η διασφάλιση της αμεροληψίας, της εμπιστευτικότητας και του απορρήτου των πληροφοριών των δοτών, καθώς και η συμμόρφωση με την ισχύουσα νομοθεσία.
2. **Διοικητικές και δομικές απαιτήσεις (Structural requirements):** Περιλαμβάνουν τη δομή διακυβέρνησης της βιοτράπεζας, τους ρόλους, τις ευθύνες (π.χ. επιστημονική επιτροπή, υπεύθυνος ποιότητας, DPO) και τον τρόπο λειτουργίας τους.

3. **Απαιτήσεις πόρων (Resource requirements):** Καλύπτουν το προσωπικό (εκπαίδευση, προσόντα), τις εγκαταστάσεις (π.χ. χώροι φύλαξης δειγμάτων), τον κατάλληλο εξοπλισμό και τις προμήθειες (π.χ. συστήματα καταγραφής θερμοκρασιών και εξοπλισμός συντήρησης δειγμάτων).
4. **Απαιτήσεις διαδικασιών (Process requirements):** Αφορούν τις διαδικασίες συλλογής, επεξεργασίας, αποθήκευσης, διανομής και καταστροφής βιολογικών δειγμάτων, καθώς και των δεδομένων που τα συνοδεύουν. Απαιτούν σαφείς και καταγεγραμμένες πολιτικές για την προστασία των δεδομένων και τη διαχείριση της συγκατάθεσης των δοτών, με πλήρη διαφάνεια.
5. **Σύστημα Διαχείρισης Ποιότητας (Quality Management System-QMS):** Προβλέπει την καθιέρωση ενός συστήματος ποιότητας που περιλαμβάνει διαδικασίες εσωτερικού ελέγχου (audits), αξιολόγηση κινδύνων, παρακολούθηση απόδοσης, διορθωτικές ενέργειες και συνεχή βελτίωση.

Το πρότυπο ISO 20387 ουσιαστικά συνδυάζει στοιχεία από άλλα πρότυπα ποιότητας (π.χ. ISO 9001) και τεχνικές απαιτήσεις αντίστοιχες με εκείνες εργαστηριακών προτύπων (ISO 17025, ISO 15189). Απαιτεί την τήρηση τυποποιημένων διαδικασιών (Standard Operating Procedures - SOPs) για τη διασφάλιση της ομοιογένειας και της αξιοπιστίας των βιολογικών δειγμάτων και δεδομένων. Η συμμόρφωση με το πρότυπο ενισχύει την αξιοπιστία της βιοτράπεζας, εξασφαλίζει την προστασία των δεδομένων και διευκολύνει τη συμμόρφωση με το GDPR.[25]

1.4 Βιβλιογραφική ανασκόπηση εφαρμογής δικτύου σε τράπεζα βιολογικών δεδομένων

Οι βιοτράπεζες αποτελούν κρίσιμες υποδομές για τη συλλογή, αποθήκευση και διάθεση βιολογικών δειγμάτων. Την τελευταία δεκαετία έχουν αναπτυχθεί μοντέλα αρχιτεκτονικής για τη δικτύωση βιοτραπεζών, με στόχο την αποτελεσματική διαμοίραση δεδομένων. Παρακάτω γίνεται ανασκόπηση τεσσάρων βασικών κατηγοριών αρχιτεκτονικής [Κεντρική (Centralized), Αποκεντρωμένη (Decentralized), Κατανεμημένη (Distributed) και Ομοσπονδιακή (Federated)] με πραγματικά παραδείγματα υλοποίησης τους.

1.4.1 UK Biobank

Αποτελεί ένα από τα μεγαλύτερα παραδείγματα βιοτράπεζας που χρησιμοποιεί κεντροποιημένη αρχιτεκτονική δικτύου. Έχει συγκεντρώσει βιολογικό υλικό και εκτεταμένα φαινοτυπικά και γενετικά δεδομένα από περίπου 500.000 συμμετέχοντες ηλικίας 40-69 ετών σε μία κεντρική βάση δεδομένων. Όλα τα δείγματα και τα δεδομένα φυλάσσονται και αναλύονται κεντρικά, επιτρέποντας ενιαία στατιστική επεξεργασία και μελέτες μεγάλης κλίμακας. Χαρακτηριστικά, η κεντρική ανάλυση γενετικών δεδομένων της UK Biobank απέδωσε περίπου 96 εκατομμύρια γενετικούς δείκτες προς διερεύνηση.[35]

1.4.2 All of Us Research Program

Ένα εθνικό πρόγραμμα των National Institutes of Health (NIH) με στόχο τη συλλογή δεδομένων από 1.000.000 ατόμων, το οποίο υιοθετεί επίσης κεντρική αρχιτεκτονική. Οι συμμετέχοντες παρέχουν βιοδείγματα, κλινικές πληροφορίες και γενετικά δεδομένα, τα οποία αποθηκεύονται σε μία κεντρική πλατφόρμα cloud (Researcher Workbench) για πρόσβαση από ερευνητές.[36]

1.4.3 EuroBioBank

Το EuroBioBank (ιδρύθηκε το 2001) είναι το πρώτο λειτουργικό δίκτυο βιοτραπεζών στην Ευρώπη, επικεντρωμένο σε σπάνιες ασθένειες. Λειτουργεί ως "εικονική" βιοτράπεζα με έναν κεντρικό online κατάλογο δειγμάτων, ενώ τα ίδια τα δείγματα παραμένουν στις τοπικές βιοτράπεζες που τα διαθέτουν. Οι ερευνητές μπορούν μέσω του καταλόγου να εντοπίσουν περίπου 130.000 διαθέσιμα δείγματα (DNA, κύτταρα, ιστοί) και να αιτηθούν πρόσβαση από την αντίστοιχη βιοτράπεζα, εξασφαλίζοντας ταχεία πρόσβαση χωρίς μεταφορά όλων των δεδομένων σε ένα κεντρικό σημείο.[37, 38]

1.4.4 AminoChain (Decentralized Biobank Platform)

Πρόσφατα εμφανίστηκαν καινοτόμες προσεγγίσεις "Web3" που αξιοποιούν την τεχνολογία blockchain για πλήρως αποκεντρωμένες βιοτράπεζες. Ένα παράδειγμα είναι το startup AminoChain (2024), που αναπτύσσει πλατφόρμα βιοτράπεζας πάνω σε blockchain Layer 2, δηλαδή σε ένα δευτερεύον δίκτυο που λειτουργεί πάνω στο βασικό πρωτόκολλο blockchain (Layer 1) για να βελτιώσει την ταχύτητα των συναλλαγών, να μειώσει τα λειτουργικά έξοδα και να διευκολύνει την κλιμάκωση. Στόχος είναι η δημιουργία ενός δικτύου βιοτραπεζών όπου τα ιατρικά ιδρύματα μπορούν να διαμοιράζονται και να συναλλάσσονται δεδομένα με ασφάλεια, χωρίς κεντρικό διακομιστή, ενώ οι ασθενείς/δότες διατηρούν τον έλεγχο και την ιδιοκτησία των δεδομένων τους και ενδεχομένως λαμβάνουν αποζημίωση για τη χρήση τους. Η αποκεντρωμένη αυτή προσέγγιση ενισχύει τη διαφάνεια και την εμπιστοσύνη, καθώς τα δεδομένα δεν μετακινούνται σε κεντρική βάση, αλλά η πρόσβαση ελέγχεται μέσω έξυπνων συμβολαίων στο blockchain.[39]

1.4.5 Δίκτυο RD-Connect για Σπάνιες Νόσους

Το RD-Connect (2015) αποτελεί ένα ευρωπαϊκό πρόγραμμα που στοχεύει στην αντιμετώπιση της διάσπαρτης ύλης σε βιοτράπεζες σπανίων νοσημάτων. Μέσω της υλοποίησης μιας κατανεμημένης πλατφόρμας, το RD-Connect συνέδεσε βιοτράπεζες, μητρώα ασθενών και βάσεις δεδομένων, προσφέροντας έναν δυναμικό online κατάλογο όπου οι ερευνητές μπορούν να εντοπίσουν βιολογικά δείγματα και τα αντίστοιχα κλινικά δεδομένα. Κάθε συμμετέχουσα βιοτράπεζα διατηρεί τα δείγματα τοπικά, ενώ τα μετα-δεδομένα τους ενοποιούνται και καθίστανται αναζητήσιμα μέσω της πλατφόρμας. Έτσι, η πλατφόρμα λειτουργεί ως γέφυρα μεταξύ διασκορπισμένων πόρων, συνδέοντας 22 ευρωπαϊκές βιοτράπεζες σπανίων παθήσεων και επιτρέποντας στους επιστήμονες την πρόσβαση σε πολύτιμα δείγματα μέσω

μιας ενιαίας διεπαφής, μέσα σε ένα ολοκληρωμένο οικοσύστημα που ενσωματώνει και τα κλινικά δεδομένα των ασθενών.[40]

1.4.6 BioSHaRE-EU - DataSHIELD

Το BioSHaRE-EU & DataSHIELD αποτελεί ένα χαρακτηριστικό παράδειγμα κατανεμημένης αρχιτεκτονικής για την κοινή ανάλυση δεδομένων από πολλές βιοτράπεζες. Στο πλαίσιο του έργου BioSHaRE-EU (2010-2015) συμμετείχαν 15 ευρωπαϊκές πληθυσμιακές βιοτράπεζες, με κύριο στόχο τη διευκόλυνση της εναρμόνισης και της ομοσπονδιακής ανάλυσης των δεδομένων τους χωρίς να μετακινούνται από τον τόπο φύλαξής τους. Για την επίτευξη αυτού του σκοπού αναπτύχθηκε μια ανοιχτή βιβλιοθήκη εργαλείων, μεταξύ των οποίων το Opal/Mica, που επιτρέπει την κατανεμημένη αποθήκευση και ευρετηρίαση των δεδομένων σε κάθε κόμβο, και το DataSHIELD, που επιτρέπει την ταυτόχρονη στατιστική ανάλυση σε πολλαπλές βιοτράπεζες χωρίς την αποκάλυψη ατομικών δεδομένων. Μέσω αυτών των εργαλείων, ένας ερευνητής μπορεί να εκτελέσει, για παράδειγμα, μια αναζήτηση σε όλες τις βιοτράπεζες και να λάβει μόνο τα συνολικά αποτελέσματα, διατηρώντας τα πρωτογενή δεδομένα τοπικά και ανώνυμα. Το BioSHaRE-EU απέδειξε ότι μια τέτοια κατανεμημένη υποδομή μπορεί να αποδώσει επιστημονικά αξιόλογα αποτελέσματα, όπως μελέτες για το άσθμα, την παχυσαρκία και τους καρδιαγγειακούς δείκτες, χωρίς την ανάγκη για κεντρική συγχώνευση των ευαίσθητων δεδομένων.[41, 42]

1.4.7 BBMRI-ERIC

Το BBMRI-ERIC είναι η μεγαλύτερη ομοσπονδιακή υποδομή βιοτραπεζών στην Ευρώπη, συνδέοντας πάνω από 20 χώρες και εκατοντάδες βιοτράπεζες σε ένα δίκτυο πολλαπλών κόμβων. Κάθε εθνικός κόμβος εγκαθιστά ένα τοπικό bridgehead server που φιλοξενεί τα σχετικά δεδομένα των δειγμάτων του, όπως οι τύποι δειγμάτων και βασικές κλινικές πληροφορίες, ενώ αυτοί οι bridgeheads συνδέονται με ένα κεντρικό σύστημα ευρετηρίων και αναζήτησης. Μέσω εργαλείων, όπως το Sample Locator του BBMRI, ένας ερευνητής μπορεί να εκτελέσει αποκεντρωμένη αναζήτηση, καθώς το ερώτημα διανέμεται σε όλους τους τοπικούς servers, οι οποίοι το επεξεργάζονται τοπικά και επιστρέφουν ανωνυμοποιημένα, συγκεντρωτικά αποτελέσματα (π.χ. ο αριθμός των κατάλληλων δειγμάτων που διατίθενται σε κάθε θέση). Σε περίπτωση που υπάρχουν αποτελέσματα, ο ερευνητής μπορεί μέσω του Negotiator tool να επικοινωνήσει και να αιτηθεί πρόσβαση στα επιμέρους δείγματα, χωρίς όμως τα δεδομένα να μετακινούνται από την τοπική υποδομή, παρά μόνο όταν εγκριθεί η πρόσβαση, διασφαλίζοντας τη συμμόρφωση με το GDPR και τις εθνικές νομοθεσίες.

Κατά την περίοδο 2021-2023, το BBMRI-ERIC αναβάθμισε την υποδομή του με την Federated Platform, η οποία ενσωματώνει τα εργαλεία Finder (για ομοσπονδιακή αναζήτηση κλινικο-γενετικών δεδομένων) και Locator (για αναζήτηση δειγμάτων). Με αυτόν τον τρόπο, το BBMRI-ERIC υλοποιεί ουσιαστικά μια "ενιαία εικονική βιοτράπεζα" σε πανευρωπαϊκό επίπεδο, χωρίς να μετακινεί ευαίσθητα δεδομένα.[43, 44]

1.4.8 Federated European Genome-Phenome Archive (FEGA)

Το Federated European Genome-Phenome Archive (FEGA) αποτελεί την ομοσπονδιακή εξέλιξη του Ευρωπαϊκού Αρχείου Γονιδιωμάτων-Φαινοτύπων. Το EGA είναι ένα μεγάλο αποθετήριο γενετικών/γονιδιωματικών δεδομένων ανθρώπου. Από το 2022, το σύστημα υλοποιείται σε ομοσπονδιακή μορφή, αντί όλα τα δεδομένα να φυλάσσονται κεντρικά στο EBI (Ηνωμένο Βασίλειο) ή στο CRG (Ισπανία), λειτουργούν πλέον εθνικοί κόμβοι EGA (π.χ. Σουηδία, Ιταλία, Φινλανδία, κ.ά.) που αποθηκεύουν τα δεδομένα σύμφωνα με τους δικούς τους νόμους, ενώ είναι διασυνδεδεμένοι σε ένα ενιαίο δίκτυο FEGA.

Το Federated EGA παρέχει ένα κεντρικό portal αναζήτησης όπου οι ερευνητές μπορούν να εντοπίσουν datasets ανεξαρτήτως του τόπου φιλοξενίας τους. Η πρόσβαση στα δεδομένα πραγματοποιείται μέσω των αντίστοιχων εθνικών κόμβων, αφού έχει προηγηθεί έγκριση, διασφαλίζοντας ότι τα δεδομένα δεν μετακινούνται εκτός της δικαιοδοσίας τους χωρίς άδεια. Μέχρι τις αρχές του 2024, το δίκτυο FEGA περιλάμβανε 7 εθνικούς κόμβους στην Ευρώπη, με επιπλέον χώρες υπό ένταξη.

Το εγχείρημα έχει καθιερωθεί ως το παγκοσμίως μεγαλύτερο ομοσπονδιακό αποθετήριο ευαίσθητων βιοϊατρικών δεδομένων, με επέκταση και εκτός Ευρώπης (π.χ. το 2025 προσχώρησε ο Canada).[45, 46]

1.4.9 GA4GH Beacon Networks

Το GA4GH Beacon Networks (Global Alliance) αποτελεί παγκόσμια πρότυπα για ομοσπονδιακή αναζήτηση γενετικών δεδομένων. Ο Παγκόσμιος Συνασπισμός για Γενετική και Υγεία (GA4GH) πρότεινε το Beacon API ως έναν ελαφρύ μηχανισμό ομοσπονδιακής διάθεσης γενετικών δεδομένων. Ένας Beacon είναι μια υπηρεσία σε έναν κόμβο (π.χ. μια βιοτράπεζα ή βάση δεδομένων) που απαντάει σε πολύ απλές ερωτήσεις, π.χ.: «Έχετε το αλληλόμορφο X σε κάποιο δείγμα»-με απάντηση yes/no. Το Beacon v1 (2015) απέδειξε ότι ιδρύματα διεθνώς μπορούσαν να συμμετέχουν σε ένα δίκτυο απαντώντας σε τέτοια αιτήματα χωρίς να αποκαλύπτουν περαιτέρω πληροφορίες, λύνοντας το πρόβλημα του διαμοιρασμού ευαίσθητων δεδομένων (ιδιαίτερα σημαντικό υπό το GDPR). Η έκδοση Beacon v2 (2022) επέκτεινε τη λειτουργικότητα ώστε να υποστηρίζονται πιο πολύπλοκα ερωτήματα με φαινοτυπικά κριτήρια και handover σε ασφαλείς πλατφόρμες για περαιτέρω ανάλυση. Επιπλέον, τα δίκτυα Beacon έχουν ήδη υλοποιηθεί σε πολλά έργα, όπως το δίκτυο ELIXIR Beacon στην Ευρώπη και το Matchmaker Exchange για σπάνιες ασθένειες, που συνδέει κλινικές βάσεις μέσω Beacon-like APIs. Αυτά τα δίκτυα λειτουργούν ως μία κοινή γλώσσα επικοινωνίας για την ενοποίηση διαφορετικών συστημάτων διαμοιρασμού δεδομένων.[47]

1.4.10 CINECA

Οι διεθνείς ομοσπονδίες μεγάλων cohorts και οι εφαρμογές ομοσπονδιακής ανάλυσης σε πληθυσμιακές μελέτες αποτελούν μια σημαντική τάση στην επιστημονική έρευνα τα τελευταία χρόνια. Παράδειγμα αποτελεί το έργο CINECA (Common Infrastructure for National Cohorts in Europe, Canada and Africa, 2019-2022), που συνέδεσε 10 μεγάλες πληθυσμιακές μελέτες, αθροίζοντας περίπου 1,4 εκατομμύρια συμμετέχοντες, σε ένα εικο-

νικό ομοσπονδιακό δίκτυο. Το CINECA αξιοποίησε πρότυπα του GA4GH (όπως Beacon, OAuth2/ELIXIR AAI, htsget κ.ά.) και εργαλεία όπως το Federated Analysis Jupyter environment, επιτρέποντας στους ερευνητές να εκτελούν αναλύσεις (π.χ. μελέτες συσχέτισης γονιδιώματος-φαινοτύπου) χωρίς να μετακινούν τα πρωτογενή δεδομένα από τις χώρες προέλευσης. Παρομοίως, η ευρωπαϊκή πρωτοβουλία 1+Million Genomes (1+MG) υιοθετεί ομοσπονδιακή φιλοσοφία για τη δικτύωση των εθνικών βάσεων γονιδιωματικών δεδομένων, ενώ δίκτυα όπως το OHDSI και το Pediatric Cancer Data Commons εφαρμόζουν ομοσπονδιακές προσεγγίσεις στην ανάλυση κλινικών δεδομένων. Όλες αυτές οι προσπάθειες συγκλίνουν στην τάση ότι η ανάλυση “πηγαίνει” στα δεδομένα, αντί να μετακινούνται τα δεδομένα προς τον αναλυτή, αυξάνοντας έτσι τη διαθεσιμότητα μεγαλύτερων και πιο ετερογενών συνόλων δεδομένων για έρευνα, ενώ διατηρείται παράλληλα η συμμόρφωση με το ρυθμιστικό πλαίσιο κάθε περιοχής.[48]

1.4.11 Συγκεντρωτικός πίνακας δικτύων σε βιοτράπεζα

Πίνακας 1.3: Συγκεντρωτικός πίνακας δικτύων σε βιοτράπεζα

UK Biobank	
Network Model	Centralized
Samples/Participants	~500,000 participants
Category	Population-based
All of Us Research Program	
Network Model	Centralized
Samples/Participants	1,000,000 (target) [>413,450 participants in 2023 [49]]
Category	Population-based
EuroBioBank	
Network Model	Decentralized
Samples/Participants	>130,000 samples
Category	Disease-specific (rare diseases)
AminoChain	
Network Model	Decentralized
Samples/Participants	-
Category	Virtual
RD-Connect	
Network Model	Federated
Samples/Participants	-
Category	Disease-specific (rare diseases)
BioSHaRE-EU & DataSHIELD	
Network Model	Federated

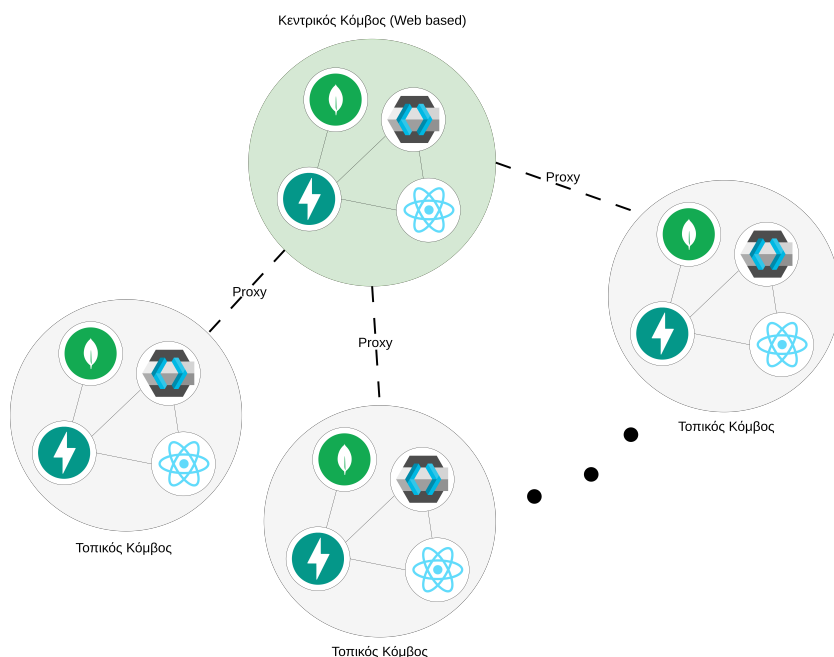
Samples/Participants	15 cohorts
Category	Population-based
BBMRI-ERIC	
Network Model	Federated
Samples/Participants	>60,000,000 samples [50]
Category	Virtual
Federated European Genome-Phenome Archive (FEGA)	
Network Model	Federated (national EGA nodes)
Samples/Participants	-
Category	Genetic (DNA/RNA)
GA4GH Beacon Network	
Network Model	Federated
Samples/Participants	>100,000 samples [51]
Category	Virtual (genomic discovery)
CINECA	
Network Model	Federated
Samples/Participants	~1,400,000 participants
Category	Virtual (federated cohort)

Κεφάλαιο 2

Πρακτικό Μέρος

2.1 Αρχιτεκτονική Δικτύου

Το σύστημα χαρακτηρίζεται ως **ομοσπονδιακό** διότι συνδυάζει την αποκεντρωμένη διαχείριση των δεδομένων με την κεντρική ενοποίηση της πληροφορίας. Κάθε φορέας διατηρεί τον δικό του τοπικό κόμβο, όπου αποθηκεύει και επεξεργάζεται τα δεδομένα του σύμφωνα με τις δικές του πολιτικές και νομικές απαιτήσεις, εξασφαλίζοντας την τοπική αυτονομία και την ασφάλεια. Παράλληλα, ένας κεντρικός κόμβος συγκεντρώνει τα metadata όλων των τοπικών κόμβων, επιτρέποντας μια ενοποιημένη πρόσβαση και αναζήτηση στα δεδομένα, χωρίς να παραβιάζεται η ανεξαρτησία κάθε φορέα. Η υιοθέτηση μηχανισμών όπως οι reverse proxies και authenticator servers διασφαλίζει ότι μόνο εξουσιοδοτημένοι χρήστες έχουν πρόσβαση, ενώ οι ασφαλείς σύνδεσμοι (π.χ. HTTPS, TLS) προσφέρουν ευελιξία στην επικοινωνία μεταξύ των ανεξάρτητων συστημάτων, χωρίς να απαιτείται κεντρική αποθήκευση των πρωτογενών δεδομένων. Μέσω αυτών των χαρακτηριστικών, το σύστημα διατηρεί την τοπική αυτονομία και παράλληλα υποστηρίζει μια ομοσπονδιακή αρχιτεκτονική με ολοκληρωμένη διαχείριση και αναζήτηση μεταδεδομένων.



Σχήμα 2.1: Περιγραφή της γενικής αρχιτεκτονικής μαζί με τις τεχνολογίες

2.1.1 Τοπικοί Κόμβοι

Κάθε συμμετέχων φορέας στο δίκτυο (π.χ. νοσοκομείο, ερευνητικό κέντρο) διατηρεί έναν τοπικό κόμβο, στον οποίο αποθηκεύονται τα βιοϊατρικά δεδομένα (π.χ. κλινικά, γενετικά, κ.λπ.). Ο συγκεκριμένος κόμβος περιλαμβάνει:

- μία τοπική βάση δεδομένων,
- τρεις διεπαφές (services) για τις διάφορες αλληλεπιδράσεις με τη βάση, και
- έναν authenticator server.

Οι τοπικές βιοτράπεζες έχουν την ευθύνη της λειτουργίας του κόμβου και διατηρούν πλήρη έλεγχο στα δεδομένα τους εφαρμόζοντας πολιτικές σύμφωνα με το GDPR (π.χ. ψευδωνυμοποίηση δεδομένων πριν από την αποθήκευση).

Οι παραπάνω τρεις διεπαφές (services) υλοποιούν ουσιαστικά:

1. μία εφαρμογή που συνδέει τη βάση δεδομένων με το ιατρικό προσωπικό του φορέα,
2. μία εφαρμογή διαχείρισης συγκαταθέσεων (δότη/ασθενούς) και
3. ένα queries service που εκτελεί ερωτήματα προς τη βάση δεδομένων.

2.1.2 Κεντρικός Κόμβος

Ο κεντρικός κόμβος αποτελεί τον σύνδεσμο του δικτύου, ο οποίος ενώνει τους ερευνητές με τους φορείς μέσω των μεταδεδομένων που έχουν καταχωρηθεί από τους τοπικούς κόμβους και μέσω της εκτέλεσης των queries service. Συγκεκριμένα, ο κεντρικός κόμβος περιλαμβάνει:

- μία κεντρική διεπαφή αναζήτησης μέσω της οποίας οι ερευνητές έχουν τη δυνατότητα να αναζητούν δείγματα και δεδομένα από όλους τους τοπικούς κόμβους,
- έναν authenticator server που εξασφαλίζει την αυθεντικοποίηση και τον έλεγχο πρόσβασης σύμφωνα με τα απαιτούμενα πρότυπα

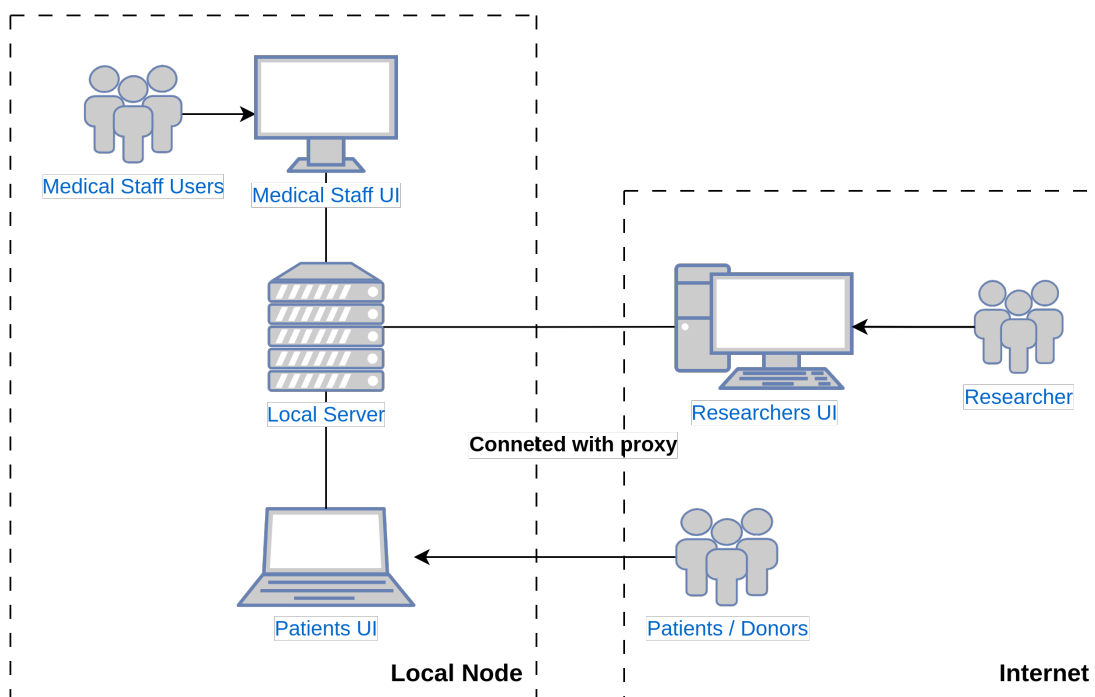
Με αυτόν τον τρόπο, ο κεντρικός κόμβος παρέχει μια ενιαία, ενοποιημένη εικόνα των δεδομένων, ενώ ταυτόχρονα διατηρείται ο τοπικός έλεγχος και η ασφάλεια των βιοϊατρικών πληροφοριών σύμφωνα με τις σχετικές πολιτικές και απαιτήσεις.

2.1.3 Μηχανισμός Σύνδεσης Τοπικού και Κεντρικού Κόμβου

Ο τρόπος σύνδεσης μεταξύ τοπικού και κεντρικού κόμβου πραγματοποιείται ως εξής:

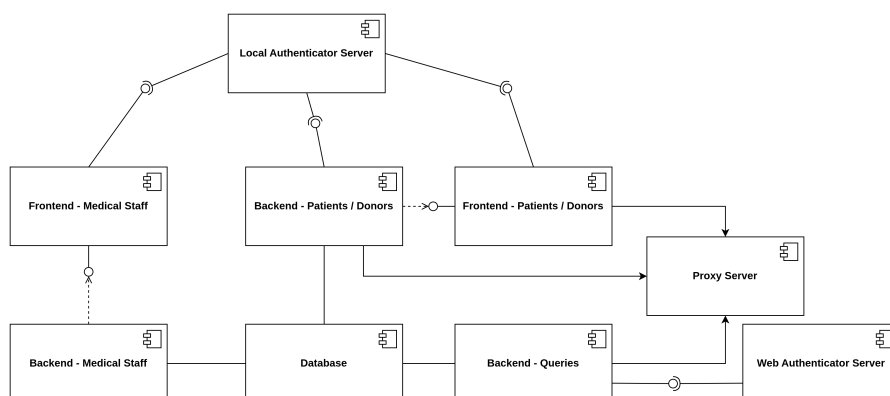
1. **Δημιουργία Connection Node:** Ο διαχειριστής του κεντρικού συστήματος χρησιμοποιεί μια ειδική διεπαφή για να δημιουργήσει μια εγγραφή σύνδεσης. Σε αυτή τη διαδικασία, καταχωρείται το URL του τοπικού κόμβου και εισάγεται ένα secret key που έχει συμφωνηθεί από κοινού με τον φορέα. Η εγγραφή αυτή αποθηκεύεται στην κεντρική βάση δεδομένων και χρησιμεύει για την επαλήθευση των αιτημάτων.

2. **Χρήση Reverse Proxy:** Οι τοπικοί κόμβοι δεν εκτίθενται απευθείας στο διαδίκτυο για λόγους ασφαλείας. Αντί αυτού, ο κάθε τοπικός κόμβος διαθέτει έναν reverse proxy ο οποίος αναλαμβάνει να προωθεί τις εξωτερικές αιτήσεις στην τοπική εφαρμογή. Αυτή η προσέγγιση διασφαλίζει ότι όλες οι επικοινωνίες μεταξύ του κεντρικού συστήματος και του τοπικού κόμβου γίνονται με αυστηρούς ελέγχους πρόσβασης, με χρήση κρυπτογράφησης και επαλήθευσης μέσω του secret key.
3. **Εκτέλεση Ερωτημάτων:** Η εφαρμογή των ερευνητών στον κεντρικό κόμβο υποβάλλει ερωτήματα (queries) μέσω της κεντρικής διεπαφής. Αυτά τα ερωτήματα μεταφέρονται μέσω των reverse proxies στους αντίστοιχους τοπικούς κόμβους, όπου το queries service τα εκτελεί τοπικά στη βάση δεδομένων. Τα αποτελέσματα επιστρέφονται στον κεντρικό κόμβο, ο οποίος τα παρουσιάζει με έναν ενιαίο τρόπο στους ερευνητές.
4. **Αυθεντικοποίηση και Ενοποίηση:** Ο κεντρικός authenticator server (για παράδειγμα, με χρήση του Keycloak) διαχειρίζεται την αυθεντικοποίηση και εξουσιοδότηση των χρηστών (ερευνητών) τόσο στο κεντρικό όσο και στους τοπικούς κόμβους.



Σχήμα 2.2: Αρχιτεκτονική ενός τοπικού κόμβου με τον κεντρικό κόμβο

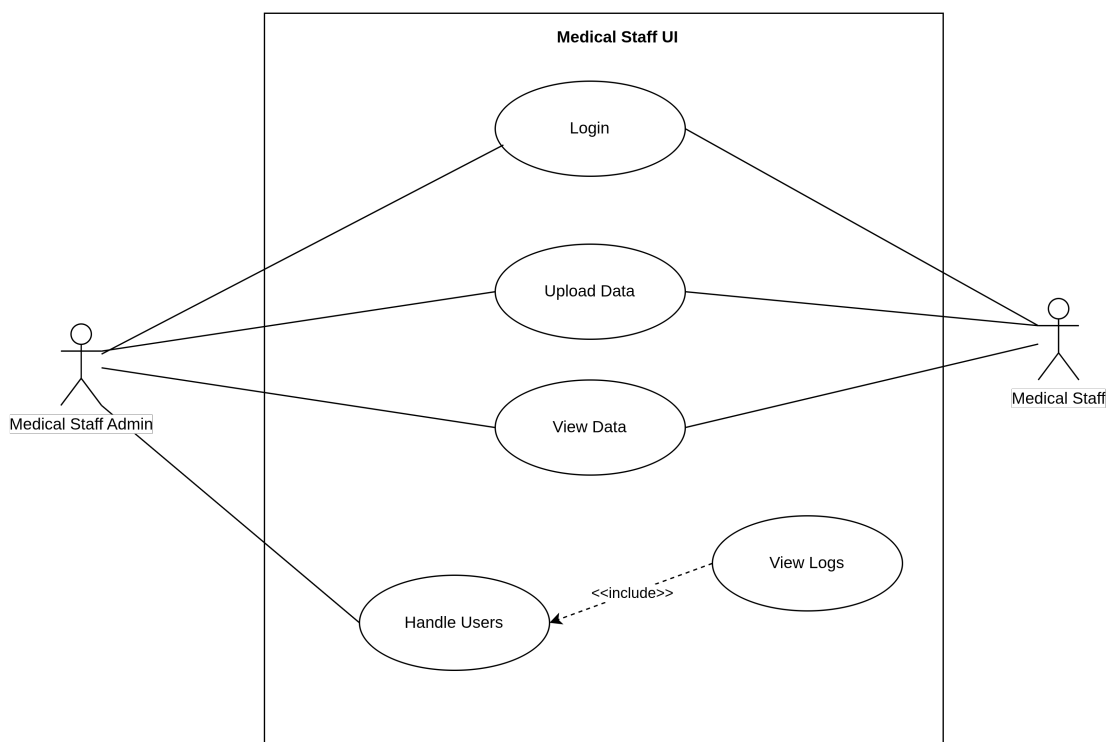
2.2 Περιγραφή Αρχιτεκτονικής Τοπικού Κόμβου



Σχήμα 2.3: Περιγραφή της αρχιτεκτονικής του τοπικού κόμβου

Η εικόνα 2.3 απεικονίζει την αρχιτεκτονική του τοπικού κόμβου, όπως έχει περιγραφεί παραπάνω, με την προσθήκη ενός Proxy Server που λειτουργεί ως ενδιάμεσος σύνδεσμος μεταξύ του τοπικού και του κεντρικού κόμβου, καθώς και ενός Web Authenticator Server που διαχειρίζεται την ταυτοποίηση των ερευνητών που συνδέονται στην web based εφαρμογή του κεντρικού κόμβου.

2.2.1 Διεπαφή Διασύνδεσης Ιατρικού Προσωπικού



Σχήμα 2.4: Λειτουργίες διεπαφής δασύνδεσης ιατρικού προσωπικού

Η εφαρμογή διασύνδεσης του ιατρικού προσωπικού αποτελεί το βασικό και πλέον καθοριστικό σύστημα για την καταχώρηση, την επεξεργασία και τη συνολική διαχείριση των βιοϊατρικών δεδομένων που σχετίζονται με τις δραστηριότητες του προσωπικού. Οι χρήστες του συστήματος έχουν τη δυνατότητα πρόσβασης μέσω του τοπικού Authenticator Server, ενός εξειδικευμένου διακομιστή ταυτοποίησης, στον οποίο έχουν προηγουμένως καταχωρηθεί τόσο τα credentials όσο και τα κρίσιμα προσωπικά τους στοιχεία από τον αρμόδιο administrator.

Κατόπιν επιτυχούς διαδικασίας ταυτοποίησης, τα μέλη του ιατρικού προσωπικού αποκτούν δικαιώματα πρόσβασης σε υπάρχουσες εγγραφές, ενώ τους παρέχεται επίσης η δυνατότητα εισαγωγής νέων δεδομένων στο σύστημα, συμβάλλοντας έτσι ενεργά στη συνεχή ενημέρωση της βάσης. Επιπρόσθετα, ο administrator, ως υπεύθυνος για τη συνολική εποπτεία του συστήματος, έχει διευρυμένα δικαιώματα και αρμοδιότητες. Συγκεκριμένα, έχει τη δυνατότητα παρακολούθησης της δραστηριότητας των υπολοίπων χρηστών (όπως για παράδειγμα των χρόνων σύνδεσης), καθώς και την ευχέρεια πλήρους διαχείρισης των επιμέρους προφίλ των μελών του προσωπικού.

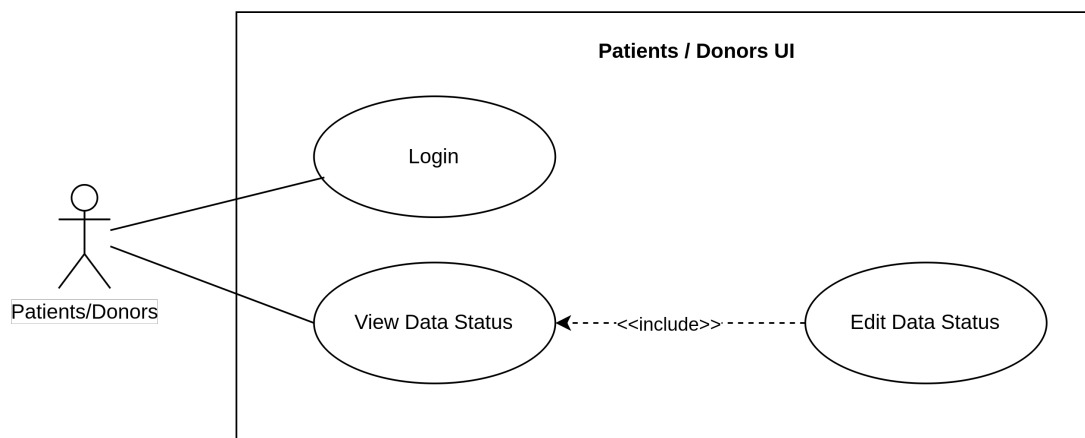
Η διαχείριση αυτή περιλαμβάνει ενέργειες όπως η δημιουργία νέων λογαριασμών, η επεξεργασία ήδη υπάρχοντων προφίλ και, εφόσον κρίνεται απαραίτητο, η απενεργοποίηση λογαριασμών, εξασφαλίζοντας την ασφάλεια και την κανονιστική συμμόρφωση της πλατφόρμας.

2.2.2 Διεπαφή Διαχείρισης Συγκαταθέσεων (Δότης/Ασθενής)

Το σύστημα διαχείρισης συγκατάθεσης έχει σχεδιαστεί ώστε να επιτρέπει την οργανωμένη και δομημένη καταγραφή, καθώς και τη συνεχή ενημέρωση των προσωπικών προτιμήσεων των συμμετεχόντων, είτε αυτοί είναι δότες είτε ασθενείς, σε ό,τι αφορά τη δευτερογενή χρήση των βιοϊατρικών και προσωπικών δεδομένων τους. Η υπηρεσία αυτή λειτουργεί σε federated επίπεδο, εξασφαλίζοντας ότι κάθε συνεργαζόμενος φορέας διατηρεί τοπικά το πλήρες σύνολο των εγγραφών συγκατάθεσης που του αντιστοιχούν, ενώ παράλληλα διασφαλίζεται ο συγχρονισμός με τον κεντρικό κόμβο. Αυτός ο συγχρονισμός ενημερώνει συνεχώς το κεντρικό σύστημα σχετικά με το τρέχον καθεστώς διαθεσιμότητας των δεδομένων, διατηρώντας έτσι συνεχή συνέπεια και διαφάνεια στην πρόσβαση και χρήση των πληροφοριών.

Η αλληλεπίδραση των χρηστών με το σύστημα πραγματοποιείται μέσω ενός web interface, η οποία φιλοξενείται στον τοπικό κόμβο και εκτίθεται στο διαδίκτυο μέσω ενός ενδιάμεσου proxy. Μέσω αυτής της διεπαφής, οι συμμετέχοντες έχουν τη δυνατότητα να τροποποιούν ανά πάσα στιγμή και σε πραγματικό χρόνο τη συγκατάθεσή τους, επιτυγχάνοντας έτσι μέγιστο έλεγχο επί των προσωπικών τους δεδομένων και των επιτρεπόμενων χρήσεών τους.

Πριν από οποιαδήποτε διαδικασία κοινοποίησης δεδομένων ή μεταδεδομένων σε τρίτους, το σύστημα διενεργεί αυστηρό έλεγχο, επιβεβαιώνοντας ότι υφίσταται έγκυρη και ενεργή συγκατάθεση για τον προβλεπόμενο σκοπό χρήσης. Με αυτόν τον τρόπο, εξασφαλίζεται η πλήρης συμμόρφωση με τις αυστηρές απαιτήσεις του GDPR, διαφυλάσσοντας παράλληλα τα δικαιώματα και την ιδιωτικότητα των χρηστών σε κάθε επίπεδο διαχείρισης.



Σχήμα 2.5: Λειτουργίες διεπαφής διαχείρισης συγκαταθέσεων (δότης/ασθενής)

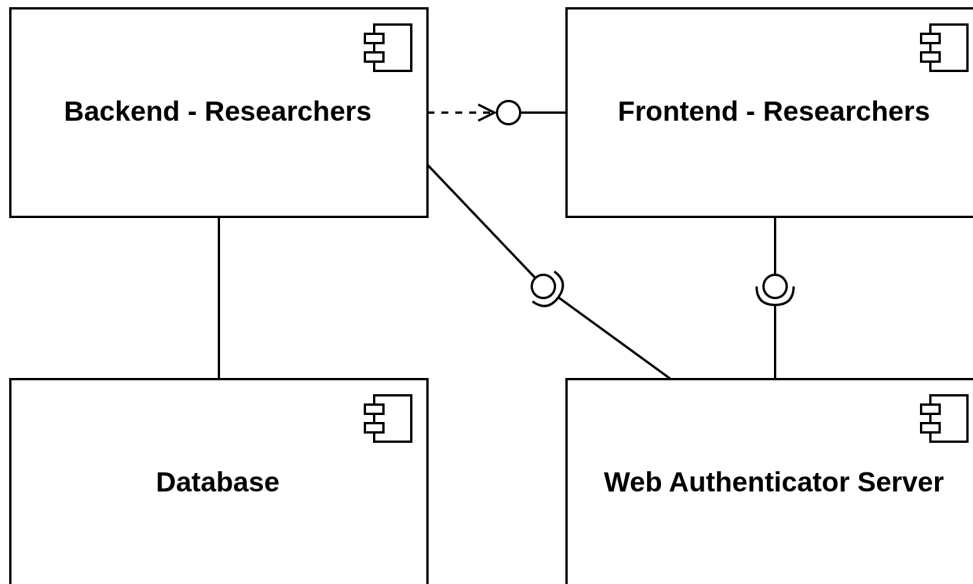
Ο δότης ή ασθενής πραγματοποιεί σύνδεση στην εφαρμογή μέσω του τοπικού Authenticator Server, ο οποίος αποτελεί τον αρμόδιο διακομιστή ταυτοποίησης. Στον Authenticator Server έχουν ήδη καταχωρηθεί τα διαπιστευτήρια (credentials) του χρήστη, καθώς και τα απαραίτητα προσωπικά δεδομένα, τα οποία συλλέγονται κατά την διαδικασία καταχώρησης δεδομένων στην βιοτράπεζα από τον αρμόδιο φορέα. Μετά την επιτυχή ολοκλήρωση της διαδικασίας σύνδεσης, στο χρήστη παρουσιάζονται όλες οι υποβληθείσες συγκαταθέσεις, είτε είναι ενεργές είτε ανενεργές, συνοδευόμενες από σαφή και ευδιάκριτη ένδειξη που υποδεικνύει την τρέχουσα κατάσταση της κάθε συγκατάθεσης.

2.2.3 Queries Service

Το συγκεκριμένο service παρέχει ένα σύνολο API endpoints τα οποία εκτελούν προκαθορισμένα και ελεγχόμενα ερωτήματα προς την τοπική βάση δεδομένων. Η πρόσβαση των ερευνητών στο σύστημα γίνεται αποκλειστικά μέσω του κεντρικού Web Authenticator Server, ενώ η επικοινωνία με τους τοπικούς κόμβους πραγματοποιείται μέσω reverse proxy. Με αυτόν τον τρόπο, τα πρωτογενή δεδομένα παραμένουν προστατευμένα και δεν εκτίθενται απευθείας εκτός του τοπικού κόμβου, εξασφαλίζοντας αυξημένα επίπεδα ασφάλειας.

Τα αποτελέσματα που επιστρέφονται στους ερευνητές περιορίζονται αποκλειστικά σε συνοπτικά ή ψευδωνυμοποιημένα στοιχεία, όπως για παράδειγμα τον αριθμό των δειγμάτων ή συγκεντρωτικά στατιστικά δεδομένα, διαφυλάσσοντας έτσι την εμπιστευτικότητα και προστατεύοντας την ιδιωτικότητα των υποκειμένων. Σε περίπτωση που απαιτείται η διαβίβαση πρωτογενών δεδομένων, η διαδικασία αυτή προϋποθέτει την προηγούμενη έγκριση τόσο από τον administrator του τοπικού κόμβου όσο και την επιβεβαίωση της συγκατάθεσης του δότη ή ασθενή.

2.3 Περιγραφή Αρχιτεκτονικής Κεντρικού Κόμβου



Σχήμα 2.6: Περιγραφή της αρχιτεκτονικής του κεντρικού κόμβου

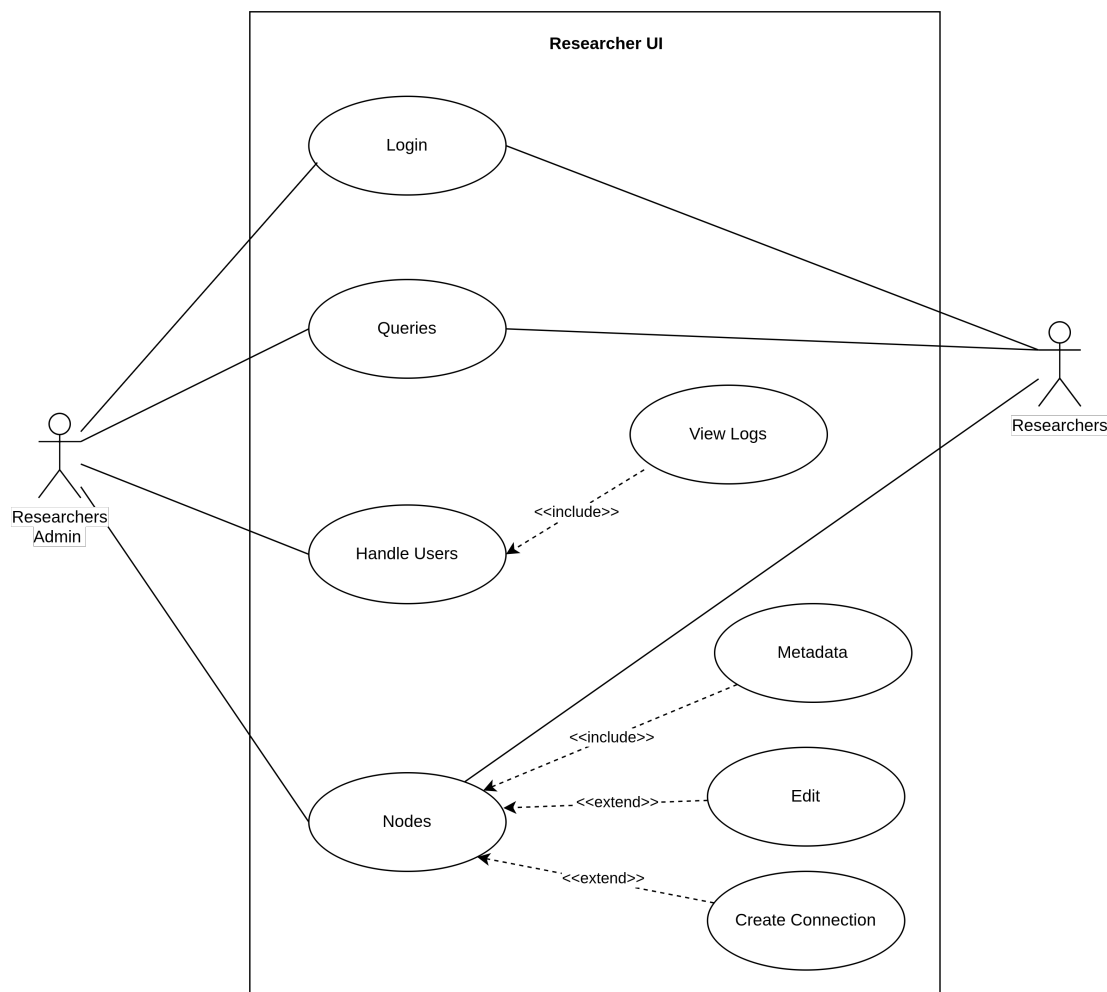
Η εικόνα 2.6 απεικονίζει με λεπτομέρεια την αρχιτεκτονική δομή του κεντρικού κόμβου, όπως έχει περιγραφεί στις προηγούμενες ενότητες. Η δομή αυτή εξασφαλίζει τη σωστή λειτουργία και διαχείριση του δικτύου, συνδυάζοντας ασφάλεια, ευελιξία και αποδοτικότητα.

2.3.1 Διεπαφή Αναζήτησης

Η πλατφόρμα προσφέρει μια κεντρική διαδικτυακή πύλη, η οποία δίνει τη δυνατότητα στους εγγεγραμμένους χρήστες, όπως επιστήμονες και κλινικούς ερευνητές, να εκτελούν ερωτήματα με στόχο την εντοπισμό διαθέσιμων δειγμάτων και δεδομένων σε ολόκληρο το δίκτυο.

Το ερώτημα αυτό διανέμεται αυτόματα σε όλους τους σχετικούς τοπικούς κόμβους του δικτύου, οι οποίοι με τη σειρά τους ανταποκρίνονται αναφέροντας αν διαθέτουν αντιστοιχίες στα αιτήματα. Η διεπαφή αναζήτησης συγκεντρώνει όλες αυτές τις απαντήσεις και τις παρουσιάζει με συνοπτικό και κατανοητό τρόπο, δίνοντας τη δυνατότητα στον ερευνητή να εντοπίσει γρήγορα και αποτελεσματικά ποιοι κόμβοι διαθέτουν το ζητούμενο υλικό.

Η συνοπτική αυτή παρουσίαση αφορά αποκλειστικά τα μεταδεδομένα (metadata), τα οποία στην υλοποίηση του συστήματος αντιστοιχούν στις πληροφορίες σχετικά με το είδος και το εύρος της πρόσβασης που έχουν εγκρίνει οι δότες. Με αυτόν τον τρόπο, τα metadata υποστηρίζουν και ενισχύουν τη συμμόρφωση με τις αρχές FAIR, καθιστώντας σαφή τα δεδομένα που είναι ευρέσιμα, προσβάσιμα και επαναχρησιμοποιήσιμα, πάντοτε με σεβασμό στις επιλογές των δοτών και στους όρους της συγκατάθεσής τους.



Σχήμα 2.7: Λειτουργίες διεπαφής αναζήτησης

Για τους administrators, η κεντρική πύλη παρέχει επιπλέον ειδικές λειτουργίες, οι οποίες περιλαμβάνουν:

- τη διαχείριση χρηστών και τον καθορισμό δικαιωμάτων πρόσβασης,
- την παρακολούθηση και καταγραφή της δραστηριότητας των χρηστών (logs),
- τη δημιουργία, συντήρηση και πλήρη διαχείριση των συνδέσεων με τους τοπικούς κόμβους.

2.4 Authenticator Servers

Για λόγους σαφούς διαχωρισμού ρόλων και ζωνών ασφαλείας, υλοποιούνται **δύο ουσιαστικά πανομοιότυποι πάροχοι ταυτότητας**, βασισμένοι στα πρωτόκολλα OAuth 2.0 και OpenID Connect:

Local Authenticator Server (LAS) Φιλοξενείται εντός του προστατευμένου δικτύου του φορέα και εξυπηρετεί αποκλειστικά το ιατρικό προσωπικό και τους ασθενείς.

Web Authenticator Server (WAS) Εκτίθεται δημόσια και απευθύνεται σε εξωτερικούς ερευνητές.

Οι δύο Authenticator Servers παρέχουν ένα κοινό σύνολο λειτουργιών:

- Έκδοση JWT access/refresh tokens με τυποποιημένα claims όπως sub, scope, role και aud.
- Self-service διαχείριση χρηστών από τους administrators, που περιλαμβάνει δημιουργία λογαριασμών, αλλαγή κωδικών και απενεργοποίηση χρηστών.
- Λεπτομερές audit logging των συνδέσεων και ενεργειών των χρηστών.

Σημείωση σχετικά με τα διαπιστευτήρια. Τα διαπιστευτήρια δεν αποθηκεύονται σε τοπικές βάσεις δεδομένων των επιμέρους υπηρεσιών. Κάθε εφαρμογή εμπιστεύεται τον αντίστοιχο Authenticator Server (LAS ή WAS) και βασίζεται αποκλειστικά σε επικυρωμένα tokens, αποφεύγοντας έτσι τον πολλαπλασιασμό και την ανεπιθύμητη διασπορά ευαίσθητων πληροφοριών.

Ροή ερευνητή

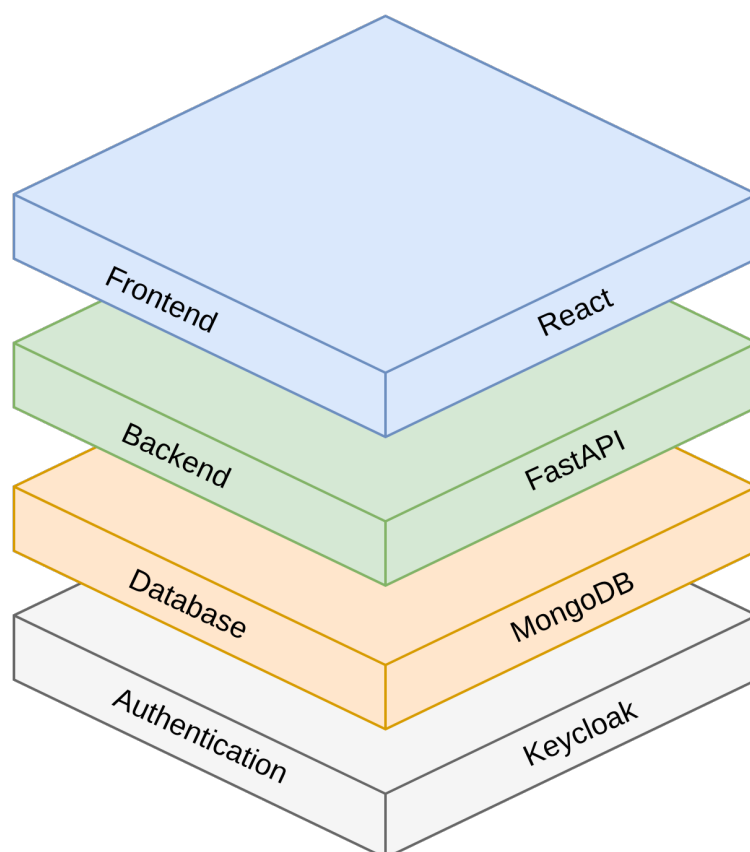
1. Ο ερευνητής συνδέεται στην κεντρική πύλη· ο WAS εκδίδει access token.
2. Η πύλη προωθεί (μέσω reverse proxy) το ερώτημα στον τοπικό Query Service, συμπεριλαμβάνοντας Authorization: Bearer <token>.
3. Ο Query Service επαληθεύει την αυθεντικότητα μέσω του token με το δημόσιο κλειδί του WAS (χωρίς να χρειάζεται κλήση στον ίδιο τον WAS).
4. Επιστρέφει συγκεντρωτικά αποτελέσματα τα οποία η πύλη παρουσιάζει στον ερευνητή.

Ροή προσωπικού/ασθενούς

1. Ο χρήστης αιτείται πρόσβαση σε εσωτερική εφαρμογή.
2. Ο LAS τον ταυτοποιεί (username/password) και εκδίδει access token.
3. Οι εσωτερικές υπηρεσίες ελέγχουν το token και αποδίδουν τα αντίστοιχα δικαιώματα.

2.5 Τεχνολογίες Ανάπτυξης Δικτύου

Οι τεχνολογίες που υιοθετήθηκαν χαρακτηρίζονται από ταχύτητα, ασφάλεια και ευκολία επεκτασιμότητας, ενώ έχουν ήδη αποδειχθεί αξιόπιστες σε πραγματικές παραγωγικές συνθήκες. Η προτεινόμενη software stack υποστηρίζει οριζόντια κλιμάκωση μέσω απλής προσθήκης νέων κόμβων, εξασφαλίζοντας έτσι ευέλικτη και αποτελεσματική ανάπτυξη του δικτύου. Επιπλέον, η αρχιτεκτονική ακολουθεί ανοιχτά πρότυπα, όπως REST APIs, που διευκολύνουν την ενσωμάτωση και τη διαλειτουργικότητα μεταξύ των συστημάτων.



Σχήμα 2.8: *Software stack του δικτύου*

2.5.1 Frontend-React

Η διεπαφή χρήστη (περιλαμβάνοντας την κεντρική πύλη, τις εφαρμογές για το ιατρικό προσωπικό και τη διαχείριση συγκαταθέσεων) υλοποιείται με React, αξιοποιώντας την component-based αρχιτεκτονική:

- Χρησιμοποιεί ένα σύγχρονο πλαίσιο για ανάπτυξη δυναμικών εφαρμογών «μίας σελίδας», όπου οι εναλλαγές οθονών γίνονται χωρίς πλήρη επαναφόρτωση του προγράμματος περιήγησης.
- Ενσωματώνει μηχανισμό δρομολόγησης που επιτρέπει καθαρό και ιεραρχικό διαχωρισμό των διαφόρων προβολών, διευκολύνοντας παράλληλα την εύκολη προσθήκη νέων λειτουργιών.
- Διαθέτει στρώμα έξυπνης ανάκτησης και προσωρινής αποθήκευσης δεδομένων (client-side caching), ώστε τα ήδη ληφθέντα αποτελέσματα από τα API να επαναχρησιμοποιούνται χωρίς περιττές κλήσεις στον διακομιστή, γεγονός που μειώνει την καθυστέρηση και την κίνηση στο δίκτυο.

2.5.2 Backend-FastAPI

Οι υπηρεσίες backend έχουν αναπτυχθεί χρησιμοποιώντας το FastAPI, ένα σύγχρονο και αποδοτικό πλαίσιο εργασίας (framework) της Python, το οποίο προσφέρει τα εξής πλεονεκτήματα :

- Υποστήριξη μηχανισμού *ασύγχρονης* επεξεργασίας αιτημάτων, που επιτρέπει την ταυτόχρονη εξυπηρέτηση πολλαπλών χρηστών με βελτιωμένη απόδοση και χαμηλό χρόνο απόκρισης, κάτι που είναι κρίσιμο για συστήματα με μεγάλο φόρτο.
- Αυτόματη δημιουργία και παροχή τεκμηρίωσης OpenAPI για όλα τα διαθέσιμα σημεία πρόσβασης (endpoints), επιτρέποντας στους προγραμματιστές και τα τρίτα συστήματα να κατανοήσουν εύκολα τη λειτουργία του backend και να δοκιμάσουν τις διεπαφές απευθείας από τον περιηγητή.
- Ενσωματωμένη αυστηρή επικύρωση και μορφοποίηση των δεδομένων εισόδου και εξόδου, γεγονός που ελαχιστοποιεί τα σφάλματα κατά την εκτέλεση και εξασφαλίζει την ορθή και ασφαλή διαχείριση των πληροφοριών.

2.5.3 Database-MongoDB

Η αποθήκευση του βιοϊατρικού περιεχομένου πραγματοποιείται με χρήση της MongoDB, μίας σύγχρονης βάσης δεδομένων τύπου NoSQL με τα εξής χαρακτηριστικά :

- Αποθηκεύει τα δεδομένα σε ευέλικτη μορφή εγγράφων JSON, επιτρέποντας διαφορετικά σχήματα και δομές ανά εγγραφή. Αυτό καθιστά τη βάση ιδανική για τη διαχείριση ετερογενών δεδομένων όπως δείγματα, συγκαταθέσεις ασθενών και αρχεία καταγραφής.
- Υποστηρίζει μηχανισμούς υψηλής διαθεσιμότητας, με αυτόματη δημιουργία αντιγράφων ασφαλείας (replication) και οριζόντια κλιμάκωση (sharding), επιτρέποντας τη διανομή μεγάλου όγκου δεδομένων σε πολλαπλούς κόμβους και διασφαλίζοντας την απρόσκοπτη λειτουργία.
- Προσφέρει εξελιγμένες δυνατότητες συγκεντρωτικής ανάλυσης (aggregation framework), που επιτρέπουν την εκτέλεση σύνθετων στατιστικών ερωτημάτων απευθείας στη βάση, χωρίς να απαιτείται μεταφορά ολόκληρων datasets στην εφαρμογή, ενισχύοντας την αποδοτικότητα και την ασφάλεια.

2.5.4 Reverse Proxy-Nginx

Ο Nginx χρησιμοποιείται ως reverse proxy για την εξυπηρέτηση των αιτημάτων προς τις υπηρεσίες, με τους εξής βασικούς ρόλους :

- Υλοποιεί ασφαλή τερματισμό κρυπτογραφημένων συνδέσεων μέσω HTTPS, διασφαλίζοντας την προστασία των δεδομένων κατά τη μεταφορά και εφαρμόζοντας βασικές πολιτικές ασφαλείας.

- Ανακατευθύνει τα εισερχόμενα αιτήματα στις κατάλληλες εσωτερικές υπηρεσίες, αποκρύπτοντας την πραγματική δομή και τοπολογία των συστημάτων, γεγονός που βελτιώνει την ασφάλεια και την ευελιξία.
- Παρέχει λειτουργίες περιορισμού ρυθμού αιτημάτων (rate limiting), προσωρινής αποθήκευσης (caching) επιλεγμένων απαντήσεων για βελτιστοποίηση της απόδοσης, καθώς και κεντρική καταγραφή συμβάντων (logging) που διευκολύνουν την παρακολούθηση και την ανάλυση της λειτουργίας του συστήματος.

2.5.5 Docker

Κάθε υπηρεσία του συστήματος πακετάρεται σε ανεξάρτητες εικόνες Docker, κάτι που προσφέρει σημαντικά οφέλη:

- Διασφαλίζει ότι το περιβάλλον εκτέλεσης (βιβλιοθήκες, ρυθμίσεις, εξαρτήσεις) είναι ακριβώς το ίδιο από το στάδιο της ανάπτυξης μέχρι και την παραγωγή, εξαλείφοντας προβλήματα ασυμβατότητας.
- Επιτυγχάνει πλήρη απομόνωση των εξαρτήσεων κάθε υπηρεσίας, αποτρέποντας συγκρούσεις λογισμικού και επιτρέποντας ευκολότερη διαχείριση και αναβάθμιση των επιμέρους συστατικών του συστήματος.

2.5.6 Keycloak

Οι Authenticator Servers (LAS και WAS) υλοποιούνται με το Keycloak, μια ολοκληρωμένη πλατφόρμα Identity & Access Management:

- Υποστηρίζει από προ-επιλογής τα πρότυπα OAuth 2.0, OpenID Connect και SAML 2.0, διευκολύνοντας τη σύνδεση με εξωτερικά συστήματα.
- Επιτρέπει λεπτομερή ορισμό ρόλων, ομάδων και scopes, ώστε κάθε χρήστης να διαθέτει μόνο τα απολύτως απαραίτητα δικαιώματα.
- Διαθέτει φιλικό Admin Console για γραφική διαχείριση χρηστών, κλειδιών και πολιτικών ασφάλειας.

Ο συνδυασμός των παραπάνω τεχνολογιών επιτρέπει την **ταχεία υλοποίηση** νέων λειτουργιών, αξιοποιώντας τις δημοφιλείς γλώσσες Python και JavaScript, ενώ ταυτόχρονα εξασφαλίζει **αυξημένη ασφάλεια** μέσω ενός ενιαίου και τυποποιημένου μηχανισμού ταυτοποίησης. Επιπλέον, η χρήση containers καθιστά εύκολη την αναπαραγωγή και εγκατάσταση κάθε υπηρεσίας στους φορείς, ενώ η MongoDB αναλαμβάνει την αυτόματη διανομή των δεδομένων σε πολλαπλούς κόμβους, επιτυγχάνοντας οριζόντια κλιμάκωση χωρίς την ανάγκη πολύπλοκων αλλαγών στην υποδομή. Τέλος, η λύση βασίζεται αποκλειστικά σε ανοιχτά πρότυπα, όπως REST, JWT και OIDC, κάτι που διευκολύνει τη διασύνδεσή της τόσο με υφιστάμενα όσο και με μελλοντικά συστήματα.

2.6 Υλοποίηση της αρχιτεκτονικής

Οι τεχνολογίες που παρουσιάστηκαν και αναλύθηκαν στο προηγούμενο κεφάλαιο αξιοποιήθηκαν ολοκληρωμένα για την κατασκευή και υλοποίηση της συνολικής αρχιτεκτονικής του συστήματος. Συγκεκριμένα, το backend του συστήματος αναπτύχθηκε με τη χρήση του σύγχρονου και αποδοτικού πλαισίου FastAPI, το οποίο λειτουργεί σε περιβάλλον προγραμματισμού Python. Από την άλλη πλευρά, το frontend υλοποιήθηκε με βάση το React, ένα δημοφιλές και ευέλικτο εργαλείο ανάπτυξης διεπαφών χρήστη, χρησιμοποιώντας JavaScript και CSS για τη δημιουργία δυναμικών και ευχάριστων οπτικά εφαρμογών. Η αποθήκευση και διαχείριση των δεδομένων πραγματοποιείται μέσω της βάσης δεδομένων MongoDB, η οποία προσφέρει ευελιξία στην αποθήκευση μη σχεσιακών δεδομένων και υποστηρίζει οριζόντια κλιμάκωση. Παράλληλα, η αυθεντικοποίηση και ο έλεγχος πρόσβασης των χρηστών υλοποιείται με τη βοήθεια του Keycloak, ενός ολοκληρωμένου συστήματος διαχείρισης ταυτότητας και δικαιωμάτων που παρέχει ασφάλεια και κεντρικό έλεγχο.

Προκειμένου να διευκολυνθεί η ανάπτυξη και η λειτουργία του συστήματος στο πλαίσιο της παρούσας διπλωματικής εργασίας, ελήφθησαν ορισμένες σημαντικές παραδοχές, οι οποίες σχετίζονται κυρίως με την δικτυακή τοπολογία και τον τρόπο με τον οποίο οργανώνονται τα components του συστήματος. Συγκεκριμένα, όλες οι επιμέρους υπηρεσίες εκτελούνται μέσα σε Docker containers που ανήκουν σε ένα κοινό εικονικό δίκτυο. Σε αυτό το δίκτυο περιλαμβάνονται επίσης η βάση δεδομένων και ο authentication server, προσφέροντας έτσι ένα ολοκληρωμένο και ενοποιημένο περιβάλλον εκτέλεσης.

Σε αντίθεση με ένα πλήρως federated περιβάλλον, όπου κάθε κόμβος θα λειτουργούσε εντελώς αυτόνομα, διαθέτοντας τη δική του βάση δεδομένων, το δικό του authentication server και όλα τα απαραίτητα components που θα ήταν απομονωμένα από τους υπόλοιπους κόμβους, εδώ η δομή είναι πιο συγκεντρωτική, καθώς όλα τα containers βρίσκονται στο ίδιο δίκτυο.

Στο εικονικό αυτό δίκτυο λειτουργούν συνολικά δύο τοπικοί κόμβοι, όπου ο καθένας φιλοξενεί το κύριο backend service του και παράλληλα ένα query service. Το κύριο service είναι υπεύθυνο για τη διαχείριση των αιτημάτων που προέρχονται από την εφαρμογή του ιατρικού προσωπικού αλλά και από τους δότες, οι οποίοι εισάγουν και ενημερώνουν δεδομένα. Το query service από την άλλη πλευρά οργανώνει και εξυπηρετεί προκαθορισμένα ερωτήματα που προέρχονται από εξωτερικές κλήσεις, συγκεκριμένα από την εφαρμογή των ερευνητών, προσφέροντας έτσι τη δυνατότητα εύκολης και ελεγχόμενης πρόσβασης σε δεδομένα της βάσης.

Επιπρόσθετα, υπάρχει και ένας ξεχωριστός web κόμβος, ο οποίος έχει σχεδιαστεί αποκλειστικά για τη χρήση από ερευνητές. Αυτός ο κόμβος διαθέτει δικό του frontend και backend και λειτουργεί ως σημείο πρόσβασης για αιτήματα που αφορούν τη δευτερογενή χρήση των δεδομένων. Ο κόμβος αυτός συνδέεται τόσο με τη βάση δεδομένων MongoDB όσο και με τον authentication server, διασφαλίζοντας ότι η πρόσβαση στα δεδομένα είναι περιορισμένη και επιτρέπεται μόνο σε χρήστες με την κατάλληλη εξουσιοδότηση.

Ο κεντρικός authentication server που χρησιμοποιείται είναι ο Keycloak, ο οποίος οργανώνει τους χρήστες και τις ομάδες σε λογικές ενότητες, γνωστές ως realms. Κάθε realm περιλαμβάνει το σύνολο των χρηστών, των ομάδων τους και των σχετικών πολιτικών ασφαλείας,

παρέχοντας έτσι έναν ευέλικτο και κεντρικό μηχανισμό ελέγχου ταυτότητας και δικαιωμάτων, που μπορεί να προσαρμοστεί ανάλογα με τις απαιτήσεις κάθε οργανισμού ή κόμβου.

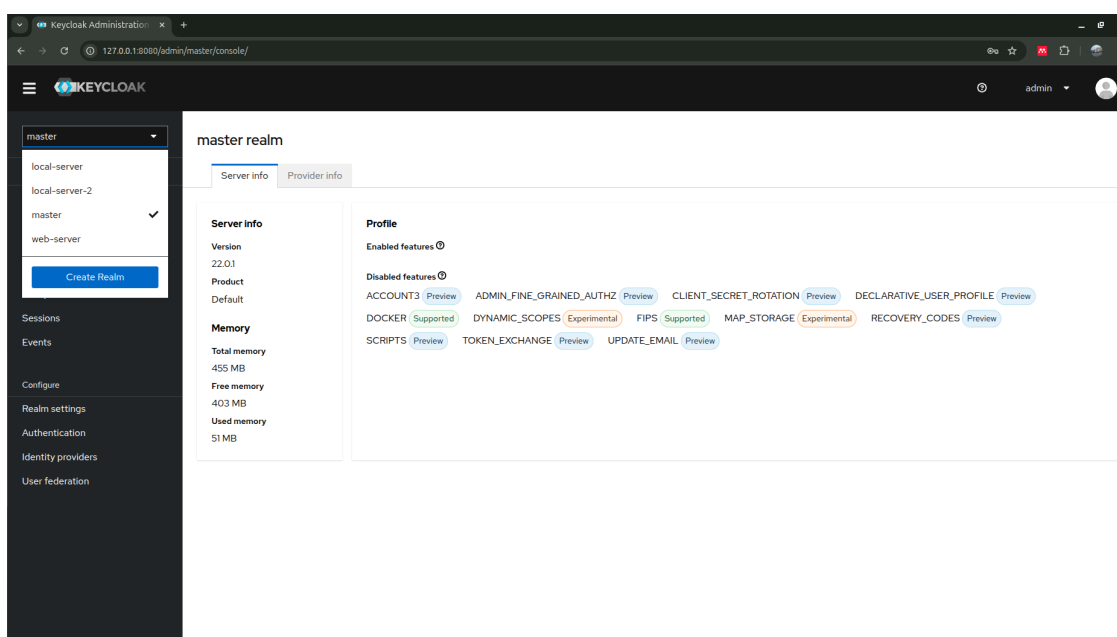
Τέλος, για να απλοποιηθεί η υλοποίηση του δικτύου και να μειωθεί η πολυπλοκότητα, δεν χρησιμοποιούνται proxy servers για την ανταλλαγή δεδομένων μεταξύ των διαφορετικών κόμβων. Αντίθετα, όλα τα containers επικοινωνούν απευθείας μέσω του εικονικού Docker δικτύου, εξασφαλίζοντας γρήγορη και αποδοτική ανταλλαγή πληροφορίας.

CONTAINER ID	IMAGE	COMMAND	CREATED	STATUS	PORTS	NAMES
5981c40eb91	federated-network-researchers-frontend	"docker-entrypoint.s..."	12 minutes ago	Up 12 minutes	0.0.0.0:3007->3007/tcp, [::]:3007->3007/tcp	researchers-frontend
9260d49b5ef7	federated-network-donors-frontend	"docker-entrypoint.s..."	12 minutes ago	Up 12 minutes	0.0.0.0:3003->3003/tcp, [::]:3003->3003/tcp	donors-frontend-1
b6e5de514ecf	federated-network-medical-staff-frontend-second	"docker-entrypoint.s..."	12 minutes ago	Up 12 minutes	0.0.0.0:3006->3006/tcp, [::]:3006->3006/tcp	medical-staff-frontend-2
911564c592fc	federated-network-medical-staff-frontend	"docker-entrypoint.s..."	12 minutes ago	Up 12 minutes	0.0.0.0:3004->3004/tcp, [::]:3004->3004/tcp	medical-staff-frontend-1
122aa00193c	federated-network-donors-frontend-second	"docker-entrypoint.s..."	12 minutes ago	Up 12 minutes	0.0.0.0:3005->3005/tcp, [::]:3005->3005/tcp	donors-frontend-2
5adb1595f7ce	federated-network-local-node-backend	"uvicorn main:app --..."	15 minutes ago	Up 12 minutes	0.0.0.0:8000->8000/tcp, [::]:8000->8000/tcp	local-node-backend-1
1e7ad53c595d	federated-network-local-node-backend-second	"uvicorn main:app --..."	15 minutes ago	Up 12 minutes	0.0.0.0:8003->8003/tcp, [::]:8003->8003/tcp	local-node-backend-2
661bf6319133	federated-network-researchers-backend	"uvicorn main:app --..."	15 minutes ago	Up 12 minutes	0.0.0.0:8002->8002/tcp, [::]:8002->8002/tcp	researchers-backend
309fa1998483	federated-network-queries-second	"uvicorn main:app --..."	15 minutes ago	Up 12 minutes	0.0.0.0:8004->8004/tcp, [::]:8004->8004/tcp	local-node-queries-2
2f2dc81c01a4	federated-network-queries	"uvicorn main:app --..."	15 minutes ago	Up 12 minutes	0.0.0.0:8001->8001/tcp, [::]:8001->8001/tcp	local-node-queries-1
98f9916190e7	quay.io/keycloak/keycloak:22.0.1	"/opt/keycloak/bin/k..."	45 minutes ago	Up 12 minutes	0.0.0.0:8080->8080/tcp, [::]:8080->8080/tcp, 8443/tcp	keycloak-auth-servers
650fb60357e	postgres:15		45 minutes ago	Up 12 minutes	5432/tcp	keycloak-db
56d5079d9c81	mongo:latest	"docker-entrypoint.s..."	45 minutes ago	Up 12 minutes	0.0.0.0:27017->27017/tcp, [::]:27017->27017/tcp	mongo-database

Εικόνα 2.1: Docker containers

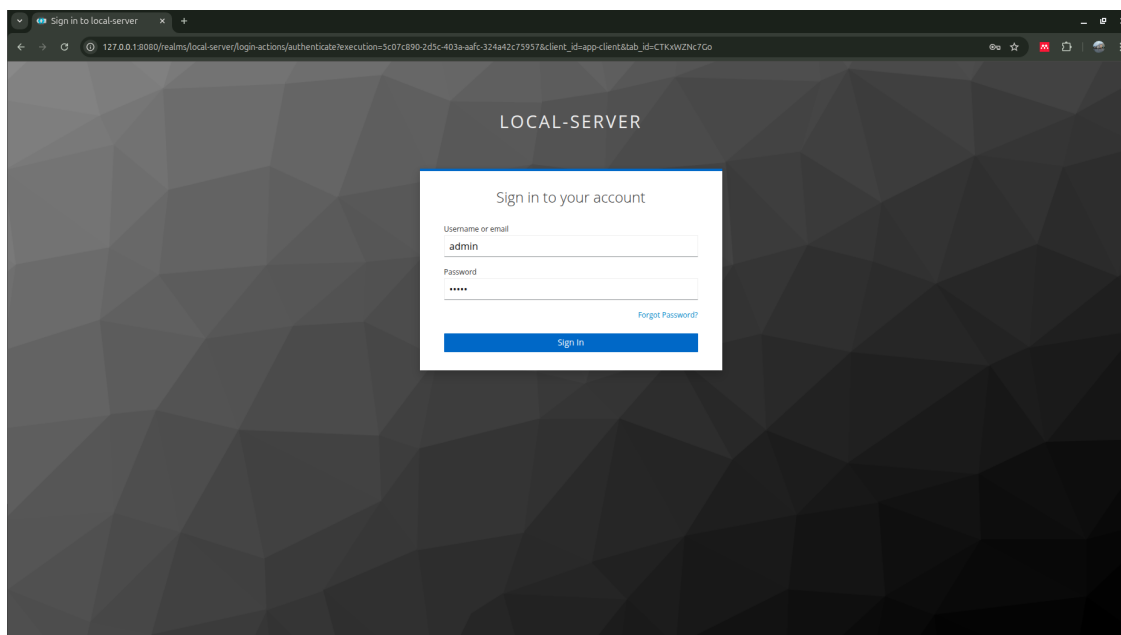
Στην εικόνα 2.1 παρουσιάζεται το πλήρες σύνολο των containers που χρησιμοποιήθηκαν για την υλοποίηση της αρχιτεκτονικής του συστήματος. Συγκεκριμένα, απεικονίζονται οι δύο τοπικοί κόμβοι, ο εξωτερικός κόμβος που εξυπηρετεί τους ερευνητές, η βάση δεδομένων καθώς και ο authentication server. Επιπλέον, στην ίδια εικόνα αποτυπώνεται και η χρήση της βάσης δεδομένων PostgreSQL, η οποία λειτουργεί ως αποθηκευτικός χώρος για τα δεδομένα του Keycloak, δηλαδή για τα στοιχεία διαχείρισης ταυτότητας και δικαιωμάτων που απαιτούνται για την ασφαλή λειτουργία του συστήματος.

Στην εικόνα 2.2 παρουσιάζεται το menu διαχείρισης των realms μέσα στο περιβάλλον του Keycloak. Όπως είναι εμφανές, για κάθε επιμέρους σύστημα έχει δημιουργηθεί ένας ξεχωριστός χώρος (realm), ο οποίος επιτρέπει την οργάνωση και διαχείριση των χρηστών με βάση τους ρόλους που τους έχουν ανατεθεί. Με αυτόν τον τρόπο εξασφαλίζεται η σαφής διαχωριστική γραμμή μεταξύ των διαφορετικών ομάδων χρηστών και η κατάλληλη ανάθεση δικαιωμάτων, ανάλογα με τις αρμοδιότητες και τα προνόμιά τους.



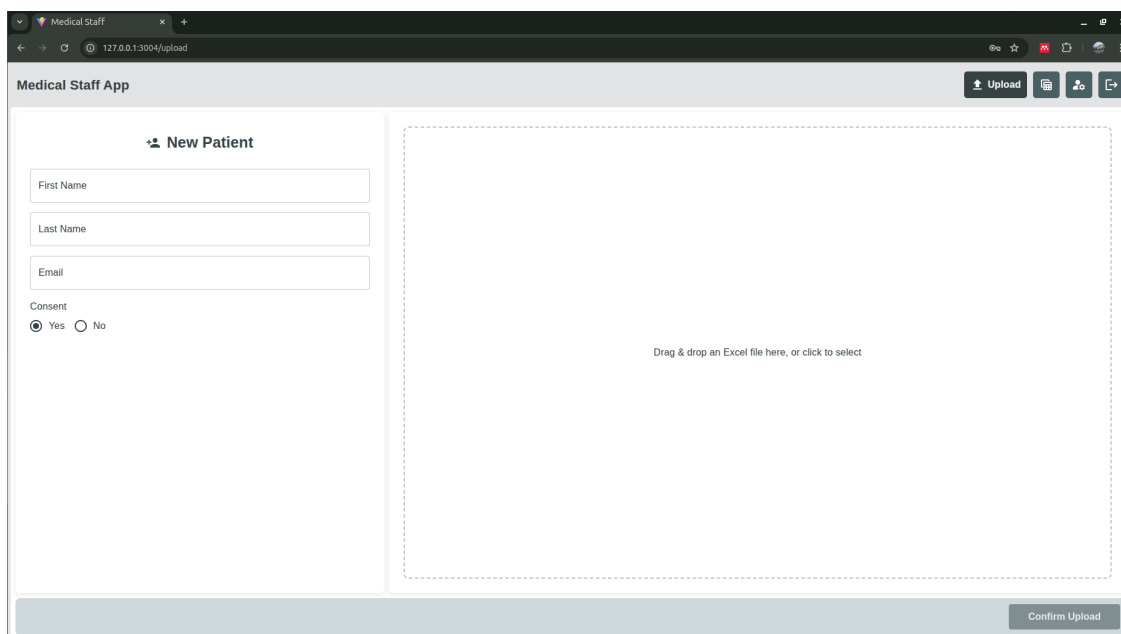
Εικόνα 2.2: διαχείριση των realms

Κατά την είσοδο στην εφαρμογή του ιατρικού προσωπικού το σύστημα μας ανακατευθύνει στον keycloak. Ο χρήστης μας είναι admin με τα αντίστοιχα permissions που του έχουμε δώσει.



Εικόνα 2.3: Σύνδεση στην εφαρμογή του ιατρικού προσωπικού ως admin

Η πρώτη οθόνη (εικόνα 2.4) που εμφανίζεται κατά την είσοδό μας στην εφαρμογή είναι για την αποθήκευση ιατρικών δεδομένων στη βάση (με όνομα Upload).



Εικόνα 2.4: Οθόνη Upload του ιατρικού προσωπικού

Στην αριστερή πλευρά της εικόνας 2.5 εμφανίζεται η φόρμα όπου εισάγονται τα στοιχεία του ασθενή ή δότη, μαζί με την αντίστοιχη συγκατάθεσή του. Συγκεκριμένα, η φόρμα αυτή επιτρέπει την καταγραφή των προσωπικών πληροφοριών, καθώς και την επιλογή της

συγκατάθεσης για τη χρήση των δεδομένων του. Για παράδειγμα, σε περίπτωση που η συγκατάθεση δηλωθεί ως "ναι", τότε τα δεδομένα του συγκεκριμένου ασθενή συμπεριλαμβάνονται στα metadata του τοπικού κόμβου. Αυτά τα metadata στη συνέχεια διαβιβάζονται στην εφαρμογή που χρησιμοποιούν οι ερευνητές, επιτρέποντας την αξιοποίηση των δεδομένων.

Στην δεξιά πλευρά της ίδιας εικόνας, παρουσιάζεται η επιλογή του αρχείου Excel, το οποίο περιέχει τα δεδομένα του ασθενή, όπως για παράδειγμα αποτελέσματα εξετάσεων ή άλλες σχετικές πληροφορίες. Μέσω αυτής της διεπαφής, το ιατρικό προσωπικό μπορεί να φορτώνει εύκολα και με ασφάλεια τα απαραίτητα δεδομένα που θα χρησιμοποιηθούν από το σύστημα.

Αξίζει να σημειωθεί ότι για τις ανάγκες της παρούσας υλοποίησης χρησιμοποιήθηκαν δεδομένα προερχόμενα από το σύνολο δεδομένων Random Healthcare Dataset (Sample), το οποίο διατίθεται μέσω της πλατφόρμας Kaggle [52]. Αυτά τα δεδομένα αξιοποιήθηκαν ώστε να προσομοιωθεί η λειτουργία του συστήματος με πραγματικά παραδείγματα υγειονομικής φύσης.

Εικόνα 2.5: Οθόνη Upload του ιατρικού προσωπικού με συμπληρωμένα τα πεδία

Στην επόμενη οθόνη, όπως φαίνεται στην εικόνα 2.6, παρουσιάζονται τα δεδομένα που είναι αυτήν τη στιγμή αποθηκευμένα στη βάση δεδομένων του τοπικού κόμβου, σε μια σελίδα με τίτλο View Data. Τα δεδομένα αυτά είναι προσβάσιμα από το ιατρικό προσωπικό και έχουν υποστεί κατάλληλη επεξεργασία, όπως για παράδειγμα ψευδωνυμοποίηση, κατά τη διαδικασία της αποθήκευσης. Η συγκεκριμένη μορφοποίηση αποσκοπεί στην ενίσχυση της προστασίας των προσωπικών πληροφοριών των ασθενών, διασφαλίζοντας ότι τα ευαίσθητα δεδομένα δεν εκτίθενται άμεσα στους χρήστες.

Στην εικόνα 2.7 παρουσιάζεται η σελίδα με όνομα Manage Users. Πρόκειται για μια λειτουργία που είναι διαθέσιμη αποκλειστικά σε χρήστες που έχουν ανατεθεί δικαιώματα admin. Η σελίδα αυτή επιτρέπει τη διαχείριση των χρηστών του συστήματος, όπως τη δημιουργία νέων λογαριασμών, την επεξεργασία στοιχείων υφιστάμενων χρηστών και τον έλεγχο

Age	Gender	Blood_Type	Height (cm)	Weight (kg)	BMI	Temperature (C)	Heart_Rate (bpm)	Blood_Pressure (...)	original_id	pseudonym_id	upload_id
32	Female	B-	185.5483514891868	110.6463010739574	32.14	36.26	70	96/89	1	0131f44a-2006-49...	3f86e99a-7cfa-49f...

Εικόνα 2.6: Οθόνη *View Data* του ιατρικού προσωπικού

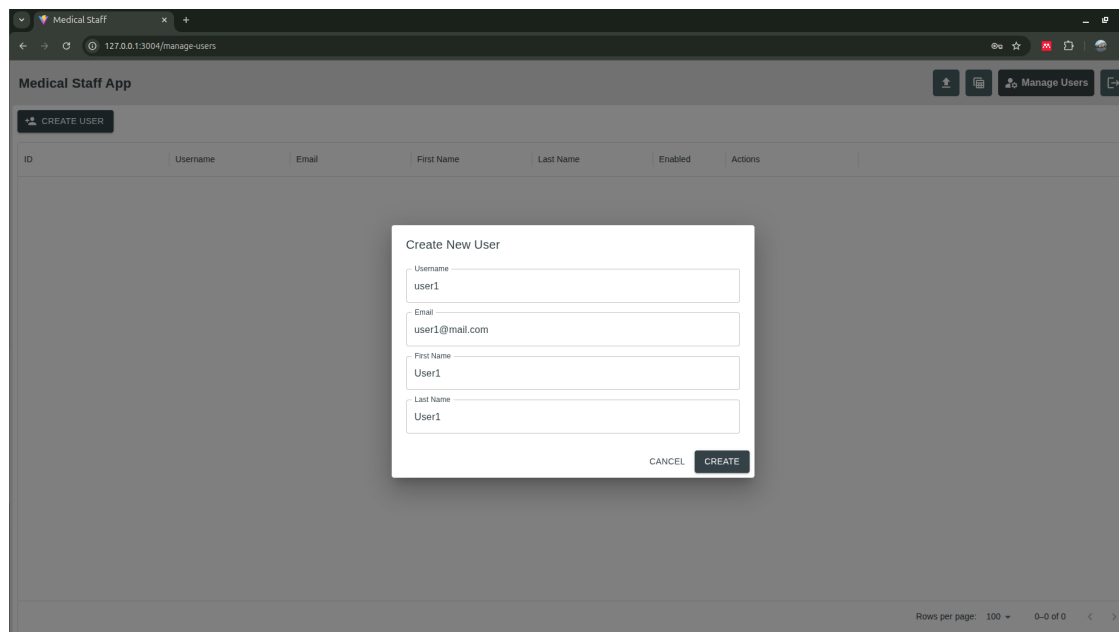
των δικαιωμάτων τους, διασφαλίζοντας έτσι την ορθή διαχείριση των προσβάσεων και την ασφάλεια του συστήματος συνολικά.

ID	Username	Email	First Name	Last Name	Enabled	Actions
No rows						

Εικόνα 2.7: Οθόνη *Manage Users* του ιατρικού προσωπικού

Ο χρήστης με δικαιώματα διαχειριστή έχει τη δυνατότητα να προσθέσει νέους χρήστες στο σύστημα, όπως φαίνεται στην εικόνα 2.8. Επιπλέον, μπορεί να επεξεργαστεί τα στοιχεία υφιστάμενων χρηστών, να τους διαγράψει ή να απενεργοποιήσει προσωρινά τους λογαριασμούς τους, όπως παρουσιάζεται στην εικόνα 2.9. Επιπρόσθετα, ο διαχειριστής έχει πρόσβαση στα logs των χρηστών, δηλαδή σε καταγραφές ενεργειών όπως η είσοδος στην εφαρμογή και άλλες σχετικές ενέργειες. Η λειτουργία αυτή θα παρουσιαστεί αναλυτικότερα σε κάποια από

τις επόμενες εικόνες.



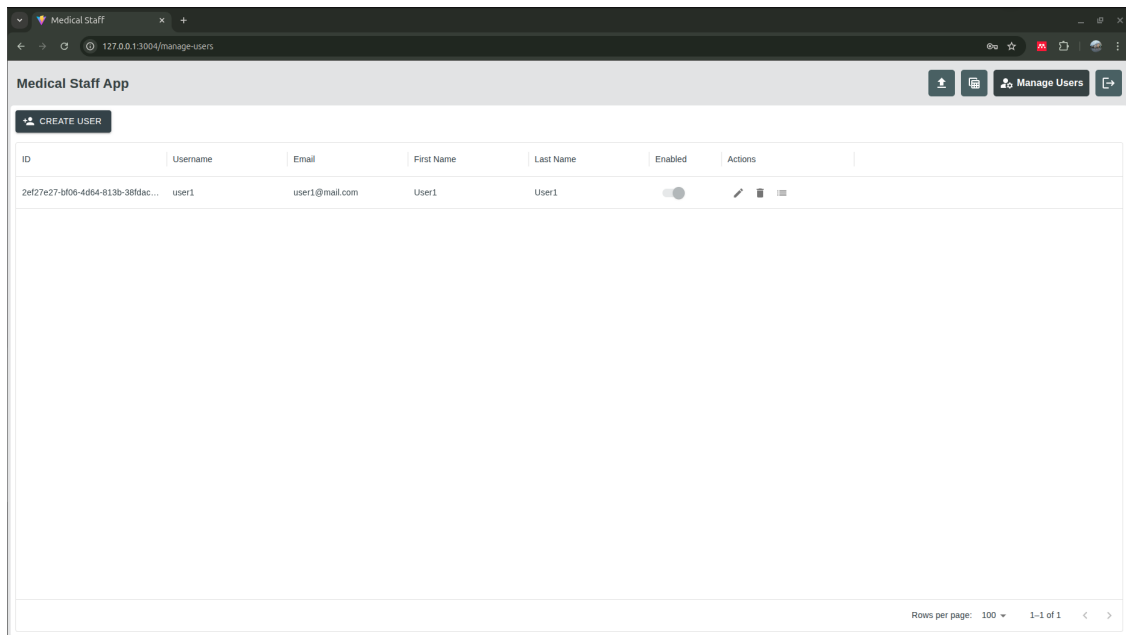
Εικόνα 2.8: Οθόνη *Manage Users* του ιατρικού προσωπικού με συμπληρωμένη την φόρμα δημιουργίας νέου χρήστη

Μετά την επιτυχή δημιουργία ενός νέου χρήστη, ακολουθεί η πρώτη σύνδεση με τα στοιχεία αυτού του χρήστη. Κατά τη διαδικασία αυτή, ο χρήστης ανακατευθύνεται ξανά στην πλατφόρμα Keycloak (όπως φαίνεται στην εικόνα 2.10). Στη συνέχεια, το σύστημα απαιτεί από τον χρήστη να ορίσει έναν νέο κωδικό πρόσβασης, διασφαλίζοντας έτσι την ασφάλεια του λογαριασμού (εικόνα 2.11). Επιπλέον, ζητείται η ενημέρωση των προσωπικών στοιχείων του χρήστη, όπως εμφανίζεται στην εικόνα 2.12, ώστε να διασφαλιστεί η επικαιροποίηση και η ακρίβεια των δεδομένων του προφίλ.

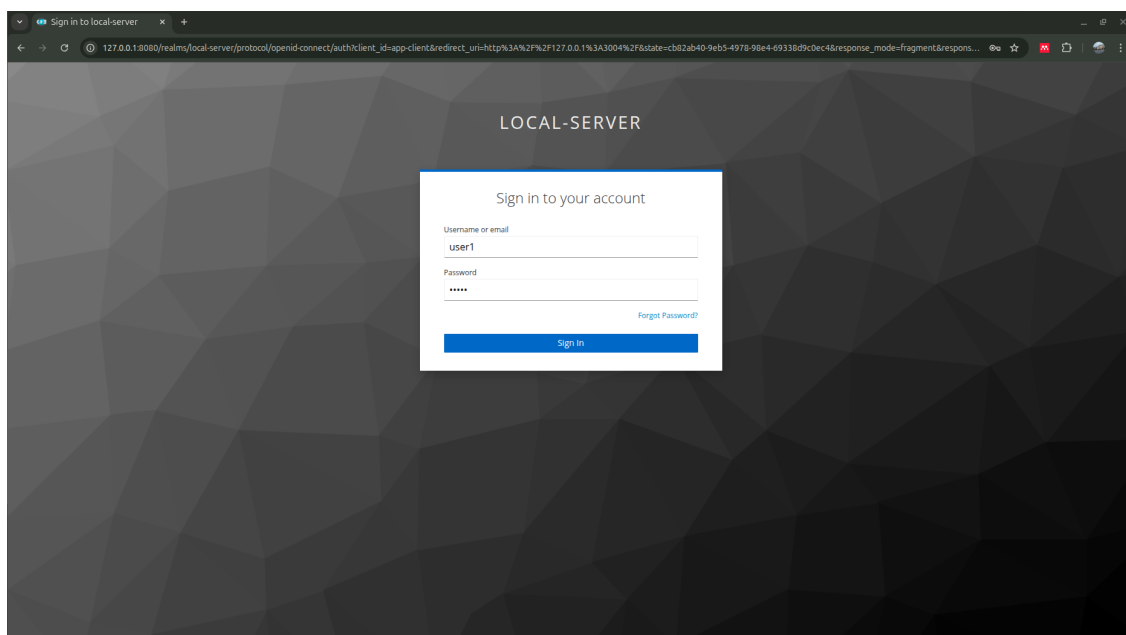
Στην εικόνα 2.13 παρουσιάζεται η αρχική οθόνη που βλέπει ο χρήστης μετά από μια επιτυχή σύνδεση στο σύστημα. Είναι σημαντικό να σημειωθεί πως, σε αυτήν την περίπτωση, η επιλογή για τη διαχείριση χρηστών (*Manage Users*) δεν είναι διαθέσιμη, καθώς ο συγκεκριμένος χρήστης δεν διαθέτει τα απαραίτητα δικαιώματα διαχειριστή για να έχει πρόσβαση σε αυτήν τη λειτουργία.

Αφού ο χρήστης εκτελέσει διάφορες ενέργειες μέσα στην εφαρμογή, ο admin έχει τη δυνατότητα να παρακολουθεί και να ελέγχει αυτές τις ενέργειες μέσω της σελίδας *Manage Users*, όπου εμφανίζονται τα σχετικά αρχεία καταγραφής (logs) των ενεργειών του χρήστη, όπως φαίνεται στην εικόνα 2.14.

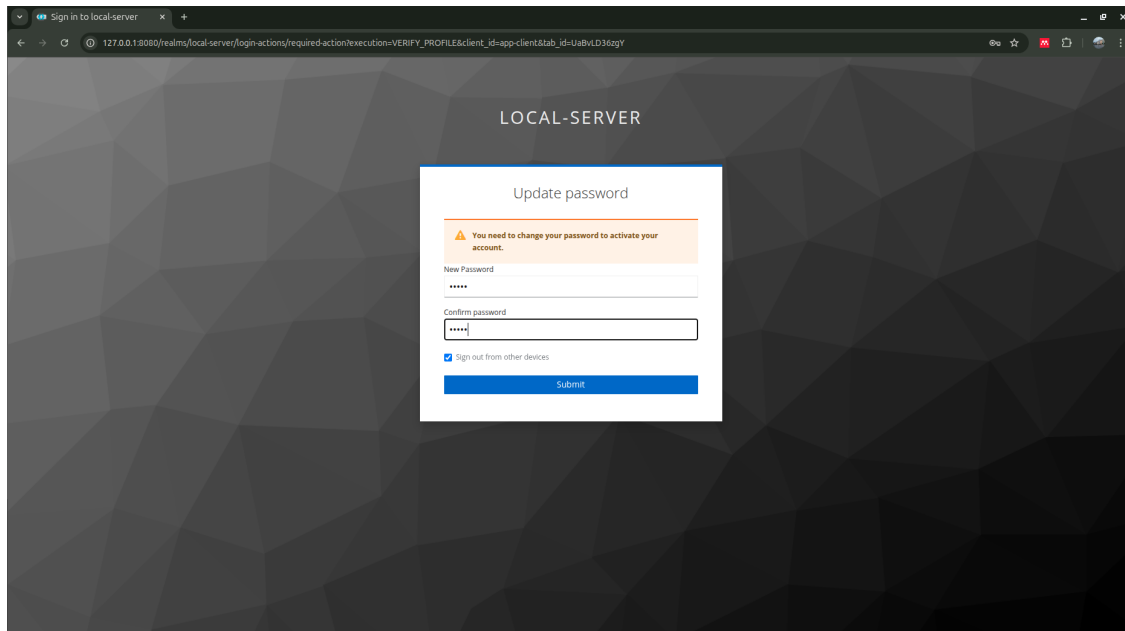
Ακριβώς την ίδια διαδικασία πρώτης σύνδεσης ακολουθεί και ο ασθενής μέσω της αντίστοιχης εφαρμογής διαχείρισης συγκαταθέσεων. Μόλις ο ασθενής ολοκληρώσει τη σύνδεσή του, του εμφανίζεται μια σελίδα όπου παρουσιάζεται η τρέχουσα κατάσταση της συγκατάθεσής του, με την αντίστοιχη επιλογή να είναι ήδη επιλεγμένη, όπως φαίνεται στην εικόνα 2.15. Ο ασθενής έχει τη δυνατότητα να ενημερώσει την επιλογή του άμεσα και σε πραγματικό χρόνο. Συγκεκριμένα, εάν στο παρελθόν είχε δώσει τη συγκατάθεσή του για τη δευτερογενή χρήση των δεδομένων του αλλά επιθυμεί να αλλάξει γνώμη, μπορεί απλώς να επιλέξει την αντίθετη επιλογή και να αποθηκεύσει την αλλαγή. Με την αποθήκευση της νέας



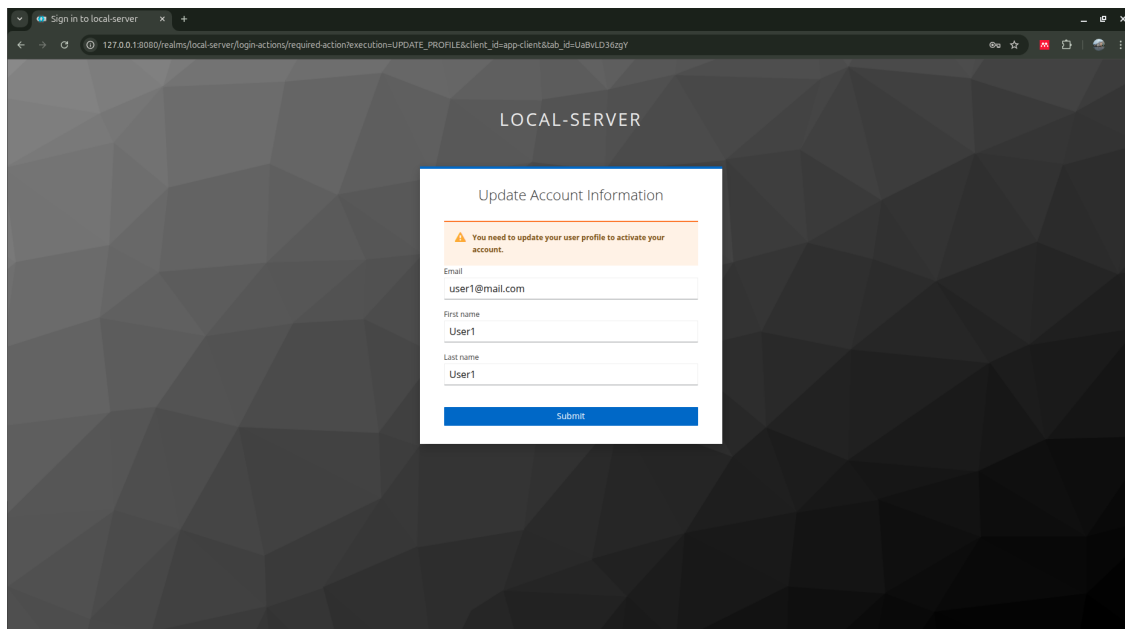
Εικόνα 2.9: Οθόνη Manage Users του ιατρικού προσωπικού με τις λειτουργίες διαχείρισης



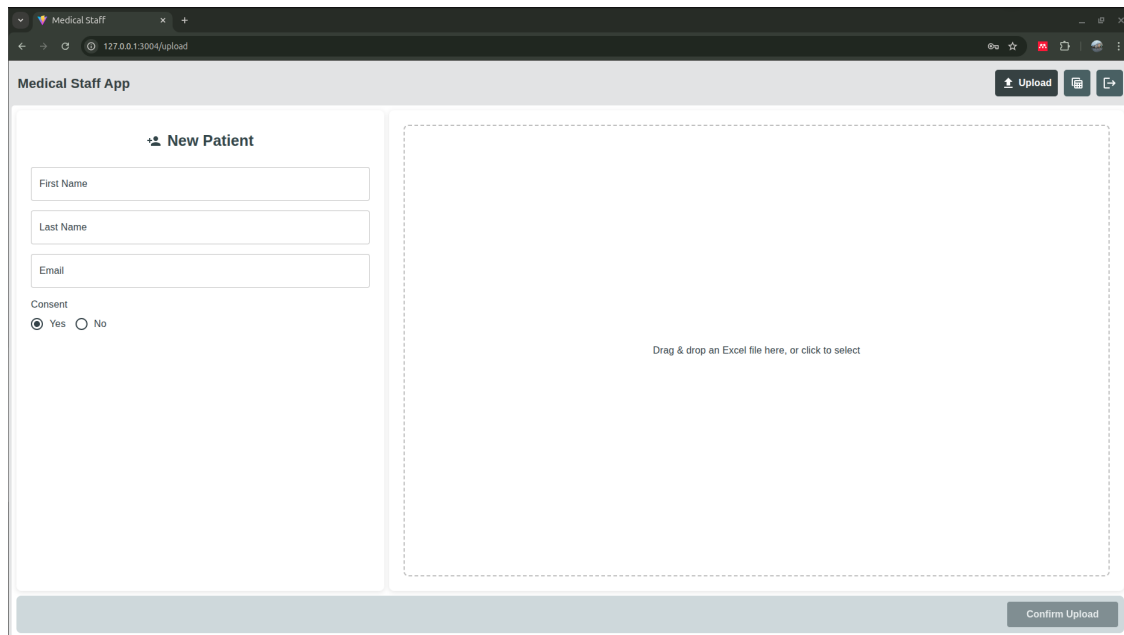
Εικόνα 2.10: Πρώτη σύνδεση στην εφαρμογή ιατρικού προσωπικού ως νέος χρήστης



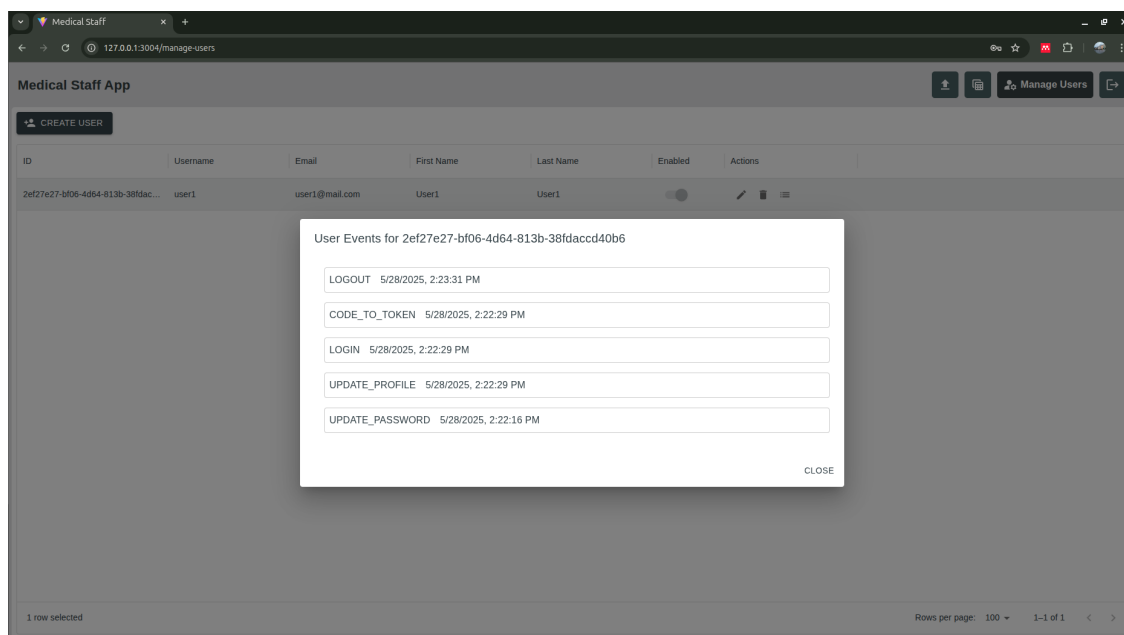
Εικόνα 2.11: Πρώτη σύνδεση στην εφαρμογή ιατρικού προσωπικού ως νέος χρήστης με αληθαγή κωδικού



Εικόνα 2.12: Πρώτη σύνδεση στην εφαρμογή ιατρικού προσωπικού ως νέος χρήστης με ενημέρωση προσωπικών στοιχείων

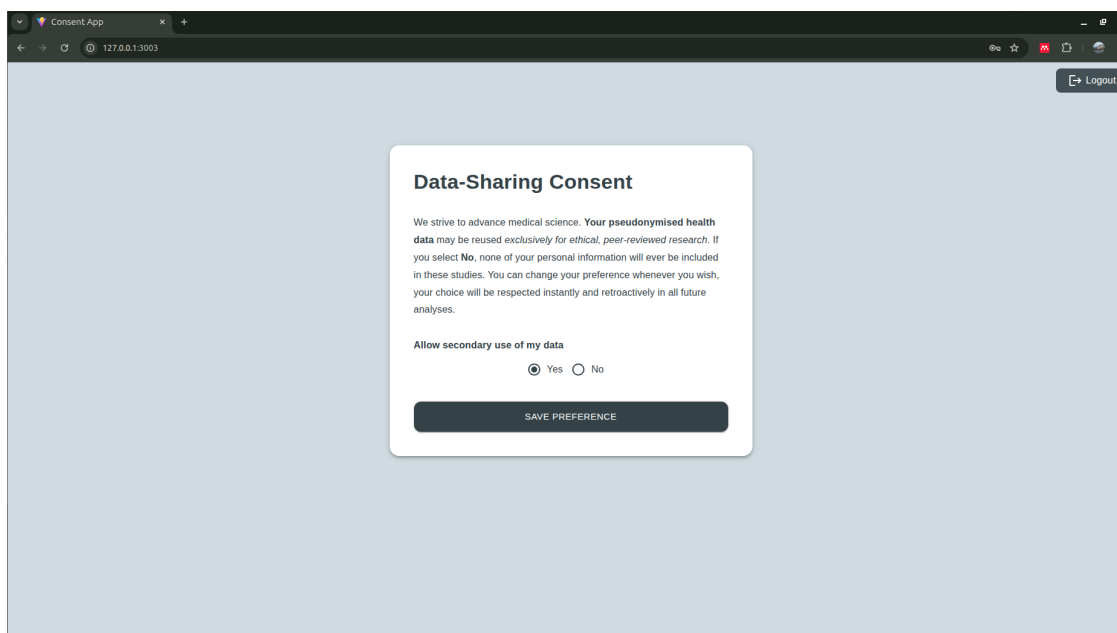


Εικόνα 2.13: Πρώτη σύνδεση στην εφαρμογή ιατρικού προσωπικού ως νέος χρήστης



Εικόνα 2.14: Τρόπος εμφάνισης των logs του χρήστη

αυτής επιλογής, το σύστημα ενημερώνει άμεσα τη βάση δεδομένων και αφαιρεί όλα τα δεδομένα που σχετίζονται με τον ασθενή από τα metadata του τοπικού κόμβου, διασφαλίζοντας έτσι την πλήρη συμμόρφωση με τις επιλογές του χρήστη.



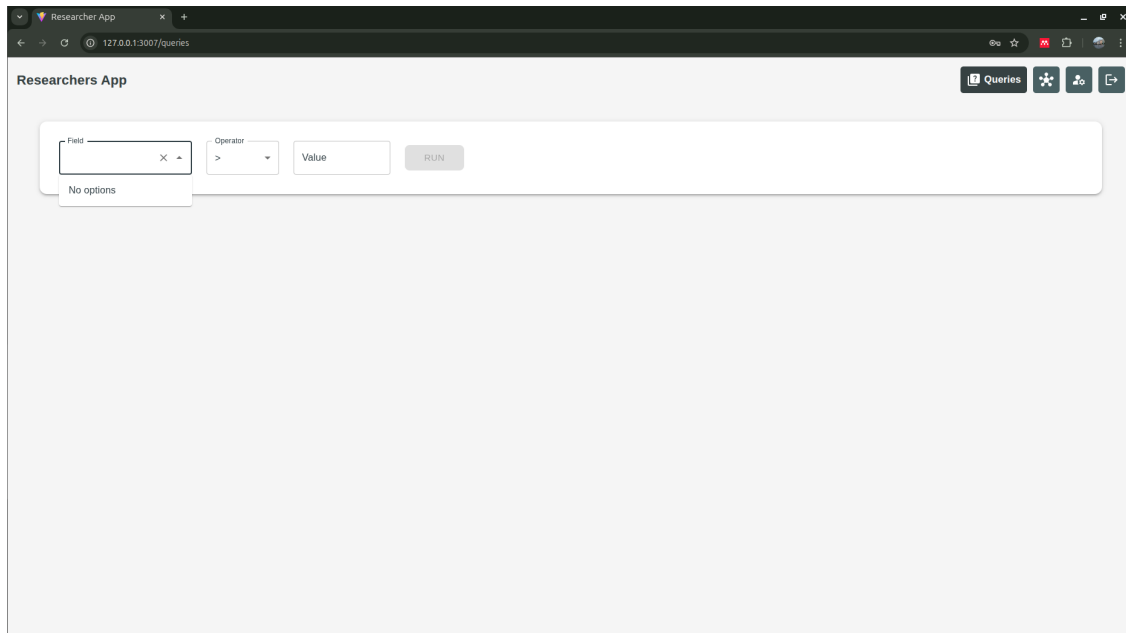
Εικόνα 2.15: Εφαρμογή για την διαχείριση συγκατάθεσης ασθενή/δότη

Σε αυτό το σημείο της διαδικασίας, προσθέτουμε επιπλέον δεδομένα στη βάση δεδομένων και στους δύο τοπικούς κόμβους, προκειμένου να διασφαλίσουμε ότι η εφαρμογή που χρησιμοποιούν οι ερευνητές διαθέτει τα απαραίτητα στοιχεία για τις μετέπειτα λειτουργίες της. Αυτή η προσθήκη δεδομένων είναι κρίσιμη, καθώς επιτρέπει τη διεξαγωγή ερωτημάτων και την αξιοποίηση των πληροφοριών από το ερευνητικό περιβάλλον.

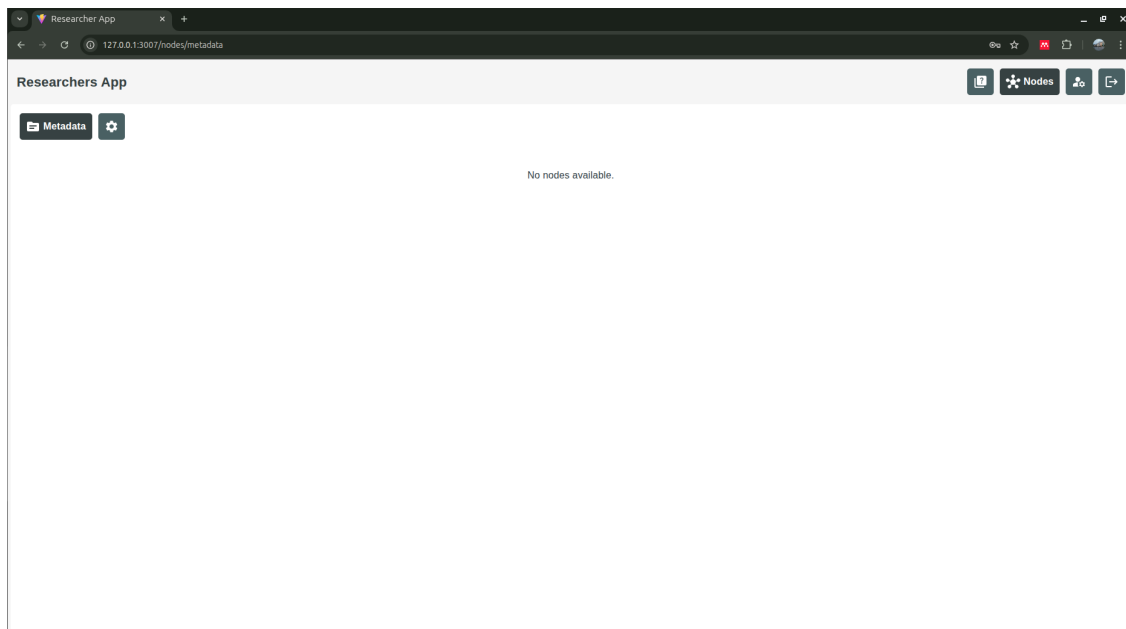
Κατά την πρώτη σύνδεση του ερευνητή στην εφαρμογή, παρουσιάζεται η αρχική σελίδα με τίτλο *Queries* (όπως φαίνεται στην εικόνα 2.16). Η σελίδα αυτή παρέχει έναν κατάλογο προκαθορισμένων API endpoints που μπορούν να χρησιμοποιηθούν από τον ερευνητή για την εξαγωγή πληροφοριών από το σύστημα. Για παράδειγμα, ο ερευνητής έχει τη δυνατότητα να επιλέξει μία μεταβλητή, έναν συγκεκριμένο έλεγχο (π.χ. μεγαλύτερο από, ίσο με) και έναν αριθμητικό παράμετρο. Το αποτέλεσμα που επιστρέφεται είναι ο συνολικός αριθμός εγγραφών που πληρούν τα κριτήρια αυτά.

Ωστόσο, προς το παρόν η επιλογή μεταβλητής στο πεδίο *Field* δεν είναι ενεργή και δεν παρέχεται προς επιλογή κάποιο συγκεκριμένο στοιχείο. Αυτό οφείλεται στο γεγονός ότι οι τοπικοί κόμβοι δεν έχουν ακόμη συνδεθεί πλήρως με το κεντρικό σύστημα και δεν έχουν διαθέσιμα τα σχετικά δεδομένα. Η παραπάνω παρατήρηση επιβεβαιώνεται και από τη σελίδα *Nodes*, όπου τόσο η προβολή των *Metadata* (όπως παρουσιάζεται στην εικόνα 2.17) όσο και η διαχείριση των τοπικών κόμβων (εικόνα 2.18) εμφανίζονται κενές, χωρίς εγγραφές ή ενεργούς κόμβους.

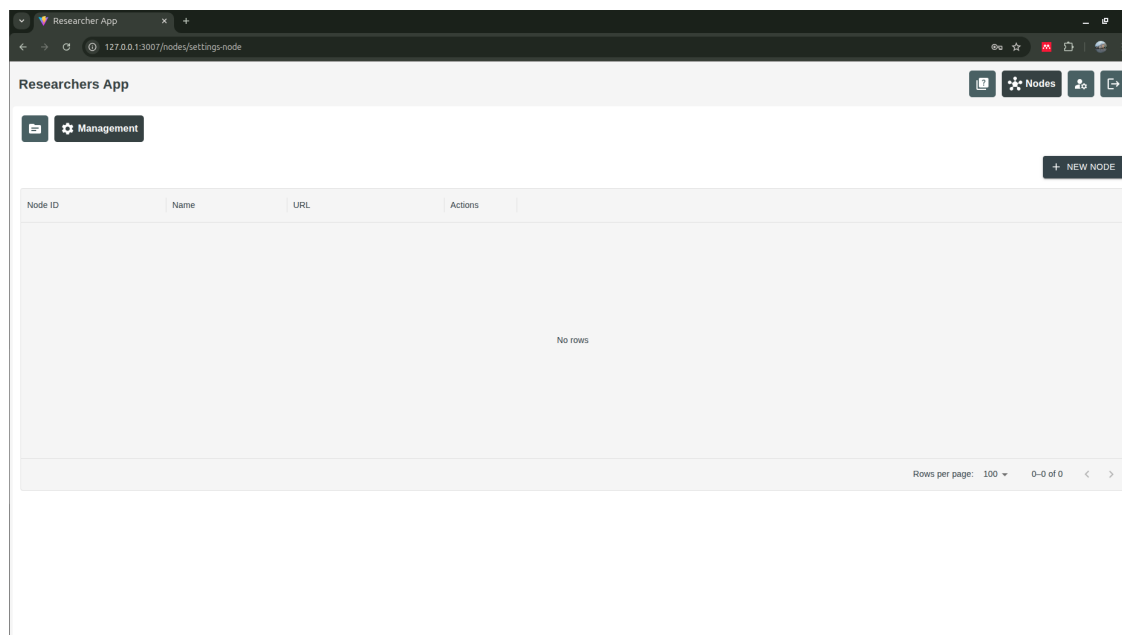
Κατά τη διαδικασία προσθήκης ενός νέου τοπικού κόμβου, όπως φαίνεται στην εικόνα 2.19, η εφαρμογή μας ζητά να εισάγουμε ορισμένες βασικές πληροφορίες που είναι απαραίτητες για την αναγνώριση και σύνδεση με τον κόμβο αυτό. Συγκεκριμένα, απαιτείται



Εικόνα 2.16: Οθόνη *Queries* χωρίς την προσθήκη τοπικών κόμβων



Εικόνα 2.17: Οθόνη *Nodes* με την εμφάνιση των *Metadata* χωρίς την προσθήκη τοπικών κόμβων



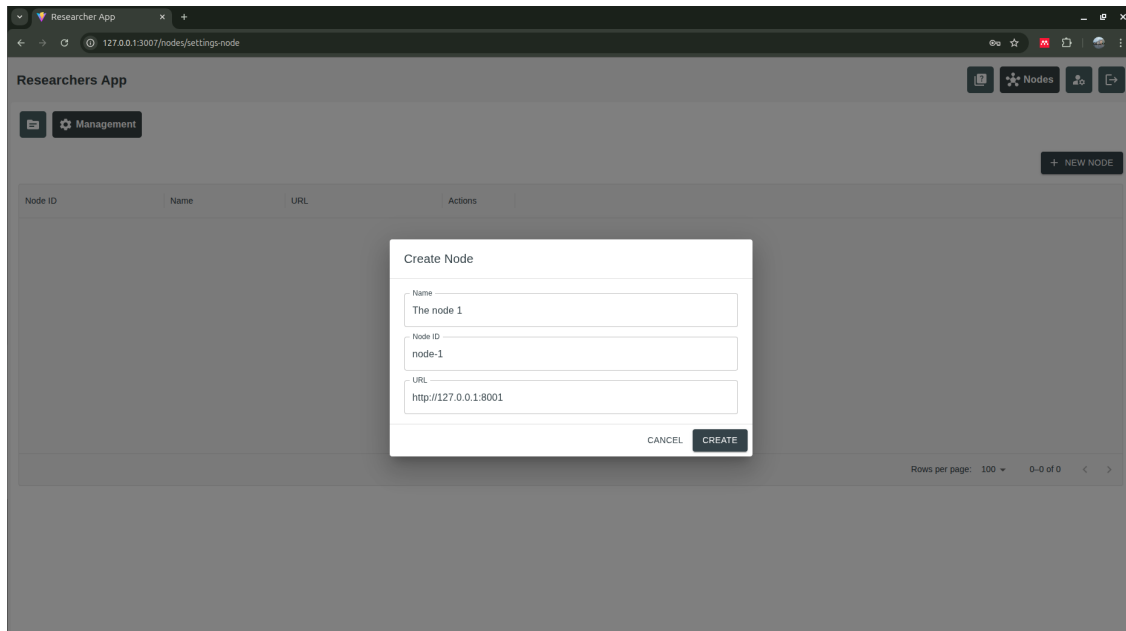
Εικόνα 2.18: Οθόνη Nodes με την εμφάνιση διαχείρισης των τοπικών κόμβων χωρίς την προσθήκη τους

να ορίσουμε ένα όνομα το οποίο θα χρησιμοποιείται στην εφαρμογή ώστε να αναγνωρίζουμε τον κόμβο με ευκολία και σαφήνεια κατά την πλοήγηση και τη διαχείρισή του. Επιπλέον, ζητείται το id του κόμβου, το οποίο στην πραγματικότητα λειτουργεί ως ένα είδος «μυστικού κλειδιού» (secret key), όπως έχει ήδη περιγραφεί στην αρχιτεκτονική του συστήματος. Αυτό το id εξασφαλίζει την ασφαλή ταυτοποίηση του κόμβου κατά τις επικοινωνίες με το κεντρικό σύστημα. Τέλος, απαιτείται η καταχώρηση της διεύθυνσης url που θα χρησιμοποιείται για τις κλήσεις προς τη βάση δεδομένων του τοπικού κόμβου. Αυτή η διεύθυνση καθορίζει το σημείο επαφής μέσω του οποίου το σύστημα θα μπορεί να αποστείλει αιτήματα και να λαμβάνει απαντήσεις από το συγκεκριμένο κόμβο, διασφαλίζοντας έτσι την ομαλή και ασφαλή λειτουργία της διασύνδεσης.

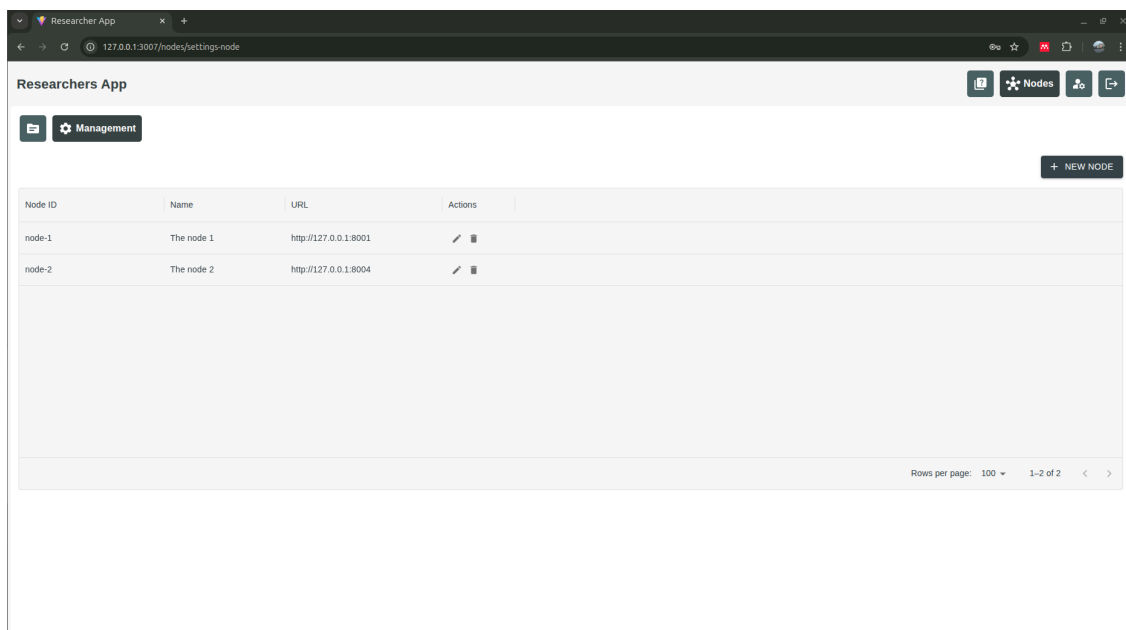
Στην εικόνα 2.20 φαίνεται η επιτυχής σύνδεση και των δύο τοπικών κόμβων που χρησιμοποιήθηκαν στο δίκτυο. Επίσης, δίνεται η επιλογή διαγραφής και επεξεργασίας της σύνδεσης.

Επιστρέφοντας στην επιλογή εμφάνισης των metadata, όπως παρουσιάζεται στην εικόνα 2.21, διακρίνουμε πλέον την πλήρη απεικόνιση των πεδίων που αφορούν τα δεδομένα των ασθενών που έχουν καταχωρηθεί στο σύστημα. Κάθε πεδίο συνοδεύεται από στατιστικές πληροφορίες ή εύρη τιμών που αναφέρονται στα αντίστοιχα δεδομένα. Για παράδειγμα, στον πρώτο τοπικό κόμβο, το πεδίο της ηλικίας παρουσιάζεται με το μέγιστο, το ελάχιστο και τον μέσο όρο τιμών, ενώ το πεδίο που αφορά το φύλο εκφράζεται με ποσοστό.

Οι τιμές αυτές προκύπτουν από τη στατιστική ανάλυση των δεδομένων που έχουν λάβει θετική συγκατάθεση από τους ασθενείς, με στόχο να δίνεται μια σαφής και κατανοητή περιγραφή των διαθέσιμων δεδομένων χωρίς να απαιτείται η άμεση αποστολή των πραγματικών δεδομένων της βάσης προς τους ερευνητές. Με αυτόν τον τρόπο, τα δεδομένα αναπαρίστανται με τη μορφή metadata, διασφαλίζοντας παράλληλα την προστασία της ιδιωτικότητας των ασθενών.

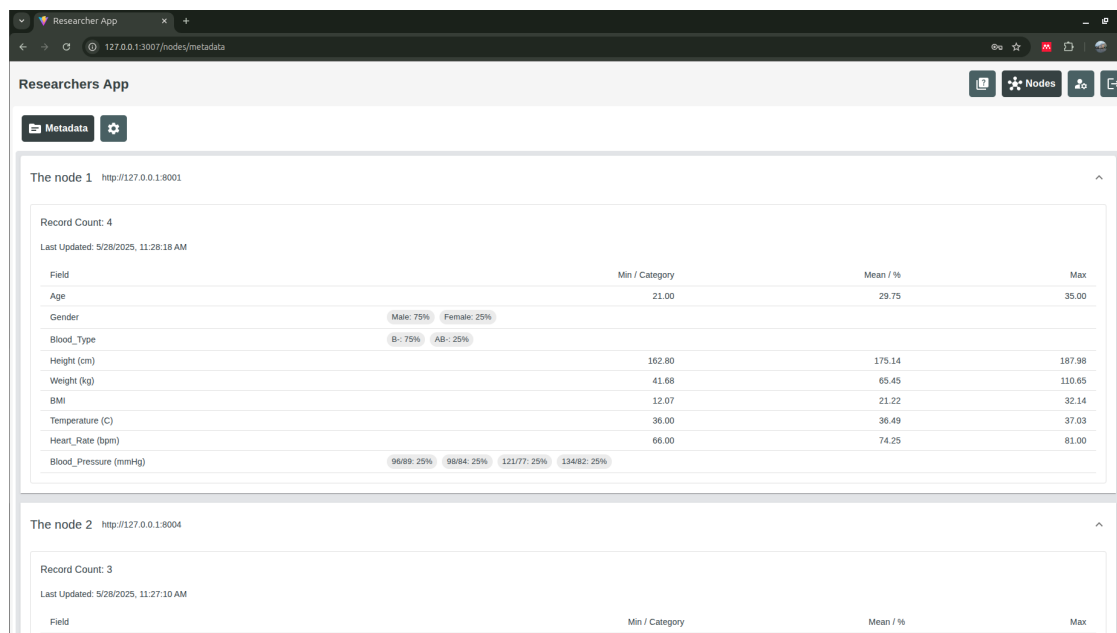


Εικόνα 2.19: Οθόνη Nodes με την εμφάνιση διαχείρισης των τοπικών κόμβων με συμπληρωμένη την φόρμα προσθήκης τους



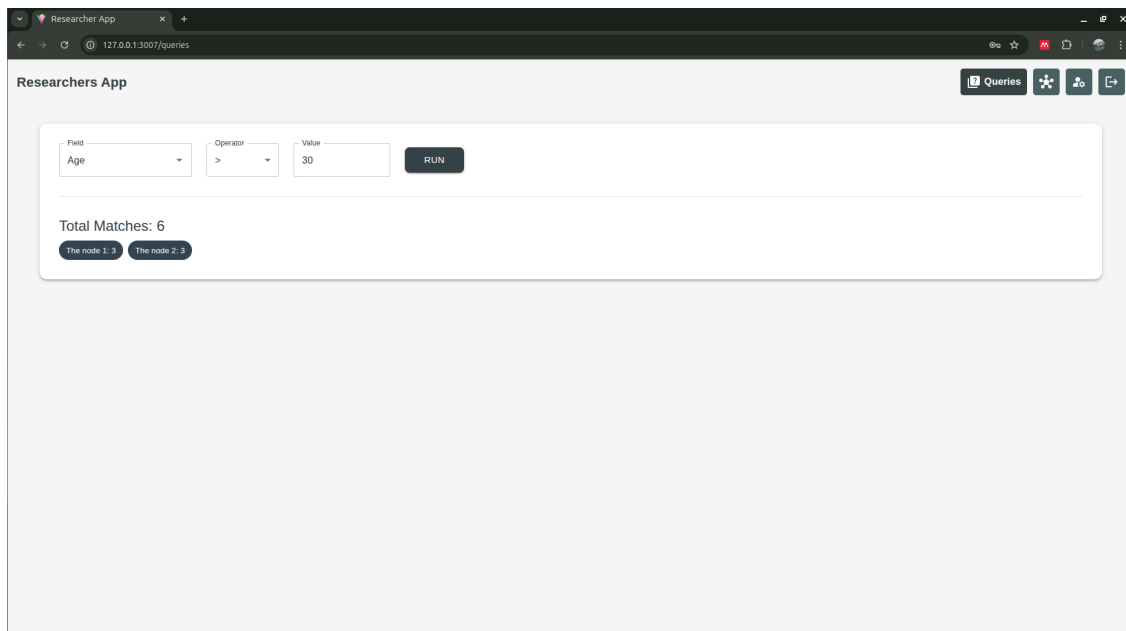
Εικόνα 2.20: Οθόνη Nodes με την εμφάνιση διαχείρισης των τοπικών κόμβων με δύο κόμβους

Σε περίπτωση που κάποιος ερευνητής επιθυμεί να αποκτήσει πρόσβαση στα πραγματικά δεδομένα ενός συγκεκριμένου τοπικού κόμβου, θα πρέπει να επικοινωνήσει απευθείας με τον αντίστοιχο φορέα που διαχειρίζεται τον κόμβο, προκειμένου να λάβει την απαραίτητη εξουσιοδότηση. Αυτή η διαδικασία εξασφαλίζει τον έλεγχο της πρόσβασης και τη συμμόρφωση με τις πολιτικές προστασίας προσωπικών δεδομένων.



Εικόνα 2.21: Οθόνη Nodes με την εμφάνιση των Metadata των δύο τοπικών κόμβων

Τέλος, όταν επιστρέφουμε στην αρχική οθόνη της εφαρμογής των ερευνητών, που ονομάζεται Queries, διακρίνουμε πλέον ότι υπάρχει η δυνατότητα επιλογής συγκεκριμένων μεταβλητών, όπως για παράδειγμα το πεδίο Age (ηλικία). Πρόκειται για μια επιλογή που δεν ήταν διαθέσιμη στην προηγούμενη εικόνα 2.16, όπου το σύστημα δεν είχε ακόμη συνδεθεί με τους τοπικούς κόμβους. Μετά την επιλογή της μεταβλητής, ο ερευνητής μπορεί να εκτελέσει προκαθορισμένους ελέγχους, οι οποίοι επιστρέφουν αποτελέσματα βασισμένα στα δεδομένα που βρίσκονται στους συνδεδεμένους τοπικούς κόμβους. Για παράδειγμα, επιλέγοντας μια συγκεκριμένη μεταβλητή και έναν αντίστοιχο έλεγχο, το σύστημα εμφανίζει το συνολικό πλήθος των εγγραφών που ικανοποιούν τα κριτήρια αυτά. Με αυτόν τον τρόπο, η εφαρμογή προσφέρει ένα εύχρηστο και αποτελεσματικό εργαλείο για τη διεξαγωγή στατιστικών αναλύσεων, χωρίς να χρειάζεται η απευθείας πρόσβαση στα ατομικά δεδομένα, διατηρώντας ταυτόχρονα την προστασία της ιδιωτικότητας και την ασφάλεια των πληροφοριών.



Εικόνα 2.22: Οθόνη Queries με την προσθήκη τοπικών κόμβων

Επίλογος

Η παρούσα διπλωματική εργασία διερεύνησε εκτενώς τις προκλήσεις και τις απαιτήσεις που προκύπτουν στη διαχείριση και ανταλλαγή δεδομένων από σύγχρονες βιοτράπεζες, παρουσιάζοντας έναν ολοκληρωμένο σχεδιασμό και υλοποίηση ενός πρωτοτύπου ομοσπονδιακού δικτύου. Το προτεινόμενο σύστημα αντιμετωπίζει αποτελεσματικά τη διαρκώς αυξανόμενη ανάγκη για ασφαλή, διαλειτουργική και ευέλικτη ανταλλαγή βιοϊατρικών δεδομένων, τηρώντας τις σύγχρονες ευρωπαϊκές κανονιστικές απαιτήσεις και τα διεθνή πρότυπα όπως ο GDPR.

Η επιλεγμένη αρχιτεκτονική δύο επιπέδων, που συνδυάζει τοπικούς κόμβους σε οργανισμούς υγείας με έναν κεντρικό κόμβο μεταδεδομένων, διασφαλίζει υψηλά επίπεδα ασφάλειας και ιδιωτικότητας, ελαχιστοποιώντας τη διακίνηση ευαίσθητων πληροφοριών. Η υλοποίηση με σύγχρονες τεχνολογίες όπως React, FastAPI, MongoDB και Keycloak επιβεβαιώνει τη δυνατότητα εύκολης εγκατάστασης, διαχείρισης και μελλοντικής επέκτασης του συστήματος.

Η συνεισφορά της εργασίας έγκειται όχι μόνο στην πρακτική υλοποίηση μιας σύγχρονης και ευέλικτης λύσης, αλλά και στη δημιουργία ενός θεωρητικού και πρακτικού υποδείγματος που μπορεί να αποτελέσει βάση για περαιτέρω ερευνητικές προσπάθειες και εφαρμογές στον τομέα των βιοτραπεζών. Μελλοντικά βήματα θα μπορούσαν να περιλαμβάνουν την ενσωμάτωση προηγμένων αναλυτικών εργαλείων, τη βελτίωση των πρωτοκόλλων διαλειτουργικότητας, καθώς και τη διεύρυνση του δικτύου σε περισσότερους οργανισμούς υγείας και ερευνητικά κέντρα, ώστε να μεγιστοποιηθεί η επιστημονική και κοινωνική αξία των δεδομένων των βιοτραπεζών.

Παράρτημα: Πηγαίος Κώδικας Διπλωματικής

Σύντομη περιγραφή αποθετηρίου

Ο πηγαίος κώδικας που χρησιμοποιήθηκε στη διπλωματική φιλοξενείται στο GitHub:

https://github.com/mardikis-konstantinos/federated_network

Οδηγίες εγκατάστασης

```
git clone https://github.com/mardikis-konstantinos/federated_network.git
cd federated_network
```

Στην συνέχεια κάθε υποφάκελος περιέχει το δικό του README.md αρχείο με τις οδηγίες χρήσης.

Path	Description
Keycloak-configs/	Ρυθμίσεις Keycloak (realms, clients, roles)
Local-Node/	Κώδικας τοπικού κόμβου που τρέχει στους φορείς
donors-frontend/	UI (ReactJS) την διαχείριση συγκατάθεσης των δω- τών/ασθενών
local-node-backend/	FastAPI REST API για όλες τις κλήσεις που αφορούν την επικοινωνία μέσα στον τοπικό κόμβο
medical-staff-frontend/	UI (ReactJS) του ιατρικού προσωπικού
queries/	FastAPI REST API για όλες τις κλήσεις που αφορούν την επικοινωνία μεταξύ τοπικού και κεντρικού κόμβου
Web-Node/	Κεντρικός κόμβος web
researchers-backend/	FastAPI REST API για όλες τις κλήσεις που αφορούν την επικοινωνία μέσα στον κεντρικό κόμβο
researchers-frontend/	UI (ReactJS) των ερευνητών
datasets/	Περιέχει διαφορετικά excel αρχεία για που αφορούν το καθένα και από έναν ασθενή. Τα δεδομένα που χρησι- μοποιήθηκαν είναι από το kaggle [52]
local-nodes-env/	Περιβάλλον Docker Compose για τοπικούς κόμβους
web-node-env/	Περιβάλλον Docker Compose για τον κεντρικό κόμβο
.gitignore	Εξαιρέσεις Git (π.χ. .env)
docker-compose.yml	Για την δημιουργία των containers

Path	Description
README.md	Λεπτομερές οδηγίες χρήσης. Κάθε φάκελος περιέχει το δικό του.

Βιβλιογραφία

- [1] A. U. Rehman, R. L. Aguiar και J. P. Barraca. *Fault-Tolerance in the Scope of Cloud Computing*. *IEEE Access*, 10:63422–63441, 2022.
- [2] Adam J Lewandowski, Martin K Rutter και Rory Collins. *Scientific and Clinical Impacts of UK Biobank in Cardiovascular Medicine*. *Circulation*, 150(24):1907–1909, 2024.
- [3] A. Bozorgchenani, D. Tarchi και G. E. Corazza. *Centralized and Distributed Architectures for Energy and Delay Efficient Fog Network-Based Edge Computing Services*. *IEEE Transactions on Green Communications and Networking*, 3(1):250–263, 2019.
- [4] F Ciancetta, B Dapice, Carmine Landi, A Pelvio και others. *Mobile sensor network architecture for environmental monitoring*, 2007.
- [5] GeeksforGeeks. *Comparison: Centralized, Decentralized and Distributed Systems*, 2021.
- [6] H. T. M. Gamage, Hesiri Weerasinghe και N. G. J. Dias. *A Survey on Blockchain Technology Concepts, Applications, and Issues*. *SN Computer Science*, 1, 2020.
- [7] R. Garratt και P. Zimmerman. *Centralized netting in financial networks*. *Journal of Banking & Finance*, 112, 2020.
- [8] L. K. Osei, Y. Cherkasova και K. M. Oware. *Unlocking the full potential of digital transformation in banking: a bibliometric review and emerging trend*. *Future Business Journal*, 9(1), 2023.
- [9] J. E. Berkes. *Decentralized Peer-to-Peer Network Architecture: Gnutella and Freenet*, 2003.
- [10] S. Mishra, R. Khatami και Y. C. Chen. *Decentralization of Energy Systems with Blockchain: Bridging Top-down and Bottom-up Management of the Electricity Grid*.
- [11] J. Moosavi, L. M. Naeni, A. M. Fathollahi-Fard και U. Fiore. *Blockchain in Supply Chain Management: A Review, Bibliometric, and Network Analysis*, 2021.
- [12] G. F. Coulouris, J. Dollimore και T. Kindberg. *Distributed Systems: Concepts and Design*. Pearson Education, 2005.
- [13] G. F. Riley και others. *A Federated Approach to Distributed Network Simulation*, 2004.

- [14] M. Wang και others. *An overview of cloud based content delivery networks: Research dimensions and state-of-the-Art. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, τόμος 9070, σελίδες 131–158, 2015.
- [15] P. Yadav και S. Kar. *A cost-efficient content distribution optimization model for fog-based content delivery networks. Journal of Cloud Computing*, 13(1), 2024.
- [16] H. Hallock και others. *Federated Networks for Distributed Analysis of Health Data. Frontiers in Public Health*, 9, 2021.
- [17] Björn Brembs, Adrian Lenardic, Peter Murray-Rust, Leslie Chan και Dasapta Irawan. *Mastodon over Mammon: towards publicly owned scholarly knowledge. Royal Society Open Science*, 10, 2023.
- [18] JE Olson, SJ Bielinski, E Ryu, EM Winkler, PY Takahashi, J Pathak και JR Cerhan. *Biobanks and personalized medicine. Clinical Genetics*, 86(1):50–55, 2014.
- [19] Yvonne G De Souza και John S Greenspan. *Biobanking past, present and future: responsibilities and benefits. AIDS*, 27(3):303–312, 2013.
- [20] A. Rebecca Reuber και Anna Morgan-Thomas. *Communicating Moral Legitimacy in Controversial Industries: The Trade in Human Tissue. Journal of Business Ethics*, 154(1):49–63, 2019.
- [21] Cathie Sudlow, John Gallacher, Naomi Allen, Valerie Beral, Paul Burton, John Danesh, Paul Downey, Paul Elliott, Jane Green, Martin Landray και others. *UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS medicine*, 12(3):e1001779, 2015.
- [22] Jasper A Bovenberg, Anna Hansell, Kees De Hoogh και Bartha Maria Knoppers. *Nature, nurture and exposure: Connecting biobank data with geographic data could yield public and individual health benefits, but risks to human rights need to be assessed. EMBO reports*, 16(4):404–406, 2015.
- [23] Committee on the Review of Omics Based Tests for Predicting Patient Outcomes in Clinical Trials, Board on Health Care Services, Board on Health Sciences Policy, Institute of Medicine, C. M. Micheel, S. J. Nass και G. S. Omenn, επιμελητές. *Evolution of Translational Omics: Lessons Learned and the Path Forward*. National Academies Press (US), Washington, DC, 2012. Chapter 2, Omics-Based Clinical Discovery: Science, Technology, and Applications. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK202165/>.
- [24] Judita Kinkorová και Ondřej Topolčan. *Biobanks in the era of big data: objectives, challenges, perspectives, and innovations for predictive, preventive, and personalised medicine. EPMA Journal*, 11(3):333–341, 2020.

- [25] Antonella Corradi, Giuseppina Bonizzi, Elham Sajjadi, Francesca Pavan, Marzia Fumagalli, Luigi Orlando Molendini, Massimo Monturano, Cristina Cassi, Camilla Rosella Musico, Luca Leoni και others. *The Regulatory Landscape of Biobanks In Europe: From Accreditation to Intellectual Property*. *Current Genomics*, 26(1):15–23, 2025.
- [26] Fotis Kitsios, Elpiniki Chatzidimitriou και Maria Kamariotou. *The ISO/IEC 27001 Information Security Management Standard: How to Extract Value from Data in the IT Sector*. *Sustainability*, 15(7), 2023.
- [27] Ramez Alkhatib και Karoline I Gaede. *Data management in biobanking: strategies, challenges, and future directions*. *BioTech*, 13(3):34, 2024.
- [28] UK Biobank Ltd. *ISO/IEC 27001:2013 Certificate of Registration*, 2021. Certificate number IS 687456.
- [29] Secfix. *Information Security Management in Healthcare: ISO 27799 under ISO 27001*, 2023.
- [30] British Standards Institution. *The Little Book of Personal Health Information (PHI): Leveraging ISO 27799 within the ISO/IEC 27000 Series*, 2024.
- [31] National Library of Medicine (NLM). *Common Data Elements (CDE) Resource Portal - Research Data Management*. <https://www.nlm.nih.gov/oet/ed/cde/tutorial/02-200.html>, 2024.
- [32] Roxana Merino-Martinez, Loreana Norlin, David van Enckevort, Gabriele Anton, Simone Schuffenhauer, Kaisa Silander, Linda Mook, Petr Holub, Raffael Bild, Morris Swertz και others. *Toward global biobank integration by implementation of the minimum information about biobank data sharing (MIABIS 2.0 Core)*. *Biopreservation and biobanking*, 14(4):298–306, 2016.
- [33] ELIXIR Europe. *RDMkit: The Research Data Management Toolkit for Life Sciences*. <https://rdmkit.elixir-europe.org/>, 2025.
- [34] ELIXIR-Greece. *Data Management Activities, Resources & Capacities*. <https://www.elixir-greece.org/el/node/269>, 2025.
- [35] Clare Bycroft, Colin Freeman, Desislava Petkova, Gavin Band, Lloyd T Elliott, Kevin Sharp, Allan Motyer, Damjan Vukcevic, Olivier Delaneau, Jared O’Connell και others. *The UK Biobank resource with deep phenotyping and genomic data*. *Nature*, 562(7726):203–209, 2018.
- [36] All of Us Research Program Investigators. *The “All of Us” research program*. *New England Journal of Medicine*, 381(7):668–676, 2019.
- [37] Marina Mora, Corrado Angelini, Fabrizia Bignami, Anne Mary Bodin, Marco Crimi, Jeanne Hélène Di Donato, Alex Felice, Cécile Jaeger, Veronika Karcagi, Yann LeCam

- και others. *The EuroBioBank Network: 10 years of hands-on experience of collaborative, transnational biobanking for rare diseases*. *European Journal of Human Genetics*, 23(9):1116–1123, 2015.
- [38] Charles Moore. *EuroBioBank: Vital Resource Treating Rare Diseases Like Friedreich’s Ataxia*, 2015.
- [39] Blockhead. *a16z Backs Decentralized Biobank Platform AminoChain*, 2024.
- [40] Lucia Monaco, Marco Crimi και Chihui Mary Wang. *The challenge for a European network of biobanks for rare diseases taken up by RD-Connect*. *Pathobiology*, 81(5-6):231–236, 2015.
- [41] Madeleine J Murtagh, Andrew Turner, Joel T Minion, Michaela Fay και Paul R Burton. *International data sharing in practice: new technologies meet old governance*. *Biopreservation and biobanking*, 14(3):231–240, 2016.
- [42] European Commission. *Biobank Standardisation and Harmonisation for Research Excellence in the European Union (BioSHaRE-EU)*, 2015.
- [43] BBMRI-ERIC. *BBMRI-ERIC Launches The Federated Platform*. <https://www.bbmri-eric.eu/news-events/bbmri-eric-launches-the-federated-platform/>, 2024. Αρχειοθετημένο: 2025-04-28.
- [44] German Biobank Node. *IT Infrastructure*. <https://www.bbmri.de/biobanking/it/infrastructure/?L=1>, 2025.
- [45] Barcelona Supercomputing Center. *Canada joins Federated EGA, marking first major expansion beyond Europe*. <https://www.bsc.es/news/bsc-news/canada-joins-federated-ega-marking-first-major-expansion-beyond-europe>, 2025.
- [46] EGA Communication. *First datasets available in the Federated EGA Network*. <https://blog.ega-archive.org/first-datasets-available-in-Federated-EGA>, 2024.
- [47] Jordi Rambla, Michael Baudis, Roberto Ariosa, Tim Beck, Lauren A Fromont, Arcadi Navarro, Rahel Paloots, Manuel Rueda, Gary Saunders, Babita Singh και others. *Beacon v2 and Beacon networks: a “lingua franca” for federated data discovery in biomedical genomics, and beyond*. *Human mutation*, 43(6):791–799, 2022.
- [48] CINECA Consortium. *Common Infrastructure for National Cohorts in Europe, Canada, and Africa (CINECA)*. <https://www.cineca-project.eu/>, 2022.
- [49] Friends of the National Institute of Mental Health. *Mental Health Data in the All of Us Research Program: Opportunities for Participants and Researchers*, 2023.
- [50] Petr Holub, Morris Swertz, Robert Reihs, David van Enckevort, Heimo Müller και Jan Eric Litton. *BBMRI-ERIC directory: 515 biobanks with over 60 million biological samples*. *Biopreservation and biobanking*, 14(6):559–562, 2016.

- [51] Global Alliance for Genomics and Health. *Extensions to the GA4GH Beacon API will enable a more powerful community resource*, 2019.
- [52] Tariq Aziz Rao. *Random Healthcare Dataset (Sample)*. <https://www.kaggle.com/datasets/raotariq/random-healthcare-dataset-sample>, 2023.