

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΕΥΦΥΕΙΣ ΤΕΧΝΟΛΟΓΙΕΣ ΔΙΑΔΙΚΤΥΟΥ»

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Οντολογική αναπαράσταση και τεκμηρίωση συμβάντων
παρακολούθησης υγείας υπό το πρίσμα του κανονισμού
EU AI Act: Μελέτη περίπτωσης κλινικών συναγερμών



Του Μεταπτυχιακού Φοιτητή
Μιχαήλ Νιόπα
Αρ. Μητρώου: 25/2024

Επιβλέπων
Περικλής Χατζημίσιος
Καθηγητής

Σεπτέμβριος 2026

Τίτλος Δ.Ε. Οντολογική αναπαράσταση και τεκμηρίωση συμβάντων παρακολούθησης υγείας υπό το πρίσμα του κανονισμού EU AI Act: Μελέτη περίπτωσης κλινικών συναγερμών

Κωδικός Δ.Ε. 25358

Ονοματεπώνυμο φοιτητή: ΜΙΧΑΗΛ ΝΙΟΠΑΣ

Ονοματεπώνυμο εισηγητή: ΠΕΡΙΚΛΗΣ ΧΑΤΖΗΜΙΣΙΟΣ

Ημερομηνία ανάληψης Δ.Ε. 30/8/2025

Ημερομηνία περάτωσης Δ.Ε. 12/9/2026

Βεβαιώνω ότι είμαι ο συγγραφέας αυτής της εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, έχω καταγράψει τις όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών, εικόνων και κειμένου, είτε αυτές αναφέρονται ακριβώς είτε παραφρασμένες. Επιπλέον, βεβαιώνω ότι αυτή η εργασία προετοιμάστηκε από εμένα προσωπικά, ειδικά ως διπλωματική εργασία, στο Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του ΔΙ.ΠΑ.Ε.

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του φοιτητή ΜΙΧΑΗΛ ΝΙΟΠΑ που την εκπόνησε. Στο πλαίσιο της πολιτικής ανοικτής πρόσβασης, ο συγγραφέας/δημιουργός εκχωρεί στο Διεθνές Πανεπιστήμιο της Ελλάδος άδεια χρήσης του δικαιώματος αναπαραγωγής, δανεισμού, παρουσίασης στο κοινό και ψηφιακής διάχυσης της εργασίας διεθνώς, σε ηλεκτρονική μορφή και σε οποιοδήποτε μέσο, για διδακτικούς και ερευνητικούς σκοπούς, άνευ ανταλλάγματος. Η ανοικτή πρόσβαση στο πλήρες κείμενο της εργασίας, δεν σημαίνει καθ' οιονδήποτε τρόπο παραχώρηση δικαιωμάτων διανοητικής ιδιοκτησίας του συγγραφέα/δημιουργού, ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αντιγραφή, πώληση, εμπορική χρήση, διανομή, έκδοση, μεταφόρτωση (downloading), ανάρτηση (uploading), μετάφραση, τροποποίηση με οποιονδήποτε τρόπο, τμηματικά ή περιληπτικά της εργασίας, χωρίς τη ρητή προηγούμενη έγγραφη συναίνεση του συγγραφέα/δημιουργού.

Η έγκριση της διπλωματικής εργασίας από το Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του Διεθνούς Πανεπιστημίου της Ελλάδος, δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του συγγραφέα, εκ μέρους του Τμήματος.

*«Στη σύζυγό μου, Νίκη, και στα δύο παιδιά μας,
που με την υπομονή, την αγάπη και την αμέριστη υποστήριξή τους
στάθηκαν δίπλα μου σε κάθε βήμα αυτής της διαδρομής.»*

Πρόλογος

Η επιλογή του θέματος της παρούσας διπλωματικής εργασίας προέκυψε από το ιδιαίτερο ενδιαφέρον που παρουσιάζει η σύγχρονη εξέλιξη της τεχνητής νοημοσύνης και, κυρίως, η ανάγκη σύνδεσης των τεχνικών δυνατοτήτων της με ζητήματα διαφάνειας, τεκμηρίωσης, ιχνηλασιμότητας και υπεύθυνης χρήσης.

Η θέσπιση του AI Act δημιουργεί ένα νέο κανονιστικό περιβάλλον για τα συστήματα τεχνητής νοημοσύνης στην Ευρωπαϊκή Ένωση, ενώ παράλληλα οι οντολογίες και οι τεχνολογίες του Σημαιολογικού Ιστού προσφέρουν δυνατότητες δομημένης και μηχανικά επεξεργάσιμης αναπαράστασης γνώσης. Η διερεύνηση της σύνδεσης αυτών των περιοχών με πραγματικά δεδομένα από healthcare monitoring αποτέλεσε το βασικό κίνητρο για την ανάπτυξη της παρούσας εργασίας.

Κατά την εκπόνησή της είχα την ευκαιρία να εμβαθύνω στον κανονισμό AI Act, στην οντολογική μοντελοποίηση και στην πρακτική ανάπτυξη οντολογιών με το Protégé. Παράλληλα, η αξιοποίηση πραγματικών VTaC alarm events επέτρεψε τη μετάβαση από τη θεωρητική μελέτη σε μια συγκεκριμένη proof-of-concept εφαρμογή.

Η διαδικασία αυτή αποτέλεσε σημαντική εμπειρία τόσο σε επιστημονικό όσο και σε τεχνικό επίπεδο και συνέβαλε στην καλύτερη κατανόηση της σχέσης μεταξύ πραγματικών δεδομένων, σημαιολογικής τεκμηρίωσης και κανονιστικών πλαισίων.

Περίληψη

Η παρούσα διπλωματική εργασία διερευνά τη δυνατότητα αξιοποίησης οντολογικής μοντελοποίησης για τη σημασιολογική αναπαράσταση και τεκμηρίωση πραγματικών συμβάντων παρακολούθησης υγείας και τη σύνδεση της σχετικής τεκμηρίωσης με επιλεγμένα σημεία αναφοράς του κανονισμού AI Act. Ως εφαρμοσμένο πεδίο χρησιμοποιείται το σύνολο δεδομένων Ventricular Tachycardia annotated alarms from Intensive Care Units (VTaC), το οποίο χρησιμοποιείται για την ανάπτυξη και αξιολόγηση αλγορίθμων που μειώνουν τους ψευδείς συναγερμούς κοιλιακής ταχυκαρδίας στις Μονάδες Εντατικής Θεραπείας (ΜΕΘ), από το οποίο επιλέχθηκαν 12 πραγματικά ventricular tachycardia alarm events, έξι true και έξι false alarms, μαζί με τα αντίστοιχα waveform records και τις διαθέσιμες expert αποφάσεις.

Αναπτύχθηκε στο Protégé μια proof-of-concept οντολογία, στην οποία τα alarm events οργανώνονται γύρω από συνδεδεμένες σημασιολογικές διαδρομές που αφορούν την προέλευση του συμβάντος, την ανθρώπινη αξιολόγηση, την τελική απόφαση και την τεκμηρίωσή του. Κεντρικό στοιχείο αποτελεί το Evidence Record, το οποίο λειτουργεί ως ενδιάμεσος τεκμηριωτικός κόμβος για το requirement-to-evidence mapping προς τέσσερα επιλεγμένα AI Act-related σημεία αναφοράς: logging, human review, traceability και risk management. Για τα false alarm cases προστέθηκε επίσης ένα lightweight risk-related επίπεδο με Risk Indicator, Risk-Relevant Event και Potential Harm.

Η τελική οντολογία περιλαμβάνει 23 classes, 12 object properties, 9 data properties και 76 individuals. Η λειτουργική επικύρωση πραγματοποιήθηκε με HermiT reasoner, 12 DL competency questions και 4 SPARQL queries. Ο reasoner δεν εντόπισε λογική ασυνέπεια στην εξεταζόμενη έκδοση, ενώ τα ερωτήματα επέτρεψαν την ανάκτηση των βασικών σχέσεων και mappings που είχαν σχεδιαστεί.

Τα αποτελέσματα δείχνουν ότι η οντολογική προσέγγιση μπορεί να οργανώσει πραγματικά healthcare monitoring events σε ένα συνεκτικό και αναζητήσιμο σημασιολογικό πλαίσιο. Η εργασία δεν αναπτύσσει μοντέλο τεχνητής νοημοσύνης, δεν χαρακτηρίζει το VTaC ως AI system και δεν αποδεικνύει νομική συμμόρφωση με τον κανονισμό AI Act αλλά εξετάζει τη δυνατότητα δομημένης τεκμηρίωσης και ιχνηλασιμότητας σε ένα περιορισμένο proof-of-concept.

«Ontology-Based Representation and Documentation of Health Monitoring Events under the EU AI Act: A Case Study of Clinical Alarms»

Michail Niopas

Abstract

This thesis investigates the use of ontology-based modelling for the semantic representation and documentation of real healthcare monitoring events and for linking the resulting documentation to selected AI Act reference points. The Ventricular Tachycardia annotated alarms from Intensive Care Units (VTaC) dataset is used as the application domain. Twelve real ventricular tachycardia alarm events were selected, including six true alarms and six false alarms, together with their corresponding waveform records and available expert decisions.

A proof-of-concept ontology was developed in Protégé, where alarm events are organised through interconnected semantic paths representing event provenance, human review, final decision, and documentation. A central element of the model is the Evidence Record, which acts as an intermediate documentation node for requirement-to-evidence mapping to four selected AI Act-related reference points: logging, human review, traceability, and risk management. For the false alarm cases, a lightweight risk-related layer was additionally introduced using Risk Indicator, Risk-Relevant Event, and Potential Harm concepts.

The final ontology contains 23 classes, 12 object properties, 9 data properties, and 76 individuals. Functional validation was performed using the HermiT reasoner, 12 DL competency questions, and 4 SPARQL queries. No logical inconsistency was detected by the reasoner in the examined ontology version, while the queries enabled the retrieval of the main relationships and mappings defined in the model.

The results indicate that an ontology-based approach can organise real healthcare monitoring events within a coherent and queryable semantic framework. The thesis does not develop an artificial intelligence model, does not classify the VTaC dataset as an AI system, and does not demonstrate legal compliance with the AI Act. Instead, it examines the potential of structured documentation, traceability, and evidence mapping within a limited proof-of-concept implementation.

Ευχαριστίες

Με την ολοκλήρωση της παρούσας διπλωματικής εργασίας θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες στον επιβλέποντα Καθηγητή μου, Δρ. Περικλή Χατζημίσιο, για την καθοδήγηση, την εμπιστοσύνη, τις ουσιαστικές παρατηρήσεις και τις πολύτιμες συμβουλές του σε όλη τη διάρκεια εκπόνησης της εργασίας. Η συμβολή του υπήρξε ιδιαίτερα σημαντική τόσο στη διαμόρφωση του ερευνητικού προσανατολισμού όσο και στην ολοκλήρωση της τελικής προσέγγισης.

Θα ήθελα επίσης να ευχαριστήσω όλους τους διδάσκοντες του Προγράμματος Μεταπτυχιακών Σπουδών «Ευφυείς Τεχνολογίες Διαδικτύου», για τις γνώσεις, τα ερεθίσματα και τις εμπειρίες που μου προσέφεραν κατά τη διάρκεια των σπουδών μου.

Ιδιαίτερη αναφορά αξίζει στους συμφοιτητές μου, με τους οποίους μοιραστήκαμε αυτό το κοινό ακαδημαϊκό ταξίδι. Οι συζητήσεις, η συνεργασία, η αλληλοϋποστήριξη, αλλά και οι δυσκολίες και οι όμορφες στιγμές που ζήσαμε μαζί, έκαναν αυτή τη διαδρομή πιο ουσιαστική και ξεχωριστή.

Τις πιο θερμές μου ευχαριστίες οφείλω στη σύζυγό μου και στα δύο παιδιά μας, για την υπομονή, την κατανόηση, την αγάπη και τη διαρκή υποστήριξή τους. Η παρουσία και η ενθάρρυνσή τους αποτέλεσαν το σημαντικότερο στήριγμα σε όλη τη διάρκεια αυτής της προσπάθειας.

Τέλος, ευχαριστώ όλους όσους, με οποιονδήποτε τρόπο, στάθηκαν δίπλα μου και συνέβαλαν στην ολοκλήρωση αυτού του σημαντικού προσωπικού και ακαδημαϊκού στόχου.

Περιεχόμενα

Πρόλογος.....	v
Περίληψη.....	vi
Abstract	vii
Ευχαριστίες	viii
Περιεχόμενα	ix
Κατάλογος Σχημάτων	xiv
Κατάλογος Πινάκων.....	xvii
Συντομογραφίες.....	xviii
Κεφάλαιο 1ο: Εισαγωγή και πλαίσιο της εργασίας	19
1.1 Εισαγωγή	19
1.2 Γενικό πλαίσιο της εργασίας	20
1.3 Πρόβλημα που εξετάζεται	21
1.4 Σκοπός της διπλωματικής εργασίας	22
1.5 Ερευνητικά ερωτήματα	23
1.6 Συμβολή της εργασίας.....	24
1.7 Δομή της εργασίας	25
1.8 Επίλογος	26
Κεφάλαιο 2ο: Θεωρητικό υπόβαθρο και σχετική βιβλιογραφία.....	28
2.1 Εισαγωγή	28
2.2 Τεχνητή νοημοσύνη και συστήματα παρακολούθησης υγείας	28
2.3 Παρακολούθηση υγείας και συμβάντα συναγερμού (alarm events)	30
2.4 True alarms, false alarms και alert fatigue	31
2.5 Διαφάνεια, τεκμηρίωση και ιχνηλασιμότητα σε συστήματα τεχνητής νοημοσύνης.....	33
2.6 Ο AI Act ως κανονιστικό πλαίσιο αναφοράς.....	35
2.7 Οντολογίες και αναπαράσταση γνώσης	40
2.8 Οντολογίες για Ethical AI / Trustworthy AI / AI governance.....	43
2.9 Κενό στη βιβλιογραφία	48
2.10 Επίλογος	50
Κεφάλαιο 3ο: Μεθοδολογία και σχεδιασμός της οντολογίας.....	52
3.1 Εισαγωγή	52
3.2 Ερευνητική προσέγγιση της εργασίας.....	53
3.3 Γιατί οντολογία και όχι μηχανική μάθηση	55

3.4	Το VTaC dataset και ο ρόλος του στην εργασία.....	58
3.5	Επιλογή των VTaC events για το proof-of-concept	61
3.6	Επιλογή των AI Act-related σημείων αναφοράς.....	64
3.7	Βασικές σημασιολογικές διαδρομές της οντολογίας.....	66
3.8	Lightweight risk-related επίπεδο για false alarm cases	68
3.9	Συνολική αρχιτεκτονική της οντολογίας.....	70
3.9.1	Operational layer	72
3.9.2	Evidence and assessment layer.....	73
3.9.3	Normative / AI Act-related layer.....	74
3.9.4	Lightweight risk layer.....	75
3.10	Σχεδιαστικές επιλογές	77
3.11	Μεθοδολογία ανάπτυξης στο Protégé	79
3.12	Όρια της μεθοδολογικής προσέγγισης	83
3.13	Επίλογος.....	85
Κεφάλαιο 4ο:	Υλοποίηση της οντολογίας στο Protégé	86
4.1	Εισαγωγή.....	86
4.2	Περιβάλλον υλοποίησης.....	86
4.3	Γενική δομή της οντολογίας.....	88
4.4	Χρήση annotations στην οντολογία	89
4.5	Classes της οντολογίας.....	91
4.5.1	Evidence and Assessment Concept	92
4.5.2	Compliance Assessment.....	92
4.5.3	Control Mechanism	93
4.5.4	Evidence Record.....	94
4.5.5	Normative Concept.....	96
4.5.6	Compliance Requirement	96
4.5.7	Human Review Requirement.....	97
4.5.8	Logging Requirement.....	98
4.5.9	Risk Management Requirement	99
4.5.10	Traceability Requirement.....	99
4.5.11	Operational Concept.....	100
4.5.12	Alert Event	101
4.5.13	Ventricular Tachycardia Alarm.....	102
4.5.14	Expert Review	103
4.5.15	Monitoring System.....	104

4.5.16	Review Decision	105
4.5.17	False Alarm	106
4.5.18	True Alarm	107
4.5.19	Waveform Record	108
4.5.20	Risk Concept	108
4.5.21	Potential Harm	109
4.5.22	Risk Indicator	110
4.5.23	Risk-Relevant Event.....	111
4.5.24	Συνολική αποτίμηση της ιεραρχίας των classes.....	112
4.6	Object properties της οντολογίας	113
4.6.1	assessesEvent.....	114
4.6.2	assessesRequirement	115
4.6.3	belongsToRecord.....	116
4.6.4	hasAlertEvent	117
4.6.5	hasDecision.....	118
4.6.6	hasReview.....	120
4.6.7	hasRiskIndicator	121
4.6.8	isClassifiedAs	122
4.6.9	isDocumentedBy	123
4.6.10	mayLeadTo	124
4.6.11	supportsRequirement.....	125
4.6.12	usesEvidence	127
4.6.13	Συνολική αποτίμηση των object properties.....	128
4.7	Data properties της οντολογίας	129
4.7.1	assessmentOutcome.....	130
4.7.2	decisionLabel.....	131
4.7.3	eventId	131
4.7.4	hasHumanReview	132
4.7.5	isLogged	133
4.7.6	isRiskRelevant	134
4.7.7	isTraceable.....	135
4.7.8	potentialHarmLabel	136
4.7.9	recordId.....	137
4.7.10	Συνολική αποτίμηση των data properties.....	138
4.8	Individuals της οντολογίας.....	139

4.8.1	Requirement individuals.....	141
4.8.2	Control mechanism individuals	142
4.8.3	Monitoring system individual.....	143
4.8.4	Waveform record individuals	144
4.8.5	VTaC alert event individuals.....	145
4.8.6	Expert review individuals	148
4.8.7	Review decision individuals	150
4.8.8	Evidence record individuals	152
4.8.9	Risk-related individuals.....	155
4.8.10	Παράδειγμα συνδεδεμένων σημασιολογικών διαδρομών για επιλεγμένο individual	158
4.8.11	Συνολική αποτίμηση των individuals.....	160
4.9	Συνολική αποτίμηση της υλοποίησης	161
4.10	Επίλογος.....	163
Κεφάλαιο 5ο:	Έλεγχος και λειτουργική επικύρωση της οντολογίας.....	164
5.1	Εισαγωγή.....	164
5.2	Σκοπός και λογική της λειτουργικής επικύρωσης.....	164
5.3	Έλεγχος συνέπειας με reasoner	166
5.4	Έλεγχος inferred γνώσης στην ιεραρχία κλάσεων	167
5.5	Έλεγχος της οντολογίας μέσω DL competency questions	169
5.5.1	DL Query 1 – Εντοπισμός των VTaC alert events	171
5.5.2	DL Query 2 – Ποια VT alarm events έχουν αξιολογηθεί ως true alarms;.....	172
5.5.3	DL Query 3 – Ποια VT alarm events έχουν αξιολογηθεί ως false alarms;	173
5.5.4	DL Query 4 – Ποια VT alarm events διαθέτουν evidence για logging;.....	174
5.5.5	DL Query 5 – Ποια VT alarm events διαθέτουν evidence για human review;	175
5.5.6	DL Query 6 – Ποια VT alarm events διαθέτουν evidence για traceability;	176
5.5.7	DL Query 7 – Ποια Evidence Records υποστηρίζουν ταυτόχρονα τα τρία βασικά requirement mappings;	177
5.5.8	DL Query 8 – Ποια false alarm events έχουν ενταχθεί στο risk layer;	178
5.5.9	DL Query 9 – Ποια risk-relevant events συνδέονται με potential harm;	179
5.5.10	DL Query 10 – Ποια Evidence Records έχουν mapping προς το req_risk_management; 180	
5.5.11	DL Query 11 – Ποια events διαθέτουν πλήρη σχεδιασμένη διαδρομή ιχνηλασιμότητας; 181	
5.5.12	DL Query 12 – Ποια waveform records συνδέονται με VT alarm events;	182
5.5.13	Συνολική αποτίμηση των DL competency questions.....	183

5.6	Έλεγχος και αναφορά αποτελεσμάτων μέσω SPARQL queries	184
5.6.1	SPARQL Query 1 – Συγκεντρωτική ανάκτηση event–record–decision–evidence	185
5.6.2	SPARQL Query 2 – Χαρτογράφηση event–evidence–requirement.....	186
5.6.3	SPARQL Query 3 – Risk-related mapping για false alarm cases.....	188
5.6.4	SPARQL Query 4 – Πλήθος VT alarm events ανά waveform record.....	189
5.6.5	Συνολική αποτίμηση των SPARQL queries	191
5.7	Συνολική αποτίμηση και επίλογος κεφαλαίου	191
Κεφάλαιο 6ο:	Συζήτηση, συμπεράσματα και μελλοντικές επεκτάσεις	194
6.1	Εισαγωγή.....	194
6.2	Συνολική ερμηνεία των αποτελεσμάτων.....	195
6.3	Απάντηση στα ερευνητικά ερωτήματα.....	197
6.4	Συμβολή της εργασίας.....	201
6.5	Περιορισμοί της εργασίας	203
6.6	Μελλοντικές επεκτάσεις.....	205
6.7	Τελικό συμπέρασμα	207
BIBΛΙΟΓΡΑΦΙΑ.....		209
ΠΑΡΑΡΤΗΜΑ A: Συγκεντρωτική τεχνική προδιαγραφή της οντολογίας		212
ΠΑΡΑΡΤΗΜΑ B: DL competency questions και SPARQL queries.....		218

Κατάλογος Σχημάτων

Σχήμα 2.1 Ενδεικτικά παραδείγματα true και false ventricular tachycardia alarm events από το VTaC dataset. (προσαρμογή από [4]).	32
Σχήμα 2.2 Το πλαίσιο των Ethics Guidelines για το Trustworthy AI [1].	34
Σχήμα 2.3 Risk-based προσέγγιση του AI Act και βασικές κατηγορίες κινδύνου. Ευρωπαϊκή Επιτροπή [18].	36
Σχήμα 2.4 Βασικές έννοιες και σχέσεις της TAIR ontology [5].	44
Σχήμα 2.5 Φάσεις χαρτογράφησης εννοιών και απαιτήσεων από κανονιστικά και τεχνικά κείμενα [5].	45
Σχήμα 2.6 Προδιαγραφή απαιτήσεων και competency questions της FRIA ontology (προσαρμογή από [17]).	46
Σχήμα 3.1 Ερευνητική λογική της οντολογικής προσέγγισης.	55
Σχήμα 3.2 Διάκριση μεταξύ μηχανικής μάθησης και οντολογικής μοντελοποίησης στην παρούσα εργασία.	58
Σχήμα 3.3 Ρόλος του VTaC dataset στην οντολογική υλοποίηση.	61
Σχήμα 3.4 Ανώτερη αρχιτεκτονική οργάνωση της οντολογίας.	71
Σχήμα 3.5 Ανάλυση του Operational Concept στην αρχιτεκτονική της οντολογίας.	73
Σχήμα 3.6 Ανάλυση του Evidence and Assessment Concept στην αρχιτεκτονική της οντολογίας.	74
Σχήμα 3.7 Ανάλυση του Normative Concept στην αρχιτεκτονική της οντολογίας.	75
Σχήμα 3.8 Ανάλυση του Risk Concept στην αρχιτεκτονική της οντολογίας.	76
Σχήμα 4.1 Γενική προβολή της οντολογίας στο Protégé.	87
Σχήμα 4.2 Παράδειγμα χρήσης annotations σε requirement individual.	90
Σχήμα 4.3 Η class Evidence and Assessment Concept και τα annotations της στο Protégé.	92
Σχήμα 4.4 Η class Evidence and Assessment Concept και οι βασικές υποκλάσεις της.	92
Σχήμα 4.5 Η class Compliance Assessment και τα annotations της στο Protégé.	93
Σχήμα 4.6 Η class Compliance Assessment χωρίς direct instances στην παρούσα proof-of-concept υλοποίηση.	93
Σχήμα 4.7 Η class Control Mechanism και τα annotations της στο Protégé.	94
Σχήμα 4.8 Η class Control Mechanism με τα individuals της.	94
Σχήμα 4.9 Η class Evidence Record και τα annotations της στο Protégé.	95
Σχήμα 4.10 Η class Evidence Record με τα individuals τεκμηρίωσης.	95
Σχήμα 4.11 Η class Normative Concept και τα annotations της στο Protégé.	96
Σχήμα 4.12 Θέση της class Normative Concept στην ιεραρχία της οντολογίας.	96
Σχήμα 4.13 Η class Compliance Requirement και τα annotations της στο Protégé.	97
Σχήμα 4.14 Υποκλάσεις της class Compliance Requirement.	97
Σχήμα 4.15 Η class Human Review Requirement και τα annotations της στο Protégé.	98
Σχήμα 4.16 Η class Human Review Requirement και το αντίστοιχο requirement individual.	98
Σχήμα 4.17 Η class Logging Requirement και τα annotations της στο Protégé.	98
Σχήμα 4.18 Η class Logging Requirement και το αντίστοιχο requirement individual.	99
Σχήμα 4.19 Η class Risk Management Requirement και τα annotations της στο Protégé.	99
Σχήμα 4.20 Η class Risk Management Requirement και το αντίστοιχο requirement individual.	99
Σχήμα 4.21 Η class Traceability Requirement και τα annotations της στο Protégé.	100
Σχήμα 4.22 Η class Traceability Requirement και το αντίστοιχο requirement individual.	100
Σχήμα 4.23 Η class Operational Concept και τα annotations της στο Protégé.	100
Σχήμα 4.24 Υποκλάσεις της class Operational Concept.	101

Σχήμα 4.25 H class Alert Event και τα annotations της στο Protégé.....	101
Σχήμα 4.26 Θέση της class Alert Event στο λειτουργικό επίπεδο της οντολογίας.....	102
Σχήμα 4.27 H class Ventricular Tachycardia Alarm και τα annotations της στο Protégé.....	102
Σχήμα 4.28 H class Ventricular Tachycardia Alarm με τα 12 πραγματικά VTaC event individuals.....	103
Σχήμα 4.29 H class Expert Review και τα annotations της στο Protégé.....	103
Σχήμα 4.30 H class Expert Review με τα 12 review individuals των VTaC events.....	104
Σχήμα 4.31 H class Monitoring System και τα annotations της στο Protégé.....	104
Σχήμα 4.32 H class Monitoring System και το αντίστοιχο individual της.....	105
Σχήμα 4.33 H class Review Decision και τα annotations της στο Protégé.....	105
Σχήμα 4.34 Υποκλάσεις της class Review Decision.....	105
Σχήμα 4.35 H class False Alarm και τα annotations της στο Protégé.....	106
Σχήμα 4.36 H class False Alarm με τα αντίστοιχα decision individuals.....	106
Σχήμα 4.37 H class True Alarm και τα annotations της στο Protégé.....	107
Σχήμα 4.38 H class True Alarm με τα αντίστοιχα decision individuals.....	107
Σχήμα 4.39 H class Waveform Record και τα annotations της στο Protégé.....	108
Σχήμα 4.40 H class Waveform Record με τα αντίστοιχα record individuals.....	108
Σχήμα 4.41 H class Risk Concept και τα annotations της στο Protégé.....	109
Σχήμα 4.42 Υποκλάσεις της class Risk Concept.....	109
Σχήμα 4.43 H class Potential Harm και τα annotations της στο Protégé.....	110
Σχήμα 4.44 H class Potential Harm και το αντίστοιχο individual vtac_harm_alert_fatigue.....	110
Σχήμα 4.45 H class Risk Indicator και τα annotations της στο Protégé.....	110
Σχήμα 4.46 H class Risk Indicator με τα αντίστοιχα indicator individuals.....	111
Σχήμα 4.47 H class Risk-Relevant Event και τα annotations της στο Protégé.....	111
Σχήμα 4.48 H class Risk-Relevant Event με τα αντίστοιχα risk event individuals.....	112
Σχήμα 4.49 Ιεραρχία των object properties της οντολογίας στο Protégé.....	114
Σχήμα 4.50 Object property assessesEvent στο Protégé μαζί με τα annotations, το domain και το range του.....	115
Σχήμα 4.51 Object property assessesRequirement στο Protégé μαζί με τα annotations, το usage, το domain και το range του.....	116
Σχήμα 4.52 Object property belongsToRecord στο Protégé μαζί με τα annotations, το usage, το domain και το range του.....	117
Σχήμα 4.53 Object property hasAlertEvent στο Protégé μαζί με τα annotations, το usage, το domain και το range του.....	118
Σχήμα 4.54 Object property hasDecision στο Protégé μαζί με τα annotations, το usage, το domain και το range του.....	119
Σχήμα 4.55 Object property hasReview στο Protégé μαζί με τα annotations, το usage, το domain και το range του.....	120
Σχήμα 4.56 Object property hasRiskIndicator στο Protégé μαζί με τα annotations, το usage, το domain και το range του.....	122
Σχήμα 4.57 Object property isClassifiedAs στο Protégé μαζί με τα annotations, το usage, το domain και το range του.....	123
Σχήμα 4.58 Object property isDocumentedBy στο Protégé μαζί με τα annotations, το usage, το domain και το range του.....	124
Σχήμα 4.59 Object property mayLeadTo στο Protégé μαζί με τα annotations, το usage, το domain και το range του.....	125

Σχήμα 4.60 Object property supportsRequirement στο Protégé μαζί με τα annotations, το usage, το domain και το range του.....	126
Σχήμα 4.61 Object property usesEvidence στο Protégé μαζί με τα annotations, το usage, το domain και το range του.....	128
Σχήμα 4.62 Ιεραρχία των data properties της οντολογίας στο Protégé.....	130
Σχήμα 4.63 Data property assessmentOutcome στο Protégé.....	130
Σχήμα 4.64 Data property decisionLabel στο Protégé.....	131
Σχήμα 4.65 Data property eventId στο Protégé.....	132
Σχήμα 4.66 Data property hasHumanReview στο Protégé.....	133
Σχήμα 4.67 Data property isLogged στο Protégé.....	134
Σχήμα 4.68 Data property isRiskRelevant στο Protégé.....	135
Σχήμα 4.69 Data property isTraceable στο Protégé.....	136
Σχήμα 4.70 Data property potentialHarmLabel στο Protégé.....	137
Σχήμα 4.71 Data property recordId στο Protégé.....	138
Σχήμα 4.72 Requirement individuals που αναπαριστούν τα επιλεγμένα AI Act-related σημεία αναφοράς της οντολογίας.....	142
Σχήμα 4.73 Individuals της class Control Mechanism στο Protégé.....	143
Σχήμα 4.74 Ενδεικτική απεικόνιση της σχέσης μεταξύ waveform record και alert events.....	144
Σχήμα 4.75 VTaC alert event individuals και ιδιότητες επιλεγμένου event στο Protégé.....	147
Σχήμα 4.76 Individuals της class Expert Review στο Protégé.....	149
Σχήμα 4.77 Αναπαράσταση των individuals της class Review Decision.....	151
Σχήμα 4.78 Evidence Record individuals και mappings προς τα επιλεγμένα AI Act-related σημεία αναφοράς στο Protégé.....	154
Σχήμα 4.79 Risk-related individuals και σύνδεση με potential harm.....	157
Σχήμα 4.80 Άμεσες object και data property assertions του vtac_event_003c13_0115.....	160
Σχήμα 5.1 Ενεργοποίηση του HermiT reasoner στο Protégé.....	166
Σχήμα 5.2 Προβολή inferred ιεραρχίας των classes μετά την εκτέλεση του HermiT reasoner.....	168
Σχήμα 5.3 DL Query 1 — Εντοπισμός των VTaC alert event individuals.....	171
Σχήμα 5.4 DL Query 2 — VT alarm events που έχουν αξιολογηθεί ως true alarms.....	172
Σχήμα 5.5 DL Query 3 — VT alarm events που έχουν αξιολογηθεί ως false alarms.....	173
Σχήμα 5.6 DL Query 4 — VT alarm events με evidence για logging.....	174
Σχήμα 5.7 DL Query 5 — VT alarm events με evidence για human review.....	175
Σχήμα 5.8 DL Query 6 — VT alarm events με evidence για traceability.....	176
Σχήμα 5.9 DL Query 7 — Evidence Records που υποστηρίζουν τα τρία βασικά requirement mappings.....	177
Σχήμα 5.10 DL Query 8 — False alarm events στο risk layer.....	178
Σχήμα 5.11 DL Query 9 — Risk-relevant events με potential harm.....	179
Σχήμα 5.12 DL Query 10 — Evidence Records με mapping προς req_risk_management.....	180
Σχήμα 5.13 DL Query 11 — Events με πλήρη σχεδιασμένη διαδρομή ιχνηλασιμότητας.....	181
Σχήμα 5.14 DL Query 12 — Waveform records με VT alarm events.....	183
Σχήμα 5.15 SPARQL Query 1 — Συγκεντρωτική ανάκτηση event-record-decision-evidence.....	186
Σχήμα 5.16 SPARQL Query 2 — Χαρτογράφηση event-evidence-requirement.....	187
Σχήμα 5.17 SPARQL Query 3 — Risk-related mapping για false alarm cases.....	189
Σχήμα 5.18 SPARQL Query 4 — Πλήθος VT alarm events ανά waveform record.....	190

Κατάλογος Πινάκων

Πίνακας 1-1 Σύνδεση ερευνητικών ερωτημάτων με τα κεφάλαια της εργασίας	24
Πίνακας 2-1 Βασικές απαιτήσεις του AI Act και σχέση με την οντολογική μοντελοποίηση	37
Πίνακας 2-2 Ενδεικτικές οντολογικές προσεγγίσεις για Ethical AI, Trustworthy AI και AI governance	47
Πίνακας 2-3 Σύνθεση βιβλιογραφικού κενού	49
Πίνακας 3-1 Waveform record individuals και πλήθος επιλεγμένων VTaC events.	62
Πίνακας 3-2 Επιλεγμένα VTaC alarm events για την proof-of-concept υλοποίηση.....	63
Πίνακας 3-3 Επιλεγμένα AI Act-related σημεία αναφοράς της οντολογίας.	64
Πίνακας 3-4 Ρόλος των στοιχείων στις βασικές σημασιολογικές διαδρομές.....	67
Πίνακας 3-5 Ρόλος των στοιχείων του lightweight risk-related επιπέδου.	70
Πίνακας 3-6 Τα τέσσερα επίπεδα της αρχιτεκτονικής της οντολογίας.	76
Πίνακας 3-7 Βασικές σχεδιαστικές επιλογές της οντολογίας.	79
Πίνακας 3-8 Μεθοδολογικά βήματα ανάπτυξης της οντολογίας στο Protégé.	82
Πίνακας 4-1 Συνοπτική δομή της υλοποιημένης οντολογίας.....	88
Πίνακας 4-2 Συγκεντρωτική παρουσίαση των βασικών individuals της οντολογίας.	140
Πίνακας 5-1 Σύνοψη ελέγχου λογικής συνέπειας με reasoner.	167
Πίνακας 5-2 DL competency questions της οντολογίας.	169
Πίνακας 5-3 SPARQL queries της οντολογίας.	184
Πίνακας 5-4 Πλήθος VT alarm events ανά waveform record.....	190

Συντομογραφίες

AI	Artificial Intelligence
AIPO	AI Principles Ontology
AIRO	AI Risk Ontology
DL	Description Logic
EU	European Union
FRIA	Fundamental Rights Impact Assessment
ICU	Intensive Care Unit
IEEE	Institute of Electrical and Electronics Engineers
OWL	Web Ontology Language
RDF	Resource Description Framework
SKOS	Simple Knowledge Organization System
SPARQL	SPARQL Protocol and RDF Query Language
TAIR	Trustworthy AI Requirements Ontology
VAIR	Vocabulary of AI Risks
VT	Ventricular Tachycardia
VTaC	Ventricular Tachycardia annotated alarms from ICUs
WFDB	Waveform Database
WHO	World Health Organization

Κεφάλαιο 1ο: Εισαγωγή και πλαίσιο της εργασίας

1.1 Εισαγωγή

Η ραγδαία ανάπτυξη της τεχνητής νοημοσύνης έχει επηρεάσει σημαντικά πολλούς τομείς της κοινωνίας, της οικονομίας και της επιστήμης. Ιδιαίτερα στον χώρο της υγείας, τα συστήματα που βασίζονται σε δεδομένα, αλγορίθμους και αυτοματοποιημένη υποστήριξη αποφάσεων αποκτούν ολοένα και μεγαλύτερη σημασία. Η χρήση τέτοιων συστημάτων μπορεί να συμβάλει στην ταχύτερη ανάλυση δεδομένων, στην υποστήριξη των επαγγελματιών υγείας και στη βελτίωση της παρακολούθησης ασθενών. Παράλληλα όμως δημιουργεί νέες απαιτήσεις σχετικά με τη διαφάνεια, την ανθρώπινη εποπτεία, την τεκμηρίωση, την ιχνηλασιμότητα και τη διαχείριση κινδύνων [1], [2].

Η ανάγκη για υπεύθυνη και αξιόπιστη τεχνητή νοημοσύνη έχει αναδειχθεί έντονα τα τελευταία χρόνια μέσα από ευρωπαϊκές και διεθνείς πρωτοβουλίες. Οι αρχές του Trustworthy AI δίνουν έμφαση όχι μόνο στην τεχνική απόδοση ενός συστήματος, αλλά και στον τρόπο με τον οποίο αυτό μπορεί να εξηγηθεί, να ελεγχθεί και να χρησιμοποιηθεί με ασφάλεια. Στο ίδιο πλαίσιο, ο AI Act της Ευρωπαϊκής Ένωσης ενισχύει τη σημασία της τεκμηρίωσης, της διαχείρισης κινδύνων, της ανθρώπινης εποπτείας και της δυνατότητας ελέγχου σε συστήματα τεχνητής νοημοσύνης, ιδιαίτερα όταν αυτά εφαρμόζονται σε ευαίσθητους τομείς [3].

Στον χώρο της υγείας, η ανάγκη αυτή γίνεται ακόμη πιο έντονη. Τα συστήματα παρακολούθησης ασθενών παράγουν συνεχώς δεδομένα και συναγεμμούς, οι οποίοι μπορεί να αφορούν κρίσιμες καταστάσεις. Ένα alarm event, όμως, δεν έχει αξία μόνο ως τεχνική εγγραφή. Για να μπορεί να αξιοποιηθεί ουσιαστικά, πρέπει να είναι σαφές από ποιο record προήλθε, πώς αξιολογήθηκε, ποια απόφαση λήφθηκε και ποια τεκμηρίωση το συνοδεύει. Η ύπαρξη false alarms, δηλαδή συναγεμμών που τελικά δεν αντιστοιχούν σε πραγματικό συμβάν, ενισχύει ακόμη περισσότερο την ανάγκη για οργανωμένη αξιολόγηση και τεκμηρίωση.

Στην παρούσα διπλωματική εργασία αξιοποιείται το VTaC dataset, ένα πραγματικό σύνολο δεδομένων που περιλαμβάνει ventricular tachycardia alarms από monitors μονάδων εντατικής θεραπείας. Το dataset αυτό είναι ιδιαίτερα κατάλληλο για τη συγκεκριμένη εργασία, επειδή περιλαμβάνει πραγματικά alarm events και τελικές ανθρώπινες αξιολογήσεις για το αν κάθε alarm είναι true ή false [4]. Με βάση αυτά τα δεδομένα, η εργασία δεν επιχειρεί να εκπαιδεύσει ένα μοντέλο μηχανικής μάθησης ούτε να προβλέψει αυτόματα την ορθότητα ενός alarm. Αντίθετα, εστιάζει στη σημασιολογική οργάνωση των διαθέσιμων στοιχείων.

Για τον σκοπό αυτό, προτείνεται και υλοποιείται μια οντολογία στο περιβάλλον Protégé. Η οντολογία χρησιμοποιείται ως μοντέλο αναπαράστασης γνώσης, το οποίο επιτρέπει τη σύνδεση πραγματικών healthcare monitoring alarm events με στοιχεία όπως waveform records, expert reviews, review decisions και evidence records, καθώς και τη χαρτογράφηση της σχετικής τεκμηρίωσης προς επιλεγμένα AI Act-related σημεία αναφοράς. Με αυτόν τον τρόπο, ένα alarm event δεν αντιμετωπίζεται ως απομονωμένη εγγραφή, αλλά ως κεντρικό στοιχείο γύρω από το οποίο οργανώνονται διαφορετικές αλλά συνδεδεμένες σημασιολογικές σχέσεις.

Κεντρική ιδέα της εργασίας είναι το requirement-to-evidence mapping, δηλαδή η σύνδεση συγκεκριμένων τεκμηρίων με επιλεγμένα AI Act-related σημεία αναφοράς που σχετίζονται με logging, human review, traceability και risk management. Η λογική αυτή συνδέεται με τη γενικότερη ερευνητική

κατεύθυνση που εξετάζει πώς έννοιες και απαιτήσεις από κανονιστικά ή τεχνικά πλαίσια μπορούν να αναπαρασταθούν οντολογικά και να γίνουν αναζητήσιμες, ελέγξιμες και επαναχρησιμοποιήσιμες [5].

Η εργασία κινείται επομένως στο σημείο συνάντησης τεσσάρων περιοχών: της διακυβέρνησης της τεχνητής νοημοσύνης στον χώρο της υγείας, της κανονιστικής τεκμηρίωσης, της οντολογικής αναπαράστασης γνώσης και της αξιοποίησης πραγματικών healthcare monitoring events. Η προσέγγιση είναι εφαρμοσμένη και proof-of-concept. Στόχος είναι να εξεταστεί πώς πραγματικά VTaC alarm events μπορούν να οργανωθούν σε ένα σημασιολογικό μοντέλο που υποστηρίζει τεκμηρίωση, ιχνηλασιμότητα και mapping προς επιλεγμένα AI Act-related σημεία αναφοράς.

1.2 Γενικό πλαίσιο της εργασίας

Η συγκεκριμένη προσέγγιση τοποθετείται στη διασταύρωση τριών ευρύτερων πεδίων: της αξιόπιστης τεχνητής νοημοσύνης, των συστημάτων παρακολούθησης υγείας και της οντολογικής αναπαράστασης γνώσης. Τα τρία αυτά πεδία συνδέονται μεταξύ τους μέσα από την ανάγκη για καλύτερη τεκμηρίωση, ιχνηλασιμότητα και ελεγχόμενη οργάνωση της πληροφορίας.

Το πρώτο πεδίο αφορά το Trustworthy AI και το κανονιστικό πλαίσιο που διαμορφώνεται γύρω από την τεχνητή νοημοσύνη. Οι σύγχρονες προσεγγίσεις δεν εξετάζουν πλέον τα συστήματα τεχνητής νοημοσύνης μόνο ως τεχνικά συστήματα, αλλά και ως συστήματα που πρέπει να είναι κατανοητά, ελέγξιμα, διαφανή και ασφαλή για τον άνθρωπο [1], [2]. Στο ίδιο πλαίσιο, ο AI Act της Ευρωπαϊκής Ένωσης αναδεικνύει τη σημασία απαιτήσεων όπως η τεκμηρίωση, η ανθρώπινη εποπτεία, η ιχνηλασιμότητα και η διαχείριση κινδύνων [3].

Το δεύτερο πεδίο αφορά τα healthcare monitoring systems. Σε τέτοια περιβάλλοντα, όπως οι μονάδες εντατικής θεραπείας, παράγονται συνεχώς alarm events από συστήματα παρακολούθησης ασθενών. Τα alarms αυτά μπορεί να σχετίζονται με κρίσιμες καταστάσεις, αλλά μπορεί επίσης να αποδειχθούν false alarms. Για τον λόγο αυτό, η απλή ύπαρξη ενός alarm δεν αρκεί. Χρειάζεται να μπορεί να τεκμηριωθεί η προέλευσή του, η αξιολόγησή του και η τελική απόφαση που το αφορά.

Το τρίτο πεδίο αφορά τις οντολογίες και τη σημασιολογική αναπαράσταση γνώσης. Οι οντολογίες επιτρέπουν την οργάνωση εννοιών, σχέσεων και πραγματικών περιστατικών σε μια δομημένη μορφή. Στο πλαίσιο της παρούσας εργασίας, αυτή η δυνατότητα αξιοποιείται για τη σημασιολογική οργάνωση πραγματικών VTaC alarm events και τη σύνδεσή τους με στοιχεία προέλευσης, ανθρώπινης αξιολόγησης και τεκμηρίωσης, καθώς και με επιλεγμένα AI Act-related σημεία αναφοράς μέσω του αντίστοιχου evidence mapping.

Η εργασία δεν αντιμετωπίζει τα δεδομένα μόνο ως εγγραφές ενός dataset. Αντίθετα, τα εξετάζει ως πραγματικά περιστατικά που μπορούν να αποκτήσουν σημασιολογική δομή. Ένα alarm event μπορεί να συνδεθεί με το waveform record από το οποίο προέρχεται, με την expert αξιολόγηση και την τελική απόφαση και, παράλληλα, με ένα Evidence Record. Το Evidence Record μπορεί στη συνέχεια να χαρτογραφείται προς επιλεγμένα AI Act-related σημεία αναφοράς. Μέσα από αυτή τη λογική, η οντολογία λειτουργεί ως ενδιάμεσο σημασιολογικό επίπεδο μεταξύ των πραγματικών δεδομένων, της τεκμηρίωσης και του requirement-to-evidence mapping.

Το VTaC dataset παρέχει το εφαρμοσμένο υπόβαθρο της εργασίας, καθώς περιλαμβάνει πραγματικά ventricular tachycardia alarm events από ICU monitors και τελικές ανθρώπινες αξιολογήσεις για το αν κάθε alarm είναι true ή false [4]. Η χρήση του επιτρέπει στην εργασία να στηριχθεί σε πραγματικά healthcare monitoring events και όχι σε υποθετικά παραδείγματα.

Μέσα σε αυτό το πλαίσιο, η παρούσα ερευνητική προσέγγιση προτείνει ένα proof-of-concept οντολογικό μοντέλο. Η βασική της κατεύθυνση συνδέεται με τη διεθνή ερευνητική προσπάθεια για ontology-based αναπαράσταση εννοιών, απαιτήσεων και σχέσεων από κανονιστικά ή τεχνικά πλαίσια [5]. Η ιδιαιτερότητα της παρούσας προσέγγισης είναι ότι η λογική αυτή εφαρμόζεται σε πραγματικά healthcare monitoring alarm events, με στόχο τη σημασιολογική τεκμηρίωση και την ιχνηλασιμότητα.

1.3 Πρόβλημα που εξετάζεται

Το βασικό πρόβλημα που εξετάζει η εργασία είναι η δυσκολία οργανωμένης τεκμηρίωσης πραγματικών healthcare monitoring alarm events με τρόπο που να υποστηρίζει ιχνηλασιμότητα και σύνδεση με επιλεγμένα AI Act-related σημεία αναφοράς. Σε ένα περιβάλλον παρακολούθησης υγείας, όπως μια μονάδα εντατικής θεραπείας, τα συστήματα παρακολούθησης παράγουν συνεχώς alarms. Τα alarms αυτά μπορεί να αντιστοιχούν σε πραγματικά κρίσιμα περιστατικά, αλλά μπορεί επίσης να αποδειχθούν false alarms.

Η απλή ύπαρξη ενός alarm event δεν αρκεί για να υποστηρίξει πλήρη κατανόηση και τεκμηρίωση του περιστατικού. Για να μπορεί το event να ενταχθεί σε μια οργανωμένη διαδικασία αναζήτησης και ελέγχου, πρέπει να είναι δυνατό να προσδιοριστεί από ποιο waveform record προήλθε, ποια expert αξιολόγηση πραγματοποιήθηκε, ποια τελική απόφαση λήφθηκε και ποια evidence συνδέεται με το συγκεκριμένο περιστατικό. Επιπλέον, πρέπει να μπορεί να εξεταστεί προς ποιο AI Act-related σημείο αναφοράς έχει χαρτογραφηθεί η αντίστοιχη τεκμηρίωση. Χωρίς τέτοια οργάνωση, τα διαφορετικά στοιχεία παραμένουν αποσπασματικά και δυσκολότερα αξιοποιήσιμα σε διαδικασίες αναζήτησης, ελέγχου και τεκμηρίωσης.

Το πρόβλημα αποκτά ιδιαίτερη σημασία στην περίπτωση των false alarms. Στα συστήματα παρακολούθησης υγείας, τα false alarms δεν αποτελούν απλώς λανθασμένες τεχνικές εγγραφές. Η συχνή παρουσία μη χρήσιμων ή μη actionable alarms έχει συσχετιστεί στη βιβλιογραφία με αυξημένο φόρτο για το κλινικό προσωπικό, μειωμένη ανταπόκριση στους συναγερμούς και φαινόμενα alarm fatigue [6]. Για τον λόγο αυτό, τα false alarm cases αποτελούν κατάλληλες περιπτώσεις για να εξεταστεί αν μια οντολογία μπορεί να δεχθεί, πέρα από τη βασική αναπαράσταση του event, και πρόσθετη risk-related σημασιολογική πληροφορία.

Η παρούσα εργασία αντιμετωπίζει το παραπάνω πρόβλημα μέσω οντολογικής αναπαράστασης. Η οντολογία χρησιμοποιείται ως ενδιάμεσο σημασιολογικό επίπεδο, το οποίο οργανώνει τα πραγματικά VTaC events και τα συνδέει με waveform records, Expert Reviews, Review Decisions και Evidence Records. Τα Evidence Records μπορούν στη συνέχεια να χαρτογραφούνται προς επιλεγμένα AI Act-related σημεία αναφοράς. Έτσι, το πρόβλημα δεν προσεγγίζεται ως πρόβλημα πρόβλεψης ή ταξινόμησης, αλλά ως πρόβλημα αναπαράστασης γνώσης, τεκμηρίωσης και ιχνηλασιμότητας.

Συνεπώς, το κεντρικό πρόβλημα μπορεί να διατυπωθεί ως εξής: πώς μπορεί ένα πραγματικό healthcare monitoring alarm event να μετατραπεί από μια απομονωμένη εγγραφή δεδομένων σε ένα οργανωμένο σημασιολογικό αντικείμενο, το οποίο διατηρεί τις συνδέσεις του με την προέλευση, την ανθρώπινη αξιολόγηση και την τεκμηρίωσή του και επιτρέπει την ελεγχόμενη χαρτογράφηση του αντίστοιχου evidence προς επιλεγμένα AI Act-related σημεία αναφοράς.

1.4 Σκοπός της διπλωματικής εργασίας

Σκοπός της παρούσας μεταπτυχιακής διπλωματικής εργασίας είναι ο σχεδιασμός, η υλοποίηση και η λειτουργική επικύρωση μιας οντολογίας για τη σημασιολογική αναπαράσταση και τεκμηρίωση πραγματικών healthcare monitoring alarm events. Η εργασία εστιάζει σε ventricular tachycardia alarms από το VTaC dataset και εξετάζει πώς τα συγκεκριμένα συμβάντα μπορούν να οργανωθούν σε ένα δομημένο μοντέλο γνώσης [4].

Η οντολογία αναπτύσσεται με στόχο να οργανώνει γύρω από κάθε alarm event διαφορετικές αλλά συνδεδεμένες κατηγορίες πληροφορίας. Ειδικότερα, το event συνδέεται με το waveform record προέλευσής του, με την expert αξιολόγηση και την τελική Review Decision και, παράλληλα, με το Evidence Record που χρησιμοποιείται για την τεκμηρίωσή του. Με αυτόν τον τρόπο, το alarm event δεν αντιμετωπίζεται ως μεμονωμένη εγγραφή δεδομένων, αλλά ως κεντρικό στοιχείο ενός ευρύτερου σημασιολογικού πλαισίου.

Κεντρικός σκοπός της εργασίας είναι η υποστήριξη της τεκμηρίωσης και της ιχνηλασιμότητας. Η προτεινόμενη δομή επιδιώκει να διατηρεί αναζητήσιμες τις σχέσεις μεταξύ του πραγματικού event, της προέλευσής του, της ανθρώπινης αξιολόγησης και του σχετικού evidence. Παράλληλα, εξετάζεται αν το Evidence Record μπορεί να λειτουργήσει ως ενδιάμεσος τεκμηριωτικός κόμβος για τη χαρτογράφηση της διαθέσιμης πληροφορίας προς επιλεγμένα AI Act-related σημεία αναφοράς.

Στην παρούσα εργασία χρησιμοποιούνται τέσσερα τέτοια σημεία αναφοράς, τα οποία σχετίζονται με logging, human review, traceability και risk management [3]. Τα σημεία αυτά δεν αντιμετωπίζονται ως πλήρης αναπαράσταση των νομικών απαιτήσεων του AI Act. Χρησιμοποιούνται στο πλαίσιο του proof-of-concept ώστε να εξεταστεί αν συγκεκριμένα Evidence Records μπορούν να συνδεθούν με κανονιστικά σχετικές έννοιες με σαφή και ελέγξιμο τρόπο. Η γενικότερη λογική της οντολογικής χαρτογράφησης concepts και normative requirements έχει επίσης εξεταστεί σε προηγούμενες ontology-based προσεγγίσεις [5].

Είναι σημαντικό να διευκρινιστεί ότι το VTaC dataset και τα healthcare monitoring alarm events που χρησιμοποιούνται στην παρούσα εργασία δεν θεωρούνται από μόνα τους σύστημα τεχνητής νοημοσύνης. Η εργασία δεν αξιολογεί ούτε επιχειρεί να ταξινομήσει το αρχικό monitoring system ως AI system σύμφωνα με τον AI Act. Το VTaC χρησιμοποιείται ως πηγή πραγματικών operational events και expert decisions, πάνω στα οποία δοκιμάζεται η προτεινόμενη οντολογική δομή. Η σύνδεση με την τεχνητή νοημοσύνη και τον AI Act αφορά το επίπεδο της σημασιολογικής τεκμηρίωσης και του requirement-to-evidence mapping που θα μπορούσε να υποστηρίξει αντίστοιχα healthcare monitoring events όταν αυτά αποτελούν μέρος ή είσοδο ενός AI-enabled συστήματος.

Ένας επιπλέον στόχος είναι να εξεταστεί αν η βασική event-centered αναπαράσταση μπορεί να εμπλουτιστεί με πρόσθετη risk-related πληροφορία. Για τον σκοπό αυτό, τα επιλεγμένα false alarm cases εντάσσονται σε ένα lightweight risk-related επίπεδο, μέσω του οποίου μπορούν να συσχετιστούν με Risk Indicators, Risk-Relevant Events και Potential Harm. Η επέκταση αυτή έχει διερευνητικό χαρακτήρα και δεν αποτελεί πλήρες σύστημα αξιολόγησης ή διαχείρισης κινδύνου.

Η εργασία δεν αποσκοπεί στην ανάπτυξη μοντέλου μηχανικής μάθησης ούτε στην αυτόματη πρόβλεψη του αν ένα alarm είναι true ή false. Οι τελικές αποφάσεις προέρχονται από τις expert αξιολογήσεις του VTaC dataset και χρησιμοποιούνται ως δεδομένα της οντολογικής αναπαράστασης [4]. Η έμφαση βρίσκεται στη σημασιολογική οργάνωση, στη σύνδεση, στην αναζήτηση και στην τεκμηρίωση της διαθέσιμης πληροφορίας.

Τελικός σκοπός είναι να αξιολογηθεί αν ένα περιορισμένο αλλά εφαρμοσμένο proof-of-concept μπορεί να οργανώσει πραγματικά healthcare monitoring events σε ένα συνεκτικό και αναζητήσιμο σημασιολογικό μοντέλο, στο οποίο η προέλευση, η ανθρώπινη αξιολόγηση, το evidence και τα επιλεγμένα AI Act-related mappings παραμένουν διακριτά αλλά συνδεδεμένα.

1.5 Ερευνητικά ερωτήματα

Με βάση το πρόβλημα και τον σκοπό της εργασίας, η παρούσα διπλωματική εργασία οργανώνεται γύρω από ένα κεντρικό ερευνητικό ερώτημα και πέντε επιμέρους ερευνητικά ερωτήματα. Τα ερωτήματα δεν αφορούν την ανάπτυξη μοντέλου μηχανικής μάθησης ή την αυτόματη πρόβλεψη της ορθότητας ενός alarm. Αντίθετα, εστιάζουν στη σημασιολογική αναπαράσταση, στην τεκμηρίωση, στην ιχνηλασιμότητα και στη σύνδεση πραγματικών healthcare monitoring events με επιλεγμένα AI Act-related σημεία αναφοράς.

Κεντρικό Ερευνητικό Ερώτημα (ΚΕΕ): Με ποιον τρόπο μπορεί μια οντολογία να υποστηρίξει τη σημασιολογική αναπαράσταση, την τεκμηρίωση και την ιχνηλασιμότητα πραγματικών healthcare monitoring alarm events, καθώς και τη σύνδεσή τους με επιλεγμένα AI Act-related σημεία αναφοράς;

Για την απάντηση του κεντρικού ερευνητικού ερωτήματος διατυπώνονται τα ακόλουθα επιμέρους ερευνητικά ερωτήματα:

ΕΕ1: Πώς μπορεί να αναπαρασταθεί οντολογικά ένα πραγματικό healthcare monitoring alarm event, ώστε να συνδέεται με το waveform record προέλευσής του, την expert αξιολόγηση, την τελική απόφαση και την τεκμηρίωσή του;

ΕΕ2: Πώς μπορούν επιλεγμένα AI Act-related σημεία αναφοράς, όπως logging, human review, traceability και risk management, να συνδεθούν με πραγματικά VTaC events μέσω Evidence Records;

ΕΕ3: Πώς μπορεί η οντολογία να υποστηρίξει την ιχνηλασιμότητα μεταξύ του waveform record, του alarm event, της expert αξιολόγησης, του Evidence Record και του requirement-to-evidence mapping;

ΕΕ4: Πώς μπορούν τα false alarm cases να αποκτήσουν πρόσθετη risk-related σημασιολογική αναπαράσταση μέσα στην οντολογία;

ΕΕ5: Με ποιον τρόπο μπορεί να ελεγχθεί λειτουργικά η προτεινόμενη οντολογία μέσω reasoner, DL competency questions και SPARQL queries;

Τα παραπάνω ερευνητικά ερωτήματα καθοδηγούν τη μεθοδολογία, τον σχεδιασμό, την υλοποίηση και τη λειτουργική επικύρωση της οντολογίας. Στα επόμενα κεφάλαια, τα ερωτήματα αυτά συνδέονται με τον σχεδιασμό του μοντέλου, την υλοποίησή του στο Protégé και τον λειτουργικό του έλεγχο μέσω reasoner, DL competency questions και SPARQL queries. Οι τελικές απαντήσεις στα ερευνητικά ερωτήματα παρουσιάζονται συνολικά στο Κεφάλαιο 6.

Πίνακας 1-1 Σύνδεση ερευνητικών ερωτημάτων με τα κεφάλαια της εργασίας

Ερευνητικό ερώτημα	Κύρια εστίαση	Κεφάλαια στα οποία εξετάζεται / απαντάται
KEE	Συνολική δυνατότητα οντολογικής αναπαράστασης, τεκμηρίωσης και ιχνηλασιμότητας πραγματικών healthcare monitoring alarm events και σύνδεσής τους με επιλεγμένα AI Act-related σημεία αναφοράς.	Κεφάλαια 3, 4, 5 και 6
EE1	Event-centered οντολογική αναπαράσταση πραγματικού alarm event και σύνδεσή του με waveform record, expert αξιολόγηση, τελική απόφαση και Evidence Record.	Κεφάλαια 3, 4 και 6
EE2	Requirement-to-evidence mapping μεταξύ πραγματικών VTaC events, Evidence Records και επιλεγμένων AI Act-related σημείων αναφοράς.	Κεφάλαια 3, 4, 5 και 6
EE3	Υποστήριξη ιχνηλασιμότητας μέσω των συνδεδεμένων σχέσεων μεταξύ waveform record, alarm event, expert αξιολόγησης, Evidence Record και requirement mapping.	Κεφάλαια 3, 4 και 5
EE4	Πρόσθετη lightweight risk-related σημασιολογική αναπαράσταση των false alarm cases.	Κεφάλαια 3, 4, 5 και 6
EE5	Λογικός και λειτουργικός έλεγχος της οντολογίας μέσω HermiT reasoner, DL competency questions και SPARQL queries.	Κεφάλαιο 5 και 6

1.6 Συμβολή της εργασίας

Η συμβολή της παρούσας διπλωματικής εργασίας εντοπίζεται στην εφαρμογή μιας οντολογικής προσέγγισης για τη σημασιολογική οργάνωση και τεκμηρίωση πραγματικών healthcare monitoring alarm events. Η προσέγγιση δεν περιορίζεται στη θεωρητική παρουσίαση εννοιών, αλλά περιλαμβάνει τον σχεδιασμό, την υλοποίηση και τη λειτουργική επικύρωση ενός proof-of-concept μοντέλου που βασίζεται σε πραγματικά δεδομένα από το VTaC dataset [4].

Σε πρώτο επίπεδο, η εργασία συμβάλλει στην event-centered αναπαράσταση πραγματικών ventricular tachycardia alarm events. Τα επιλεγμένα events δεν αντιμετωπίζονται ως απλές εγγραφές δεδομένων, αλλά ως σημασιολογικά αντικείμενα που μπορούν να συνδεθούν με waveform records, Expert Reviews, Review Decisions και Evidence Records. Με αυτόν τον τρόπο, διαφορετικές πληροφορίες που σχετίζονται με το ίδιο περιστατικό οργανώνονται σε μια ενιαία και αναζητήσιμη δομή γνώσης.

Σε δεύτερο επίπεδο, η συμβολή αφορά τη σύνδεση της event-level τεκμηρίωσης με επιλεγμένα AI Act-related σημεία αναφοράς. Η εργασία χρησιμοποιεί τα req_logging, req_human_review, req_traceability και req_risk_management ως εφαρμοσμένα σημεία αναφοράς μέσα στο proof-of-concept. Η σύνδεση δεν πραγματοποιείται απευθείας από το πραγματικό event προς το κανονιστικό σημείο, αλλά μέσω του Evidence Record, ώστε να διατηρείται σαφής διάκριση μεταξύ πραγματικού περιστατικού, διαθέσιμης τεκμηρίωσης και κανονιστικού mapping [3], [5].

Ιδιαίτερη σημασία έχει επομένως ο ρόλος του Evidence Record ως ενδιάμεσου τεκμηριωτικού κόμβου. Μέσω αυτού οργανώνεται η μετάβαση από το operational επίπεδο των πραγματικών alarm events προς το normative επίπεδο των επιλεγμένων AI Act-related σημείων αναφοράς. Η σχεδιαστική αυτή επιλογή υποστηρίζει το requirement-to-evidence mapping και επιτρέπει να εξεταστεί ποιο evidence έχει χαρτογραφηθεί προς ποιο σημείο αναφοράς, χωρίς το mapping να αποτελεί από μόνο του απόδειξη νομικής συμμόρφωσης.

Σε τρίτο επίπεδο, η εργασία συμβάλλει στην αναπαράσταση της ιχνηλασιμότητας ως συνδυασμού διαφορετικών σημασιολογικών σχέσεων γύρω από το ίδιο Alert Event. Η προέλευση του event, η ανθρώπινη αξιολόγηση, η τελική απόφαση και η τεκμηρίωση παραμένουν διακριτές πληροφορίες, αλλά μπορούν να αναζητηθούν και να συνδυαστούν μέσα στο ίδιο μοντέλο.

Επιπλέον, η οντολογία περιλαμβάνει ένα lightweight risk-related επίπεδο για τα false alarm cases. Η επέκταση αυτή επιτρέπει στα επιλεγμένα false alarms να συνδέονται με Risk Indicators, Risk-Relevant Events και Potential Harm, χωρίς να μεταβάλλεται η αρχική expert απόφαση του dataset και χωρίς να δηλώνεται ότι ένα συγκεκριμένο false alarm προκάλεσε πραγματική βλάβη. Με αυτόν τον τρόπο εξετάζεται η δυνατότητα εμπλουτισμού της βασικής event representation με πρόσθετη risk-related πληροφορία.

Μια ακόμη διάσταση της συμβολής είναι ότι η προτεινόμενη οντολογία σχεδιάζεται ως λειτουργικά ελέγξιμο μοντέλο και όχι μόνο ως στατική ταξινόμηση εννοιών. Η χρήση reasoner, DL competency questions και SPARQL queries επιτρέπει να εξεταστεί αν οι βασικές σημασιολογικές σχέσεις μπορούν να ανακτηθούν σύμφωνα με τον σχεδιασμό της οντολογίας.

Σε σχέση με τη σχετική βιβλιογραφία, η εργασία αξιοποιεί τη γενικότερη λογική ontology-based representation και mapping κανονιστικών concepts και requirements, όπως αυτή έχει παρουσιαστεί σε σχετικές προσεγγίσεις [5], και τη μεταφέρει σε ένα συγκεκριμένο healthcare monitoring use case με πραγματικά VTaC events και event-level evidence. Με βάση τις πηγές που εξετάστηκαν στην παρούσα εργασία, δεν εντοπίστηκε προσέγγιση που να συνδυάζει με τον ίδιο ακριβώς τρόπο πραγματικά VTaC alarm events, expert decisions, Evidence Records, AI Act-related mappings και lightweight risk-related representation.

Συνολικά, η συμβολή της διπλωματικής βρίσκεται στον εφαρμοσμένο συνδυασμό πραγματικών δεδομένων, οντολογικής αναπαράστασης, τεκμηρίωσης και κανονιστικά σχετικού mapping μέσα σε ένα ενιαίο proof-of-concept. Η εργασία δεν προτείνει ένα πλήρες σύστημα νομικής συμμόρφωσης ούτε ένα μοντέλο τεχνητής νοημοσύνης, αλλά εξετάζει πώς πραγματικά healthcare monitoring events μπορούν να οργανωθούν σημασιολογικά και να συνδεθούν με evidence και επιλεγμένα AI Act-related σημεία αναφοράς με τρόπο σαφή, αναζητήσιμο και λειτουργικά ελέγξιμο.

1.7 Δομή της εργασίας

Η διπλωματική εργασία οργανώνεται σε έξι κεφάλαια, τα οποία ακολουθούν τη μετάβαση από το θεωρητικό και κανονιστικό πλαίσιο προς τον σχεδιασμό, την υλοποίηση, τη λειτουργική επικύρωση και την τελική αποτίμηση της προτεινόμενης οντολογικής προσέγγισης.

Στο Κεφάλαιο 1 παρουσιάζεται το εισαγωγικό πλαίσιο της εργασίας. Ορίζονται το πρόβλημα που εξετάζεται, ο σκοπός της διπλωματικής, τα ερευνητικά ερωτήματα και η βασική συμβολή της προτεινόμενης προσέγγισης.

Στο Κεφάλαιο 2 παρουσιάζεται το θεωρητικό υπόβαθρο και η σχετική βιβλιογραφία. Εξετάζονται η τεχνητή νοημοσύνη στον χώρο της υγείας, τα healthcare monitoring alarm events, τα true και false alarms, το alert fatigue, καθώς και ζητήματα διαφάνειας, τεκμηρίωσης και ιχνηλασιμότητας. Παράλληλα, παρουσιάζεται ο AI Act ως κανονιστικό πλαίσιο αναφοράς και αναλύεται ο ρόλος των οντολογιών και συναφών ontology-based προσεγγίσεων στην αναπαράσταση και οργάνωση γνώσης.

Στο Κεφάλαιο 3 αναπτύσσονται η μεθοδολογία και ο σχεδιασμός της οντολογίας. Παρουσιάζεται ο ρόλος του VTaC dataset, αιτιολογείται η επιλογή της οντολογικής προσέγγισης αντί της μηχανικής μάθησης και περιγράφεται η επιλογή των πραγματικών VTaC events που χρησιμοποιούνται στο proof-of-concept. Στη συνέχεια παρουσιάζονται τα επιλεγμένα AI Act-related σημεία αναφοράς, οι συνδεδεμένες σημασιολογικές διαδρομές του μοντέλου, το lightweight risk-related επίπεδο, η συνολική αρχιτεκτονική και οι βασικές σχεδιαστικές επιλογές.

Στο Κεφάλαιο 4 παρουσιάζεται η πρακτική υλοποίηση της οντολογίας στο Protégé. Αναλύονται οι classes, τα object properties, τα data properties, τα individuals και τα annotations του μοντέλου, καθώς και ο τρόπος με τον οποίο τα πραγματικά VTaC events συνδέονται με πληροφορίες προέλευσης, expert αξιολόγησης, τεκμηρίωσης, AI Act-related mappings και risk-related στοιχεία.

Στο Κεφάλαιο 5 πραγματοποιείται ο έλεγχος και η λειτουργική επικύρωση της οντολογίας. Εξετάζεται η λογική συνέπεια του μοντέλου με τον HermiT reasoner και ελέγχεται η δυνατότητα ανάκτησης των βασικών σημασιολογικών σχέσεων μέσω DL competency questions και SPARQL queries.

Στο Κεφάλαιο 6 παρουσιάζονται η συνολική συζήτηση και ερμηνεία των αποτελεσμάτων, οι απαντήσεις στα ερευνητικά ερωτήματα, η αποτίμηση της συμβολής και των περιορισμών της εργασίας και οι βασικές κατευθύνσεις για μελλοντική επέκταση. Το κεφάλαιο ολοκληρώνεται με το τελικό συμπέρασμα της διπλωματικής.

Τέλος, ακολουθούν η βιβλιογραφία και δύο παραρτήματα. Στο Παράρτημα Α παρουσιάζεται συγκεντρωτικά η τεχνική προδιαγραφή της οντολογίας ενώ στο Παράρτημα Β συγκεντρώνονται τα DL competency questions και τα SPARQL queries που χρησιμοποιήθηκαν για τη λειτουργική επικύρωσή της.

1.8 Επίλογος

Στο πρώτο κεφάλαιο οριοθετήθηκε το αντικείμενο της διπλωματικής εργασίας και παρουσιάστηκαν το πρόβλημα, ο σκοπός, τα ερευνητικά ερωτήματα και η βασική κατεύθυνση της προτεινόμενης οντολογικής προσέγγισης. Η εργασία τοποθετήθηκε στο πλαίσιο της σημασιολογικής οργάνωσης πραγματικών healthcare monitoring events και της σύνδεσης της σχετικής τεκμηρίωσης με επιλεγμένα AI Act-related σημεία αναφοράς. Παράλληλα, αποφασήθηκε ότι η προσέγγιση δεν αφορά πρόβλεψη true ή false alarms ούτε αξιολόγηση του VTaC monitoring environment ως AI system. Το VTaC χρησιμοποιείται ως πηγή πραγματικών alarm events και expert decisions, πάνω στα οποία αναπτύσσεται και ελέγχεται το proof-of-concept οντολογικό μοντέλο.

Το επόμενο κεφάλαιο παρουσιάζει το θεωρητικό και βιβλιογραφικό υπόβαθρο που υποστηρίζει τον σχεδιασμό της εργασίας. Εξετάζονται ειδικότερα τα healthcare monitoring alarms, τα false alarms και το alert fatigue, οι έννοιες της τεκμηρίωσης και της ιχνηλασιμότητας, ο AI Act ως κανονιστικό πλαίσιο αναφοράς και οι βασικές ontology-based προσεγγίσεις που σχετίζονται με Ethical AI, Trustworthy AI και AI governance.

Κεφάλαιο 2ο: Θεωρητικό υπόβαθρο και σχετική βιβλιογραφία

2.1 Εισαγωγή

Το παρόν κεφάλαιο παρουσιάζει το θεωρητικό υπόβαθρο και τη σχετική βιβλιογραφία πάνω στα οποία βασίζεται η διπλωματική εργασία. Στόχος του κεφαλαίου είναι να αποσαφηνίσει τις βασικές έννοιες που συνδέονται με το αντικείμενο της μελέτης και να δείξει πώς αυτές υποστηρίζουν τον σχεδιασμό της οντολογίας που αναπτύσσεται στα επόμενα κεφάλαια.

Η ανάλυση ξεκινά από τη χρήση της τεχνητής νοημοσύνης στον χώρο της υγείας και ειδικότερα από τα συστήματα παρακολούθησης ασθενών. Σε τέτοια περιβάλλοντα, τα δεδομένα και οι συναγερμοί που παράγονται από monitoring systems μπορούν να έχουν κρίσιμη σημασία για την παρακολούθηση της κατάστασης ενός ασθενούς. Για τον λόγο αυτό, η αξιολόγηση, η τεκμηρίωση και η δυνατότητα αναζήτησης της σχετικής πληροφορίας αποτελούν βασικές ανάγκες.

Στη συνέχεια εξετάζονται τα healthcare monitoring alarm events, με έμφαση στη διάκριση ανάμεσα σε true alarms και false alarms. Η διάκριση αυτή είναι σημαντική για τη συγκεκριμένη εργασία, επειδή τα false alarms δεν αντιμετωπίζονται μόνο ως λανθασμένες εγγραφές, αλλά και ως περιστατικά που μπορούν να συνδεθούν με ζητήματα όπως το alert fatigue και η διαχείριση κινδύνου.

Ακολουθεί η παρουσίαση βασικών εννοιών που σχετίζονται με το Trustworthy AI και το Ethical AI. Ιδιαίτερη σημασία δίνεται στη διαφάνεια, την τεκμηρίωση, την ανθρώπινη εποπτεία, τη λογοδοσία και την ιχνηλασιμότητα, καθώς οι έννοιες αυτές συνδέονται άμεσα με τον τρόπο με τον οποίο ένα σύστημα τεχνητής νοημοσύνης μπορεί να γίνει πιο κατανοητό, ελέγξιμο και αξιόπιστο [1], [2].

Το κεφάλαιο εξετάζει επίσης τον AI Act ως κανονιστικό πλαίσιο αναφοράς. Η ανάλυση δεν έχει στόχο την πλήρη νομική ερμηνεία του κανονισμού, αλλά την ανάδειξη επιλεγμένων απαιτήσεων που είναι σχετικές με τη λογική της οντολογίας, όπως logging, human review, traceability και risk management [3].

Στη συνέχεια παρουσιάζονται οι οντολογίες ως μέσο αναπαράστασης γνώσης. Η οντολογική προσέγγιση είναι κεντρική για τη διπλωματική εργασία, επειδή επιτρέπει την οργάνωση εννοιών, σχέσεων και πραγματικών περιστατικών σε ένα δομημένο και αναζητήσιμο μοντέλο. Μέσα από αυτή τη λογική, πραγματικά VTaC alarm events μπορούν να συνδεθούν με waveform records, expert reviews, review decisions και evidence records, ενώ η σχετική τεκμηρίωση μπορεί να χαρτογραφείται προς επιλεγμένα AI Act-related σημεία αναφοράς [4].

Τέλος, εξετάζονται βιβλιογραφικές προσεγγίσεις που αξιοποιούν οντολογίες για Ethical AI, Trustworthy AI και AI governance. Ιδιαίτερη σημασία έχουν οι εργασίες που επιχειρούν να αναπαραστήσουν έννοιες και απαιτήσεις από κανονιστικά ή τεχνικά κείμενα σε μορφή γνώσης που μπορεί να αναζητηθεί και να ελεγχθεί [5]. Με βάση αυτή τη βιβλιογραφική βάση, στο τέλος του κεφαλαίου εντοπίζεται το κενό που επιχειρεί να καλύψει η συγκεκριμένη μελέτη εφαρμογής.

2.2 Τεχνητή νοημοσύνη και συστήματα παρακολούθησης υγείας

Η τεχνητή νοημοσύνη έχει αποκτήσει σημαντικό ρόλο στον χώρο της υγείας, καθώς μπορεί να αξιοποιηθεί για την ανάλυση μεγάλου όγκου δεδομένων, την υποστήριξη κλινικών αποφάσεων, την αναγνώριση προτύπων και τη βελτίωση διαδικασιών παρακολούθησης. Η αυξανόμενη διαθεσιμότητα ιατρικών δεδομένων, όπως ηλεκτρονικοί φάκελοι υγείας, ιατρικές εικόνες, σήματα από αισθητήρες και

δεδομένα από συστήματα monitoring, έχει δημιουργήσει νέες δυνατότητες για την εφαρμογή τεχνικών τεχνητής νοημοσύνης στην ιατρική πράξη [7], [8].

Στη βιβλιογραφία, η τεχνητή νοημοσύνη στην υγεία παρουσιάζεται συχνά ως τεχνολογία που μπορεί να ενισχύσει την ικανότητα των επαγγελματιών υγείας και όχι απαραίτητα να τους αντικαταστήσει. Η λογική αυτή είναι ιδιαίτερα σημαντική, επειδή σε ευαίσθητα περιβάλλοντα, όπως η κλινική φροντίδα, η ανθρώπινη κρίση εξακολουθεί να έχει κεντρικό ρόλο. Η τεχνητή νοημοσύνη μπορεί να υποστηρίξει την ανάλυση και την ερμηνεία δεδομένων, αλλά η αξιοποίηση των αποτελεσμάτων της πρέπει να συνδέεται με ανθρώπινη εποπτεία, σαφή τεκμηρίωση και κατάλληλους μηχανισμούς ελέγχου [7], [9].

Ένα από τα πεδία όπου η τεχνητή νοημοσύνη μπορεί να έχει ιδιαίτερη σημασία είναι η παρακολούθηση της κατάστασης των ασθενών. Τα συστήματα παρακολούθησης υγείας συλλέγουν και επεξεργάζονται δεδομένα που σχετίζονται με ζωτικά σημεία, φυσιολογικά σήματα και πιθανές μεταβολές στην κατάσταση ενός ασθενούς. Σε περιβάλλοντα αυξημένης κρισιμότητας, όπως οι μονάδες εντατικής θεραπείας, τέτοια συστήματα μπορούν να παράγουν συναγερμούς όταν εντοπίζονται ενδείξεις που απαιτούν προσοχή.

Η αξία αυτών των συστημάτων δεν βρίσκεται μόνο στην παραγωγή δεδομένων ή συναγερμών, αλλά και στον τρόπο με τον οποίο η παραγόμενη πληροφορία μπορεί να αξιολογηθεί και να αξιοποιηθεί. Ένα alarm event μπορεί να αποτελεί ένδειξη πραγματικού προβλήματος, αλλά μπορεί επίσης να είναι λανθασμένο ή κλινικά μη σημαντικό. Επομένως, η σωστή διαχείριση των alarm events απαιτεί σύνδεση του συμβάντος με την πηγή του, την αξιολόγησή του, την τελική απόφαση και την αντίστοιχη τεκμηρίωση.

Η χρήση τεχνητής νοημοσύνης σε τέτοια περιβάλλοντα δημιουργεί πρόσθετες ανάγκες. Όταν ένα σύστημα επεξεργάζεται δεδομένα υγείας ή υποστηρίζει διαδικασίες που μπορεί να επηρεάσουν την κλινική πράξη, δεν αρκεί να είναι τεχνικά αποδοτικό. Πρέπει επίσης να είναι κατανοητό, ελέγξιμο και κατάλληλα τεκμηριωμένο. Η διεθνής βιβλιογραφία για την AI in health and medicine αναδεικνύει την ανάγκη για συστήματα που συνδυάζουν τεχνική ακρίβεια, ανθρώπινη συνεργασία, ασφάλεια και υπεύθυνη ενσωμάτωση στην πράξη [8], [9].

Στο ίδιο πλαίσιο, οι αρχές του Trustworthy AI τονίζουν ότι η χρήση τεχνητής νοημοσύνης πρέπει να συνοδεύεται από απαιτήσεις όπως ανθρώπινη εποπτεία, τεχνική αξιοπιστία, διαφάνεια, λογοδοσία και διαχείριση κινδύνων [1]. Οι απαιτήσεις αυτές αποκτούν ιδιαίτερη βαρύτητα στον χώρο της υγείας, επειδή τα δεδομένα και οι αποφάσεις που σχετίζονται με ασθενείς δεν έχουν μόνο τεχνική, αλλά και ανθρώπινη, δεοντολογική και οργανωτική σημασία.

Για τη διπλωματική αυτή, τα συστήματα παρακολούθησης υγείας δεν εξετάζονται ως γενικά πληροφοριακά συστήματα, αλλά ως περιβάλλοντα όπου παράγονται συγκεκριμένα alarm events. Το ενδιαφέρον εστιάζει στη δυνατότητα οργάνωσης αυτών των events με τρόπο που να επιτρέπει ιχνηλασιμότητα και τεκμηρίωση. Αντί το alarm να αντιμετωπίζεται μόνο ως τεχνικό αποτέλεσμα ενός συστήματος monitoring, εξετάζεται ως συμβάν που μπορεί να ενταχθεί σε ένα δομημένο σύνολο συνδεδεμένων σημασιολογικών σχέσεων.

Η προσέγγιση αυτή είναι συμβατή με τη γενικότερη κατεύθυνση της υπεύθυνης τεχνητής νοημοσύνης στην υγεία. Ο WHO επισημαίνει ότι η AI for health πρέπει να αναπτύσσεται και να χρησιμοποιείται με τρόπο που λαμβάνει υπόψη ζητήματα ethics, governance, ασφάλειας και ανθρώπινων δικαιωμάτων [9]. Επομένως, η τεχνική αξιοποίηση των δεδομένων υγείας χρειάζεται να συνοδεύεται από τρόπους ελέγχου και τεκμηρίωσης της πληροφορίας που παράγεται.

Σε αυτό το πλαίσιο, τα healthcare monitoring alarm events αποτελούν κατάλληλο πεδίο μελέτης. Πρόκειται για συμβάντα που έχουν πραγματική λειτουργική σημασία, συνδέονται με δεδομένα ασθενών και απαιτούν αξιολόγηση. Η ύπαρξη πραγματικών datasets, όπως το VTaC, επιτρέπει την εξέταση αυτών των συμβάντων όχι μόνο θεωρητικά, αλλά και μέσα από συγκεκριμένα παραδείγματα [4].

Συνολικά, η τεχνητή νοημοσύνη και τα συστήματα παρακολούθησης υγείας δημιουργούν ένα πεδίο όπου η τεχνική δυνατότητα πρέπει να συνδυάζεται με τεκμηρίωση, ανθρώπινη αξιολόγηση και δυνατότητα ελέγχου. Αυτή η ανάγκη αποτελεί βασικό υπόβαθρο για την οντολογική προσέγγιση που ακολουθείται στη συνέχεια της εργασίας. Η οντολογία χρησιμοποιείται όχι για να αντικαταστήσει την κλινική κρίση ή να προβλέψει αυτόματα την ορθότητα ενός alarm, αλλά για να οργανώσει τη γνώση γύρω από τα alarm events και να υποστηρίξει τη σύνδεσή τους με τεκμήρια και απαιτήσεις.

2.3 Παρακολούθηση υγείας και συμβάντα συναγερμού (alarm events)

Τα συστήματα παρακολούθησης υγείας αποτελούν βασικό μέρος της σύγχρονης κλινικής πρακτικής, ιδιαίτερα σε περιβάλλοντα αυξημένης κρισιμότητας, όπως οι μονάδες εντατικής θεραπείας. Τα συστήματα αυτά συλλέγουν και επεξεργάζονται δεδομένα από φυσιολογικά σήματα, ζωτικά σημεία και ιατρικές συσκευές, με στόχο την παρακολούθηση της κατάστασης του ασθενούς και την έγκαιρη αναγνώριση πιθανών προβλημάτων.

Ένα από τα βασικά αποτελέσματα της λειτουργίας τέτοιων συστημάτων είναι η παραγωγή alarm events. Τα alarm events ενεργοποιούνται όταν το σύστημα εντοπίζει τιμές, μοτίβα ή ενδείξεις που θεωρεί ότι χρειάζονται προσοχή. Σε ένα ICU περιβάλλον, ένας συναγερμός μπορεί να αφορά κρίσιμες μεταβολές στην καρδιακή λειτουργία, στην αναπνοή, στην οξυγόνωση ή σε άλλα παρακολουθούμενα μεγέθη. Επομένως, τα alarms έχουν σημαντικό ρόλο στην κλινική παρακολούθηση.

Ωστόσο, η ύπαρξη ενός alarm δεν σημαίνει από μόνη της ότι υπάρχει πραγματικό κλινικό συμβάν. Ένα alarm μπορεί να είναι αληθές, όταν αντιστοιχεί σε πραγματική κατάσταση που απαιτεί προσοχή, αλλά μπορεί επίσης να είναι ψευδές ή κλινικά μη σημαντικό. Για τον λόγο αυτό, τα alarm events χρειάζονται αξιολόγηση, τεκμηρίωση και σύνδεση με το πλαίσιο μέσα στο οποίο παράγονται [10], [11].

Η βιβλιογραφία για τα clinical alarms αναδεικνύει ότι το πρόβλημα δεν περιορίζεται στην τεχνική παραγωγή ενός συναγερμού. Το κρίσιμο ζήτημα είναι η ποιότητα, η ακρίβεια, η κλινική σημασία και η δυνατότητα διαχείρισης των alarms. Σε περιβάλλοντα όπου παράγεται μεγάλος αριθμός συναγερμών, η διάκριση ανάμεσα σε σημαντικά και μη σημαντικά alarms γίνεται ιδιαίτερα δύσκολη. Αυτό μπορεί να επιβαρύνει το προσωπικό και να επηρεάσει τη συνολική λειτουργία του monitoring περιβάλλοντος [6], [10].

Για να μπορεί ένα alarm event να αξιοποιηθεί ουσιαστικά, δεν αρκεί να υπάρχει μόνο ως απλή καταγραφή. Πρέπει να μπορεί να συνδεθεί με την πηγή του, δηλαδή με το record ή το σήμα από το οποίο προήλθε. Πρέπει επίσης να συνδεθεί με την αξιολόγηση που πραγματοποιήθηκε και με την τελική απόφαση σχετικά με το αν ήταν πραγματικό ή ψευδές. Η σύνδεση αυτή είναι απαραίτητη για την ιχνηλασιμότητα και την οργανωμένη τεκμηρίωση.

Στο πλαίσιο της διπλωματικής εργασίας, ιδιαίτερη σημασία έχει το VTaC dataset. Το VTaC αποτελεί δημόσια διαθέσιμο dataset στο PhysioNet και περιλαμβάνει ventricular tachycardia alarm events από ICU monitors [4]. Η κοιλιακή ταχυκαρδία αποτελεί σοβαρή καρδιακή αρρυθμία και οι συναγερμοί που σχετίζονται με αυτήν έχουν ιδιαίτερο ενδιαφέρον σε περιβάλλοντα εντατικής παρακολούθησης.

Το VTaC είναι κατάλληλο για τη συγκεκριμένη μελέτη, επειδή δεν παρέχει μόνο γενικές πληροφορίες για alarms, αλλά πραγματικά alarm events με ανθρώπινη αξιολόγηση. Κάθε alarm event συνδέεται με waveform recording και συνοδεύεται από τελική απόφαση, η οποία δηλώνει αν το alarm χαρακτηρίστηκε ως true ή false. Αυτή η δομή επιτρέπει τη μετάβαση από την απλή εγγραφή δεδομένων στη σημασιολογική αναπαράσταση του event μέσα στην οντολογία [4].

Η ύπαρξη record identifiers, event identifiers και decision labels είναι ιδιαίτερα σημαντική για την οντολογική μοντελοποίηση. Το record identifier μπορεί να χρησιμοποιηθεί για την αναπαράσταση του waveform record, το event identifier για την αναπαράσταση του alarm event και το decision label για τη σύνδεση με την τελική review decision. Έτσι, τα διαθέσιμα στοιχεία του dataset μπορούν να μετατραπούν σε classes, individuals και σχέσεις μέσα στο οντολογικό μοντέλο.

Στην παρούσα μελέτη, τα healthcare monitoring alarm events δεν εξετάζονται ως αντικείμενο πρόβλεψης, αλλά ως αντικείμενο τεκμηρίωσης. Αυτό σημαίνει ότι το ζητούμενο δεν είναι να αποφασίσει η οντολογία αν ένα alarm είναι true ή false. Η απόφαση αυτή υπάρχει ήδη στο dataset. Το ζητούμενο είναι να οργανωθεί το event με τέτοιο τρόπο ώστε να μπορεί να συνδεθεί με την προέλευσή του, την expert αξιολόγηση, την τελική απόφαση και το Evidence Record που το τεκμηριώνει.

Με βάση αυτή τη λογική, το alarm event λειτουργεί ως κεντρικό σημείο γύρω από το οποίο οργανώνονται διαφορετικές αλλά συνδεδεμένες σημασιολογικές διαδρομές. Το event συνδέεται με το waveform record προέλευσής του, με το Expert Review και τη Review Decision και, παράλληλα, με το Evidence Record. Το Evidence Record μπορεί στη συνέχεια να χαρτογραφείται προς επιλεγμένα AI Act-related σημεία αναφοράς, όπως logging, human review και traceability.

Συνολικά, τα healthcare monitoring alarm events αποτελούν κατάλληλο πεδίο για οντολογική αναπαράσταση, επειδή συνδυάζουν τεχνικά δεδομένα, ανθρώπινη αξιολόγηση και ανάγκη τεκμηρίωσης. Η αξιοποίηση του VTaC dataset επιτρέπει να εξεταστεί αυτή η λογική σε πραγματικά ICU alarm events και να δημιουργηθεί η βάση για το requirement-to-evidence mapping που αναπτύσσεται στα επόμενα κεφάλαια.

2.4 True alarms, false alarms και alert fatigue

Η διάκριση ανάμεσα σε true alarms και false alarms αποτελεί βασικό ζήτημα στα συστήματα παρακολούθησης υγείας. Ένα true alarm αντιστοιχεί σε πραγματικό συμβάν ή σε κατάσταση που έχει κλινική σημασία και απαιτεί προσοχή. Αντίθετα, ένα false alarm ενεργοποιείται χωρίς να υπάρχει αντίστοιχο πραγματικό ή κλινικά σημαντικό περιστατικό. Η διάκριση αυτή είναι κρίσιμη, επειδή επηρεάζει τον τρόπο με τον οποίο το προσωπικό αντιλαμβάνεται, αξιολογεί και διαχειρίζεται τους συναγερμούς.

Στα περιβάλλοντα εντατικής θεραπείας, τα monitoring systems παράγουν μεγάλο αριθμό alarms. Τα alarms αυτά είναι απαραίτητα για την ασφάλεια των ασθενών, επειδή μπορούν να ειδοποιήσουν έγκαιρα για μεταβολές σε ζωτικά σημεία ή φυσιολογικά σήματα. Ωστόσο, όταν μεγάλο μέρος των alarms είναι τεχνικά λανθασμένο ή κλινικά μη σημαντικό, τότε η συνεχής έκθεση του προσωπικού σε συναγερμούς μπορεί να δημιουργήσει σοβαρό πρόβλημα διαχείρισης [6], [10]. Το φαινόμενο αυτό περιγράφεται στη βιβλιογραφία ως alert fatigue ή alarm fatigue. Ο όρος αναφέρεται στη σταδιακή κόπωση, εξοικείωση ή μειωμένη ευαισθησία των επαγγελματιών υγείας απέναντι σε συχνούς, επαναλαμβανόμενους ή μη σημαντικούς συναγερμούς. Όταν το προσωπικό εκτίθεται σε μεγάλο αριθμό false ή non-actionable

alarms, μπορεί να μειωθεί η προσοχή απέναντι σε πραγματικά κρίσιμα alarms, με πιθανές συνέπειες για την ασφάλεια των ασθενών [11], [13].

Η βιβλιογραφία δείχνει ότι το πρόβλημα των false alarms δεν είναι απλώς τεχνικό. Δεν αφορά μόνο το αν ένας αισθητήρας ή ένας αλγόριθμος ενεργοποίησε σωστά ή λανθασμένα έναν συναγερμό. Αφορά και τον τρόπο με τον οποίο οι συναγερμοί επηρεάζουν την καθημερινή λειτουργία μιας μονάδας, τον φόρτο εργασίας του προσωπικού και την αξιοπιστία του συνολικού monitoring περιβάλλοντος [12].

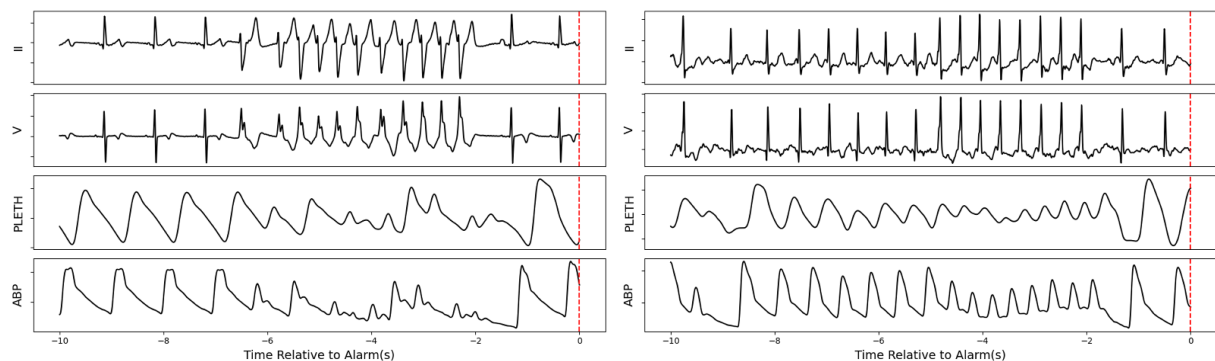
Για τον λόγο αυτό, η αξιολόγηση ενός alarm δεν μπορεί να περιορίζεται στην απλή καταγραφή της ετικέτας true ή false. Χρειάζεται να είναι σαφές από ποια πηγή προήλθε το alarm, ποιος ή ποια διαδικασία το αξιολόγησε, ποια απόφαση λήφθηκε και πώς αυτή η πληροφορία τεκμηριώνεται. Η ύπαρξη οργανωμένης τεκμηρίωσης επιτρέπει τη μελλοντική αναζήτηση, τον έλεγχο και τη σύνδεση των alarms με ευρύτερες απαιτήσεις αξιοπιστίας και διαχείρισης κινδύνου.

Στο VTaC dataset, κάθε ventricular tachycardia alarm event συνοδεύεται από τελική απόφαση σχετικά με το αν χαρακτηρίζεται ως true ή false alarm [4]. Αυτό καθιστά το dataset ιδιαίτερα κατάλληλο για την παρούσα οντολογική μελέτη, επειδή επιτρέπει να αναπαρασταθεί όχι μόνο το ίδιο το alarm event, αλλά και η αξιολόγηση που το συνοδεύει. Η πληροφορία αυτή μπορεί να οργανωθεί σε ξεχωριστές έννοιες, όπως Expert Review και Review Decision.

Στο οντολογικό μοντέλο, η διάκριση true/false δεν αντιμετωπίζεται ως απλή τιμή που αποθηκεύεται πάνω στο event. Αντίθετα, η απόφαση αναπαρίσταται ως ξεχωριστό στοιχείο της διαδρομής ανθρώπινης αξιολόγησης. Έτσι, το alarm event συνδέεται με το Expert Review και, μέσω αυτού, με την αντίστοιχη Review Decision. Η επιλογή αυτή επιτρέπει πιο καθαρή αναπαράσταση της διαδικασίας και υποστηρίζει την ιχνηλασιμότητα.

Τα false alarm cases έχουν ιδιαίτερη σημασία για την εργασία, επειδή συνδέονται με τη λογική του risk layer. Στη βιβλιογραφία, τα συχνά false ή μη χρήσιμα alarms αναφέρονται ως παράγοντας που μπορεί να ενισχύσει το alarm fatigue και να επηρεάσει την ανταπόκριση του προσωπικού [6], [12], [13]. Για τον λόγο αυτό, στην οντολογία τα false alarms δεν αναπαρίστανται μόνο ως λανθασμένες αποφάσεις, αλλά συνδέονται επιπλέον με risk indicators, risk-relevant events και potential harm.

Η σύνδεση αυτή δεν σημαίνει ότι κάθε false alarm προκαλεί υποχρεωτικά βλάβη ή ότι αποδεικνύεται πραγματική επίπτωση σε συγκεκριμένο περιστατικό. Αντίθετα, σημαίνει ότι το false alarm μπορεί να αποκτήσει πρόσθετη σημασιολογική σημασία ως περιστατικό που σχετίζεται με πιθανό κίνδυνο. Η αναπαράσταση αυτή είναι χρήσιμη, επειδή επιτρέπει στο μοντέλο να διακρίνει τα false alarm cases από τα true alarm cases όχι μόνο ως προς την απόφαση, αλλά και ως προς τη σχέση τους με τη διαχείριση κινδύνου.



Σχήμα 2.1 Ενδεικτικά παραδείγματα true και false ventricular tachycardia alarm events από το VTaC dataset. (προσαρμογή από [4]).

Η έννοια του alert fatigue συνδέεται επομένως με τη συνολική ανάγκη για τεκμηρίωση και ιχνηλασιμότητα. Όταν ένα σύστημα παράγει μεγάλο αριθμό alarms, η δυνατότητα να εντοπιστούν, να ταξινομηθούν και να συσχετιστούν με evidence records αποκτά ιδιαίτερη σημασία. Η οντολογική αναπαράσταση μπορεί να συμβάλει σε αυτή την κατεύθυνση, επειδή οργανώνει τα alarm events και τις σχετικές αποφάσεις σε δομημένη μορφή.

Συνολικά, η διάκριση ανάμεσα σε true alarms, false alarms και alarm fatigue αποτελεί βασικό θεωρητικό υπόβαθρο για τη συγκεκριμένη προσέγγιση. Τα true alarms υποστηρίζουν τη βασική διαδρομή τεκμηρίωσης, ενώ τα false alarms επιτρέπουν την πρόσθετη risk-related αναπαράσταση. Αυτή η διαφοροποίηση είναι σημαντική για το οντολογικό μοντέλο, επειδή δείχνει ότι η ίδια δομή μπορεί να καλύψει τόσο την τεκμηρίωση των πραγματικών alarms όσο και την οργανωμένη αποτύπωση των false alarm cases ως περιστατικών με πιθανή σημασία για risk management.

2.5 Διαφάνεια, τεκμηρίωση και ιχνηλασιμότητα σε συστήματα τεχνητής νοημοσύνης

Η διαφάνεια, η τεκμηρίωση και η ιχνηλασιμότητα αποτελούν βασικές έννοιες στη συζήτηση γύρω από την υπεύθυνη και αξιόπιστη τεχνητή νοημοσύνη. Οι έννοιες αυτές δεν αφορούν μόνο το αν ένα σύστημα λειτουργεί τεχνικά σωστά, αλλά και το αν οι διαδικασίες, τα δεδομένα και τα αποτελέσματά του μπορούν να γίνουν κατανοητά, να ελεγχθούν και να συσχετιστούν με συγκεκριμένα τεκμήρια.

Στις κατευθυντήριες οδηγίες για το Trustworthy AI, η διαφάνεια εμφανίζεται ως μία από τις βασικές απαιτήσεις για την αξιόπιστη ανάπτυξη και χρήση συστημάτων τεχνητής νοημοσύνης [1]. Η έννοια αυτή συνδέεται με τη δυνατότητα κατανόησης του τρόπου λειτουργίας ενός συστήματος, των δεδομένων που χρησιμοποιεί, των περιορισμών του και των αποτελεσμάτων που παράγει. Σε περιβάλλοντα υψηλής κρισιμότητας, όπως ο χώρος της υγείας, η διαφάνεια αποκτά ιδιαίτερη σημασία, επειδή τα αποτελέσματα ενός συστήματος μπορεί να επηρεάζουν διαδικασίες παρακολούθησης, αξιολόγησης και λήψης αποφάσεων.

Η διεθνής βιβλιογραφία για το AI ethics συνδέει τη διαφάνεια με έννοιες όπως explainability, interpretability και understandability. Παράλληλα, η διαφάνεια δεν αφορά μόνο την τεχνική εξήγηση ενός αλγορίθμου, αλλά και τη δυνατότητα αιτιολόγησης της διαδικασίας σχεδιασμού, υλοποίησης και χρήσης ενός συστήματος [15]. Αυτό σημαίνει ότι ένα σύστημα δεν αρκεί να παράγει ένα αποτέλεσμα. Πρέπει να υπάρχει τρόπος να εξηγηθεί σε ποιο πλαίσιο δημιουργήθηκε αυτό το αποτέλεσμα, ποια δεδομένα το υποστήριζαν και πώς μπορεί να αξιολογηθεί.

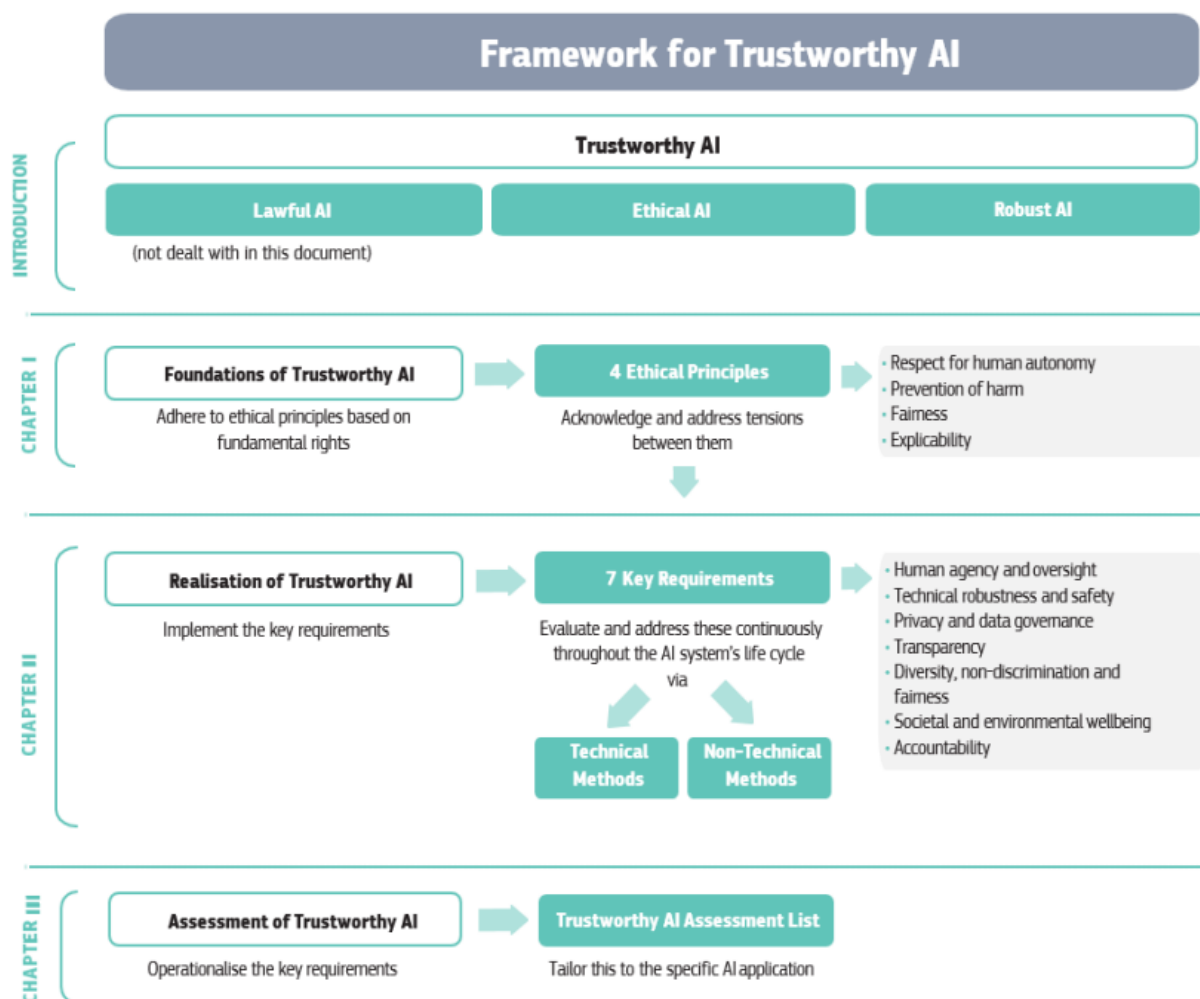
Η τεκμηρίωση αποτελεί πρακτική προϋπόθεση για τη διαφάνεια. Χωρίς οργανωμένη τεκμηρίωση, η πληροφορία παραμένει αποσπασματική και δύσκολα ελέγξιμη. Σε ένα σύστημα τεχνητής νοημοσύνης ή σε ένα AI-supported περιβάλλον, η τεκμηρίωση μπορεί να αφορά τα δεδομένα εισόδου, τις διαδικασίες αξιολόγησης, τις αποφάσεις, τα τεχνικά χαρακτηριστικά του συστήματος, τους περιορισμούς και τα μέτρα ελέγχου. Η ύπαρξη τέτοιων στοιχείων επιτρέπει σε διαφορετικούς εμπλεκόμενους να κατανοήσουν καλύτερα τι συνέβη και με βάση ποια πληροφορία.

Η ιχνηλασιμότητα συνδέεται άμεσα με την τεκμηρίωση. Αφορά τη δυνατότητα να ακολουθηθεί η διαδρομή μιας πληροφορίας από την πηγή της μέχρι το τελικό αποτέλεσμα ή την τεκμηριωτική της χρήση. Στο πλαίσιο των healthcare monitoring alarm events, η ιχνηλασιμότητα σημαίνει ότι ένα alarm πρέπει να μπορεί να συνδεθεί με το waveform record προέλευσής του, με την αξιολόγηση που πραγματοποιήθηκε, με την τελική απόφαση και με το evidence record που το τεκμηριώνει.

Η σημασία της ιχνηλασιμότητας αναδεικνύεται ιδιαίτερα όταν ένα σύστημα χρησιμοποιείται σε κρίσιμα περιβάλλοντα. Οι κατευθυντήριες οδηγίες για το Trustworthy AI τονίζουν την ανάγκη διευκόλυνσης της traceability και της auditability, ιδιαίτερα σε κρίσιμες καταστάσεις [1]. Η auditability προϋποθέτει ότι υπάρχουν επαρκή στοιχεία ώστε να μπορεί να εξεταστεί εκ των υστέρων τι συνέβη, ποια δεδομένα χρησιμοποιήθηκαν και πώς λήφθηκε ή υποστηρίχθηκε μια απόφαση.

Αντίστοιχα, το IEEE Ethically Aligned Design αναδεικνύει την ανάγκη τα intelligent and autonomous systems να ευθυγραμμίζονται με ανθρώπινες αξίες και ηθικές αρχές, δίνοντας έμφαση σε ζητήματα όπως accountability και transparency [2]. Η προσέγγιση αυτή είναι σημαντική, επειδή μεταφέρει τη συζήτηση πέρα από την καθαρά τεχνική απόδοση. Ένα σύστημα μπορεί να είναι λειτουργικό, αλλά να μην είναι επαρκώς ελέγξιμο ή κατανοητό. Σε τέτοια περίπτωση, η αξιοπιστία του παραμένει περιορισμένη.

Η λογοδοσία, ή accountability, συνδέεται στενά με τη διαφάνεια και την ιχνηλασιμότητα. Για να μπορεί κάποιος να αποδώσει ή να αναλάβει ευθύνη για τη λειτουργία ενός συστήματος, πρέπει να υπάρχουν διαθέσιμα τεκμήρια. Η βιβλιογραφία για το Responsible AI επισημαίνει ότι η υπευθυνότητα συνδέεται με μηχανισμούς auditability και accountability κατά τον σχεδιασμό, την ανάπτυξη και τη χρήση ενός AI-based συστήματος [14]. Συνεπώς, η λογοδοσία δεν μπορεί να στηριχθεί μόνο σε γενικές δηλώσεις, αλλά χρειάζεται οργανωμένη τεκμηρίωση και δυνατότητα ελέγχου.



Σχήμα 2.2 Το πλαίσιο των Ethics Guidelines για το Trustworthy AI [1].

Στον χώρο της υγείας, οι παραπάνω έννοιες αποκτούν πιο πρακτικό χαρακτήρα. Ένα alarm event που παράγεται από ένα monitoring system δεν πρέπει να αντιμετωπίζεται μόνο ως στιγμιαίο τεχνικό αποτέλεσμα. Χρειάζεται να μπορεί να αναζητηθεί, να συνδεθεί με την πηγή του, να αξιολογηθεί ως true ή false, και να ενταχθεί σε ένα πλαίσιο τεκμηρίωσης. Η ανάγκη αυτή γίνεται ακόμη πιο έντονη όταν εξετάζονται false alarms, επειδή η σημασία τους μπορεί να συνδέεται με ευρύτερα ζητήματα διαχείρισης κινδύνου και alarm fatigue.

Η οντολογική αναπαράσταση μπορεί να υποστηρίξει αυτές τις ανάγκες, επειδή επιτρέπει τη ρητή σύνδεση εννοιών και περιστατικών. Μέσα σε μια οντολογία, οι έννοιες δεν παραμένουν αόριστες. Ένα alarm event μπορεί να συνδεθεί με ένα waveform record, ένα expert review, μία review decision και ένα evidence record. Έτσι, η τεκμηρίωση δεν είναι απλώς κείμενο ή συνοδευτική περιγραφή, αλλά μέρος μιας δομημένης σημασιολογικής διαδρομής.

Στο πλαίσιο της διπλωματικής εργασίας, η διαφάνεια, η τεκμηρίωση και η ιχνηλασιμότητα λειτουργούν ως θεωρητική βάση για τον σχεδιασμό της οντολογίας. Το μοντέλο που αναπτύσσεται στα επόμενα κεφάλαια επιχειρεί να αποτυπώσει αυτές τις έννοιες με πρακτικό τρόπο, συνδέοντας πραγματικά VTaC alarm events με Evidence Records και, μέσω αυτών, με επιλεγμένα AI Act-related σημεία αναφοράς. Έτσι, η οντολογία λειτουργεί ως μηχανισμός οργάνωσης της πληροφορίας και υποστήριξης του requirement-to-evidence mapping.

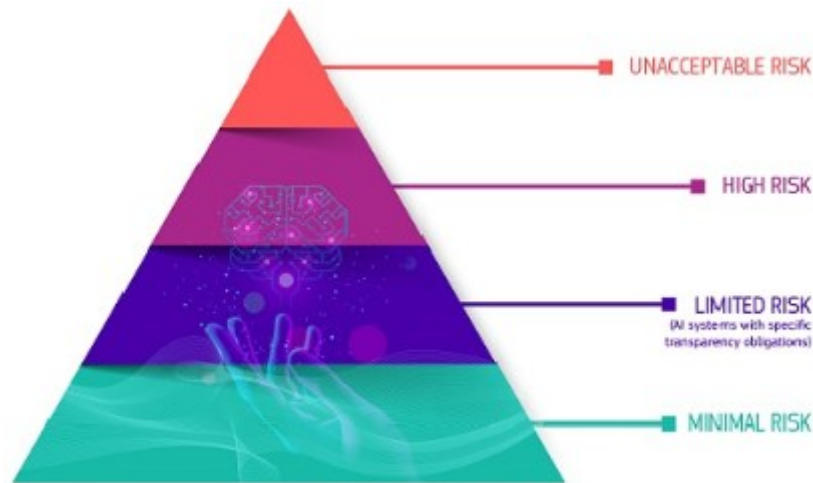
Η παρούσα ενότητα εστίασε στις γενικές έννοιες της διαφάνειας, της τεκμηρίωσης, της ιχνηλασιμότητας, της λογοδοσίας και της auditability. Η κανονιστική τους διάσταση, καθώς και η σύνδεσή τους με συγκεκριμένες απαιτήσεις του AI Act, παρουσιάζονται στην επόμενη ενότητα.

2.6 Ο AI Act ως κανονιστικό πλαίσιο αναφοράς

Ο AI Act της Ευρωπαϊκής Ένωσης, δηλαδή ο Κανονισμός (EU) 2024/1689, αποτελεί το πρώτο ολοκληρωμένο οριζόντιο κανονιστικό πλαίσιο της Ευρωπαϊκής Ένωσης για την τεχνητή νοημοσύνη. Ο βασικός του στόχος είναι η θέσπιση εναρμονισμένων κανόνων για την ανάπτυξη, τη διάθεση και τη χρήση συστημάτων τεχνητής νοημοσύνης, με τρόπο που να προωθεί την καινοτομία, αλλά ταυτόχρονα να προστατεύει την υγεία, την ασφάλεια και τα θεμελιώδη δικαιώματα [3].

Η σημασία του AI Act δεν περιορίζεται στο ότι εισάγει ένα νέο νομικό πλαίσιο. Η μεγαλύτερη ιδιαιτερότητά του είναι ότι επιχειρεί να συνδέσει την τεχνητή νοημοσύνη με συγκεκριμένες απαιτήσεις ελέγχου, τεκμηρίωσης και διακυβέρνησης. Με άλλα λόγια, δεν αντιμετωπίζει τα AI systems μόνο ως τεχνολογικά προϊόντα, αλλά ως συστήματα που πρέπει να μπορούν να ελεγχθούν, να τεκμηριωθούν, να εποπτευθούν και να αξιολογηθούν ως προς τους κινδύνους που ενδέχεται να δημιουργούν.

Κεντρικό χαρακτηριστικό του κανονισμού είναι η risk-based προσέγγιση. Σύμφωνα με αυτή τη λογική, οι υποχρεώσεις που ισχύουν για ένα σύστημα τεχνητής νοημοσύνης δεν είναι ίδιες σε όλες τις περιπτώσεις, αλλά εξαρτώνται από το επίπεδο κινδύνου που μπορεί να προκαλέσει το σύστημα. Όσο μεγαλύτερη είναι η πιθανή επίδραση ενός AI system στην ασφάλεια, στην υγεία ή στα θεμελιώδη δικαιώματα, τόσο αυστηρότερες είναι οι απαιτήσεις που πρέπει να ικανοποιούνται [3], [18].



Σχήμα 2.3 Risk-based προσέγγιση του AI Act και βασικές κατηγορίες κινδύνου. Ευρωπαϊκή Επιτροπή [18].

Η risk-based λογική του AI Act συνδέεται άμεσα με τη γενικότερη συζήτηση γύρω από το Trustworthy AI. Οι προηγούμενες κατευθυντήριες οδηγίες για αξιόπιστη τεχνητή νοημοσύνη είχαν ήδη αναδείξει απαιτήσεις όπως ανθρώπινη εποπτεία, τεχνική αξιοπιστία, διαφάνεια, λογοδοσία και διαχείριση κινδύνων [1]. Ο AI Act μεταφέρει μεγάλο μέρος αυτής της συζήτησης σε πιο δεσμευτικό κανονιστικό πλαίσιο, ιδίως για συστήματα που θεωρούνται υψηλού κινδύνου.

Ιδιαίτερη βαρύτητα δίνεται στα high-risk AI systems. Τα συστήματα αυτά δεν απαγορεύονται, αλλά υπόκεινται σε αυξημένες απαιτήσεις. Ο κανονισμός προβλέπει ότι τέτοια συστήματα πρέπει να σχεδιάζονται και να λειτουργούν με τρόπο που επιτρέπει τη διαχείριση κινδύνων, την ύπαρξη τεχνικής τεκμηρίωσης, την καταγραφή λειτουργιών, τη διαφάνεια προς τους χρήστες, την ανθρώπινη εποπτεία και την επίτευξη κατάλληλου επιπέδου ακρίβειας, ανθεκτικότητας και κυβερνοασφάλειας [3].

Στο πλαίσιο των high-risk AI systems, το άρθρο 9 του AI Act αφορά το risk management system. Η απαίτηση αυτή συνδέεται με τη συστηματική αναγνώριση, εκτίμηση και αντιμετώπιση κινδύνων που μπορεί να προκύψουν από τη χρήση ενός συστήματος τεχνητής νοημοσύνης. Η έννοια του κινδύνου δεν περιορίζεται μόνο σε τεχνική αστοχία, αλλά περιλαμβάνει πιθανές επιπτώσεις στην υγεία, στην ασφάλεια και στα θεμελιώδη δικαιώματα [3], [16].

Το άρθρο 10 αφορά τα δεδομένα και τη διακυβέρνησή τους. Η απαίτηση αυτή είναι ιδιαίτερα σημαντική σε συστήματα τεχνητής νοημοσύνης που βασίζονται σε training, validation ή testing datasets. Θέτει ζητήματα ποιότητας, καταλληλότητας, αντιπροσωπευτικότητας και διαχείρισης δεδομένων. Παρότι η συγκεκριμένη διπλωματική δεν εκπαιδεύει μοντέλο μηχανικής μάθησης, η απαίτηση αυτή αναφέρεται βιβλιογραφικά ως βασικό στοιχείο του AI Act, επειδή δείχνει τη σημασία που δίνει ο κανονισμός στην ποιότητα και στην προέλευση της πληροφορίας [3], [14].

Το άρθρο 11 αφορά την τεχνική τεκμηρίωση. Η τεχνική τεκμηρίωση επιτρέπει την κατανόηση του σκοπού, της λειτουργίας, των χαρακτηριστικών και των περιορισμών ενός AI system. Χωρίς οργανωμένη τεκμηρίωση, η αξιολόγηση ενός συστήματος γίνεται δύσκολη, τόσο από τεχνική όσο και από κανονιστική πλευρά. Η απαίτηση αυτή συνδέεται άμεσα με τη λογική της εργασίας, επειδή η οντολογία που αναπτύσσεται έχει ως βασικό στόχο τη σημασιολογική οργάνωση και τεκμηρίωση πραγματικών healthcare monitoring events [3], [5].

Το άρθρο 12 αφορά το record-keeping και προβλέπει ότι τα high-risk AI systems πρέπει να διαθέτουν τεχνική δυνατότητα αυτόματης καταγραφής γεγονότων (logs) κατά τη διάρκεια της λειτουργίας τους.

Οι δυνατότητες αυτές συνδέονται άμεσα με την ιχνηλασιμότητα, καθώς ο κανονισμός προβλέπει ότι το logging πρέπει να επιτρέπει κατάλληλο επίπεδο traceability της λειτουργίας του συστήματος, ανάλογα με τον σκοπό χρήσης του [3]. Στο πλαίσιο της παρούσας εργασίας, η απαίτηση αυτή χρησιμοποιείται ως κανονιστικό σημείο αναφοράς για την οργάνωση στοιχείων καταγραφής, όπως τα event identifiers και record identifiers του VTaC dataset. Τα στοιχεία αυτά αξιοποιούνται στο proof-of-concept μοντέλο για την αναζήτηση και τη σύνδεση των alarm events με την προέλευσή τους, χωρίς να θεωρούνται από μόνα τους απόδειξη πλήρους συμμόρφωσης ενός πραγματικού high-risk AI system με το άρθρο 12. Το άρθρο 13 αφορά τη διαφάνεια και την παροχή πληροφοριών στους deployers. Η διαφάνεια δεν σημαίνει μόνο ότι ένα σύστημα είναι τεχνικά προσβάσιμο ή ότι παρέχει κάποιο αποτέλεσμα. Σημαίνει ότι υπάρχουν επαρκείς πληροφορίες ώστε οι χρήστες ή οι υπεύθυνοι λειτουργίας να μπορούν να κατανοήσουν το πλαίσιο χρήσης, τις δυνατότητες και τους περιορισμούς του συστήματος. Η βιβλιογραφία για Trustworthy AI συνδέει τη διαφάνεια με την εξηγησιμότητα, την κατανόηση και τη δυνατότητα ελέγχου [1], [14], [15].

Το άρθρο 14 αφορά την ανθρώπινη εποπτεία (human oversight) και προβλέπει ότι τα high-risk AI systems πρέπει να σχεδιάζονται και να αναπτύσσονται με τρόπο που να επιτρέπει την αποτελεσματική εποπτεία τους από φυσικά πρόσωπα κατά τη χρήση τους. Σκοπός της εποπτείας είναι, μεταξύ άλλων, η πρόληψη ή ο περιορισμός κινδύνων για την υγεία, την ασφάλεια και τα θεμελιώδη δικαιώματα [3]. Στην παρούσα εργασία, η ανθρώπινη διάσταση αναπαρίσταται σε εφαρμοσμένο επίπεδο μέσω των εννοιών Expert Review και Review Decision, οι οποίες αποτυπώνουν την expert αξιολόγηση των VTaC alarm events. Η αναπαράσταση αυτή είναι συναφής με την ευρύτερη λογική της ανθρώπινης εποπτείας, αλλά δεν πρέπει να ερμηνεύεται ως πλήρης υλοποίηση ή αξιολόγηση της συμμόρφωσης με όλες τις απαιτήσεις του άρθρου 14.

Το άρθρο 15 αναφέρεται στην ακρίβεια, την ανθεκτικότητα και την κυβερνοασφάλεια. Οι απαιτήσεις αυτές αφορούν τη συνολική τεχνική αξιοπιστία ενός AI system. Ένα σύστημα υψηλού κινδύνου πρέπει να λειτουργεί με κατάλληλο επίπεδο ακρίβειας, να είναι ανθεκτικό σε σφάλματα και να προστατεύεται από απειλές που μπορεί να επηρεάσουν τη λειτουργία ή τα αποτελέσματά του [3]. Παρότι τα στοιχεία αυτά δεν μοντελοποιούνται αναλυτικά στην οντολογία της εργασίας, αποτελούν βασικό μέρος του ευρύτερου πλαισίου του AI Act.

Ο παρακάτω πίνακας συνοψίζει τις βασικές απαιτήσεις του AI Act για high-risk AI systems και δείχνει ποιες από αυτές αξιοποιούνται άμεσα στην οντολογική υλοποίηση της εργασίας.

Πίνακας 2-1 Βασικές απαιτήσεις του AI Act και σχέση με την οντολογική μοντελοποίηση

Απαίτηση / σχετική διάσταση του AI Act	Βασική σημασία	Σχέση με τη διπλωματική εργασία
Risk management	Αναγνώριση, εκτίμηση και αντιμετώπιση κινδύνων	Συνδέεται σε περιορισμένο επίπεδο με το lightweight risk-related επίπεδο για τα false alarm cases
Data and data governance	Ποιότητα, καταλληλότητα και διαχείριση δεδομένων	Αναφέρεται βιβλιογραφικά, αλλά δεν μοντελοποιείται πλήρως, επειδή δεν γίνεται εκπαίδευση ML μοντέλου
Technical documentation	Τεκμηρίωση σκοπού, λειτουργίας, χαρακτηριστικών και περιορισμών συστήματος	Συνδέεται με τη γενική λογική του evidence record και της τεκμηριωμένης αναπαράστασης

Record-keeping / logging	Καταγραφή γεγονότων και λειτουργιών για μελλοντικό έλεγχο	Συσχετίζεται στο proof-of-concept με event identifiers, record identifiers και Evidence Records, χωρίς να τεκμηριώνεται πλήρης logging capability του άρθρου 12
Transparency	Παροχή επαρκών πληροφοριών για κατανόηση και χρήση	Υποστηρίζεται έμμεσα μέσω ρητών σχέσεων και δυνατότητας ερωτημάτων
Human oversight	Δυνατότητα ανθρώπινης εποπτείας, αξιολόγησης και παρέμβασης	Συσχετίζεται σε εφαρμοσμένο επίπεδο με τις έννοιες Expert Review και Review Decision
Accuracy, robustness and cybersecurity	Τεχνική αξιοπιστία, ανθεκτικότητα και ασφάλεια	Αναγνωρίζεται ως σημαντική απαίτηση, αλλά δεν αποτελεί αντικείμενο υλοποίησης

Από τον πίνακα φαίνεται ότι η εργασία δεν αγνοεί το ευρύτερο πλαίσιο του AI Act. Αντίθετα, το παρουσιάζει συνολικά και στη συνέχεια επιλέγει συγκεκριμένες απαιτήσεις που μπορούν να υποστηριχθούν από τα διαθέσιμα δεδομένα και από τη σχεδιασμένη οντολογική δομή.

Η επιλογή των τεσσάρων AI Act-related σημείων αναφοράς που αξιοποιούνται στην οντολογία έγινε με βάση τη δυνατότητα σύνδεσής τους με πραγματικά στοιχεία του VTaC use case. Στο οντολογικό μοντέλο τα σημεία αυτά αναπαρίστανται μέσω των requirement individuals req_logging, req_human_review, req_traceability και req_risk_management. Οι ονομασίες αυτές αποτελούν εφαρμοσμένες έννοιες της οντολογίας και δεν πρέπει να θεωρούνται σε όλες τις περιπτώσεις αυτούσιες ονομασίες ανεξάρτητων νομικών απαιτήσεων του AI Act. Το req_logging συνδέεται κυρίως με τη λογική του record-keeping του άρθρου 12, το req_human_review με τη διάσταση της ανθρώπινης αξιολόγησης που είναι συναφής με το human oversight του άρθρου 14, το req_traceability με την απαίτηση ιχνηλασιμότητας που ενσωματώνεται ιδιαίτερα στο άρθρο 12, ενώ το req_risk_management συνδέεται με το πλαίσιο διαχείρισης κινδύνου του άρθρου 9 [3].

Η συγκεκριμένη επιλογή είναι μεθοδολογικά σημαντική. Αν επιχειρούνταν η πλήρης μοντελοποίηση όλων των απαιτήσεων του AI Act, η εργασία θα μετακινούνταν από ένα εστιασμένο proof-of-concept μοντέλο σε μια πολύ ευρύτερη προσπάθεια κανονιστικής συμμόρφωσης. Κάτι τέτοιο θα απαιτούσε περισσότερα δεδομένα, περισσότερα τεχνικά στοιχεία για το σύστημα, λεπτομερή ανάλυση του κύκλου ζωής του AI system, καθώς και στοιχεία που δεν υπάρχουν στο VTaC dataset.

Για παράδειγμα, η απαίτηση για data governance είναι ιδιαίτερα σημαντική, αλλά αφορά σε μεγάλο βαθμό τη διαχείριση των δεδομένων που χρησιμοποιούνται για την ανάπτυξη, εκπαίδευση, επικύρωση και δοκιμή ενός AI system. Στη συγκεκριμένη εργασία, το VTaC dataset δεν χρησιμοποιείται για εκπαίδευση μοντέλου. Χρησιμοποιείται ως πηγή πραγματικών alarm events και expert labels, τα οποία αναπαρίστανται οντολογικά. Επομένως, η πλήρης ανάλυση data governance θα αποτελούσε διαφορετικό αντικείμενο.

Αντίστοιχα, οι απαιτήσεις για accuracy, robustness και cybersecurity δεν μπορούν να αξιολογηθούν πλήρως στο πλαίσιο της συγκεκριμένης υλοποίησης, επειδή δεν αναπτύσσεται ούτε δοκιμάζεται ένα σύστημα πρόβλεψης. Η οντολογία δεν αποφασίζει αν ένα alarm είναι true ή false. Η απόφαση αυτή λαμβάνεται από το dataset και εισάγεται στο μοντέλο ως διαθέσιμη πληροφορία. Συνεπώς, οι απαιτήσεις αυτές αναγνωρίζονται βιβλιογραφικά, αλλά δεν μετατρέπονται σε κύριο μέρος της οντολογικής υλοποίησης.

Η απαίτηση για τεχνική τεκμηρίωση συνδέεται περισσότερο άμεσα με την προσέγγιση της εργασίας. Παρότι δεν δημιουργείται πλήρης φάκελος τεχνικής τεκμηρίωσης ενός AI system, η οντολογία οργανώνει τεκμήρια γύρω από τα alarm events. Το Evidence Record λειτουργεί ως κόμβος που συγκεντρώνει την τεκμηρίωση του event, ενώ μέσω του supportsRequirement το Evidence Record μπορεί να χαρτογραφείται προς επιλεγμένα AI Act-related σημεία αναφοράς. Με αυτόν τον τρόπο, η τεκμηρίωση δεν παραμένει απλή περιγραφή, αλλά αποκτά δομημένη μορφή.

Η απαίτηση για logging είναι από τις πιο κατάλληλες για τη συγκεκριμένη μελέτη εφαρμογής. Το VTaC dataset περιλαμβάνει αναγνωρίσιμα στοιχεία για records και events, τα οποία μπορούν να αναπαρασταθούν ως data properties και individuals μέσα στην οντολογία [4]. Η ύπαρξη τέτοιων identifiers επιτρέπει την αναζήτηση και τη σύνδεση των alarm events με το record προέλευσής τους, στοιχείο που υποστηρίζει τη λογική της ιχνηλασιμότητας. Επίσης, αξιοποιείται ως διαθέσιμο τεκμήριο στο proof-of-concept και δεν ισοδυναμεί με επιβεβαίωση της πλήρους τεχνικής logging capability που απαιτεί το άρθρο 12.

Η απαίτηση για human oversight συνδέεται με τον τρόπο με τον οποίο αξιοποιούνται οι expert decisions του VTaC dataset. Στην οντολογία, η ανθρώπινη αξιολόγηση δεν εμφανίζεται ως απλή ετικέτα πάνω στο event. Αντίθετα, αναπαρίσταται μέσω των εννοιών Expert Review και Review Decision. Αυτό επιτρέπει να φανεί ότι το alarm event έχει περάσει από αξιολόγηση και ότι η απόφαση true ή false είναι μέρος μιας τεκμηριωμένης σημασιολογικής διαδρομής. Η Expert Review του VTaC αποτελεί ανθρώπινη αξιολόγηση των alarm events και όχι πλήρη μηχανισμό human oversight ενός high-risk AI system κατά την έννοια του άρθρου 14.

Η traceability αποτελεί ένα από τα κεντρικά σημεία σύνδεσης ανάμεσα στον AI Act και στην οντολογική προσέγγιση της εργασίας. Το άρθρο 12 συνδέει τις δυνατότητες καταγραφής γεγονότων με την εξασφάλιση κατάλληλου επιπέδου traceability της λειτουργίας ενός high-risk AI system [3]. Στην παρούσα proof-of-concept υλοποίηση, η ιχνηλασιμότητα δεν αναπαρίσταται ως μία μοναδική γραμμική object-property αλυσίδα. Οργανώνεται μέσω διαφορετικών αλλά συνδεδεμένων σχέσεων γύρω από το ίδιο Alert Event: της σύνδεσης του event με το Waveform Record προέλευσής του, της διαδρομής Alert Event → Expert Review → Review Decision και της διαδρομής Alert Event → Evidence Record → Compliance Requirement. Με τον τρόπο αυτό μπορούν να ανακτηθούν διαφορετικές όψεις της προέλευσης, της αξιολόγησης και της τεκμηρίωσης του ίδιου περιστατικού.

Η απαίτηση για risk management αξιοποιείται κυρίως για τα false alarm cases. Η βιβλιογραφία για τα clinical alarms δείχνει ότι τα false και non-actionable alarms μπορούν να συνδεθούν με alarm fatigue και με επιπτώσεις στη λειτουργία ενός κλινικού περιβάλλοντος [6], [12], [13]. Για τον λόγο αυτό, το risk layer της οντολογίας συνδέει τα false alarms με Risk Indicator, Risk-Relevant Event και Potential Harm. Η σύνδεση αυτή δεν αποδεικνύει πραγματική βλάβη σε συγκεκριμένο περιστατικό, αλλά αποτυπώνει τη risk-related σημασία του false alarm. Το lightweight risk layer της παρούσας οντολογίας δεν αναπαριστά το πλήρες risk management system του άρθρου 9, το οποίο αποτελεί πολύ ευρύτερη διαδικασία σε όλο τον κύκλο ζωής ενός high-risk AI system. Αντίθετα, οργανώνει περιορισμένες risk-related πληροφορίες που είναι διαθέσιμες για τα false alarm cases του συγκεκριμένου proof-of-concept.

Η βιβλιογραφία γύρω από AI risk ontologies υποστηρίζει ότι οι έννοιες κινδύνου μπορούν να αναπαρασταθούν με οντολογικό τρόπο. Το AIRO, για παράδειγμα, αναπτύχθηκε για την αναπαράσταση AI risks με βάση την προτεινόμενη τότε μορφή του AI Act και πρότυπα διαχείρισης κινδύνου [16]. Παρότι η συγκεκριμένη διπλωματική δεν επαναχρησιμοποιεί πλήρως το AIRO, αξιοποιεί την ίδια γενική κατεύθυνση: οι κίνδυνοι και οι σχετικές έννοιες μπορούν να οργανωθούν σημασιολογικά και να γίνουν αντικείμενο ερωτημάτων.

Παράλληλα, η TAIR ontology δείχνει ότι οι απαιτήσεις από κανονισμούς και standards μπορούν να εξαχθούν, να οργανωθούν και να δημοσιευθούν σε μορφή που είναι traceable, queryable και navigable [5]. Η εργασία αυτή είναι σημαντική για τη διπλωματική, επειδή στηρίζει τη βασική ιδέα ότι οι κανονιστικές απαιτήσεις δεν χρειάζεται να παραμένουν μόνο σε φυσικό κείμενο. Μπορούν να μετατραπούν σε δομημένα στοιχεία γνώσης και να συνδεθούν με τεκμήρια.

Αντίστοιχα, η εργασία για την οντολογική αναπαράσταση των Fundamental Rights Impact Assessments δείχνει ότι ο AI Act δημιουργεί ανάγκη για machine-readable information, structured documentation και εργαλεία που μπορούν να υποστηρίξουν διαδικασίες αξιολόγησης και διακυβέρνησης [17]. Η προσέγγιση αυτή ενισχύει τη βιβλιογραφική βάση της διπλωματικής, επειδή δείχνει ότι η οντολογική μοντελοποίηση είναι κατάλληλη για τη γεφύρωση ανάμεσα σε κανονιστικές απαιτήσεις και τεχνικές ή οργανωτικές διαδικασίες.

Είναι επίσης σημαντικό να διευκρινιστεί ότι το VTaC dataset και τα alarm events που χρησιμοποιούνται στην εργασία δεν αποτελούν από μόνα τους σύστημα τεχνητής νοημοσύνης και δεν ταξινομούνται στην παρούσα μελέτη ως high-risk AI system. Το VTaC χρησιμοποιείται ως πηγή πραγματικών healthcare monitoring events και expert decisions πάνω στα οποία δοκιμάζεται η προτεινόμενη οντολογική δομή. Κατά συνέπεια, ο AI Act αξιοποιείται ως κανονιστικό πλαίσιο αναφοράς για το requirement-to-evidence mapping και όχι ως βάση για τη δήλωση νομικής συμμόρφωσης του ίδιου του dataset.

Η σύνδεση του AI Act με την οντολογία της εργασίας δεν πρέπει να ερμηνευθεί ως πλήρης αξιολόγηση συμμόρφωσης ενός πραγματικού high-risk AI system. Η υλοποίηση δεν αξιολογεί εμπορικό προϊόν, δεν εξετάζει πλήρη κύκλο ζωής AI system και δεν εκτελεί διαδικασία πιστοποίησης. Αντίθετα, προτείνει ένα περιορισμένο αλλά συγκεκριμένο proof-of-concept μοντέλο, το οποίο δείχνει πώς επιλεγμένα AI Act-related σημεία αναφοράς μπορούν να συνδεθούν με την τεκμηρίωση πραγματικών healthcare monitoring alarm events μέσω evidence records.

Με αυτόν τον τρόπο, ο AI Act λειτουργεί ως κανονιστικό πλαίσιο αναφοράς και όχι ως αντικείμενο πλήρους νομικής συμμόρφωσης. Η εργασία αξιοποιεί τις απαιτήσεις του για να ορίσει τι είδους τεκμηρίωση έχει σημασία, ποιες σχέσεις πρέπει να αποτυπωθούν και πώς μπορεί να δημιουργηθεί μια ελέγξιμη διαδρομή από το πραγματικό event μέχρι το requirement. Η αξία της προσέγγισης βρίσκεται ακριβώς σε αυτή τη μετάφραση: από το επίπεδο των γενικών απαιτήσεων στο επίπεδο της σημασιολογικής αναπαράστασης συγκεκριμένων περιστατικών.

Συνολικά, ο AI Act παρέχει το θεωρητικό και κανονιστικό υπόβαθρο για τις έννοιες που αξιοποιεί η οντολογία: risk management, logging, human review, traceability και documentation. Η ενότητα αυτή έδειξε ότι τα τέσσερα AI Act-related σημεία αναφοράς που χρησιμοποιούνται στην υλοποίηση δεν επιλέχθηκαν αποσπασματικά, αλλά εντάσσονται σε ένα ευρύτερο σύνολο απαιτήσεων για high-risk AI systems. Η επόμενη ενότητα εξετάζει τις οντολογίες και την αναπαράσταση γνώσης, δηλαδή το τεχνικό και σημασιολογικό μέσο μέσα από το οποίο οι παραπάνω έννοιες οργανώνονται στο προτεινόμενο μοντέλο.

2.7 Οντολογίες και αναπαράσταση γνώσης

Οι οντολογίες αποτελούν βασικό εργαλείο της αναπαράστασης γνώσης και του Σημασιολογικού Ιστού. Σε γενικό επίπεδο, μια οντολογία μπορεί να θεωρηθεί ως δομημένη περιγραφή ενός πεδίου γνώσης, η οποία ορίζει έννοιες, σχέσεις και περιορισμούς με τρόπο που μπορεί να γίνει κατανοητός από ανθρώπους και να υποβληθεί σε επεξεργασία από υπολογιστικά συστήματα. Η κλασική προσέγγιση του

Gruber περιγράφει την οντολογία ως ρητή προδιαγραφή μιας εννοιολόγησης, δηλαδή ως έναν σαφή τρόπο αποτύπωσης των βασικών εννοιών και σχέσεων ενός πεδίου [19].

Η σημασία των οντολογιών δεν περιορίζεται στον ορισμό εννοιών. Οι οντολογίες επιτρέπουν την οργάνωση γνώσης σε μορφή που μπορεί να επαναχρησιμοποιηθεί, να επεκταθεί και να ελεγχθεί. Στο πλαίσιο της γνώσης και των πληροφοριακών συστημάτων, η οντολογική προσέγγιση βοηθά στη δημιουργία κοινής κατανόησης ανάμεσα σε διαφορετικά συστήματα, χρήστες και εφαρμογές [20], [21].

Ένα βασικό χαρακτηριστικό των οντολογιών είναι ότι προσφέρουν ρητή σημασιολογική δομή. Αυτό σημαίνει ότι οι έννοιες δεν εμφανίζονται απλώς ως λέξεις ή ετικέτες, αλλά συνδέονται μεταξύ τους μέσα από καθορισμένες σχέσεις. Για παράδειγμα, μια έννοια μπορεί να ανήκει σε μια ευρύτερη κατηγορία, να συνδέεται με άλλη έννοια μέσω ιδιότητας ή να περιορίζεται από συγκεκριμένους κανόνες. Με αυτόν τον τρόπο, η γνώση αποκτά μορφή που μπορεί να αναζητηθεί και να ερμηνευθεί συστηματικά.

Στην ανάπτυξη μιας οντολογίας, οι βασικές δομικές μονάδες είναι οι classes, οι properties και τα individuals. Οι classes περιγράφουν τις βασικές κατηγορίες εννοιών ενός πεδίου. Οι properties εκφράζουν σχέσεις ή χαρακτηριστικά, ενώ τα individuals αναπαριστούν συγκεκριμένες πραγματικές περιπτώσεις. Η διάκριση αυτή είναι σημαντική, επειδή επιτρέπει τη μετάβαση από το γενικό επίπεδο της θεωρητικής περιγραφής στο ειδικό επίπεδο των πραγματικών δεδομένων [22].

Στη συγκεκριμένη διπλωματική εργασία, οι classes χρησιμοποιούνται για την αναπαράσταση εννοιών όπως Alert Event, Waveform Record, Expert Review, Review Decision, Evidence Record και Compliance Requirement. Τα object properties χρησιμοποιούνται για να συνδέσουν αυτές τις έννοιες μεταξύ τους, ενώ τα data properties αποτυπώνουν τιμές όπως event identifiers, record identifiers και decision labels. Τα individuals αντιστοιχούν σε πραγματικά VTaC events, records, decisions, evidence records και requirements.

Η οντολογική αναπαράσταση διαφοροποιείται από μια απλή βάση δεδομένων. Σε μια απλή δομή δεδομένων, οι εγγραφές μπορεί να αποθηκεύονται σε πίνακες ή αρχεία χωρίς να είναι πάντα εμφανής η σημασιολογική τους σχέση. Αντίθετα, σε μια οντολογία, η σύνδεση ανάμεσα στις έννοιες είναι μέρος του ίδιου του μοντέλου. Έτσι, ένα alarm event δεν εμφανίζεται μόνο ως γραμμή δεδομένων, αλλά ως αντικείμενο που συνδέεται με την πηγή του, την αξιολόγηση, την απόφαση και το σχετικό τεκμήριο.

Η τεχνολογική βάση των σύγχρονων οντολογιών στον Σημασιολογικό Ιστό συνδέεται στενά με το RDF και το OWL. Το RDF παρέχει ένα γράφο-κεντρικό μοντέλο αναπαράστασης πληροφορίας, στο οποίο οι δηλώσεις εκφράζονται συνήθως με τη μορφή subject–predicate–object [23]. Με αυτή τη λογική, οι οντότητες και οι τιμές αναπαρίστανται ως κόμβοι, ενώ οι σχέσεις μεταξύ τους αποτυπώνονται ως ακμές του γράφου.

Το OWL παρέχει πλουσιότερες δυνατότητες σημασιολογικής και λογικής αναπαράστασης, επιτρέποντας τον ορισμό classes, properties, ιεραρχιών, αξιωμάτων και περιορισμών [24]. Μέσω αυτών των μηχανισμών, μια οντολογία μπορεί να περιγράψει όχι μόνο ποια δεδομένα υπάρχουν, αλλά και τη σημασία των σχέσεων μεταξύ τους. Η δυνατότητα αυτή επιτρέπει επίσης τη χρήση λογικών μηχανισμών reasoning για τον έλεγχο της συνέπειας και την εξαγωγή γνώσης που προκύπτει από τα δηλωμένα axioms.

Η δυνατότητα λογικού ελέγχου είναι ιδιαίτερα σημαντική. Με τη χρήση reasoner, μια οντολογία μπορεί να εξεταστεί ως προς τη λογική συνέπειά της και να συναγάγει σχέσεις ή ταξινομήσεις που προκύπτουν από τα δηλωμένα axioms. Με αυτόν τον τρόπο, το μοντέλο δεν παραμένει απλώς μια στατική

ταξινόμηση εννοιών, αλλά μπορεί να υποστηρίξει ελέγχους που βοηθούν στον εντοπισμό ασυνεπειών και στην επαλήθευση της οντολογικής δομής.

Η γνώση που περιέχεται σε μια οντολογία μπορεί να αναζητηθεί και να ελεγχθεί με διαφορετικούς μηχανισμούς. Στο Protégé, τα DL Queries επιτρέπουν την υποβολή λογικών εκφράσεων πάνω στις classes και στις σχέσεις της οντολογίας, αξιοποιώντας τη γνώση που είναι διαθέσιμη μέσω του reasoner. Το SPARQL, αντίθετα, αποτελεί γλώσσα ερωτημάτων για δεδομένα RDF και επιτρέπει πιο σύνθετη ανάκτηση και συνδυασμό πληροφοριών από τον γράφο γνώσης [25]. Στην παρούσα εργασία, τα DL Queries χρησιμοποιούνται κυρίως για τον έλεγχο συγκεκριμένων σημασιολογικών συνθηκών και την ανάκτηση individuals, ενώ τα SPARQL queries χρησιμοποιούνται για την παρουσίαση συνδυασμένων σχέσεων και συγκεντρωτικών αποτελεσμάτων.

Στο πλαίσιο της εργασίας, τα DL Queries και τα SPARQL Queries δεν χρησιμοποιούνται ως απλή τεχνική επίδειξη. Αποτελούν μηχανισμό λειτουργικής επικύρωσης της οντολογίας. Μέσω αυτών εξετάζεται αν οι βασικές σημασιολογικές διαδρομές μπορούν να ανακτηθούν με τον τρόπο που σχεδιάστηκαν. Για παράδειγμα, μπορεί να ελεγχθεί αν ένα VTaC alarm event συνδέεται με waveform record, expert review, review decision και evidence record, καθώς και αν το evidence record υποστηρίζει συγκεκριμένη απαίτηση.

Η ανάπτυξη οντολογιών απαιτεί επίσης μεθοδολογική οργάνωση. Η βιβλιογραφία για ontology engineering επισημαίνει ότι ο σχεδιασμός πρέπει να ξεκινά από το πεδίο ενδιαφέροντος, τους βασικούς όρους, τις σχέσεις και τα ερωτήματα που η οντολογία πρέπει να μπορεί να απαντήσει [22]. Αυτά τα ερωτήματα λειτουργούν ως competency questions και βοηθούν στον έλεγχο του αν το μοντέλο καλύπτει τις ανάγκες για τις οποίες δημιουργήθηκε.

Η έννοια των competency questions είναι ιδιαίτερα χρήσιμη στη συγκεκριμένη εργασία. Τα ερευνητικά και λειτουργικά ερωτήματα που τέθηκαν κατά τον σχεδιασμό της οντολογίας μεταφράστηκαν αργότερα σε DL Queries και SPARQL Queries. Έτσι, ο έλεγχος της οντολογίας δεν έγινε αυθαίρετα, αλλά βασίστηκε σε συγκεκριμένες ερωτήσεις που συνδέονται με τη μεθοδολογία και τον σκοπό της υλοποίησης.

Για την πρακτική ανάπτυξη της οντολογίας χρησιμοποιήθηκε το Protégé. Το Protégé αποτελεί ευρέως χρησιμοποιούμενη πλατφόρμα ανοικτού κώδικα για τη δημιουργία, επεξεργασία και ανάλυση οντολογιών [26]. Μέσα από το περιβάλλον αυτό είναι δυνατός ο ορισμός classes, properties, individuals, annotations και logical axioms, καθώς και η χρήση reasoners και εργαλείων οπτικοποίησης.

Η επιλογή του Protégé συνδέεται με τον εφαρμοσμένο χαρακτήρα της εργασίας. Η οντολογία δεν παραμένει σε θεωρητικό επίπεδο, αλλά υλοποιείται πρακτικά σε περιβάλλον που υποστηρίζει OWL modeling, έλεγχο συνέπειας και ερωτήματα. Αυτό επιτρέπει την τεκμηρίωση της πορείας από τον σχεδιασμό μέχρι τη λειτουργική επικύρωση του μοντέλου.

Ιδιαίτερη σημασία έχει η δυνατότητα μιας οντολογίας να συνδέει διαφορετικά επίπεδα πληροφορίας. Στην περίπτωση της διπλωματικής, το πρώτο επίπεδο αφορά πραγματικά healthcare monitoring events. Το δεύτερο αφορά την ανθρώπινη αξιολόγηση και την τελική απόφαση. Το τρίτο αφορά τα evidence records. Το τέταρτο αφορά τα επιλεγμένα AI Act-related σημεία αναφοράς προς τα οποία χαρτογραφείται η σχετική τεκμηρίωση. Η οντολογία επιτρέπει αυτά τα επίπεδα να αποτυπωθούν σε ενιαίο σημασιολογικό μοντέλο.

Με αυτόν τον τρόπο, η οντολογική αναπαράσταση λειτουργεί ως γέφυρα ανάμεσα στα δεδομένα και στις απαιτήσεις. Το VTaC alarm event αποτελεί πραγματικό περιστατικό, ενώ το requirement αποτελεί

έννοια που προέρχεται από κανονιστικό πλαίσιο. Το evidence record επιτρέπει τη σύνδεση αυτών των δύο επιπέδων, ενώ οι σχέσεις της οντολογίας καθιστούν τη σύνδεση ρητή, αναζητήσιμη και ελέγξιμη.

Συνολικά, οι οντολογίες προσφέρουν ένα κατάλληλο πλαίσιο για την αναπαράσταση γνώσης σε περιπτώσεις όπου υπάρχουν πολλές έννοιες, σχέσεις και επίπεδα τεκμηρίωσης. Για τη διπλωματική εργασία, η αξία τους βρίσκεται στη δυνατότητα να οργανώσουν πραγματικά VTaC alarm events και να τα συνδέσουν με αξιολογήσεις, αποφάσεις και τεκμήρια και, μέσω αυτών, με επιλεγμένα AI Act-related σημεία αναφοράς. Η επόμενη ενότητα εξετάζει πιο ειδικά πώς οι οντολογίες έχουν χρησιμοποιηθεί στη βιβλιογραφία για Ethical AI, Trustworthy AI και AI governance.

2.8 Οντολογίες για Ethical AI / Trustworthy AI / AI governance

Οι οντολογίες έχουν αρχίσει να χρησιμοποιούνται όλο και περισσότερο στη βιβλιογραφία που αφορά το Ethical AI, το Trustworthy AI και το AI governance. Ο λόγος είναι ότι οι έννοιες που εμφανίζονται σε αυτά τα πεδία είναι συχνά γενικές, πολυεπίπεδες και δύσκολες να εφαρμοστούν με άμεσο τεχνικό τρόπο. Έννοιες όπως transparency, accountability, human oversight, risk, fairness, responsibility και compliance έχουν μεγάλη σημασία, αλλά χρειάζονται πιο σαφή αναπαράσταση για να μπορούν να ενταχθούν σε πληροφοριακά συστήματα, μηχανισμούς ελέγχου και διαδικασίες τεκμηρίωσης.

Η σχετική βιβλιογραφία δείχνει ότι υπάρχει συχνά απόσταση ανάμεσα στις γενικές αρχές της ηθικής τεχνητής νοημοσύνης και στην πρακτική εφαρμογή τους. Πολλά κείμενα πολιτικής και κατευθυντήριες οδηγίες περιγράφουν σημαντικές αρχές, όμως αυτές οι αρχές δεν είναι πάντα άμεσα αξιοποιήσιμες σε τεχνικό επίπεδο. Για παράδειγμα, η απαίτηση για διαφάνεια μπορεί να σημαίνει διαφορετικά πράγματα ανάλογα με το πεδίο εφαρμογής, τον χρήστη, το είδος του συστήματος και το είδος της πληροφορίας που χρειάζεται να τεκμηριωθεί [1], [2], [14], [15].

Σε αυτό το σημείο οι οντολογίες προσφέρουν ένα σημαντικό πλεονέκτημα. Μπορούν να μετατρέψουν γενικές έννοιες σε δομημένα στοιχεία γνώσης, να ορίσουν σχέσεις ανάμεσα σε αυτές τις έννοιες και να υποστηρίξουν ερωτήματα πάνω στο μοντέλο. Έτσι, η συζήτηση δεν μένει μόνο στο επίπεδο των αρχών, αλλά μπορεί να περάσει στο επίπεδο της σημασιολογικής αναπαράστασης, της τεκμηρίωσης και της αναζήτησης.

Μια χαρακτηριστική κατεύθυνση στη βιβλιογραφία είναι η οντολογική αναπαράσταση των αρχών του Ethical AI. Η εργασία για την AI Principles Ontology, γνωστή ως AIPO, αναπτύχθηκε για να αντιμετωπίσει το πρόβλημα της μεγάλης αύξησης των AI ethics principles και των σχετικών πολιτικών κειμένων [28]. Η βασική ιδέα είναι ότι οι αρχές αυτές μπορούν να οργανωθούν ως έννοιες σε knowledge graph, ώστε να είναι πιο εύκολη η σύγκριση, η αναζήτηση και η επαναχρησιμοποίησή τους.

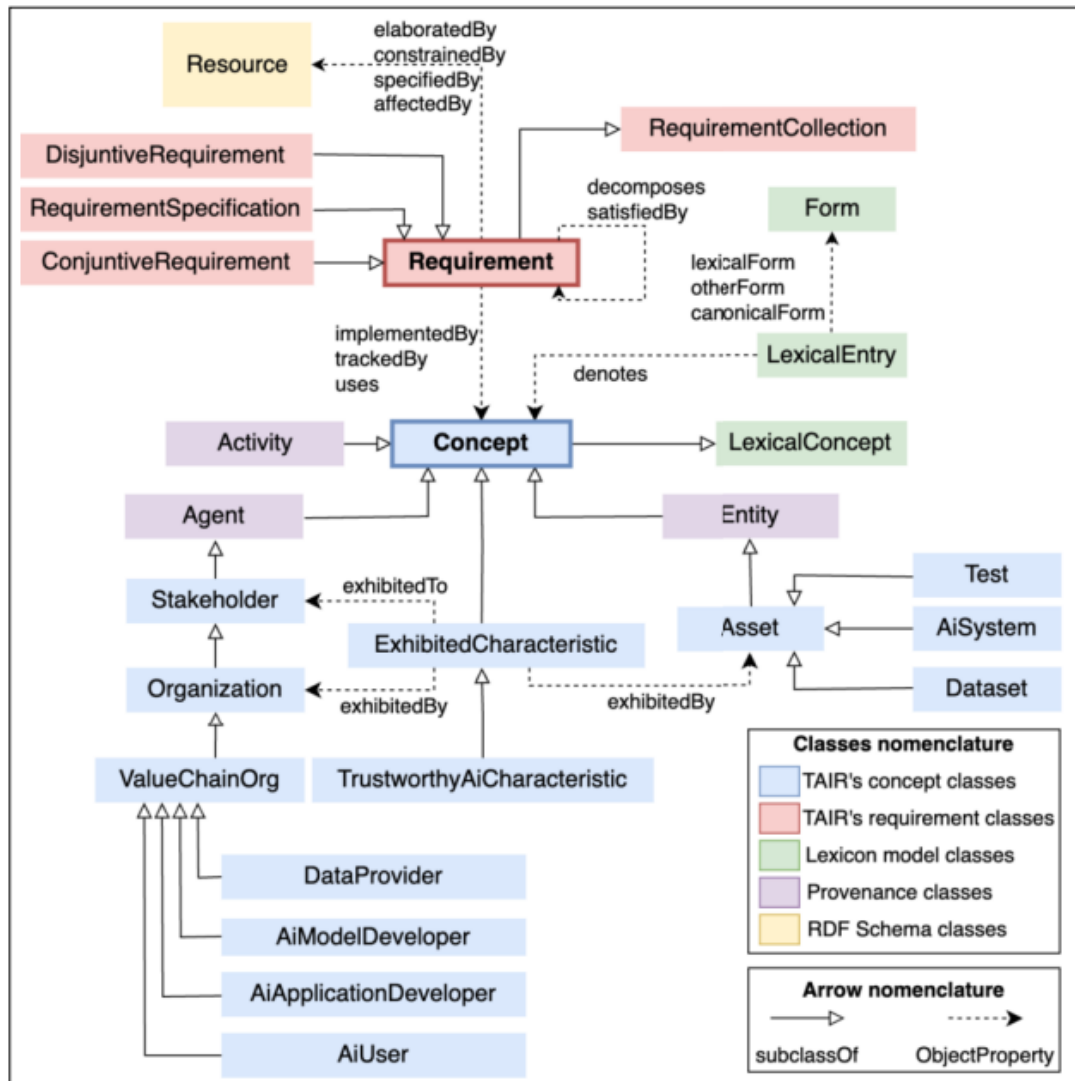
Η AIPO είναι σημαντική για τη συγκεκριμένη βιβλιογραφική ενότητα, επειδή δείχνει πώς οι ηθικές αρχές μπορούν να αναπαρασταθούν από στατικά κείμενα σε δομημένη σημασιολογική μορφή. Η προσέγγιση αυτή συνδέεται με τη γενικότερη ανάγκη να μη μένουν οι αρχές του Ethical AI μόνο σε επίπεδο διακηρύξεων, αλλά να αποκτούν μορφή που μπορεί να αξιοποιηθεί από υπολογιστικά συστήματα. Αν και η AIPO δεν ασχολείται ειδικά με healthcare monitoring events, προσφέρει σημαντικό υπόβαθρο για τη χρήση οντολογιών σε ethical AI contexts.

Μια δεύτερη σημαντική κατεύθυνση αφορά την ontology-based requirements engineering προσέγγιση για ethicality requirements. Η εργασία των Guizzardi, Amaral, Guizzardi και Mylopoulos αξιοποιεί τη μέθοδο Ontology-based Requirements Engineering (ObRE), με στόχο την οντολογική αποσαφήνιση εννοιών που σχετίζονται με την ηθικότητα και την υποστήριξη της διαδικασίας εντοπισμού και

ανάλυσης αντίστοιχων απαιτήσεων συστήματος [27]. Η προσέγγιση εξετάζει ethicality principles όπως Beneficence, Non-maleficence, Explicability και Autonomy και δείχνει πώς η οντολογική ανάλυση μπορεί να χρησιμοποιηθεί ώστε οι έννοιες αυτές να αποκτήσουν πιο σαφή θέση στη διαδικασία requirements engineering.

Η σημασία της προσέγγισης για την παρούσα διπλωματική βρίσκεται κυρίως στη μεθοδολογική της λογική. Δείχνει ότι γενικές ethical ή trustworthy έννοιες μπορούν να αποσαφηνιστούν οντολογικά και να συνδεθούν με συγκεκριμένες απαιτήσεις και στοιχεία ενός συστήματος. Η παρούσα εργασία δεν εφαρμόζει τη μέθοδο ObRE ούτε μοντελοποιεί ethicality requirements με τον ίδιο τρόπο, αλλά αξιοποιεί την ευρύτερη ιδέα ότι η οντολογική αναπαράσταση μπορεί να λειτουργήσει ως γέφυρα ανάμεσα σε αφηρημένες αρχές ή απαιτήσεις και σε πιο συγκεκριμένα στοιχεία τεκμηρίωσης.

Ιδιαίτερα κοντά στο αντικείμενο της εργασίας βρίσκεται η Trustworthy AI Requirements Ontology, γνωστή ως TAIR. Η TAIR προτείνει έναν τρόπο χαρτογράφησης εννοιών και απαιτήσεων από κανονιστικά και τεχνικά κείμενα, με έμφαση στον AI Act και σε διεθνή standards [5]. Η βασική συμβολή της είναι ότι αντιμετωπίζει τις απαιτήσεις όχι ως απλό κείμενο, αλλά ως στοιχεία που μπορούν να εξαχθούν, να οργανωθούν και να δημοσιευθούν σε μορφή open knowledge graph.



Σχήμα 2.4 Βασικές έννοιες και σχέσεις της TAIR ontology [5].

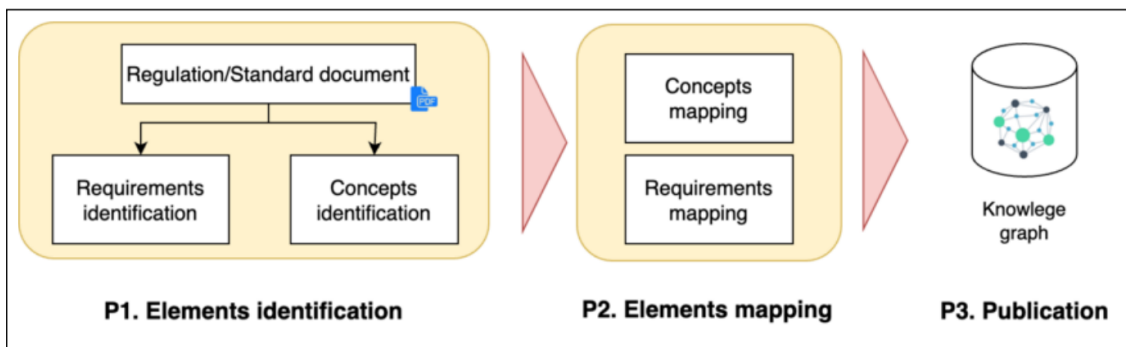
Η TAIR είναι πολύ σημαντική για τη διπλωματική, επειδή δείχνει ότι οι απαιτήσεις του AI Act μπορούν να γίνουν traceable, queryable και navigable. Αυτό σημαίνει ότι μια απαίτηση μπορεί να συνδεθεί με έννοιες, κείμενα, standards και άλλα στοιχεία με τρόπο που επιτρέπει αναζήτηση και έλεγχο. Η λογική αυτή βρίσκεται πολύ κοντά στη δική μας προσέγγιση, όπου Evidence Records που συνδέονται με πραγματικά VTaC alarm events χαρτογραφούνται προς επιλεγμένα AI Act-related σημεία αναφοράς.

Ωστόσο, υπάρχει μια βασική διαφορά. Η TAIR εστιάζει κυρίως στη χαρτογράφηση απαιτήσεων και εννοιών από regulatory και standardization documents. Η δική μας εργασία κινείται σε πιο εφαρμοσμένο επίπεδο healthcare monitoring. Δεν επιχειρεί να δημιουργήσει γενική οντολογία για ολόκληρο τον AI Act, αλλά να δείξει πώς επιλεγμένα AI Act-related σημεία αναφοράς μπορούν να συνδεθούν, μέσω Evidence Records, με πραγματικά alarm events, ανθρώπινη αξιολόγηση, αποφάσεις και τεκμήρια.

Πρέπει ωστόσο να ληφθεί υπόψη ότι η TAIR αναπτύχθηκε σε περίοδο κατά την οποία η τελική μορφή του AI Act δεν είχε ακόμη ολοκληρωθεί. Για τον λόγο αυτό, η συγκεκριμένη εργασία αξιοποιείται στην παρούσα διπλωματική κυρίως ως μεθοδολογικό παράδειγμα οντολογικής εξαγωγής, οργάνωσης και χαρτογράφησης κανονιστικών απαιτήσεων και όχι ως πηγή για την ακριβή σημερινή διατύπωση των υποχρεώσεων του Κανονισμού. Για την κανονιστική ερμηνεία των επιλεγμένων απαιτήσεων χρησιμοποιείται ως κύρια πηγή ο τελικός Κανονισμός (EU) 2024/1689 [3].

Μια ακόμη σημαντική προσέγγιση είναι το AI Risk Ontology, γνωστό ως AIRO. Το AIRO αναπτύχθηκε ως OWL ontology για τη σημασιολογική αναπαράσταση κινδύνων τεχνητής νοημοσύνης, με βάση μια πρόωμη προτεινόμενη μορφή του AI Act και σχετικά πρότυπα διαχείρισης κινδύνου [16]. Το μοντέλο οργανώνει έννοιες που σχετίζονται με AI risks, stakeholders, AI processes και risk management και αποτελεί επίσης ανώτερο οντολογικό επίπεδο για το Vocabulary of AI Risks (VAIR). Επειδή βασίζεται σε προγενέστερη μορφή του AI Act, αξιοποιείται στην παρούσα εργασία ως βιβλιογραφικό και οντολογικό παράδειγμα αναπαράστασης risk-related γνώσης και όχι ως πηγή για την τελική κανονιστική διατύπωση του Κανονισμού.

Το AIRO είναι σχετικό με τη διπλωματική εργασία κυρίως λόγω του risk layer. Στο μοντέλο που αναπτύσσεται, τα false alarm cases συνδέονται με risk indicators, risk-relevant events και potential harm. Η εργασία δεν επαναχρησιμοποιεί πλήρως το AIRO ούτε επιχειρεί να αναπαραστήσει συνολικά το risk management framework του AI Act. Παρ' όλα αυτά, αξιοποιεί την ίδια βασική ιδέα: ότι οι έννοιες κινδύνου μπορούν να οργανωθούν οντολογικά και να χρησιμοποιηθούν για πιο καθαρή αναπαράσταση περιστατικών που έχουν πιθανή risk-related σημασία.



Σχήμα 2.5 Φάσεις χαρτογράφησης εννοιών και απαιτήσεων από κανονιστικά και τεχνικά κείμενα [5].

Στο ίδιο ερευνητικό πεδίο εντάσσεται και η οντολογία για τα Fundamental Rights Impact Assessments, δηλαδή η FRIA ontology. Η συγκεκριμένη εργασία εξετάζει πώς η διαδικασία FRIA του AI Act μπορεί να αναπαρασταθεί ως machine-readable και interoperable ontology [17]. Η προσέγγιση αυτή είναι σημαντική, επειδή συνδέει άμεσα τον AI Act με την ανάγκη για δομημένη τεκμηρίωση, αυτοματοποιημένα εργαλεία, competency questions και σημασιολογικά standards.

Η FRIA ontology δείχνει ότι συγκεκριμένες κανονιστικές διαδικασίες μπορούν να αναπαρασταθούν ως machine-readable, interoperable και δομημένες σημασιολογικές διαδικασίες, ώστε να υποστηρίζονται η τεκμηρίωση, η διαχείριση της σχετικής πληροφορίας και η ανάπτυξη εργαλείων υποστήριξης. Η διπλωματική εργασία ακολουθεί ανάλογη λογική, αλλά σε μικρότερο και πιο συγκεκριμένο πεδίο: τα healthcare monitoring alarm events.

Οι παραπάνω προσεγγίσεις έχουν κοινό χαρακτηριστικό ότι προσπαθούν να γεφυρώσουν το χάσμα ανάμεσα σε γενικές αρχές, κανονιστικές απαιτήσεις και τεχνική αναπαράσταση. Η AIPO ξεκινά από ethical AI principles. Η ontology-based ethicality requirements προσέγγιση ξεκινά από ethicality requirements και requirements engineering. Το AIRO ξεκινά από AI risks. Η TAIR ξεκινά από απαιτήσεις και concepts του AI Act και των standards. Η FRIA ontology ξεκινά από μια συγκεκριμένη διαδικασία του AI Act. Όλες όμως δείχνουν ότι η οντολογική μοντελοποίηση μπορεί να χρησιμοποιηθεί για τη μετάβαση από το κείμενο και τις αρχές σε δομημένη γνώση.

Από τον πίνακα προκύπτει ότι η διπλωματική εργασία δεν αναπτύσσεται απομονωμένα από τη βιβλιογραφία. Αντίθετα, εντάσσεται σε μια ευρύτερη ερευνητική κατεύθυνση που χρησιμοποιεί οντολογίες για να οργανώσει έννοιες, απαιτήσεις, κινδύνους και διαδικασίες γύρω από το Ethical AI και το Trustworthy AI.

1. Purpose	
The purpose of this ontology is to model the FRIA as an information process.	
2. Scope	
The scope of this ontology is limited to modelling the FRIA as defined in AI Act Article 27.	
3. Implementation Language	
W3C semantic web standards - OWL, RDFS, SKOS	
4. Intended End-Users	
Organisations who create and use FRIA i.e. deployers of the AI systems and AI Act authorities	
5. Intended Uses	
Use 1. To document obligations regarding FRIA. Use 2. To document outcomes of FRIA. Use 3. To notify authorities regarding FRIA.	
6. Ontology Requirements	
a. Non-Functional Requirements	
NFR 1. Interoperability: The ontology must extend existing legal compliance ontologies (e.g. DPV) NFR 2. Scalability: The ontology should be adaptable/extensible for future developments. NFR 3. Usability: The ontology should support use by legal and non-legal stakeholders.	
b. Functional Requirements: Groups of Competency Questions	
CQG1. Related to AI Act obligations	CQG2. Related to Organisational Governance
CQ1. When was the FRIA conducted? CQ2. What is the Intended Purpose of the AI system? CQ3. What are the risks, consequences, impacts? CQ4. What are the mitigation measures?	CQ5. What is the outcome of the FRIA process? CQ6. What fundamental rights are affected? CQ7. What authorities are notified for the FRIA? CQ8. What documentation/tools were used for FRIA?

Σχήμα 2.6 Προδιαγραφή απαιτήσεων και competency questions της FRIA ontology (προσαρμογή από [17])

Πίνακας 2-2 Ενδεικτικές οντολογικές προσεγγίσεις για Ethical AI, Trustworthy AI και AI governance

Προσέγγιση	Κύρια εστίαση	Σχέση με τη διπλωματική εργασία
AIPO	Οντολογική αναπαράσταση AI ethics principles	Δείχνει πώς γενικές αρχές Ethical AI μπορούν να οργανωθούν ως σημασιολογική γνώση
Ontology-based ethicality requirements	Οντολογική αποσαφήνιση και ανάλυση ethicality requirements	Υποστηρίζει τη λογική ότι οι ethical/trustworthy απαιτήσεις μπορούν να αναλυθούν και να τεκμηριωθούν
AIRO / VAIR	Οντολογική αναπαράσταση AI risks και risk-related εννοιών	Συνδέεται με το lightweight risk layer της εργασίας για false alarm cases
TAIR	Χαρτογράφηση concepts και requirements από AI Act και standards	Στηρίζει μεθοδολογικά τη σημασιολογική αναπαράσταση, χαρτογράφηση και αναζήτηση κανονιστικών requirements, πάνω στην οποία επεκτείνεται η λογική του requirement-to-evidence mapping της παρούσας εργασίας
FRIA ontology	Οντολογική αναπαράσταση Fundamental Rights Impact Assessments	Δείχνει πώς διαδικασίες του AI Act μπορούν να γίνουν machine-readable και τεκμηριώσιμες

Παράλληλα, η προσέγγιση της εργασίας διαφοροποιείται από τις παραπάνω πηγές ως προς το πεδίο εφαρμογής. Οι περισσότερες σχετικές οντολογίες κινούνται σε γενικό επίπεδο αρχών, απαιτήσεων, regulatory clauses, standards ή impact assessment procedures. Η δική μας εφαρμογή εστιάζει σε πραγματικά healthcare monitoring alarm events. Αυτό σημαίνει ότι η οντολογία δεν οργανώνει μόνο κανονιστικές έννοιες, αλλά συνδέει αυτές τις έννοιες με πραγματικά περιστατικά, όπως VTaC events, waveform records, expert reviews, review decisions και evidence records.

Η διαφοροποίηση αυτή είναι σημαντική, επειδή μεταφέρει τη συζήτηση από το γενικό επίπεδο της AI governance στο επίπεδο της τεκμηρίωσης συγκεκριμένων δεδομένων. Ένα requirement δεν παραμένει απλώς μια γενική κανονιστική απαίτηση. Συνδέεται με evidence record, το οποίο με τη σειρά του συνδέεται με συγκεκριμένο alarm event και με την αξιολόγησή του. Έτσι, η σχέση ανάμεσα στο κανονιστικό επίπεδο και στο επίπεδο των πραγματικών δεδομένων γίνεται πιο σαφής.

Η βιβλιογραφία για Ethical AI και AI governance δείχνει επίσης ότι η απλή ύπαρξη αρχών δεν αρκεί. Χρειάζονται μηχανισμοί που επιτρέπουν να ελεγχθεί αν και πώς αυτές οι αρχές συνδέονται με πρακτικές διαδικασίες, δεδομένα και τεκμήρια. Η οντολογική αναπαράσταση μπορεί να συμβάλει σε αυτό, επειδή επιτρέπει τη δημιουργία σχέσεων που μπορούν να ανακτηθούν μέσω ερωτημάτων. Στη διπλωματική εργασία, αυτό ελέγχεται πρακτικά με DL Queries και SPARQL queries. Επιπλέον, οι σχετικές οντολογικές προσεγγίσεις αναδεικνύουν τη σημασία της επαναχρησιμοποίησης γνώσης και της διαλειτουργικότητας. Πολλές από αυτές αξιοποιούν Semantic Web standards, όπως RDF, OWL, SKOS και SPARQL, ώστε η γνώση να μπορεί να οργανωθεί με τρόπο επεκτάσιμο και συμβατό με άλλες πηγές [17], [28]. Αν και η συγκεκριμένη υλοποίηση είναι μικρής κλίμακας, ακολουθεί την ίδια γενική λογική, καθώς υλοποιείται σε OWL περιβάλλον και ελέγχεται με σημασιολογικά ερωτήματα.

Συνολικά, η βιβλιογραφία για οντολογίες σε Ethical AI, Trustworthy AI και AI governance παρέχει ισχυρή θεωρητική και μεθοδολογική βάση για τη συγκεκριμένη εργασία. Δείχνει ότι οι οντολογίες μπορούν να χρησιμοποιηθούν για την αναπαράσταση αρχών, απαιτήσεων, κινδύνων, τεκμηρίων και

διαδικασιών συμμόρφωσης. Η συμβολή της διπλωματικής βρίσκεται στο ότι εφαρμόζει αυτή τη λογική σε ένα συγκεκριμένο healthcare monitoring use case, χρησιμοποιώντας πραγματικά VTaC alarm events και συνδεδεσθέντά τα με Evidence Records και επιλεγμένα AI Act-related σημεία αναφοράς.

2.9 Κενό στη βιβλιογραφία

Η βιβλιογραφική ανασκόπηση που προηγήθηκε δείχνει ότι η διπλωματική εργασία στηρίζεται σε ήδη διαμορφωμένες ερευνητικές περιοχές. Η τεχνητή νοημοσύνη στην υγεία, τα healthcare monitoring systems, τα clinical alarms, το Trustworthy AI, ο AI Act και οι οντολογίες αποτελούν πεδία με σημαντική βιβλιογραφική ανάπτυξη. Ωστόσο, μέσα στις πηγές που εξετάστηκαν στο πλαίσιο της παρούσας εργασίας, δεν εντοπίστηκε προσέγγιση που να συνδυάζει σε ένα ενιαίο εφαρμοσμένο οντολογικό μοντέλο πραγματικά VTaC healthcare monitoring alarm events, expert αξιολόγηση, evidence records και χαρτογράφηση προς επιλεγμένα AI Act-related σημεία αναφοράς. Το βιβλιογραφικό κενό που εξετάζεται επομένως δεν αφορά την απουσία έρευνας στα επιμέρους πεδία, αλλά τον συγκεκριμένο τρόπο σύνδεσής τους.

Στη βιβλιογραφία για την τεχνητή νοημοσύνη στην υγεία, η έμφαση δίνεται συχνά στις δυνατότητες των AI systems να υποστηρίζουν διάγνωση, πρόγνωση, ανάλυση ιατρικών δεδομένων και λήψη αποφάσεων [7], [8]. Παράλληλα, αναδεικνύεται η ανάγκη για υπεύθυνη χρήση της τεχνητής νοημοσύνης, ιδιαίτερα σε περιβάλλοντα όπου οι αποφάσεις επηρεάζουν ανθρώπους και κλινικές διαδικασίες [9]. Στις πηγές που εξετάστηκαν για την AI στην υγεία, δεν αποτελεί κύρια εστίαση η οντολογική τεκμηρίωση πραγματικών monitoring alarm events και η σύνδεσή τους με κανονιστικές απαιτήσεις.

Αντίστοιχα, οι μελέτες για clinical alarms, false alarms και alarm fatigue αναδεικνύουν τη σημασία των συναγεμίων στα συστήματα παρακολούθησης υγείας. Η σχετική βιβλιογραφία δείχνει ότι τα false ή non-actionable alarms μπορούν να επιβαρύνουν το προσωπικό και να συνδεθούν με φαινόμενα alarm fatigue [6], [10], [11], [12], [13]. Οι πηγές που εξετάστηκαν τεκμηριώνουν κυρίως το κλινικό και λειτουργικό πρόβλημα των alarms και του alarm fatigue, χωρίς να αναπτύσσουν τον συγκεκριμένο συνδυασμό οντολογικής αναπαράστασης alarm events, evidence records και AI Act-related σημείων αναφοράς που αποτελεί αντικείμενο της παρούσας εργασίας.

Παράλληλα, η βιβλιογραφία για το Trustworthy AI και το Ethical AI αναδεικνύει έννοιες όπως διαφάνεια, λογοδοσία, ανθρώπινη εποπτεία, τεκμηρίωση, αξιοπιστία και διαχείριση κινδύνων [1], [2], [14], [15]. Οι έννοιες αυτές είναι κρίσιμες για την ανάπτυξη και χρήση συστημάτων τεχνητής νοημοσύνης. Ωστόσο, σε αρκετές περιπτώσεις παραμένουν σε επίπεδο αρχών ή γενικών απαιτήσεων. Το ζητούμενο είναι πώς μπορούν αυτές οι έννοιες να συνδεθούν με συγκεκριμένα δεδομένα, πραγματικά περιστατικά και τεκμήρια.

Ο AI Act ενισχύει ακόμη περισσότερο την ανάγκη για τέτοια σύνδεση. Ο κανονισμός εισάγει απαιτήσεις για high-risk AI systems, όπως risk management, data governance, technical documentation, record-keeping, transparency, human oversight, accuracy, robustness και cybersecurity [3]. Παρότι το κανονιστικό πλαίσιο είναι πλέον σαφέστερο, η πρακτική μετάβαση από τις απαιτήσεις του κανονισμού σε τεχνικά και σημασιολογικά μοντέλα τεκμηρίωσης παραμένει σύνθετο ζήτημα.

Οι ontology-based προσεγγίσεις για AI governance επιχειρούν να καλύψουν μέρος της απόστασης μεταξύ κανονιστικών απαιτήσεων και τεχνικής αναπαράστασης. Η TAIR οργανώνει concepts και normative requirements από τον AI Act και σχετικά standards σε σημασιολογική μορφή που μπορεί να

είναι traceable, queryable και navigable [5]. Αντίστοιχα, το AIRO οργανώνει έννοιες που σχετίζονται με AI risks και risk management [16], ενώ η FRIA ontology αναπαριστά σημασιολογικά τη διαδικασία Fundamental Rights Impact Assessment του AI Act [17]. Οι προσεγγίσεις αυτές τεκμηριώνουν ότι κανονιστικές έννοιες, απαιτήσεις, κίνδυνοι και διαδικασίες μπορούν να αναπαρασταθούν με οντολογικό τρόπο. Στις εργασίες που εξετάστηκαν, ωστόσο, η κύρια εστίαση βρίσκεται στο κανονιστικό, σημασιολογικό ή governance επίπεδο και όχι στη σύνδεση πραγματικών VTaC healthcare monitoring alarm events με expert decisions, evidence records και επιλεγμένα AI Act-related requirement mappings.

Ωστόσο, οι περισσότερες από αυτές τις προσεγγίσεις κινούνται κυρίως στο επίπεδο των κανονιστικών εννοιών, των απαιτήσεων, των principles ή των διαδικασιών governance. Δεν εστιάζουν ειδικά στη σύνδεση αυτών των απαιτήσεων με πραγματικά healthcare monitoring alarm events. Επομένως, παραμένει ένα εφαρμοσμένο κενό: πώς μπορεί ένα πραγματικό alarm event από dataset υγείας να συνδεθεί με αξιολόγηση, απόφαση, evidence record και επιλεγμένα AI Act-related σημεία αναφοράς.

Στην άλλη πλευρά, datasets όπως το VTaC παρέχουν πραγματικά ventricular tachycardia alarm events από ICU monitors, μαζί με expert labels true ή false [4]. Αυτά τα δεδομένα έχουν μεγάλη αξία για την έρευνα σε clinical alarms και false alarm detection. Ωστόσο δεν έχει σχεδιαστεί ως σημασιολογικό ή κανονιστικό μοντέλο και επομένως δεν περιλαμβάνει από μόνο του οντολογικές σχέσεις μεταξύ record, event, expert review, evidence και AI Act-related requirement.

Πίνακας 2-3 Σύνθεση βιβλιογραφικού κενού

Πεδίο βιβλιογραφίας	Τι καλύπτει	Τι παραμένει ανοιχτό για τη συγκεκριμένη εργασία
AI στην υγεία	Εφαρμογές AI σε διάγνωση, πρόγνωση, ανάλυση δεδομένων και υποστήριξη αποφάσεων	Στις πηγές που εξετάστηκαν δεν αποτελεί κύρια εστίαση η οντολογική τεκμηρίωση πραγματικών monitoring alarm events και η σύνδεσή τους με AI Act-related σημεία αναφοράς.
Clinical alarms και alarm fatigue	Σημασία true/false alarms, προβλήματα false alarms και επιπτώσεις στο προσωπικό	Οι εξεταζόμενες πηγές επικεντρώνονται κυρίως στο πρόβλημα των alarms και των επιπτώσεών τους και όχι σε event–evidence–requirement οντολογική αναπαράσταση.
Trustworthy AI / Ethical AI	Αρχές όπως διαφάνεια, λογοδοσία, ανθρώπινη εποπτεία και αξιοπιστία	Οι γενικές αρχές χρειάζονται εφαρμοσμένους μηχανισμούς σύνδεσης με συγκεκριμένα δεδομένα και τεκμήρια.
AI Act	Κανονιστικές απαιτήσεις για high-risk AI systems	Παρέχει το κανονιστικό πλαίσιο και τις απαιτήσεις, όχι από μόνος του ένα εφαρμοσμένο οντολογικό μοντέλο σύνδεσης requirement–evidence–event.
Οντολογίες για AI governance	Αναπαράσταση αρχών, απαιτήσεων, κινδύνων και κανονιστικών διαδικασιών	Στις προσεγγίσεις που εξετάστηκαν, η έμφαση βρίσκεται κυρίως σε principles, requirements, risks ή regulatory processes και όχι σε πραγματικά VTaC alarm events.
VTaC dataset	Πραγματικά VT alarm events με expert labels true/false	Παρέχει πραγματικά VT alarm events και expert decisions, αλλά δεν έχει σχεδιαστεί ως requirement-to-evidence οντολογικό μοντέλο.

Επομένως, το βασικό βιβλιογραφικό κενό βρίσκεται στη διασταύρωση των παραπάνω πεδίων. Υπάρχει βιβλιογραφία για AI στην υγεία, αλλά όχι απαραίτητα για οντολογική τεκμηρίωση πραγματικών alarm events. Υπάρχει βιβλιογραφία για alarm fatigue, αλλά όχι απαραίτητα για σύνδεση false alarms με AI Act-related requirements. Υπάρχει βιβλιογραφία για AI Act ontologies και AI governance, αλλά όχι ειδικά εφαρμοσμένη σε πραγματικά VTaC healthcare monitoring events.

Το κενό αυτό είναι το σημείο στο οποίο τοποθετείται η διπλωματική εργασία. Η προτεινόμενη οντολογία δεν επιχειρεί να αντικαταστήσει τις υπάρχουσες οντολογίες AI governance ούτε να δημιουργήσει πλήρες μοντέλο συμμόρφωσης με τον AI Act. Αντίθετα, αξιοποιεί τη λογική της οντολογικής αναπαράστασης για να συνδέσει πραγματικά healthcare monitoring alarm events με στοιχεία τεκμηρίωσης και επιλεγμένα AI Act-related σημεία αναφοράς.

Η βασική διαφοροποίηση της προσέγγισης είναι ότι ξεκινά από πραγματικά περιστατικά. Τα VTaC events δεν χρησιμοποιούνται ως υποθετικά παραδείγματα, αλλά ως πραγματικές εγγραφές με identifiers, waveform records και expert decisions. Πάνω σε αυτά δημιουργείται σημασιολογική δομή που επιτρέπει, αφενός, τη σύνδεση του event με το waveform record προέλευσής του και με την expert αξιολόγηση και την τελική απόφαση και, αφετέρου, την παράλληλη σύνδεσή του με το αντίστοιχο Evidence Record.

Μέσω του evidence record, η εργασία επιχειρεί να γεφυρώσει το operational επίπεδο των healthcare monitoring events με το normative επίπεδο των επιλεγμένων AI Act-related σημείων αναφοράς. Αυτό είναι σημαντικό, επειδή οι απαιτήσεις όπως logging, human review, traceability και risk management δεν εμφανίζονται απλώς ως γενικές έννοιες. Συνδέονται με συγκεκριμένα τεκμήρια που προκύπτουν από πραγματικά alarm events.

Επιπλέον, η ενσωμάτωση ενός lightweight risk layer για false alarm cases επιτρέπει να αναπαρασταθεί η πρόσθετη σημασία των false alarms σε σχέση με alert fatigue και risk management. Με αυτόν τον τρόπο, τα false alarms δεν αντιμετωπίζονται μόνο ως αρνητικές αποφάσεις, αλλά ως περιστατικά που μπορούν να έχουν risk-related σημασία μέσα στο οντολογικό μοντέλο.

Με βάση το σύνολο της βιβλιογραφίας που εξετάστηκε, το ειδικό κενό στο οποίο τοποθετείται η παρούσα εργασία αφορά τη σύνδεση τεσσάρων επιπέδων σε ένα ενιαίο proof-of-concept μοντέλο: (α) πραγματικών healthcare monitoring alarm events από το VTaC dataset, (β) της expert αξιολόγησης και της τελικής απόφασης για τα events, (γ) της οντολογικής οργάνωσης της σχετικής τεκμηρίωσης μέσω evidence records και (δ) της χαρτογράφησης των τεκμηρίων αυτών προς επιλεγμένα AI Act-related σημεία αναφοράς. Στις πηγές που εξετάστηκαν δεν εντοπίστηκε εφαρμογή που να συνδυάζει με αυτόν ακριβώς τον τρόπο τα τέσσερα επίπεδα. Η διαπίστωση αυτή δεν συνεπάγεται ότι δεν υπάρχουν ή δεν μπορούν να υπάρξουν συναφείς προσεγγίσεις, αλλά οριοθετεί το συγκεκριμένο εφαρμοσμένο κενό στο οποίο εστιάζει η διπλωματική εργασία.

2.10 Επίλογος

Στο δεύτερο κεφάλαιο παρουσιάστηκε το θεωρητικό υπόβαθρο και η σχετική βιβλιογραφία που υποστηρίζουν τη διπλωματική εργασία. Η ανάλυση οργανώθηκε γύρω από τις βασικές έννοιες που είναι απαραίτητες για την κατανόηση του προτεινόμενου οντολογικού μοντέλου: την τεχνητή νοημοσύνη στην υγεία, τα συστήματα παρακολούθησης ασθενών, τα healthcare monitoring alarm events, τα true και false alarms, το alert fatigue, το Trustworthy AI, τον AI Act και τις οντολογίες ως μέσο αναπαράστασης γνώσης.

Αρχικά εξετάστηκε η χρήση της τεχνητής νοημοσύνης και των συστημάτων παρακολούθησης στον χώρο της υγείας. Ιδιαίτερη έμφαση δόθηκε στα clinical alarm events και στη διάκριση μεταξύ true και false alarms, καθώς η ύπαρξη συχνών false alarms μπορεί να συνδεθεί με αυξημένο φόρτο για το προσωπικό και με φαινόμενα alarm fatigue. Η ανάλυση αυτή θεμελιώνει τη σημασία της οργανωμένης αξιολόγησης, τεκμηρίωσης και ιχνηλασιμότητας των alarm events.

Στη συνέχεια παρουσιάστηκαν οι έννοιες της διαφάνειας, της τεκμηρίωσης, της ιχνηλασιμότητας, της λογοδοσίας και της ανθρώπινης εποπτείας στο πλαίσιο του Trustworthy AI. Οι έννοιες αυτές συνδέθηκαν με τον AI Act, ο οποίος χρησιμοποιείται στην παρούσα εργασία ως κανονιστικό πλαίσιο αναφοράς. Η ανάλυση εστίασε ιδιαίτερα σε διαστάσεις σχετικές με το record-keeping και logging, την ανθρώπινη εποπτεία, την ιχνηλασιμότητα και τη διαχείριση κινδύνων. Παράλληλα, αποσαφηνίστηκε ότι τα αντίστοιχα requirement individuals της οντολογίας αποτελούν εφαρμοσμένες AI Act-related έννοιες του proof-of-concept μοντέλου και δεν συνιστούν από μόνα τους πλήρη αναπαράσταση ή απόδειξη νομικής συμμόρφωσης με τον Κανονισμό.

Ακολούθως εξετάστηκαν οι οντολογίες και οι βασικές τεχνολογίες αναπαράστασης γνώσης που χρησιμοποιούνται στην εργασία, όπως RDF, OWL, reasoning, competency questions και SPARQL. Η οντολογική προσέγγιση επιτρέπει τη ρητή οργάνωση εννοιών, σχέσεων και πραγματικών περιστατικών και παρέχει τη βάση για τη δημιουργία ενός ελέγξιμου και αναζητήσιμου σημασιολογικού μοντέλου.

Ιδιαίτερη έμφαση δόθηκε σε σχετικές οντολογικές προσεγγίσεις για Ethical AI, Trustworthy AI και AI governance. Εργασίες όπως η AIPO, η ontology-based προσέγγιση για ethicality requirements, το AIRO, η TAIR και η FRIA ontology δείχνουν διαφορετικούς τρόπους με τους οποίους ethical principles, requirements, risks και κανονιστικές διαδικασίες μπορούν να οργανωθούν σε σημασιολογική μορφή. Οι προσεγγίσεις αυτές παρέχουν σημαντική θεωρητική και μεθοδολογική βάση, αλλά έχουν διαφορετικό πεδίο εφαρμογής από την παρούσα διπλωματική.

Με βάση τη βιβλιογραφία που εξετάστηκε, το ειδικό κενό στο οποίο τοποθετείται η εργασία αφορά τον συνδυασμό πραγματικών VTaC healthcare monitoring alarm events, expert αξιολόγησης και review decisions, οντολογικής τεκμηρίωσης μέσω evidence records και σύνδεσης των τεκμηρίων αυτών με επιλεγμένα AI Act-related σημεία αναφοράς. Η διαπίστωση αυτή δεν αφορά την απουσία έρευνας στα επιμέρους πεδία, αλλά τον συγκεκριμένο τρόπο συνδυασμού τους σε ένα ενιαίο εφαρμοσμένο proof-of-concept μοντέλο.

Το επόμενο κεφάλαιο παρουσιάζει τη μεθοδολογία και τον σχεδιασμό της οντολογίας. Εξηγείται πώς οι έννοιες που θεμελιώθηκαν στο παρόν κεφάλαιο μετατρέπονται σε συγκεκριμένη οντολογική δομή, πώς αξιοποιούνται τα πραγματικά VTaC events και πώς σχεδιάζονται οι βασικές συνδεδεμένες σημασιολογικές διαδρομές, το requirement-to-evidence mapping και το lightweight risk-related επίπεδο για τα false alarm cases.

Κεφάλαιο 3ο: Μεθοδολογία και σχεδιασμός της οντολογίας

3.1 Εισαγωγή

Το παρόν κεφάλαιο παρουσιάζει τη μεθοδολογική προσέγγιση και τον σχεδιασμό της οντολογίας που αναπτύχθηκε στο πλαίσιο της διπλωματικής εργασίας. Η οντολογία υλοποιήθηκε με χρήση του περιβάλλοντος Protégé [26] και έχει ως στόχο τη σημασιολογική οργάνωση πραγματικών healthcare monitoring alarm events, καθώς και τη σύνδεσή τους με επιλεγμένα AI Act-related σημεία αναφοράς [3].

Η προσέγγιση που ακολουθείται δεν βασίζεται στη μηχανική μάθηση. Δεν επιχειρείται η εκπαίδευση ενός μοντέλου, η πρόβλεψη της ορθότητας ενός alarm ή η στατιστική αξιολόγηση της απόδοσης κάποιου αλγορίθμου. Αντίθετα, η εργασία επικεντρώνεται στη μοντελοποίηση γνώσης. Δηλαδή, εξετάζει πώς μπορούν να οργανωθούν έννοιες, σχέσεις και πραγματικά περιστατικά με τέτοιο τρόπο ώστε να δημιουργείται μια ελέγξιμη, αναζητήσιμη και κατανοητή σημασιολογική δομή.

Η βασική ανάγκη που επιχειρεί να καλύψει η οντολογία είναι η μετάβαση από επιμέρους τεχνικά δεδομένα σε οργανωμένη σημασιολογική τεκμηρίωση. Σε ένα σύστημα παρακολούθησης υγείας, ένα alarm event από μόνο του δεν επαρκεί για να υποστηρίξει την ιχνηλασιμότητα της σχετικής πληροφορίας. Απαιτείται να είναι δυνατή η σύνδεσή του με το record προέλευσης, την αξιολόγηση, την τελική απόφαση και το αντίστοιχο evidence record.

Για τον λόγο αυτό, η οντολογία οργανώθηκε γύρω από ένα σύνολο συνδεδεμένων σημασιολογικών διαδρομών που συναντώνται στο ίδιο Alert Event. Οι βασικές διαδρομές είναι:

Alert Event --belongsToRecord→ Waveform Record

Alert Event --hasReview→ Expert Review --hasDecision→ Review Decision

Alert Event --isDocumentedBy→ Evidence Record --supportsRequirement→ Compliance Requirement

Οι διαδρομές αυτές επιτρέπουν να διατηρούνται χωριστά αλλά συνδεδεμένα η προέλευση του event, η ανθρώπινη αξιολόγηση, η τελική απόφαση και η τεκμηρίωση. Δεν αποτελούν μία μοναδική γραμμική object-property αλυσίδα. Το Alert Event λειτουργεί ως κοινό σημείο γύρω από το οποίο οργανώνονται οι διαφορετικές πληροφορίες του proof-of-concept.

Η χρήση οντολογίας επιτρέπει την αναπαράσταση αυτών των σχέσεων με ρητό και μηχανικά επεξεργάσιμο τρόπο. Οι έννοιες και οι σχέσεις της εργασίας αποτυπώνονται ως classes, object properties, data properties και individuals στο OWL μοντέλο [24]. Με αυτόν τον τρόπο, η υλοποιημένη γνώση μπορεί στη συνέχεια να ελεγχθεί με reasoner, να εξεταστεί μέσω DL Queries και να ανακτηθεί ή να παρουσιαστεί μέσω SPARQL queries [25].

Η επιλογή του Protégé ως περιβάλλοντος υλοποίησης συνδέεται με τη δυνατότητα δημιουργίας και επεξεργασίας OWL ontologies, ορισμού classes και properties, εισαγωγής individuals και αξιοποίησης reasoning μηχανισμών [26]. Στην παρούσα εργασία το περιβάλλον χρησιμοποιείται τόσο για την κατασκευή του μοντέλου όσο και για τον λειτουργικό του έλεγχο, ο οποίος παρουσιάζεται αναλυτικά στο Κεφάλαιο 5.

Ιδιαίτερη σημασία έχει ότι η οντολογία δεν επιδιώκει να αποδείξει πλήρη νομική συμμόρφωση με τον AI Act. Ο Κανονισμός (EU) 2024/1689 περιλαμβάνει ένα ευρύτερο σύνολο νομικών, οργανωτικών και τεχνικών υποχρεώσεων, οι οποίες δεν μπορούν να αξιολογηθούν συνολικά με βάση τα διαθέσιμα

δεδομένα της συγκεκριμένης μελέτης [3]. Η παρούσα εργασία προτείνει ένα περιορισμένο proof-of-concept σημασιολογικό μοντέλο, το οποίο εξετάζει πώς πραγματικά VTaC alarm events και οι διαθέσιμες expert αποφάσεις τους [4] μπορούν να οργανωθούν σε evidence records και να συνδεθούν με επιλεγμένα AI Act-related σημεία αναφοράς.

Με βάση τα παραπάνω, το κεφάλαιο δεν περιορίζεται στην απλή καταγραφή των τεχνικών στοιχείων της οντολογίας. Παρουσιάζει τη λογική πίσω από τις βασικές σχεδιαστικές επιλογές. Εξηγείται γιατί επιλέχθηκε η οντολογική προσέγγιση αντί της μηχανικής μάθησης, πώς αξιοποιείται το VTaC dataset [4], γιατί επιλέχθηκαν συγκεκριμένα AI Act-related σημεία αναφοράς, πώς οργανώθηκαν οι βασικές συνδεδεμένες σημασιολογικές διαδρομές και γιατί τα false alarm cases συνδέθηκαν με ένα lightweight risk-related επίπεδο.

3.2 Ερευνητική προσέγγιση της εργασίας

Η ανάπτυξη της οντολογίας βασίστηκε σε έναν συγκεκριμένο ερευνητικό προβληματισμό: πώς μπορούν γενικές αρχές και απαιτήσεις που σχετίζονται με το Ethical και Trustworthy AI να συνδεθούν με μια πιο συγκεκριμένη, ελέγξιμη και τεχνικά οργανωμένη μορφή τεκμηρίωσης. Η υπεύθυνη ανάπτυξη και χρήση συστημάτων τεχνητής νοημοσύνης δεν συνδέεται μόνο με την τεχνική τους απόδοση, αλλά και με ζητήματα όπως η ανθρώπινη εποπτεία, η διαφάνεια, η λογοδοσία, η ασφάλεια και η διαχείριση κινδύνων [1], [14].

Οι αρχές αυτές συχνά διατυπώνονται σε σχετικά υψηλό επίπεδο αφαίρεσης και απαιτούν περαιτέρω εξειδίκευση προκειμένου να συνδεθούν με συγκεκριμένα στοιχεία ενός συστήματος. Στη σχετική βιβλιογραφία έχουν προταθεί οντολογικές προσεγγίσεις για την οργάνωση Ethical AI principles και για την αποσαφήνιση και ανάλυση ethicality requirements [27], [28]. Οι προσεγγίσεις αυτές υποστηρίζουν τη γενικότερη ιδέα ότι η οντολογική μοντελοποίηση μπορεί να λειτουργήσει ως μηχανισμός μετάβασης από γενικές έννοιες και αρχές σε περισσότερο δομημένες σημασιολογικές αναπαραστάσεις.

Σε αυτό το πλαίσιο προκύπτει ο ρόλος της οντολογίας στην παρούσα εργασία. Η οντολογική μοντελοποίηση επιτρέπει τη ρητή αναπαράσταση εννοιών και σχέσεων, ώστε στοιχεία που διαφορετικά θα παρέμεναν αποσπασματικά να μπορούν να οργανωθούν σε ενιαίο σημασιολογικό μοντέλο [20], [22]. Για παράδειγμα, η έννοια της ιχνηλασιμότητας δεν αντιμετωπίζεται μόνο ως γενική αρχή, αλλά αποτυπώνεται μέσα από συγκεκριμένες σχέσεις που επιτρέπουν να εντοπιστεί από ποιο waveform record προέρχεται ένα event, ποια αξιολόγηση πραγματοποιήθηκε, ποια απόφαση λήφθηκε και ποιο evidence record συνδέεται με το περιστατικό.

Η ερευνητική αφετηρία, επομένως, δεν ήταν η απλή δημιουργία ενός τεχνικού σχήματος στο Protégé, αλλά η ανάγκη γεφύρωσης διαφορετικών επιπέδων πληροφορίας. Από τη μία πλευρά βρίσκονται οι αρχές του Trustworthy AI και το κανονιστικό πλαίσιο του AI Act. Από την άλλη πλευρά βρίσκονται τα πραγματικά healthcare monitoring alarm events και οι διαθέσιμες expert αποφάσεις του VTaC dataset. Η οντολογία σχεδιάστηκε ως ενδιάμεσο σημασιολογικό επίπεδο που επιτρέπει την οργανωμένη σύνδεση αυτών των επιπέδων.

Ο AI Act χρησιμοποιείται στην εργασία ως κανονιστικό πλαίσιο αναφοράς [3]. Η υλοποίηση δεν επιχειρεί να μοντελοποιήσει το σύνολο του Κανονισμού ούτε να αξιολογήσει πλήρη νομική συμμόρφωση. Αντίθετα, εστιάζει σε επιλεγμένα AI Act-related σημεία αναφοράς που παρουσιάζουν συνάφεια με τα διαθέσιμα στοιχεία του healthcare monitoring use case. Στο οντολογικό μοντέλο τα

σημεία αυτά αναπαρίστανται μέσω των requirement individuals req_logging, req_human_review, req_traceability και req_risk_management.

Η επιλογή αυτή ακολουθεί τη βασική αρχή ότι κάθε κανονιστική έννοια που εισάγεται στο proof-of-concept πρέπει να μπορεί να συνδεθεί με κάποιο συγκεκριμένο στοιχείο της υλοποίησης. Η σχετική βιβλιογραφία δείχνει ότι concepts και normative requirements από κανονιστικά και τεχνικά κείμενα μπορούν να εξαχθούν, να οργανωθούν σημασιολογικά και να γίνουν αναζητήσιμα μέσω οντολογικών και knowledge-graph προσεγγίσεων [5]. Στην παρούσα εργασία η λογική αυτή μεταφέρεται σε εφαρμοσμένο επίπεδο, καθώς τα επιλεγμένα AI Act-related σημεία αναφοράς συνδέονται με τεκμήρια που προκύπτουν από πραγματικά healthcare monitoring events.

Κεντρική σχεδιαστική επιλογή αποτέλεσε η χρήση του Evidence Record ως ενδιάμεσου κόμβου. Το Evidence Record λειτουργεί ως σημείο σύνδεσης μεταξύ του operational επιπέδου των πραγματικών events και του normative επιπέδου των επιλεγμένων requirement individuals. Ένα alarm event δεν συνδέεται απευθείας με ένα requirement individual. Αντίθετα, το event τεκμηριώνεται μέσω ενός evidence record και το evidence record συνδέεται στη συνέχεια με ένα ή περισσότερα επιλεγμένα requirements. Με αυτόν τον τρόπο, το requirement-to-evidence mapping γίνεται ρητό και αναζητήσιμο μέσα στην οντολογία.

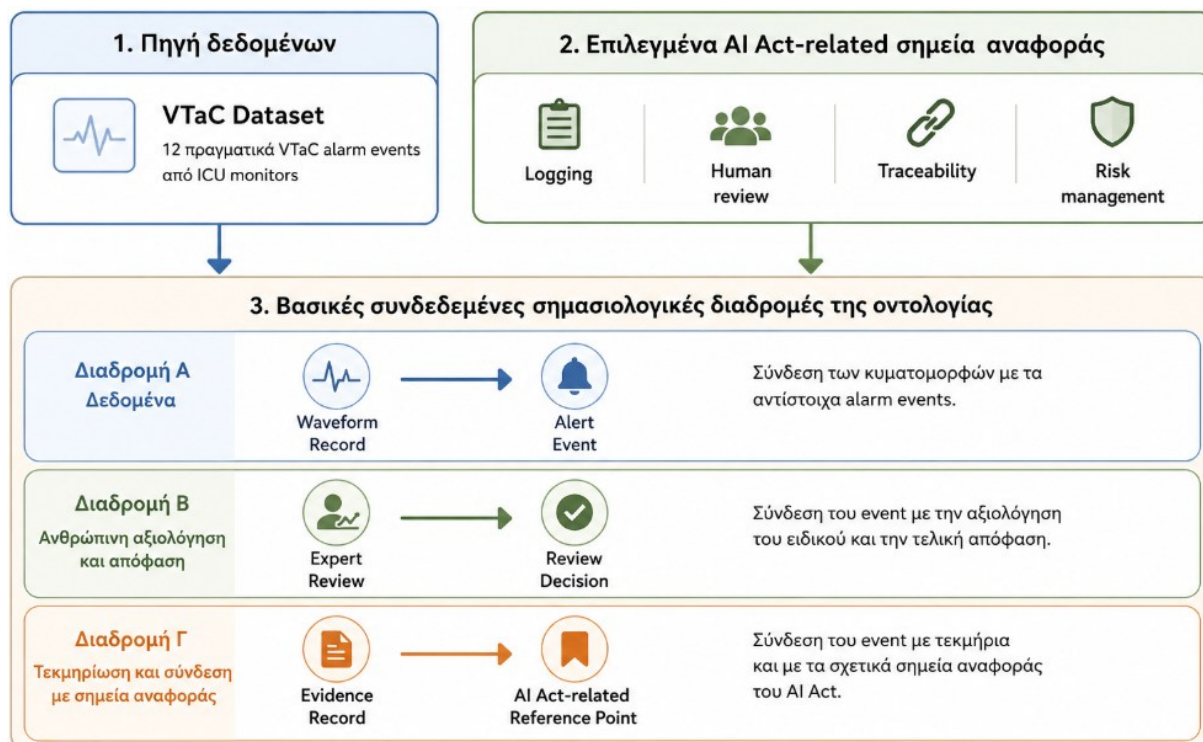
Η χρήση του Evidence Record αποτελεί σχεδιαστική επιλογή της παρούσας εργασίας και δεν αποτελεί αυτούσια εφαρμογή κάποιας από τις οντολογίες της βιβλιογραφίας. Η TAIR, για παράδειγμα, παρέχει μεθοδολογικό προηγούμενο για τη σημασιολογική αναπαράσταση και χαρτογράφηση κανονιστικών concepts και requirements [5], αλλά η σύνδεση πραγματικών VTaC events με evidence records και requirement individuals αποτελεί μέρος του συγκεκριμένου proof-of-concept που αναπτύσσεται στη διπλωματική.

Αντίστοιχα, η ερευνητική λογική επηρέασε την απόφαση να συμπεριληφθεί ένα ξεχωριστό lightweight risk-related επίπεδο για τα false alarm cases. Η βιβλιογραφία για τα clinical alarms δείχνει ότι τα συχνά false ή μη χρήσιμα alarms μπορούν να συνδέονται με φαινόμενα alarm fatigue και με πρόσθετο φόρτο για το προσωπικό [6], [13]. Για τον λόγο αυτό, τα false alarm cases χρησιμοποιούνται στην οντολογία για να εξεταστεί πώς μπορεί να αναπαρασταθεί πρόσθετη risk-related πληροφορία μέσω Risk Indicator, Risk-Relevant Event και Potential Harm. Η αναπαράσταση αυτή δεν αποδεικνύει πραγματική βλάβη σε συγκεκριμένο περιστατικό και δεν αποτελεί πλήρες risk management system κατά την έννοια του άρθρου 9 του AI Act. Αποτελεί περιορισμένο σημασιολογικό επίπεδο για την οργάνωση risk-related πληροφορίας στο πλαίσιο του proof-of-concept.

Η συνολική ερευνητική πορεία της εργασίας μπορεί να συνοψιστεί εννοιολογικά ως εξής, χωρίς η αλληλουχία αυτή να αντιστοιχεί σε μία μοναδική object-property διαδρομή της οντολογίας:

Ethical / Trustworthy AI principles → AI Act-related σημεία αναφοράς → Ontological modelling → Real healthcare monitoring events → Evidence records → Requirement mappings

Η αλληλουχία αυτή δείχνει ότι το μοντέλο δεν ξεκινά ούτε ολοκληρώνεται στο dataset. Το VTaC αποτελεί το πραγματικό πεδίο εφαρμογής και παρέχει τα alarm events και τις expert αποφάσεις που χρησιμοποιούνται στην υλοποίηση [4]. Η θεωρητική και κανονιστική βάση βρίσκεται στις αρχές του Trustworthy AI και στον AI Act, ενώ η οντολογία αποτελεί το σημασιολογικό μέσο με το οποίο τα πραγματικά περιστατικά οργανώνονται και συνδέονται με τεκμήρια και επιλεγμένα requirement mappings.



Σχήμα 3.1 Ερευνητική λογική της οντολογικής προσέγγισης.

Η συγκεκριμένη προσέγγιση καθοδηγεί τις σχεδιαστικές επιλογές που παρουσιάζονται στις επόμενες ενότητες. Αρχικά αποφασίζεται γιατί η εργασία ακολουθεί οντολογική προσέγγιση και όχι προσέγγιση μηχανικής μάθησης. Στη συνέχεια παρουσιάζεται αναλυτικά ο ρόλος του VTaC dataset, η επιλογή των πραγματικών alarm events, η επιλογή των AI Act-related σημείων αναφοράς και ο σχεδιασμός των βασικών συνδεδεμένων σημασιολογικών διαδρομών του μοντέλου.

3.3 Γιατί οντολογία και όχι μηχανική μάθηση

Ένα κρίσιμο σημείο για την κατανόηση της παρούσας εργασίας είναι η σαφής διάκριση ανάμεσα στη μηχανική μάθηση και στην οντολογική μοντελοποίηση. Η ύπαρξη ενός πραγματικού dataset, όπως το VTaC, θα μπορούσε αρχικά να δημιουργήσει την εντύπωση ότι ο βασικός σκοπός της εργασίας είναι η εκπαίδευση ή η αξιολόγηση ενός μοντέλου μηχανικής μάθησης. Ωστόσο, η παρούσα διπλωματική ακολουθεί διαφορετική ερευνητική κατεύθυνση.

Η μηχανική μάθηση επικεντρώνεται στην ανάπτυξη υπολογιστικών μοντέλων που βελτιώνουν την απόδοσή τους σε μια συγκεκριμένη εργασία μέσα από την αξιοποίηση δεδομένων και εμπειρίας [29]. Ανάλογα με το πρόβλημα, ένα μοντέλο μπορεί να χρησιμοποιηθεί για ταξινόμηση, πρόβλεψη ή αναγνώριση προτύπων. Σε ένα healthcare alarm use case, για παράδειγμα, μια προσέγγιση μηχανικής μάθησης θα μπορούσε να επιχειρήσει να προβλέψει αν ένα νέο alarm event είναι true ή false.

Η οντολογική μοντελοποίηση έχει διαφορετικό σκοπό. Μια οντολογία οργανώνει έννοιες, σχέσεις, περιορισμούς και συγκεκριμένα instances ενός πεδίου γνώσης με ρητό και δομημένο τρόπο [19], [20], [22]. Δεν εκπαιδεύεται πάνω σε δεδομένα με στόχο την προγνωστική γενίκευση. Αντίθετα, επιτρέπει την αναπαράσταση του τι σημαίνουν οι οντότητες ενός πεδίου και πώς συνδέονται μεταξύ τους.

Στο πλαίσιο της παρούσας εργασίας, το βασικό ερευνητικό ερώτημα δεν είναι αν μπορεί να προβλεφθεί αυτόματα αν ένα alarm είναι true ή false. Η πληροφορία αυτή είναι ήδη διαθέσιμη στο VTaC dataset μέσω των expert decisions [4]. Το ερευνητικό ενδιαφέρον βρίσκεται στο πώς ένα πραγματικό alarm event μπορεί να οργανωθεί σημασιολογικά και να συνδεθεί με το waveform record προέλευσής του, την expert αξιολόγηση, την τελική απόφαση, το evidence record και τα επιλεγμένα AI Act-related requirement individuals.

Η διαφορετική λογική των δύο προσεγγίσεων μπορεί να αποδοθεί συνοπτικά ως εξής:

Machine Learning:

Training data → Model → Prediction / Classification → Performance evaluation

Παρούσα οντολογική προσέγγιση:

Alert Event → Waveform Record

Alert Event → Expert Review → Review Decision

Alert Event → Evidence Record → Requirement Mapping

Στην πρώτη περίπτωση, το dataset αποτελεί κυρίως πηγή παραδειγμάτων από τα οποία ένα μοντέλο επιχειρεί να μάθει πρότυπα και να γενικεύσει σε νέες περιπτώσεις [29]. Στη δεύτερη περίπτωση, τα πραγματικά δεδομένα χρησιμοποιούνται ως instances ενός σημασιολογικού μοντέλου, στο οποίο διαφορετικές πληροφορίες οργανώνονται μέσω συνδεδεμένων αλλά διακριτών σχέσεων γύρω από το ίδιο event.

Η επιλογή της οντολογίας είναι επομένως άμεσα συνδεδεμένη με το πρόβλημα που εξετάζεται. Η εργασία δεν αντιμετωπίζει την ταξινόμηση των alarms ως ανοικτό πρόβλημα πρόβλεψης. Αντίθετα, εξετάζει ένα πρόβλημα αναπαράστασης γνώσης και τεκμηρίωσης: πώς μπορεί η πληροφορία γύρω από ένα alarm event να οργανωθεί με τρόπο που να υποστηρίζει αναζήτηση, ιχνηλασιμότητα και σύνδεση με evidence και AI Act-related σημεία αναφοράς.

Η διάκριση αυτή είναι ιδιαίτερα σημαντική και για τον ρόλο των expert decisions του VTaC. Η οντολογία δεν αποφασίζει αν ένα alarm είναι true ή false. Η τελική απόφαση προέρχεται από το dataset και εισάγεται στην οντολογία ως διαθέσιμη πληροφορία [4]. Επομένως, οι classes True Alarm και False Alarm, καθώς και τα αντίστοιχα Review Decision individuals, δεν αποτελούν αποτέλεσμα προγνωστικού αλγορίθμου της παρούσας εργασίας αλλά σημασιολογική αναπαράσταση των διαθέσιμων expert labels.

Η οντολογία επίσης δεν πρέπει να θεωρηθεί ανταγωνιστική προς τη μηχανική μάθηση. Οι δύο προσεγγίσεις μπορούν να λειτουργήσουν συμπληρωματικά. Σε ένα μελλοντικό πραγματικό σύστημα, ένα μοντέλο μηχανικής μάθησης θα μπορούσε να παράγει ή να υποστηρίζει την ταξινόμηση ενός alarm, ενώ μια οντολογική δομή θα μπορούσε να οργανώνει το παραγόμενο event, την αξιολόγηση, την απόφαση, τα τεκμήρια και τις σχετικές κανονιστικές έννοιες. Με αυτή τη λογική, η μηχανική μάθηση αφορά κυρίως το προγνωστικό μέρος, ενώ η οντολογία μπορεί να υποστηρίζει το σημασιολογικό και τεκμηριωτικό επίπεδο.

Η χρήση των 12 VTaC events πρέπει επίσης να ερμηνευθεί μέσα στο συγκεκριμένο μεθοδολογικό πλαίσιο. Αν η εργασία είχε ως στόχο την ανάπτυξη και αξιολόγηση μοντέλου μηχανικής μάθησης, ένα τόσο περιορισμένο και σκόπιμα επιλεγμένο σύνολο περιστατικών δεν θα αποτελούσε κατάλληλη βάση για ισχυρισμούς στατιστικής γενίκευσης σχετικά με την προγνωστική απόδοση ενός μοντέλου. Στην παρούσα εργασία, όμως, τα 12 events δεν χρησιμοποιούνται για εκπαίδευση ή προγνωστική

αξιολόγηση. Χρησιμοποιούνται ως proof-of-concept instances, δηλαδή ως συγκεκριμένες πραγματικές περιπτώσεις για να ελεγχθεί αν η οντολογική δομή μπορεί να αναπαραστήσει και να ανακτήσει τις βασικές σημασιολογικές διαδρομές που έχουν σχεδιαστεί.

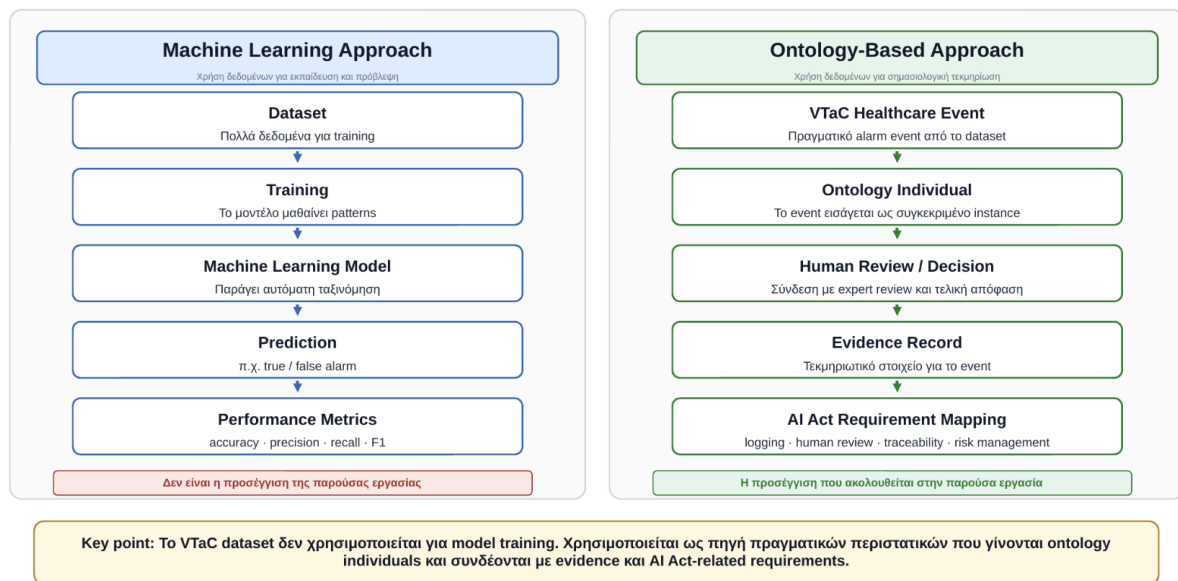
Η επιλογή των 12 events επιτρέπει να δοκιμαστούν διαφορετικές περιπτώσεις μέσα στο ίδιο μοντέλο. Περιλαμβάνονται true και false alarms, events από διαφορετικά waveform records και περιπτώσεις όπου ένα waveform record συνδέεται με περισσότερα από ένα events. Με αυτόν τον τρόπο, το proof-of-concept δεν αξιολογείται ως προς τη στατιστική αντιπροσωπευτικότητα του δείγματος, αλλά ως προς το αν μπορεί να αναπαραστήσει διαφορετικές μορφές σχέσεων και να υποστηρίξει τις σχετικές ερωτήσεις ελέγχου.

Η διαφορετική φύση της εργασίας επηρεάζει και τον τρόπο αξιολόγησης. Σε μια εφαρμογή μηχανικής μάθησης, η αξιολόγηση θα μπορούσε να βασίζεται σε μετρικές προγνωστικής απόδοσης, όπως accuracy, precision, recall ή άλλες κατάλληλες μετρικές [29]. Στην παρούσα εργασία δεν αξιολογείται προγνωστικό μοντέλο. Η λειτουργική επικύρωση της οντολογίας πραγματοποιείται μέσω ελέγχου λογικής συνέπειας με reasoner, DL competency questions και SPARQL queries. Τα εργαλεία αυτά χρησιμοποιούνται για να διαπιστωθεί αν οι βασικές σχέσεις της οντολογίας μπορούν να ανακτηθούν και αν το μοντέλο συμπεριφέρεται σύμφωνα με τον σχεδιασμό του.

Επομένως, η αξιολόγηση της παρούσας υλοποίησης αφορά τη σημασιολογική λειτουργικότητα και όχι την προγνωστική ακρίβεια. Εξετάζεται αν μπορούν να ανακτηθούν τα πραγματικά VTaC events, αν αυτά συνδέονται με τα σωστά records, reviews και decisions, αν τα evidence records συνδέονται με τα επιλεγμένα requirement individuals και αν τα false alarm cases μπορούν να αναπαρασταθούν στο lightweight risk-related επίπεδο.

Η διάκριση αυτή είναι απαραίτητη και για την αποφυγή υπερβολικών ισχυρισμών. Η παρούσα εργασία δεν ισχυρίζεται ότι αναπτύχθηκε μοντέλο τεχνητής νοημοσύνης για την αυτόματη ανίχνευση ή ταξινόμηση ventricular tachycardia alarms. Δεν αξιολογείται βελτίωση της κλινικής ακρίβειας και δεν επιχειρείται αυτόματη απόφαση συμμόρφωσης με τον AI Act. Αντίθετα, αναπτύσσεται και ελέγχεται ένα οντολογικό proof-of-concept για τη σημασιολογική οργάνωση πραγματικών healthcare monitoring events, την τεκμηρίωσή τους και τη σύνδεσή τους με επιλεγμένα AI Act-related σημεία αναφοράς.

Η επιλογή της οντολογίας συνδέεται τελικά με τον ίδιο τον σκοπό της διπλωματικής. Η μηχανική μάθηση θα αποτελούσε κατάλληλη προσέγγιση αν το κύριο ερευνητικό ερώτημα αφορούσε την αυτόματη πρόβλεψη ή ταξινόμηση των alarms. Η οντολογική προσέγγιση είναι κατάλληλη για την παρούσα εργασία επειδή το ερευνητικό ερώτημα αφορά την αναπαράσταση γνώσης, την τεκμηρίωση, την ιχνηλασιμότητα, την αναζήτηση και τη σύνδεση πραγματικών περιστατικών με evidence και επιλεγμένες κανονιστικές έννοιες.



Σχήμα 3.2 Διάκριση μεταξύ μηχανικής μάθησης και οντολογικής μοντελοποίησης στην παρούσα εργασία.

3.4 Το VTaC dataset και ο ρόλος του στην εργασία

Το VTaC dataset, με πλήρη τίτλο “VTaC: A Benchmark Dataset of Ventricular Tachycardia Alarms from ICU Monitors”, αποτελεί δημόσια διαθέσιμο dataset στο PhysioNet και περιλαμβάνει δεδομένα συναγεμίων κοιλιακής ταχυκαρδίας από monitors μονάδων εντατικής θεραπείας [4]. Η επιλογή του συγκεκριμένου dataset είναι ιδιαίτερα σημαντική για την παρούσα εργασία, επειδή παρέχει πραγματικά healthcare monitoring alarm events και όχι υποθετικά ή τεχνητά παραδείγματα.

Το VTaC dataset δεν αξιοποιείται στην εργασία ως σύνολο εκπαίδευσης για μοντέλο μηχανικής μάθησης. Δεν χρησιμοποιείται για να εκπαιδευτεί αλγόριθμος, να παραχθεί πρόβλεψη ή να υπολογιστούν μετρικές απόδοσης. Αντίθετα, χρησιμοποιείται ως πηγή πραγματικών περιστατικών, τα οποία μπορούν να εισαχθούν στην οντολογία ως individuals και να συνδεθούν με σχέσεις που υποστηρίζουν ιχνηλασιμότητα, τεκμηρίωση και requirement-to-evidence mapping.

Η επιλογή του VTaC είναι κατάλληλη για την παρούσα υλοποίηση, επειδή το dataset δεν περιλαμβάνει μόνο τεχνικά σήματα ή αφηρημένες εγγραφές. Περιλαμβάνει πραγματικά alarm events από περιβάλλον ICU, μαζί με τελική αξιολόγηση για το αν κάθε alarm ήταν αληθές ή ψευδές. Αυτό επιτρέπει τη σύνδεση ενός τεχνικού συμβάντος με διαδικασία αξιολόγησης και τεκμηρίωσης, κάτι που είναι απαραίτητο για το οντολογικό μοντέλο.

Το πρόβλημα που αναδεικνύει το dataset είναι ιδιαίτερα σημαντικό στον χώρο της υγείας. Τα false arrhythmia alarms αποτελούν σοβαρό ζήτημα σε μονάδες εντατικής θεραπείας, καθώς μπορούν να συμβάλουν σε alarm fatigue, να αυξήσουν το γνωστικό φορτίο των κλινικών και να επηρεάσουν την ασφάλεια των ασθενών. Το VTaC εστιάζει ειδικά στους συναγεμμούς κοιλιακής ταχυκαρδίας, οι οποίοι αποτελούν κρίσιμη κατηγορία alarm events σε συστήματα παρακολούθησης ασθενών.

Σε επίπεδο περιεχομένου, το VTaC περιλαμβάνει περισσότερα από 5.000 annotated ventricular tachycardia alarm events. Σύμφωνα με την περιγραφή του dataset, η τελική έκδοση περιέχει 5.037 annotated VT alarm events, από τα οποία 1.441, δηλαδή 28,61%, είναι true alarms. Τα alarms

δημιουργήθηκαν αυτόματα από commercial bedside monitors σε ICUs τριών μεγάλων νοσοκομείων στις Ηνωμένες Πολιτείες.

Η αρχική διαδικασία συλλογής ήταν ακόμη μεγαλύτερης κλίμακας. Οι δημιουργοί του dataset αναφέρουν ότι εξήγαγαν 18.465 VT alarm waveform events από 2.376 μοναδικά patient waveform records. Για λόγους ισορροπημένης δειγματοληψίας και αποφυγής υπερεκπροσώπησης συγκεκριμένων records, επιλέχθηκαν έως πέντε alarm events από κάθε patient record, οδηγώντας σε 5.742 events για annotation.

Κάθε waveform recording σχετίζεται με το χρονικό σημείο εμφάνισης ενός VT alarm. Στην τρέχουσα έκδοση, κάθε record περιλαμβάνει ένα εξάλεπτο τμήμα σήματος, το οποίο καλύπτει πέντε λεπτά πριν από την έναρξη του alarm και ένα λεπτό μετά. Το alarm onset βρίσκεται στο τέλος του πέμπτου λεπτού. Αυτή η πληροφορία είναι σημαντική, επειδή δείχνει ότι κάθε alarm event δεν είναι απλώς μία ετικέτα, αλλά συνδέεται με συγκεκριμένο χρονικό παράθυρο φυσιολογικών σημάτων.

Η διαδικασία annotation αποτελεί βασικό στοιχείο της αξιοπιστίας του dataset. Κάθε VT alarm event αξιολογήθηκε ανεξάρτητα από τουλάχιστον δύο ανθρώπινους expert annotators. Οι διαθέσιμες επιλογές αξιολόγησης ήταν True, όταν ο annotator θεωρούσε ότι το alarm ήταν σωστό, False, όταν θεωρούσε ότι το alarm ήταν λανθασμένο, Uncertain, όταν δεν μπορούσε να αποφασίσει με βεβαιότητα, και Reject, όταν το σήμα ήταν μη αναγνώσιμο λόγω θορύβου, artifacts ή άλλου προβλήματος.

Όταν οι δύο annotators διαφωνούσαν, εφαρμόστηκε διαδικασία adjudication. Οι διαφωνίες επιλύθηκαν είτε μέσω άμεσης συζήτησης μεταξύ των annotators είτε μέσω ψήφου adjudicator. Η διαδικασία αυτή έχει σημασία για την παρούσα εργασία, επειδή δείχνει ότι η τελική απόφαση true ή false δεν αποτελεί αυθαίρετη τιμή, αλλά προέρχεται από ανθρώπινη διαδικασία αξιολόγησης.

Από τα 5.742 events που αξιολογήθηκαν, σε 4.534 events, δηλαδή 78,96%, οι δύο independent annotators συμφώνησαν. Σε 1.208 events, δηλαδή 21,04%, υπήρξε σύγκρουση μεταξύ των δύο αρχικών αποφάσεων. Από αυτά, 816 adjudicated events επιλύθηκαν, ενώ 392 non-adjudicated events αφαιρέθηκαν. Μετά την αφαίρεση των περιπτώσεων Rejected και Uncertain, η τελική έκδοση περιλαμβάνει 5.037 events.

Η δομή των αρχείων του VTaC είναι επίσης σημαντική για την οντολογική αξιοποίηση. Τα waveform files βρίσκονται στον φάκελο waveforms και είναι αποθηκευμένα σε WFDB format, το οποίο χρησιμοποιείται συχνά στο PhysioNet για waveform recordings. Τα δεδομένα οργανώνονται με βάση patient waveform records. Κάθε sub-folder αντιστοιχεί σε ένα patient waveform record και μπορεί να περιέχει περισσότερα από ένα waveform records, καθένα από τα οποία αντιστοιχεί σε ένα VT alarm event.

Οι τελικές αποφάσεις των annotators αποθηκεύονται στο αρχείο event_labels.csv. Κάθε γραμμή του αρχείου αντιστοιχεί σε ένα VT alarm. Η στήλη record περιέχει το όνομα του patient waveform record, η στήλη event περιέχει το μοναδικό όνομα του alarm event και η στήλη decision περιέχει την τελική ανθρώπινη απόφαση για το alarm. Τα ονόματα των records και των events είναι surrogate names, ώστε να προστατεύεται η ταυτότητα των ασθενών.

Για τη συγκεκριμένη οντολογική υλοποίηση, τα σημαντικότερα πεδία του dataset είναι τα εξής:

record: χρησιμοποιείται για τη δημιουργία ή αναφορά σε individual της κλάσης Waveform Record.

event: χρησιμοποιείται για τη δημιουργία individual της κλάσης Alert Event και ειδικότερα της κλάσης Ventricular Tachycardia Alarm.

decision: χρησιμοποιείται για τη σύνδεση του event με αντίστοιχη Review Decision, δηλαδή με απόφαση τύπου True Alarm ή False Alarm.

Με αυτόν τον τρόπο, το event_labels.csv λειτουργεί ως βασικό σημείο μετάβασης από το dataset στην οντολογία. Δεν αξιοποιείται ολόκληρο το waveform signal στο Protégé. Αντίθετα, αξιοποιούνται τα αναγνωριστικά και οι τελικές αποφάσεις που επιτρέπουν τη δημιουργία δομημένων individuals και object property assertions.

Στο πλαίσιο της παρούσας εργασίας, το VTaC dataset δεν χρησιμοποιείται εξαντλητικά. Από το συνολικό dataset επιλέχθηκαν 12 πραγματικά alarm events για την υλοποίηση του proof-of-concept. Η επιλογή περιλαμβάνει 6 true alarms και 6 false alarms, ώστε να υπάρχουν παραδείγματα και για τις δύο βασικές κατηγορίες απόφασης. Η ισορροπία αυτή δεν στοχεύει σε στατιστική αντιπροσώπευση του dataset, αλλά σε λειτουργικό έλεγχο της οντολογικής δομής.

Κάθε επιλεγμένο alarm event μετατράπηκε σε individual στο Protégé. Το αντίστοιχο record εισήχθη ως individual της κλάσης Waveform Record, ενώ το alarm εισήχθη ως individual της κλάσης Ventricular Tachycardia Alarm. Η τελική απόφαση του dataset αξιοποιήθηκε για τη δημιουργία ή σύνδεση με individual της κλάσης Review Decision. Έπειτα προστέθηκε Evidence Record, ώστε το event να μπορεί να συσχετιστεί με επιλεγμένες AI Act-related requirements. Σύμφωνα με τον οδηγό υλοποίησης, από τα 5.037 annotated VT alarm events επιλέχθηκαν 12 πραγματικά alarm events, 6 true και 6 false, ενώ τα πεδία record, event και decision αξιοποιήθηκαν για τη δημιουργία individuals και σχέσεων στο Protégé.

Η χαρτογράφηση από το dataset στην οντολογία ακολουθεί την παρακάτω λογική:

record → Waveform Record individual

event → Alert Event / Ventricular Tachycardia Alarm individual

expert αξιολόγηση → Expert Review individual

decision → Review Decision individual

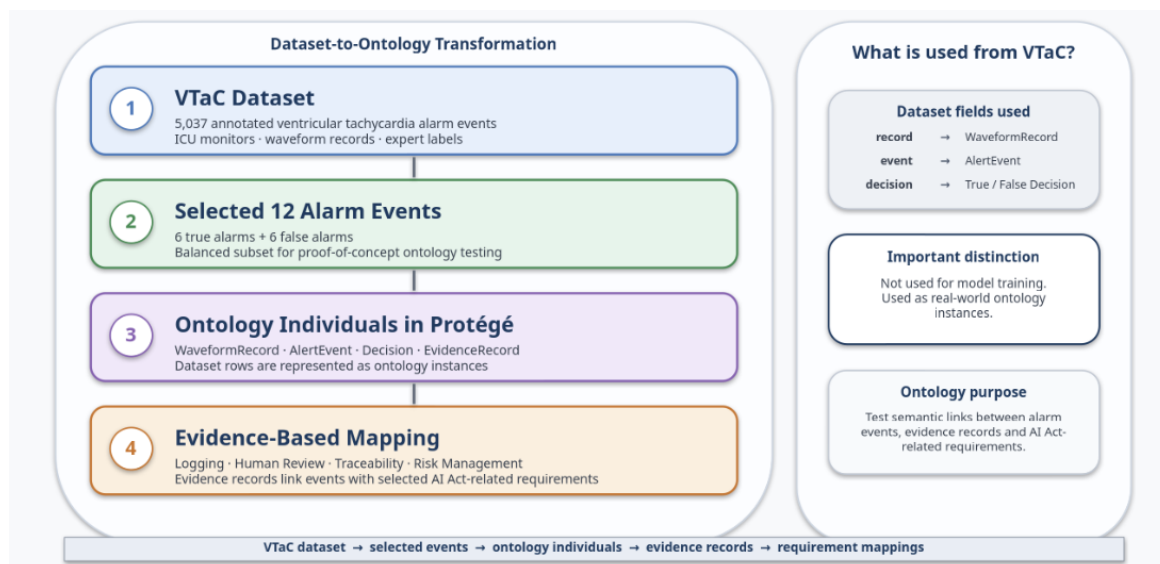
event → Evidence Record μέσω του isDocumentedBy

Evidence Record → επιλεγμένα AI Act-related σημεία αναφοράς μέσω του supportsRequirement

Η παραπάνω μετατροπή αποτελεί κρίσιμο σημείο της μεθοδολογίας. Το dataset παρέχει τις πραγματικές εγγραφές, ενώ η οντολογία προσθέτει το σημασιολογικό πλαίσιο που επιτρέπει τη σύνδεση κάθε event με το waveform record προέλευσής του, την expert αξιολόγηση, την τελική απόφαση και το αντίστοιχο Evidence Record. Οι πληροφορίες αυτές οργανώνονται μέσω διακριτών αλλά συνδεδεμένων σημασιολογικών σχέσεων γύρω από το ίδιο Alert Event και δεν αποτελούν μία μοναδική γραμμική object-property αλυσίδα.

Η ύπαρξη true και false alarms επέτρεψε τη διαφοροποίηση της οντολογικής αναπαράστασης. Και στις δύο περιπτώσεις διατηρούνται οι βασικές σχέσεις προέλευσης, αξιολόγησης, απόφασης και τεκμηρίωσης. Στα false alarm cases προστίθεται επιπλέον το lightweight risk-related επίπεδο, ώστε να εξεταστεί η δυνατότητα αναπαράστασης πρόσθετης risk-related πληροφορίας, όπως Risk Indicator, Risk-Relevant Event και Potential Harm.

Η χρήση του VTaC έχει και έναν πρακτικό περιορισμό. Το dataset δεν συνοδεύεται από λεπτομερή clinical information για κάθε waveform recording. Αυτό σημαίνει ότι η οντολογία δεν μπορεί να υποστηρίξει πλήρη κλινική ερμηνεία κάθε περιστατικού. Η αξιοποίησή του περιορίζεται στα alarm events, στα waveform record identifiers και στις τελικές αποφάσεις των annotators. Ο περιορισμός αυτός είναι συμβατός με τον σκοπό της εργασίας, επειδή η εργασία δεν επιδιώκει κλινική διάγνωση, αλλά οντολογική τεκμηρίωση και σημασιολογική οργάνωση.



Σχήμα 3.3 Ρόλος του VTaC dataset στην οντολογική υλοποίηση.

Το VTaC είναι επίσης κατάλληλο από πλευράς δεοντολογίας και χρήσης, επειδή τα δεδομένα έχουν υποστεί de-identification. Τα ονόματα των records και των events είναι surrogate identifiers και δεν παραπέμπουν άμεσα σε ταυτότητα ασθενούς. Αυτό επιτρέπει τη χρήση του dataset σε ερευνητικό πλαίσιο, χωρίς η παρούσα εργασία να διαχειρίζεται αναγνωρίσιμα προσωπικά δεδομένα ασθενών.

Συνολικά, το VTaC προσφέρει τρία στοιχεία που είναι απαραίτητα για την οντολογική υλοποίηση. Πρώτον, παρέχει πραγματικά ICU alarm events. Δεύτερον, προσφέρει τελικές ανθρώπινες αποφάσεις true ή false. Τρίτον, οργανώνει τα δεδομένα με record και event identifiers, τα οποία μπορούν να μετατραπούν σε individuals και object property assertions στο Protégé. Με βάση αυτά τα στοιχεία, το dataset λειτουργεί ως πραγματικό υπόβαθρο για την εφαρμογή και τον έλεγχο του μοντέλου.

3.5 Επιλογή των VTaC events για το proof-of-concept

Για την υλοποίηση της οντολογίας δεν χρησιμοποιήθηκε το σύνολο των VTaC alarm events. Αντίθετα, επιλέχθηκε ένα περιορισμένο σύνολο 12 πραγματικών ventricular tachycardia alarm events από το VTaC dataset [4], με σκοπό τη λειτουργική δοκιμή της προτεινόμενης οντολογικής δομής στο πλαίσιο ενός proof-of-concept.

Η επιλογή των 12 events δεν αποσκοπεί σε στατιστική αντιπροσώπευση του συνολικού VTaC dataset και δεν χρησιμοποιείται για την εξαγωγή συμπερασμάτων σχετικά με τη συνολική συχνότητα ή κατανομή των true και false alarms. Ο αριθμός και η σύνθεση των events επιλέχθηκαν με βάση τις ανάγκες λειτουργικού ελέγχου της οντολογίας. Στόχος ήταν να υπάρχουν πραγματικές περιπτώσεις που να επιτρέπουν τον έλεγχο της βασικής σημασιολογικής αλυσίδας, της διάκρισης true και false alarm cases, των σχέσεων μεταξύ waveform records και events και της πρόσθετης risk-related αναπαράστασης των false alarms.

Το επιλεγμένο σύνολο περιλαμβάνει 6 events με τελική απόφαση true alarm και 6 events με τελική απόφαση false alarm. Η ισόρροπη αυτή κατανομή αποτελεί σκόπιμη σχεδιαστική επιλογή του proof-of-concept και δεν αντανάκλα την πραγματική αναλογία true και false alarms στο συνολικό VTaC dataset. Οι τελικές αποφάσεις true ή false δεν δημιουργήθηκαν από την οντολογία, αλλά προέρχονται από τα διαθέσιμα expert labels του VTaC [4].

Η παρουσία και των δύο κατηγοριών είναι σημαντική για τη λειτουργική επικύρωση του μοντέλου. Τα true και false alarm cases επιτρέπουν τον έλεγχο των βασικών συνδεδεμένων σημασιολογικών διαδρομών γύρω από το ίδιο Alert Event:

Alert Event → belongsToRecord → Waveform Record

Alert Event → hasReview → Expert Review → hasDecision → Review Decision

Alert Event → isDocumentedBy → Evidence Record → supportsRequirement → Compliance Requirement

Τα false alarm cases χρησιμοποιούνται επιπλέον για τη δοκιμή του lightweight risk-related επιπέδου μέσω δύο παράλληλων διαδρομών:

Alert Event → hasRiskIndicator → Risk Indicator

Alert Event → isClassifiedAs → Risk-Relevant Event → mayLeadTo → Potential Harm

Με αυτόν τον τρόπο, τα επιλεγμένα instances επιτρέπουν τον έλεγχο τόσο της βασικής σημασιολογικής δομής όσο και της πρόσθετης risk-related αναπαράστασης, χωρίς να υπονοείται μία μοναδική γραμμική object-property αλυσίδα.

Τα 12 events προέρχονται από έξι διαφορετικά waveform records. Η επιλογή περιλαμβάνει τόσο περιπτώσεις όπου ένα waveform record συνδέεται με ένα μόνο επιλεγμένο event όσο και περιπτώσεις όπου το ίδιο waveform record συνδέεται με περισσότερα events. Η διαφοροποίηση αυτή επιτρέπει να εξεταστεί αν η οντολογία μπορεί να αναπαραστήσει σωστά σχέσεις ενός record με ένα ή περισσότερα alarm events.

Η κατανομή αυτή είναι χρήσιμη για τον έλεγχο της ιχνηλασιμότητας, επειδή επιτρέπει τα ερωτήματα της οντολογίας να λειτουργούν και προς τις δύο κατευθύνσεις: από ένα συγκεκριμένο alarm event προς το waveform record προέλευσής του και από ένα waveform record προς τα events που συνδέονται με αυτό.

Πίνακας 3-1 Waveform record individuals και πλήθος επιλεγμένων VTaC events.

Waveform record individual	Πλήθος επιλεγμένων events
vtac_record_003c13	2
vtac_record_004bad	4
vtac_record_0060de	1
vtac_record_015f23	1
vtac_record_0179fd	3
vtac_record_01a952	1

Πίνακας 3-2 Επιλεγμένα VTaC alarm events για την proof-of-concept υλοποίηση.

Κατηγορία απόφασης	VTaC alert event individual	Waveform record
False alarm	vtac_event_003c13_0115	vtac_record_003c13
False alarm	vtac_event_003c13_0126	vtac_record_003c13
False alarm	vtac_event_004bad_0015	vtac_record_004bad
False alarm	vtac_event_004bad_1115	vtac_record_004bad
False alarm	vtac_event_004bad_1426	vtac_record_004bad
False alarm	vtac_event_004bad_2348	vtac_record_004bad
True alarm	vtac_event_0060de_0560	vtac_record_0060de
True alarm	vtac_event_015f23_0512	vtac_record_015f23
True alarm	vtac_event_0179fd_1194	vtac_record_0179fd
True alarm	vtac_event_0179fd_1265	vtac_record_0179fd
True alarm	vtac_event_0179fd_2773	vtac_record_0179fd
True alarm	vtac_event_01a952_1067	vtac_record_01a952

Ο Πίνακας 3-2 παρουσιάζει τα 12 πραγματικά VTaC alarm events που χρησιμοποιήθηκαν ως individuals στην οντολογία. Τα συγκεκριμένα events αναπαρίστανται ως instances της class Ventricular Tachycardia Alarm, η οποία αποτελεί ειδικότερη class της Alert Event. Για κάθε event διατηρείται η σύνδεση με το αντίστοιχο waveform record και με την τελική expert απόφαση του dataset [4].

Για κάθε επιλεγμένο event δημιουργήθηκε επίσης αντίστοιχη σημασιολογική διαδρομή που περιλαμβάνει Expert Review, Review Decision και Evidence Record. Η επιλογή αυτή αποτελεί μέρος του σχεδιασμού της παρούσας οντολογίας και επιτρέπει η τελική ετικέτα true ή false να μην αντιμετωπίζεται απλώς ως μεμονωμένη τιμή, αλλά ως μέρος μιας οργανωμένης αλυσίδας τεκμηρίωσης.

Ιδιαίτερη σημασία έχουν τα έξι false alarm cases, καθώς χρησιμοποιούνται επιπλέον για τη δοκιμή του lightweight risk-related επιπέδου. Στην οντολογία τα συγκεκριμένα events μπορούν να συνδεθούν με risk indicators, risk-relevant events και το potential harm που σχετίζεται με το alert fatigue. Η σύνδεση αυτή δεν σημαίνει ότι αποδεικνύεται πραγματική βλάβη για κάποιο συγκεκριμένο VTaC περιστατικό. Αποτελεί σημασιολογική αναπαράσταση της πιθανής risk-related σημασίας των false alarms, βάσει της λογικής που έχει υιοθετηθεί στο proof-of-concept.

Συνολικά, το σύνολο των 12 VTaC events επιλέχθηκε ώστε να παρέχει επαρκή ποικιλία πραγματικών περιπτώσεων για τον λειτουργικό έλεγχο των βασικών μηχανισμών της οντολογίας. Η αξία του δεν βρίσκεται στο μέγεθος ή στη στατιστική αντιπροσωπευτικότητα του δείγματος, αλλά στη δυνατότητα να ελεγχθούν με πραγματικά instances οι βασικές συνδεδεμένες σημασιολογικές διαδρομές, οι σχέσεις record–event, η διάκριση true/false, το requirement-to-evidence mapping και το lightweight risk-related επίπεδο. Η λειτουργική αυτή αξιολόγηση πραγματοποιείται αναλυτικά στο Κεφάλαιο 5 μέσω reasoner, DL competency questions και SPARQL queries.

3.6 Επιλογή των AI Act-related σημείων αναφοράς

Η οντολογία που αναπτύχθηκε στην παρούσα εργασία δεν επιχειρεί να μοντελοποιήσει το σύνολο του AI Act. Μια πλήρης αναπαράσταση όλων των άρθρων, υποχρεώσεων, εμπλεκόμενων φορέων και διαδικασιών του Κανονισμού θα απαιτούσε σημαντικά ευρύτερη νομική και τεχνική ανάλυση, η οποία υπερβαίνει το πλαίσιο της συγκεκριμένης proof-of-concept υλοποίησης. Για τον λόγο αυτό, επιλέχθηκε ένα περιορισμένο σύνολο AI Act-related σημείων αναφοράς που παρουσιάζουν σαφή συνάφεια με τα διαθέσιμα στοιχεία του healthcare monitoring use case [3].

Βασικό κριτήριο επιλογής ήταν η δυνατότητα σύνδεσης κάθε σημείου αναφοράς με συγκεκριμένη πληροφορία ή τεκμήριο που μπορεί να αναπαρασταθεί στο οντολογικό μοντέλο. Η επιλογή αυτή ακολουθεί τη γενικότερη μεθοδολογική λογική σύμφωνα με την οποία κανονιστικές έννοιες και απαιτήσεις μπορούν να οργανωθούν σημασιολογικά και να γίνουν αναζητήσιμες και ιχνηλατήσιμες μέσα από ontology-based προσεγγίσεις [5]. Στην παρούσα εργασία, όμως, η χαρτογράφηση δεν αφορά ολόκληρο τον AI Act, αλλά συγκεκριμένα στοιχεία που μπορούν να συνδεθούν με τα διαθέσιμα VTaC events και τα evidence records του proof-of-concept.

Στην οντολογία χρησιμοποιούνται τέσσερα requirement individuals:

- **req_logging**
- **req_human_review**
- **req_traceability**
- **req_risk_management**

Οι ονομασίες αυτές αποτελούν εφαρμοσμένες έννοιες της οντολογίας και δεν πρέπει να θεωρούνται σε όλες τις περιπτώσεις αυτούσιες ονομασίες ανεξάρτητων νομικών απαιτήσεων του AI Act. Το req_logging συνδέεται κυρίως με τη λογική του record-keeping του άρθρου 12, το req_human_review με την ανθρώπινη αξιολόγηση ως εφαρμοσμένη διάσταση συναφής με το human oversight του άρθρου 14, το req_traceability με την ιχνηλασιμότητα που συνδέεται ιδιαίτερα με τις logging capabilities του άρθρου 12, ενώ το req_risk_management σχετίζεται με το ευρύτερο πλαίσιο διαχείρισης κινδύνου του άρθρου 9 [3].

Πίνακας 3-3 Επιλεγμένα AI Act-related σημεία αναφοράς της οντολογίας.

Requirement individual	Περιγραφή	Ρόλος στην οντολογία
req_logging	Σχετίζεται κυρίως με το record-keeping και την καταγραφή γεγονότων του άρθρου 12.	Συνδέει evidence records με διαθέσιμα στοιχεία καταγραφής και αναγνώρισης των alarm events.
req_human_review	Εφαρμοσμένη έννοια ανθρώπινης αξιολόγησης, συναφής με την ευρύτερη απαίτηση human oversight του άρθρου 14.	Συνδέεται με την αναπαράσταση των Expert Review και Review Decision.
req_traceability	Σχετίζεται με την ιχνηλασιμότητα, η οποία υποστηρίζεται ιδιαίτερα από τις απαιτήσεις record-keeping του άρθρου 12.	Συνδέεται με τη δυνατότητα ανάκτησης των σχέσεων προέλευσης, αξιολόγησης και τεκμηρίωσης που οργανώνονται γύρω από το alarm event.
req_risk_management	Σχετίζεται με το ευρύτερο risk management framework του άρθρου 9.	Χρησιμοποιείται στο lightweight risk-related επίπεδο για τα false alarm cases.

Το req_logging επιλέχθηκε επειδή το VTaC παρέχει αναγνωρίσιμα στοιχεία για τα waveform records και τα alarm events, τα οποία μπορούν να διατηρηθούν στην οντολογία μέσω των αντίστοιχων identifiers [4]. Τα στοιχεία αυτά επιτρέπουν την αναζήτηση και την οργανωμένη σύνδεση ενός συγκεκριμένου event με το record προέλευσής του. Στο proof-of-concept χρησιμοποιούνται επομένως ως διαθέσιμα στοιχεία τεκμηρίωσης που είναι συναφή με τη λογική της καταγραφής. Η ύπαρξή τους, ωστόσο, δεν ισοδυναμεί με επιβεβαίωση της πλήρους τεχνικής logging capability που απαιτεί το άρθρο 12 από ένα πραγματικό high-risk AI system [3].

Το req_human_review επιλέχθηκε επειδή η ανθρώπινη expert αξιολόγηση αποτελεί πραγματικό στοιχείο του VTaC dataset. Οι τελικές αποφάσεις true ή false προέρχονται από ανθρώπινη διαδικασία annotation και χρησιμοποιούνται στην οντολογία μέσω των εννοιών Expert Review και Review Decision [4]. Η επιλογή του συγκεκριμένου requirement individual επιτρέπει την οργάνωση αυτής της ανθρώπινης αξιολόγησης στο σημασιολογικό μοντέλο. Ωστόσο, το Expert Review του VTaC δεν αποτελεί πλήρη μηχανισμό human oversight ενός high-risk AI system κατά την έννοια του άρθρου 14. Για τον λόγο αυτό, το req_human_review αντιμετωπίζεται ως εφαρμοσμένη οντολογική έννοια συναφής με την ανθρώπινη διάσταση της εποπτείας και όχι ως απόδειξη συμμόρφωσης με το σύνολο των απαιτήσεων του άρθρου 14 [3].

Το req_traceability επιλέχθηκε επειδή η ιχνηλασιμότητα αποτελεί κεντρικό στόχο της προτεινόμενης οντολογικής δομής. Το άρθρο 12 συνδέει τις δυνατότητες logging με την εξασφάλιση κατάλληλου επιπέδου traceability της λειτουργίας ενός high-risk AI system [3]. Στην παρούσα εργασία, η έννοια αυτή αποτυπώνεται εφαρμοσμένα μέσω της δυνατότητας να ακολουθηθεί μια σαφής σημασιολογική διαδρομή από το waveform record προς το alarm event, το expert review, τη review decision και το evidence record. Το req_traceability χρησιμοποιείται επομένως ως requirement individual που επιτρέπει να καταγραφεί και να αναζητηθεί αυτή η διάσταση μέσα στο proof-of-concept.

Το req_risk_management επιλέχθηκε για να υποστηρίξει την πρόσθετη αναπαράσταση των false alarm cases. Ο AI Act προβλέπει στο άρθρο 9 ένα ευρύτερο risk management system για high-risk AI systems [3]. Στην παρούσα οντολογία δεν επιχειρείται η αναπαράσταση ή αξιολόγηση ολόκληρης αυτής της διαδικασίας. Αντίθετα, υλοποιείται ένα lightweight risk-related επίπεδο στο οποίο επιλεγμένα false alarm cases μπορούν να συνδεθούν με Risk Indicator, Risk-Relevant Event και Potential Harm. Η συγκεκριμένη επιλογή βασίζεται επίσης στη βιβλιογραφική τεκμηρίωση ότι τα συχνά false ή μη χρήσιμα alarms μπορούν να σχετίζονται με το φαινόμενο alarm fatigue [6], [13].

Η σύνδεση των τεσσάρων requirement individuals με τα πραγματικά events δεν πραγματοποιείται απευθείας. Κεντρική σχεδιαστική επιλογή της οντολογίας είναι η χρήση του Evidence Record ως ενδιάμεσου κόμβου:

Alert Event → Evidence Record → Requirement individual

Με αυτή τη δομή, ένα requirement individual δεν συνδέεται απευθείας με ένα event. Το event συνδέεται πρώτα με ένα evidence record που οργανώνει τη σχετική πληροφορία και στη συνέχεια το evidence record συνδέεται μέσω του supportsRequirement με ένα ή περισσότερα requirement individuals. Η επιλογή αυτή αποτελεί τη βάση του requirement-to-evidence mapping της παρούσας εργασίας.

Στην υλοποίηση, τα evidence records των true alarm cases μπορούν να υποστηρίζουν τα req_logging, req_human_review και req_traceability. Τα false alarm cases ακολουθούν την ίδια βασική λογική τεκμηρίωσης και, επιπλέον, συμμετέχουν στο lightweight risk-related επίπεδο, επιτρέποντας τη σύνδεση των αντίστοιχων evidence records και με το req_risk_management. Η διαφοροποίηση αυτή

αποτελεί σχεδιαστική επιλογή του proof-of-concept και δεν πρέπει να ερμηνεύεται ως κανονιστικός κανόνας σύμφωνα με τον οποίο κάθε false alarm συνεπάγεται αυτομάτως απαίτηση risk management.

Είναι σημαντικό να επαναληφθεί ότι το VTaC dataset και τα alarm events που χρησιμοποιούνται στην εργασία δεν αποτελούν από μόνα τους σύστημα τεχνητής νοημοσύνης ούτε ταξινομούνται στο πλαίσιο της παρούσας μελέτης ως high-risk AI system. Το dataset χρησιμοποιείται ως πηγή πραγματικών healthcare monitoring events και expert decisions πάνω στα οποία δοκιμάζεται η προτεινόμενη οντολογική δομή [4]. Ο AI Act λειτουργεί επομένως ως κανονιστικό πλαίσιο αναφοράς για την επιλογή και τη σημασιολογική οργάνωση σχετικών σημείων τεκμηρίωσης και όχι ως βάση για δήλωση νομικής συμμόρφωσης του VTaC dataset.

Συνολικά, τα τέσσερα requirement individuals επιλέχθηκαν επειδή επιτρέπουν να δοκιμαστεί με ελεγχόμενο τρόπο η βασική λογική της εργασίας: η μετάβαση από πραγματικό healthcare monitoring event σε οργανωμένο evidence και από εκεί σε επιλεγμένο AI Act-related σημείο αναφοράς. Με αυτόν τον τρόπο, το proof-of-concept παραμένει περιορισμένο ως προς το κανονιστικό του εύρος, αλλά αρκετά εκφραστικό ώστε να υποστηρίξει το requirement-to-evidence mapping που αποτελεί κεντρικό στοιχείο της οντολογικής προσέγγισης.

3.7 Βασικές σημασιολογικές διαδρομές της οντολογίας

Η σημασιολογική δομή της οντολογίας δεν οργανώνεται ως μία μοναδική γραμμική ακολουθία. Αντίθετα, διαφορετικές αλλά συνδεδεμένες διαδρομές οργανώνονται γύρω από το ίδιο Alert Event και καλύπτουν την προέλευση του συμβάντος, την ανθρώπινη αξιολόγηση και την τεκμηρίωσή του. Οι τρεις βασικές διαδρομές είναι:

Alert Event --belongsToRecord→ Waveform Record

Alert Event --hasReview→ Expert Review --hasDecision→ Review Decision

Alert Event --isDocumentedBy→ Evidence Record --supportsRequirement→ Requirement individual

Η διάκριση αυτή αποτυπώνει την πραγματική δομή της υλοποίησης και αποφεύγει την εντύπωση ότι το Waveform Record, η Review Decision και το Evidence Record συνδέονται διαδοχικά μεταξύ τους μέσω μίας ενιαίας object-property αλυσίδας.

Η σύνδεση του waveform record με το alarm event αναπαρίσταται με δύο συμπληρωματικά object properties. Από την πλευρά του record χρησιμοποιείται το hasAlertEvent, ενώ από την πλευρά του event χρησιμοποιείται το belongsToRecord. Με τον τρόπο αυτό, μπορεί να αναζητηθεί τόσο το waveform record από το οποίο προέρχεται ένα event όσο και το σύνολο των επιλεγμένων events που συνδέονται με ένα συγκεκριμένο record.

Η σχέση μεταξύ του Alert Event και της ανθρώπινης αξιολόγησης αποτυπώνεται μέσω του object property hasReview. Κάθε επιλεγμένο VTaC alarm event συνδέεται με αντίστοιχο Expert Review individual. Η επιλογή αυτή επιτρέπει η διαδικασία αξιολόγησης να αναπαρίσταται ως ξεχωριστό σημασιολογικό στοιχείο και όχι απλώς ως μια τιμή αποθηκευμένη πάνω στο event.

Το Expert Review συνδέεται στη συνέχεια με το αντίστοιχο Review Decision μέσω του object property hasDecision. Η τελική απόφαση αναπαρίσταται ως ξεχωριστό individual και ταξινομείται στην κατάλληλη κατηγορία True Alarm ή False Alarm. Με αυτόν τον τρόπο, η απόφαση αποκτά ανεξάρτητη

θέση στο μοντέλο και παραμένει διακριτή τόσο από το ίδιο το alarm event όσο και από τη διαδικασία αξιολόγησης.

Παράλληλα με τη διαδρομή αξιολόγησης, το Alert Event συνδέεται με το αντίστοιχο Evidence Record μέσω του object property `isDocumentedBy`. Το Evidence Record λειτουργεί ως τεκμηριωτικός κόμβος που οργανώνει τη διαθέσιμη πληροφορία για το συγκεκριμένο event και επιτρέπει τη μετάβαση από το operational επίπεδο προς το επίπεδο των επιλεγμένων AI Act-related σημείων αναφοράς.

Το Evidence Record συνδέεται με ένα ή περισσότερα requirement individuals μέσω του object property `supportsRequirement`. Στο σημείο αυτό υλοποιείται το requirement-to-evidence mapping της εργασίας. Το mapping δεν δηλώνει ότι το event συμμορφώνεται νομικά με συγκεκριμένη απαίτηση. Δηλώνει ότι το συγκεκριμένο evidence record έχει μοντελοποιηθεί ώστε να υποστηρίζει τεκμηριωτικά ένα επιλεγμένο AI Act-related σημείο αναφοράς μέσα στο proof-of-concept.

Σε επίπεδο object properties, οι αντίστοιχες σχέσεις αποτυπώνονται ως εξής:

Waveform Record --hasAlertEvent→ Alert Event

και αντίστροφα:

Alert Event --belongsToRecord→ Waveform Record

Για την ανθρώπινη αξιολόγηση:

Alert Event --hasReview→ Expert Review --hasDecision→ Review Decision

Για την τεκμηρίωση και το requirement mapping:

Alert Event --isDocumentedBy→ Evidence Record --supportsRequirement→ Requirement individual

Η διάκριση αυτή είναι σημαντική, επειδή αποτυπώνει με ακρίβεια την πραγματική υλοποίηση της οντολογίας. Η ιχνηλασιμότητα προκύπτει από τη συνδυασμένη αξιοποίηση των παραπάνω σχέσεων γύρω από το ίδιο Alert Event και όχι από μία μοναδική γραμμική σειρά άμεσων object-property συνδέσεων.

Πίνακας 3-4 Ρόλος των στοιχείων στις βασικές σημασιολογικές διαδρομές.

Στοιχείο του μοντέλου	Ρόλος στο μοντέλο
Waveform Record	Αναπαριστά το record προέλευσης του alarm event.
Alert Event	Αποτελεί το κεντρικό πραγματικό περιστατικό γύρω από το οποίο οργανώνονται οι υπόλοιπες σχέσεις.
Expert Review	Αναπαριστά την ανθρώπινη expert αξιολόγηση του event.
Review Decision	Αναπαριστά την τελική απόφαση true ή false που προκύπτει από την expert αξιολόγηση.
Evidence Record	Λειτουργεί ως τεκμηριωτικός κόμβος για το event και ως γέφυρα προς το requirement mapping.
Compliance Requirement	Αναπαριστά ένα επιλεγμένο AI Act-related σημείο αναφοράς που υποστηρίζεται τεκμηριωτικά από το evidence record.

Η συγκεκριμένη δομή υποστηρίζει τη βασική ιχνηλασιμότητα της οντολογίας. Για ένα δεδομένο VTaC alarm event μπορεί να εντοπιστεί το waveform record προέλευσής του, το Expert Review που συνδέεται με αυτό, η αντίστοιχη Review Decision και το Evidence Record που το τεκμηριώνει. Από το Evidence Record μπορεί στη συνέχεια να ανακτηθεί το σύνολο των requirement individuals προς τα οποία έχει πραγματοποιηθεί το αντίστοιχο evidence mapping.

Για τα true alarm cases, η παραπάνω δομή χρησιμοποιείται για τον έλεγχο της βασικής διαδρομής αξιολόγησης και τεκμηρίωσης. Για τα false alarm cases, η ίδια δομή διατηρείται και επεκτείνεται επιπλέον με το lightweight risk-related επίπεδο που παρουσιάζεται στην επόμενη ενότητα.

Η βασική σημασιολογική δομή έχει άμεση σχέση με τη λειτουργική επικύρωση που πραγματοποιείται στο Κεφάλαιο 5. Τα DL competency questions χρησιμοποιούνται για να ελεγχθούν επιμέρους σημασιολογικές συνθήκες και σύνολα individuals, ενώ τα SPARQL queries επιτρέπουν τη συγκεντρωτική ανάκτηση και παρουσίαση συνδυασμένων σχέσεων, όπως event–record–decision–evidence και event–evidence–requirement. Με τον τρόπο αυτό ελέγχεται αν οι σχέσεις που σχεδιάστηκαν στη μεθοδολογία έχουν αποτυπωθεί και μπορούν να ανακτηθούν από την υλοποιημένη οντολογία.

Συνολικά, η βασική σημασιολογική δομή λειτουργεί ως γέφυρα ανάμεσα στα πραγματικά δεδομένα του VTaC, στην ανθρώπινη αξιολόγηση, στην τεκμηρίωση και στα επιλεγμένα AI Act-related σημεία αναφοράς. Η αξία της βρίσκεται στο ότι ένα alarm event δεν παραμένει απομονωμένη εγγραφή, αλλά αποκτά σαφές πλαίσιο προέλευσης, αξιολόγησης, απόφασης και τεκμηρίωσης μέσα σε ένα ενιαίο και αναζητήσιμο μοντέλο γνώσης.

3.8 Lightweight risk-related επίπεδο για false alarm cases

Πέρα από τη βασική σημασιολογική δομή τεκμηρίωσης, η οντολογία περιλαμβάνει ένα πρόσθετο lightweight risk-related επίπεδο για τα false alarm cases. Η επιλογή αυτή βασίζεται στη διαφορετική σημασία που μπορούν να αποκτήσουν τα συχνά false ή μη χρήσιμα alarms σε ένα healthcare monitoring περιβάλλον, καθώς στη σχετική βιβλιογραφία συνδέονται με προβλήματα όπως το alarm fatigue και ο αυξημένος φόρτος για το κλινικό προσωπικό [6], [13].

Η ενσωμάτωση του risk-related επιπέδου δεν σημαίνει ότι κάθε false alarm προκαλεί πραγματική βλάβη ούτε ότι πραγματοποιείται πλήρης αξιολόγηση κινδύνου για κάθε περιστατικό. Η οντολογία δεν διαθέτει τα κλινικά και οργανωτικά δεδομένα που θα απαιτούνταν για έναν τέτοιο ισχυρισμό. Αντίθετα, το συγκεκριμένο επίπεδο λειτουργεί ως περιορισμένη σημασιολογική επέκταση του proof-of-concept, ώστε τα false alarm cases να μπορούν να οργανωθούν ως περιστατικά με πρόσθετη risk-related σημασία.

Στην παρούσα υλοποίηση, τα έξι false alarm events που έχουν επιλεγεί από το VTaC dataset [4] συμμετέχουν στο lightweight risk-related επίπεδο. Για κάθε ένα από αυτά έχουν δημιουργηθεί αντίστοιχα Risk Indicator και Risk-Relevant Event individuals. Παράλληλα, χρησιμοποιείται το Potential Harm individual vtac_harm_alert_fatigue, το οποίο αναπαριστά την πιθανή σύνδεση των false alarm cases με το φαινόμενο του alert fatigue.

Η risk-related δομή δεν είναι αυστηρά γραμμική. Το Alert Event αποτελεί το κεντρικό σημείο από το οποίο ξεκινούν δύο διαφορετικές σημασιολογικές συνδέσεις:

Alert Event --hasRiskIndicator→ Risk Indicator

και

Alert Event --isClassifiedAs→ Risk-Relevant Event

Στη συνέχεια, το Risk-Relevant Event συνδέεται με το Potential Harm:

Risk-Relevant Event --mayLeadTo→ Potential Harm

Το object property `hasRiskIndicator` συνδέει επομένως ένα false alarm event με την αντίστοιχη ένδειξη κινδύνου. Ο Risk Indicator χρησιμοποιείται για να αποτυπώσει ότι το συγκεκριμένο event έχει χαρακτηριστικό που δικαιολογεί την πρόσθετη risk-related αναπαράστασή του μέσα στο proof-of-concept.

Το object property `isClassifiedAs` συνδέει το ίδιο Alert Event με ένα ξεχωριστό Risk-Relevant Event individual. Με αυτή τη σχέση, το αρχικό πραγματικό alarm event δεν μετατρέπεται σε risk event ούτε αλλάζει την κύρια οντολογική του ταυτότητα. Αντίθετα, συνδέεται με μια πρόσθετη οντολογική αναπαράσταση που επιτρέπει να εξεταστεί η risk-related διάστασή του.

Τέλος, το object property `mayLeadTo` συνδέει το Risk-Relevant Event με το Potential Harm. Στην παρούσα υλοποίηση, τα risk-relevant events των false alarm cases συνδέονται με το `vtac_harm_alert_fatigue`. Η χρήση του όρου `mayLeadTo` είναι σκόπιμη, καθώς εκφράζει πιθανή σχέση και όχι βεβαιότητα αιτιώδους επίπτωσης. Η οντολογία επομένως δεν δηλώνει ότι ένα συγκεκριμένο false alarm προκάλεσε alert fatigue, αλλά ότι το συγκεκριμένο case έχει μοντελοποιηθεί ως περιστατικό που μπορεί να σχετίζεται με αυτή την πιθανή επίπτωση.

Η risk-related πληροφορία συμπληρώνεται και από το data property `isRiskRelevant`, το οποίο χρησιμοποιείται ως boolean ένδειξη για το αν ένα Alert Event συμμετέχει στο lightweight risk-related επίπεδο. Το συγκεκριμένο data property λειτουργεί συμπληρωματικά προς τις object-property σχέσεις και δεν αντικαθιστά τη σημασιολογική αναπαράσταση μέσω Risk Indicator, Risk-Relevant Event και Potential Harm.

Η συνολική risk-related αναπαράσταση μπορεί επομένως να αποδοθεί μέσω δύο διακριτών αλλά συνδεδεμένων διαδρομών που ξεκινούν από το ίδιο Alert Event:

Alert Event → hasRiskIndicator → Risk Indicator**Alert Event → isClassifiedAs → Risk-Relevant Event → mayLeadTo → Potential Harm**

και συμπληρωματικά μέσω του data property:

Alert Event → isRiskRelevant → true

Οι δύο object-property διαδρομές λειτουργούν παράλληλα γύρω από το ίδιο Alert Event και δεν αποτελούν μία μοναδική γραμμική αλυσίδα. Το `isRiskRelevant` χρησιμοποιείται ως συμπληρωματική boolean ένδειξη ότι το συγκεκριμένο event έχει ενταχθεί στο lightweight risk-related επίπεδο.

Πίνακας 3-5 Ρόλος των στοιχείων του lightweight risk-related επιπέδου.

Στοιχείο του risk-related επιπέδου	Ρόλος στο μοντέλο
False alarm case	Πραγματικό VTaC alarm event με τελική expert απόφαση false, το οποίο επιλέγεται για πρόσθετη risk-related αναπαράσταση.
Risk Indicator	Αναπαριστά ένδειξη που συνδέεται με τη risk-related σημασία του συγκεκριμένου false alarm case.
Risk-Relevant Event	Αποτελεί ξεχωριστή οντολογική αναπαράσταση της risk-related διάστασης του αρχικού alarm event.
Potential Harm	Αναπαριστά πιθανή αρνητική επίπτωση που μπορεί να σχετίζεται με το risk-relevant case.
isRiskRelevant	Boolean ένδειξη ότι το Alert Event συμμετέχει στο lightweight risk-related επίπεδο.

Η συγκεκριμένη δομή συνδέεται και με το requirement-to-evidence mapping. Τα false alarm cases ακολουθούν την ίδια βασική διαδρομή τεκμηρίωσης με τα true alarms και μπορούν να συνδέονται μέσω των evidence records με τα req_logging, req_human_review και req_traceability. Επιπλέον, στο πλαίσιο του proof-of-concept, τα evidence records των false alarm cases συνδέονται και με το req_risk_management. Η σύνδεση αυτή αποτελεί σχεδιαστική επιλογή της παρούσας οντολογίας και δεν σημαίνει ότι κάθε false alarm ενεργοποιεί αυτομάτως ή ικανοποιεί κάποια συγκεκριμένη νομική υποχρέωση του άρθρου 9 του AI Act [3].

Το lightweight risk-related επίπεδο πρέπει επομένως να διακρίνεται από ένα πλήρες risk management system. Το άρθρο 9 του AI Act αφορά πολύ ευρύτερη διαδικασία διαχείρισης κινδύνων για high-risk AI systems [3]. Η παρούσα υλοποίηση δεν επιχειρεί να αναπαραστήσει ολόκληρη αυτή τη διαδικασία. Χρησιμοποιεί μόνο ένα περιορισμένο σύνολο risk-related concepts για να δείξει ότι η οντολογία μπορεί να επεκτείνει τη βασική τεκμηρίωση ενός event όταν υπάρχει πρόσθετη πληροφορία σχετική με πιθανό κίνδυνο.

Η λειτουργικότητα της συγκεκριμένης δομής ελέγχεται στο Κεφάλαιο 5. Τα DL Queries εξετάζουν ποια false alarm events έχουν ενταχθεί στο risk-related επίπεδο, ποια Risk-Relevant Event individuals συνδέονται με Potential Harm και ποια evidence records υποστηρίζουν το req_risk_management. Επιπλέον, το αντίστοιχο SPARQL query ανακτά συγκεντρωτικά για κάθε σχετικό event το Risk Indicator, το Risk-Relevant Event και το Potential Harm.

Συνολικά, η προσθήκη του lightweight risk-related επιπέδου επιτρέπει στην οντολογία να διαφοροποιήσει τα false alarm cases χωρίς να αποδίδει σε αυτά μη τεκμηριωμένες κλινικές συνέπειες. Το επίπεδο αυτό αποτελεί περιορισμένη αλλά λειτουργική επέκταση του βασικού μοντέλου και δείχνει πώς ένα πραγματικό healthcare monitoring event μπορεί να αποκτήσει πρόσθετη σημασιολογική αναπαράσταση όταν υπάρχουν λόγοι να εξεταστεί και από πλευράς πιθανής risk-related σημασίας.

3.9 Συνολική αρχιτεκτονική της οντολογίας

Η συνολική αρχιτεκτονική της οντολογίας σχεδιάστηκε με σκοπό να διαχωρίζει με σαφή τρόπο τις διαφορετικές κατηγορίες γνώσης που απαιτούνται για την αναπαράσταση των πραγματικών healthcare

monitoring events, της τεκμηρίωσής τους, των επιλεγμένων AI Act-related σημείων αναφοράς και της πρόσθετης risk-related πληροφορίας.

Σε επίπεδο ιεραρχίας classes, κάτω από τη γενική OWL κλάση owl:Thing οργανώνονται τέσσερις βασικές θεματικές κλάσεις:

- **Operational Concept**
- **Evidence and Assessment Concept**
- **Normative Concept**
- **Risk Concept**

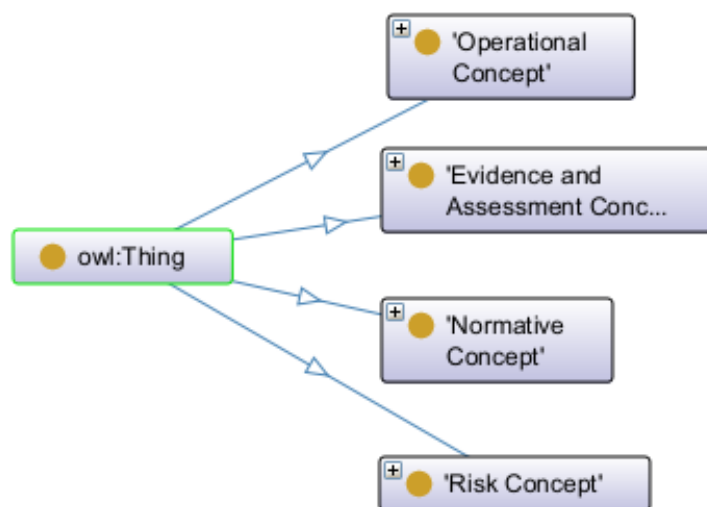
Η owl:Thing αποτελεί την ανώτερη προκαθορισμένη κλάση στην OWL και λειτουργεί ως τεχνική ρίζα της ιεραρχίας [24]. Οι τέσσερις παραπάνω κλάσεις αποτελούν τις κύριες εννοιολογικές ομάδες του μοντέλου.

Στην παρούσα εργασία οι ομάδες αυτές περιγράφονται και ως τέσσερα εννοιολογικά layers:

- **Operational layer**
- **Evidence and assessment layer**
- **Normative / AI Act-related layer**
- **Lightweight risk layer**

Ο όρος *layer* χρησιμοποιείται για να διευκολύνει την κατανόηση της αρχιτεκτονικής και του διαφορετικού ρόλου των εννοιών. Δεν υποδηλώνει ότι έχουν υλοποιηθεί τέσσερις ανεξάρτητες οντολογίες, τεχνικά modules ή απομονωμένα υποσυστήματα. Όλες οι classes, properties και individuals αποτελούν μέρος του ίδιου OWL μοντέλου και συνδέονται μεταξύ τους μέσω object και data properties.

Το Σχήμα 3.4 παρουσιάζει την πρώτη βαθμίδα της ιεραρχίας των θεματικών classes. Κάτω από την owl:Thing εμφανίζονται οι τέσσερις κύριες εννοιολογικές ομάδες: Operational Concept, Evidence and Assessment Concept, Normative Concept και Risk Concept.



Σχήμα 3.4 Ανώτερη αρχιτεκτονική οργάνωση της οντολογίας.

Ο διαχωρισμός αυτός επιτρέπει να διακρίνονται καθαρά:

- τα πραγματικά healthcare monitoring στοιχεία,
- η τεκμηρίωση και οι σχετικές έννοιες αξιολόγησης,
- τα επιλεγμένα AI Act-related σημεία αναφοράς,
- και η πρόσθετη αναπαράσταση risk-related πληροφορίας.

Τα επίπεδα δεν λειτουργούν ανεξάρτητα. Η σημασιολογική αξία του μοντέλου προκύπτει κυρίως από τις σχέσεις μεταξύ τους. Ένα πραγματικό VTaC event του operational επιπέδου μπορεί, για παράδειγμα, να συνδέεται με Evidence Record του evidence and assessment επιπέδου και το Evidence Record να υποστηρίζει requirement individuals του normative επιπέδου. Αν πρόκειται για false alarm case που έχει ενταχθεί στο lightweight risk-related μέρος της υλοποίησης, το ίδιο event μπορεί επιπλέον να συνδέεται με τις σχετικές risk concepts.

Με αυτόν τον τρόπο, η αρχιτεκτονική οργανώνει διαφορετικά είδη γνώσης χωρίς να τα απομονώνει και υποστηρίζει τη βασική λογική της εργασίας:

πραγματικό event → αξιολόγηση και τεκμηρίωση → requirement mapping → πρόσθετη risk-related αναπαράσταση όπου εφαρμόζεται.

3.9.1 Operational layer

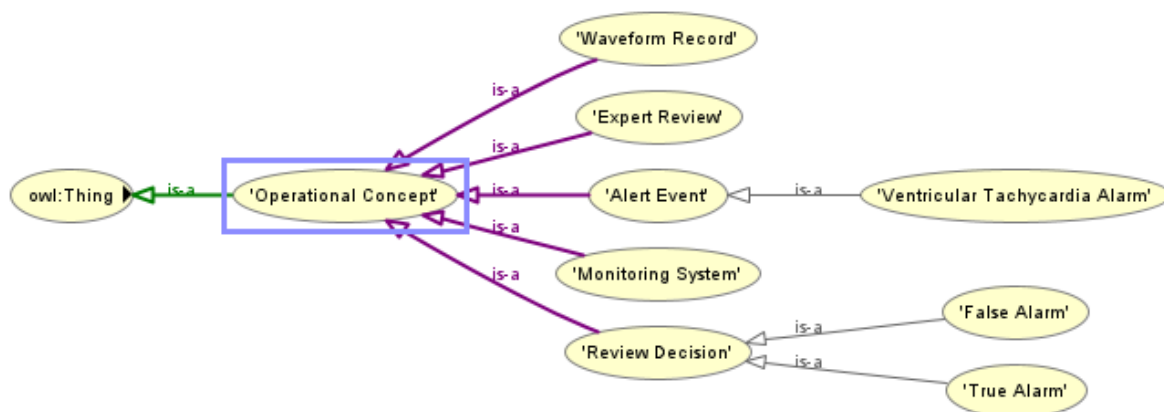
Το Operational layer περιλαμβάνει τις έννοιες που σχετίζονται άμεσα με το healthcare monitoring scenario και τα πραγματικά VTaC data instances [4]. Η ανώτερη class του συγκεκριμένου επιπέδου είναι η Operational Concept.

Κάτω από αυτήν οργανώνονται οι classes:

- Monitoring System
- Waveform Record
- Alert Event
- Ventricular Tachycardia Alarm
- Expert Review
- Review Decision
- True Alarm
- False Alarm

Η Monitoring System αναπαριστά το γενικό πλαίσιο του συστήματος παρακολούθησης από το οποίο προέρχονται τα alarm events.

Η Waveform Record αναπαριστά το record προέλευσης των πραγματικών VTaC events. Μέσω των σχέσεων hasAlertEvent και belongsToRecord διατηρείται η σύνδεση μεταξύ waveform record και alarm event.



Σχήμα 3.5 Ανάλυση του Operational Concept στην αρχιτεκτονική της οντολογίας.

Η Alert Event αποτελεί τη γενικότερη έννοια για ένα συμβάν συναγερμού, ενώ η Ventricular Tachycardia Alarm αποτελεί ειδικότερη class για τα ventricular tachycardia alarms που χρησιμοποιούνται στην παρούσα εφαρμογή.

Η Expert Review αναπαριστά την ανθρώπινη αξιολόγηση του πραγματικού event, ενώ η Review Decision αναπαριστά την τελική απόφαση της αξιολόγησης. Οι True Alarm και False Alarm αποτελούν ειδικότερες κατηγορίες της Review Decision και χρησιμοποιούνται για την αναπαράσταση των αντίστοιχων τελικών expert labels του VTaC [4].

Παρότι οι Expert Review και Review Decision σχετίζονται σημασιολογικά με την έννοια της αξιολόγησης, στην πραγματική ιεραρχία της παρούσας οντολογίας έχουν ενταχθεί στο Operational Concept. Η επιλογή αυτή αντανακλά το γεγονός ότι οι συγκεκριμένες αξιολογήσεις και αποφάσεις αποτελούν μέρος των πραγματικών πληροφοριών που συνοδεύουν τα VTaC events και όχι κανονιστικές αξιολογήσεις συμμόρφωσης.

Το Operational layer απαντά ουσιαστικά στο ερώτημα:

Τι πραγματικό περιστατικό αναπαρίσταται και ποιες πληροφορίες συνοδεύουν την αξιολόγησή του;

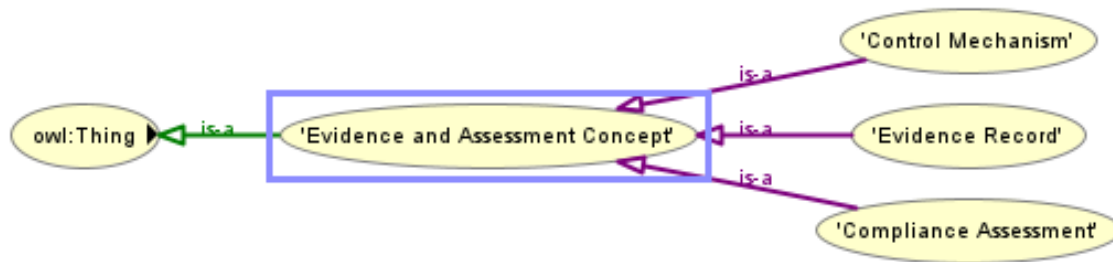
Αποτελεί επομένως το επίπεδο στο οποίο εισάγονται τα πραγματικά instances πάνω στα οποία δοκιμάζεται το υπόλοιπο οντολογικό μοντέλο.

3.9.2 Evidence and assessment layer

Το Evidence and assessment layer οργανώνεται κάτω από την class Evidence and Assessment Concept και περιλαμβάνει τις classes:

- **Evidence Record**
- **Compliance Assessment**
- **Control Mechanism**

Κεντρική θέση στο συγκεκριμένο επίπεδο έχει η Evidence Record. Τα evidence record individuals χρησιμοποιούνται για τη σημασιολογική τεκμηρίωση των επιλεγμένων events και λειτουργούν ως γέφυρα μεταξύ του operational και του normative επιπέδου.



Σχήμα 3.6 Ανάλυση του Evidence and Assessment Concept στην αρχιτεκτονική της οντολογίας.

Η σύνδεση ενός Alert Event με Evidence Record πραγματοποιείται μέσω του object property isDocumentedBy. Στη συνέχεια, το Evidence Record μπορεί να συνδεθεί με ένα ή περισσότερα requirement individuals μέσω του supportsRequirement.

Το επίπεδο περιλαμβάνει επίσης τις classes Compliance Assessment και Control Mechanism, οι οποίες παρέχουν τη δυνατότητα αναπαράστασης εννοιών σχετικών με αξιολόγηση τεκμηρίων και μηχανισμούς ελέγχου μέσα στο ευρύτερο proof-of-concept μοντέλο.

Είναι σημαντικό να διακρίνεται το Compliance Assessment από το Expert Review. Το Expert Review αφορά την ανθρώπινη αξιολόγηση του ίδιου του VTaC alarm event και για τον λόγο αυτό έχει τοποθετηθεί στο Operational Concept. Αντίθετα, το Compliance Assessment ανήκει στο επίπεδο τεκμηρίωσης και κανονιστικής αξιολόγησης της οντολογίας.

Το Evidence and assessment layer δεν πρέπει να ερμηνεύεται ως μηχανισμός που παράγει αυτόματα αποδεικτικά στοιχεία ή αποφασίζει αν υπάρχει νομική συμμόρφωση. Ο ρόλος του είναι να οργανώνει σημασιολογικά τα evidence-related στοιχεία που έχουν εισαχθεί στο proof-of-concept.

Το βασικό ερώτημα στο οποίο απαντά είναι επομένως:

Πώς οργανώνεται και συνδέεται η διαθέσιμη τεκμηρίωση ενός event μέσα στο οντολογικό μοντέλο;

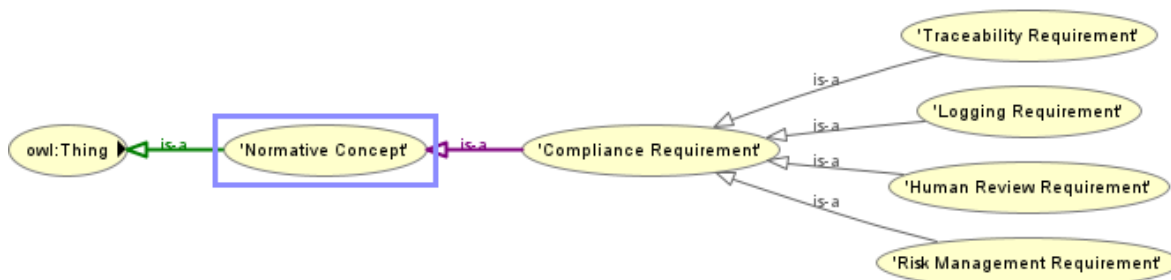
3.9.3 Normative / AI Act-related layer

Το Normative / AI Act-related layer οργανώνεται κάτω από την class Normative Concept και περιλαμβάνει την class Compliance Requirement και τις ειδικότερες classes:

- **Logging Requirement**
- **Human Review Requirement**
- **Traceability Requirement**
- **Risk Management Requirement**

Οι classes αυτές παρέχουν την οντολογική δομή μέσα στην οποία έχουν τοποθετηθεί τα τέσσερα requirement individuals:

- req_logging
- req_human_review
- req_traceability
- req_risk_management



Σχήμα 3.7 Ανάλυση του Normative Concept στην αρχιτεκτονική της οντολογίας.

Όπως αποσαφηνίστηκε στην Ενότητα 3.6, τα συγκεκριμένα individuals αποτελούν εφαρμοσμένα AI Act-related σημεία αναφοράς του proof-of-concept και δεν πρέπει να θεωρούνται τέσσερις αυτούσιες και ανεξάρτητες νομικές απαιτήσεις του Κανονισμού.

Η σύνδεση του normative layer με τα πραγματικά events πραγματοποιείται κυρίως μέσω των Evidence Records. Τα requirement individuals δεν συνδέονται απευθείας με ένα event ως δήλωση νομικής συμμόρφωσης. Αντίθετα, το Evidence Record που τεκμηριώνει ένα event μπορεί να συνδέεται μέσω supportsRequirement με ένα ή περισσότερα requirement individuals.

Η επιλογή αυτή επιτρέπει να διαχωρίζεται:

το πραγματικό περιστατικό

από

το διαθέσιμο evidence

και από

το κανονιστικό σημείο αναφοράς με το οποίο το evidence έχει χαρτογραφηθεί.

Το Normative / AI Act-related layer απαντά επομένως στο ερώτημα:

Με ποια επιλεγμένα AI Act-related σημεία αναφοράς έχει χαρτογραφηθεί το διαθέσιμο evidence;

Ο AI Act χρησιμοποιείται στο επίπεδο αυτό ως κανονιστικό πλαίσιο αναφοράς [3]. Η ύπαρξη ενός mapping δεν αποτελεί απόδειξη πλήρους νομικής συμμόρφωσης.

3.9.4 Lightweight risk layer

Το τέταρτο εννοιολογικό επίπεδο είναι το Lightweight risk layer και οργανώνεται κάτω από την class Risk Concept.

Οι βασικές classes του είναι:

- **Risk Indicator**
- **Risk-Relevant Event**
- **Potential Harm**

Το επίπεδο αυτό χρησιμοποιείται στην παρούσα υλοποίηση για την πρόσθετη αναπαράσταση των έξι false alarm cases. Η σχετική επιλογή συνδέεται με τη βιβλιογραφικά τεκμηριωμένη σημασία των false και μη χρήσιμων alarms για φαινόμενα όπως το alarm fatigue [6], [13].

Η δομή του συγκεκριμένου επιπέδου βασίζεται στις σχέσεις:

Alert Event --hasRiskIndicator→ Risk Indicator

Alert Event --isClassifiedAs→ Risk-Related Event

Risk-Related Event --mayLeadTo→ Potential Harm

Στην παρούσα υλοποίηση το βασικό Potential Harm individual είναι το:

vtac_harm_alert_fatigue

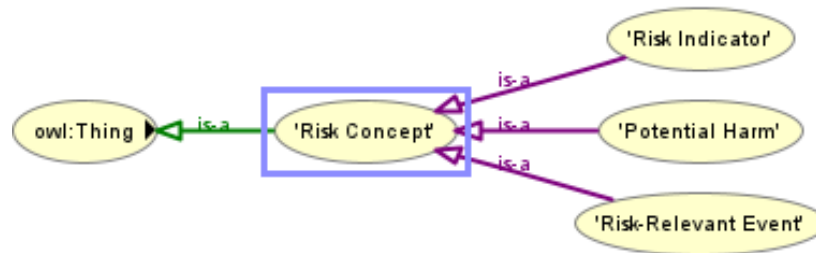
Το individual αυτό δεν δηλώνει ότι ένα συγκεκριμένο VTaC false alarm προκάλεσε πραγματικά alert fatigue. Χρησιμοποιείται για να αποτυπώσει μια πιθανή risk-related επίπτωση με την οποία έχουν συνδεθεί σημασιολογικά τα συγκεκριμένα false alarm cases.

Το Lightweight risk layer απαντά επομένως στο ερώτημα:

Πώς μπορεί ένα επιλεγμένο false alarm case να αποκτήσει πρόσθετη risk-related σημασιολογική αναπαράσταση;

Το συγκεκριμένο layer δεν αποτελεί πλήρες risk management system σύμφωνα με το άρθρο 9 του AI Act [3]. Αποτελεί περιορισμένη επέκταση της βασικής οντολογικής δομής για τη δοκιμή της αναπαράστασης risk-related πληροφορίας.

Η συνολική σχέση των τεσσάρων επιπέδων παρουσιάζεται συνοπτικά στον επόμενο πίνακα.



Σχήμα 3.8 Ανάλυση του Risk Concept στην αρχιτεκτονική της οντολογίας.

Πίνακας 3-6 Τα τέσσερα επίπεδα της αρχιτεκτονικής της οντολογίας.

Εννοιολογικό επίπεδο	Βασικός ρόλος	Ενδεικτικές έννοιες
Operational layer	Αναπαράσταση πραγματικών VTaC records, alarm events, reviews και decisions.	Waveform Record, Alert Event, Expert Review, Review Decision, Monitoring System
Evidence and assessment layer	Οργάνωση evidence records και συναφών εννοιών αξιολόγησης και ελέγχου.	Evidence Record, Compliance Assessment, Control Mechanism
Normative / AI Act-related layer	Αναπαράσταση των επιλεγμένων AI Act-related σημείων αναφοράς και του requirement mapping.	Compliance Requirement, req_logging, req_human_review, req_traceability, req_risk_management
Lightweight risk layer	Πρόσθετη αναπαράσταση risk-related πληροφορίας για τα επιλεγμένα false alarm cases.	Risk Indicator, Risk-Related Event, Potential Harm

Η τετραμερής αυτή οργάνωση επιτρέπει να διατηρούνται διακριτοί οι διαφορετικοί σημασιολογικοί ρόλοι χωρίς να διασπάται η οντολογία σε ανεξάρτητα συστήματα. Το Operational layer παρέχει τα πραγματικά περιστατικά, το Evidence and assessment layer οργανώνει την τεκμηρίωσή τους, το Normative layer παρέχει τα επιλεγμένα κανονιστικά σημεία αναφοράς και το Lightweight risk layer προσθέτει risk-related πληροφορία όπου αυτό έχει προβλεφθεί από τον σχεδιασμό.

Η λειτουργικότητα του μοντέλου προκύπτει τελικά από τη διασύνδεση των τεσσάρων επιπέδων μέσω των object και data properties και όχι μόνο από την ιεραρχική ταξινόμηση των classes. Με αυτόν τον τρόπο, η αρχιτεκτονική παρέχει ένα ενιαίο πλαίσιο μέσα στο οποίο πραγματικά VTaC events, expert decisions, evidence records, requirement mappings και risk-related concepts μπορούν να οργανωθούν και να ανακτηθούν με συνεκτικό τρόπο.

3.10 Σχεδιαστικές επιλογές

Η ανάπτυξη της οντολογίας βασίστηκε σε ένα σύνολο σχεδιαστικών επιλογών που καθόρισαν το πεδίο εφαρμογής, τη δομή και τον τρόπο λειτουργικού ελέγχου του μοντέλου. Στην ontology engineering βιβλιογραφία, ο καθορισμός του domain και του scope, η επιλογή των κατάλληλων classes και properties και η δημιουργία συγκεκριμένων instances αποτελούν βασικά στοιχεία της διαδικασίας ανάπτυξης μιας οντολογίας [22], [28]. Στην παρούσα εργασία οι γενικές αυτές αρχές προσαρμόστηκαν στις ανάγκες του συγκεκριμένου healthcare monitoring proof-of-concept.

Η πρώτη σχεδιαστική επιλογή ήταν ο **σαφής περιορισμός του πεδίου εφαρμογής**. Η οντολογία δεν επιχειρεί να αναπαραστήσει όλα τα είδη healthcare monitoring events, όλα τα πιθανά συστήματα τεχνητής νοημοσύνης ή το σύνολο του AI Act. Εστιάζει σε ventricular tachycardia alarm events και σε ένα περιορισμένο σύνολο εννοιών που είναι απαραίτητες για την εξέταση του συγκεκριμένου ερευνητικού προβλήματος.

Η δεύτερη επιλογή ήταν η **χρήση πραγματικών δεδομένων** αντί για υποθετικά περιστατικά. Τα 12 επιλεγμένα VTaC events προέρχονται από πραγματικές εγγραφές του dataset και συνοδεύονται από διαθέσιμες expert αποφάσεις [4]. Με αυτόν τον τρόπο, η οντολογία δοκιμάζεται πάνω σε πραγματικά instances, χωρίς όμως το περιορισμένο αυτό σύνολο να αντιμετωπίζεται ως στατιστικά αντιπροσωπευτικό δείγμα.

Η τρίτη επιλογή ήταν ο **διαχωρισμός μεταξύ event, review και decision**. Η πληροφορία true ή false δεν αποθηκεύεται απλώς ως μοναδικό χαρακτηριστικό του Alert Event. Το event, η διαδικασία ανθρώπινης αξιολόγησης και η τελική απόφαση αναπαρίστανται ως διακριτές έννοιες, ώστε να διατηρείται σαφέστερη σημασιολογική δομή.

Η τέταρτη επιλογή ήταν η **χρήση του Evidence Record ως διακριτού τεκμηριωτικού κόμβου**. Το Alert Event συνδέεται με Evidence Record μέσω του isDocumentedBy, ενώ το Evidence Record συνδέεται με επιλεγμένα requirement individuals μέσω του supportsRequirement. Με αυτόν τον τρόπο διαχωρίζεται το πραγματικό περιστατικό από την τεκμηρίωση και από το κανονιστικό σημείο αναφοράς με το οποίο αυτή έχει χαρτογραφηθεί.

Η πέμπτη επιλογή ήταν ο **περιορισμός του normative μέρους σε τέσσερα requirement individuals**: req_logging, req_human_review, req_traceability και req_risk_management. Όπως αναλύθηκε στην Ενότητα 3.6, τα συγκεκριμένα individuals χρησιμοποιούνται ως εφαρμοσμένα AI Act-related σημεία αναφοράς του proof-of-concept και όχι ως πλήρης ή αυτούσια αναπαράσταση του συνόλου των νομικών απαιτήσεων του AI Act [3].

Η έκτη επιλογή ήταν η **προσθήκη lightweight risk-related αναπαράστασης για τα false alarm cases**. Η επέκταση αυτή επιτρέπει να εξεταστεί αν ένα πραγματικό false alarm event μπορεί να συνδεθεί με Risk Indicator, Risk-Relevant Event και Potential Harm, χωρίς να υποστηρίζεται ότι αποδεικνύεται πραγματική βλάβη ή ότι υλοποιείται πλήρες risk management system.

Η έβδομη επιλογή ήταν η **οργάνωση των classes σε τέσσερις κύριες εννοιολογικές ομάδες**: Operational Concept, Evidence and Assessment Concept, Normative Concept και Risk Concept. Η οργάνωση αυτή διευκολύνει την κατανόηση των διαφορετικών ρόλων των εννοιών, χωρίς να σημαίνει ότι πρόκειται για τέσσερις ανεξάρτητες οντολογίες ή τεχνικά απομονωμένα modules.

Η όγδοη επιλογή ήταν η **αναπαράσταση των πραγματικών στοιχείων ως individuals**. Waveform records, alarm events, expert reviews, review decisions, evidence records, requirement instances και risk-related στοιχεία αποκτούν συγκεκριμένες instantiations μέσα στην οντολογία. Με αυτόν τον τρόπο, το μοντέλο δεν παραμένει μόνο σε επίπεδο αφηρημένων classes, αλλά εφαρμόζεται σε συγκεκριμένες περιπτώσεις και μπορεί να ελεγχθεί μέσω queries.

Η ένατη επιλογή ήταν η **χρήση object properties για τη ρητή αναπαράσταση των σημασιολογικών σχέσεων**. Properties όπως belongsToRecord, hasReview, hasDecision, isDocumentedBy, supportsRequirement, hasRiskIndicator, isClassifiedAs και mayLeadTo συνδέουν μεταξύ τους τα διαφορετικά στοιχεία του μοντέλου. Οι σχέσεις αυτές δεν σχηματίζουν απαραίτητα μία μοναδική γραμμική αλυσίδα, αλλά από τη σύνθεσή τους προκύπτουν οι διαδρομές ιχνηλασιμότητας και risk-related πληροφορίας που εξετάζονται στην εργασία.

Η δέκατη επιλογή ήταν η **χρήση data properties για συγκεκριμένες περιγραφικές τιμές**. Properties όπως eventId, recordId, decisionLabel, hasHumanReview, isLogged, isTraceable, isRiskRelevant και potentialHarmLabel επιτρέπουν την αποθήκευση identifiers, labels και boolean ενδείξεων. Οι ενδείξεις αυτές αποτελούν στοιχεία του οντολογικού proof-of-concept και δεν πρέπει από μόνες τους να ερμηνεύονται ως απόδειξη ότι ένα πραγματικό εξωτερικό σύστημα ικανοποιεί αντίστοιχη τεχνική ή νομική απαίτηση.

Η ενδέκατη επιλογή ήταν ο **σχεδιασμός της οντολογίας με δυνατότητα λειτουργικού ελέγχου**. Η υλοποίηση δεν σχεδιάστηκε μόνο για στατική παρουσίαση της ιεραρχίας των classes. Η δομή των classes, properties και individuals οργανώθηκε ώστε να μπορούν να εξεταστούν η λογική συνέπεια και οι βασικές σημασιολογικές διαδρομές μέσω reasoner, DL competency questions και SPARQL queries. Η χρήση SPARQL για τη σημασιολογική ανάκτηση και εξέταση οντολογικά οργανωμένης γνώσης αποτελεί καθιερωμένη δυνατότητα αντίστοιχων semantic approaches [25], [28].

Η δωδέκατη και ιδιαίτερα σημαντική επιλογή ήταν η **αποφυγή υπερβολικών ισχυρισμών σχετικά με τις δυνατότητες του μοντέλου**. Η οντολογία δεν παρουσιάζεται ως σύστημα αυτόματης ιατρικής διάγνωσης, ως μοντέλο πρόβλεψης true και false alarms ή ως μηχανισμός που αποδεικνύει πλήρη νομική συμμόρφωση με τον AI Act. Παρουσιάζεται ως εφαρμοσμένο proof-of-concept για τη σημασιολογική αναπαράσταση, τεκμηρίωση, ιχνηλασιμότητα και αναζήτηση σχέσεων μεταξύ πραγματικών healthcare monitoring events, διαθέσιμων evidence records και επιλεγμένων AI Act-related σημείων αναφοράς.

Οι βασικές σχεδιαστικές επιλογές συνοψίζονται στον Πίνακα 3-7.

Πίνακας 3-7 Βασικές σχεδιαστικές επιλογές της οντολογίας.

Σχεδιαστική επιλογή	Σκοπός
Περιορισμός του πεδίου εφαρμογής	Εστίαση σε συγκεκριμένο και ελέγξιμο VTaC healthcare monitoring use case.
Χρήση πραγματικών VTaC events	Εφαρμογή του μοντέλου σε πραγματικά και όχι υποθετικά περιστατικά.
Διαχωρισμός event, review και decision	Σαφής διάκριση μεταξύ περιστατικού, ανθρώπινης αξιολόγησης και τελικής απόφασης.
Evidence Record ως τεκμηριωτικός κόμβος	Διαχωρισμός event, evidence και requirement mapping.
Τέσσερα AI Act-related σημεία αναφοράς	Περιορισμένο και ελέγξιμο normative scope.
Lightweight risk-related αναπαράσταση	Πρόσθετη σημασιολογική αναπαράσταση των false alarm cases.
Τετραμερής εννοιολογική οργάνωση	Διάκριση operational, evidence/assessment, normative και risk concepts.
Αναπαράσταση μέσω individuals	Εφαρμογή της οντολογίας σε συγκεκριμένες πραγματικές περιπτώσεις.
Χρήση object properties	Ρητή αναπαράσταση σημασιολογικών σχέσεων και διαδρομών.
Χρήση data properties	Αποθήκευση identifiers, labels και περιγραφικών/boolean τιμών.
Σχεδιασμός με δυνατότητα ελέγχου	Υποστήριξη reasoner, DL competency questions και SPARQL queries.
Περιορισμός των ισχυρισμών	Σαφής διάκριση της σημασιολογικής τεκμηρίωσης από πρόβλεψη, ιατρική διάγνωση και πλήρη νομική συμμόρφωση.

Οι επιλογές αυτές λειτουργούν συνδυαστικά και καθορίζουν τον χαρακτήρα της υλοποίησης. Το αποτέλεσμα είναι ένα περιορισμένο αλλά λειτουργικά ελέγξιμο οντολογικό μοντέλο, στο οποίο πραγματικά VTaC events μπορούν να συνδεθούν με την προέλευση και την expert αξιολόγησή τους, να τεκμηριωθούν μέσω Evidence Records, να χαρτογραφηθούν σε επιλεγμένα AI Act-related σημεία αναφοράς και, για τα false alarm cases, να αποκτήσουν πρόσθετη risk-related αναπαράσταση.

3.11 Μεθοδολογία ανάπτυξης στο Protégé

Η ανάπτυξη της οντολογίας πραγματοποιήθηκε στο περιβάλλον Protégé, το οποίο χρησιμοποιήθηκε ως βασικό εργαλείο σχεδίασης, υλοποίησης, τεκμηρίωσης και λειτουργικού ελέγχου του μοντέλου [26]. Η διαδικασία δεν περιορίστηκε στην απλή εισαγωγή όρων, αλλά ακολούθησε μια οργανωμένη και σταδιακή πορεία από τον καθορισμό του πεδίου εφαρμογής μέχρι τη δημιουργία πραγματικών instances και τη λειτουργική επικύρωση της οντολογίας.

Η διαδικασία ανάπτυξης διαμορφώθηκε με βάση γενικές αρχές ontology engineering, σύμφωνα με τις οποίες ο σκοπός και το πεδίο μιας οντολογίας, οι βασικές έννοιες του domain, η ιεραρχία των classes, οι properties και τα instances πρέπει να καθορίζονται με βάση τις ανάγκες της εφαρμογής [22]. Δεν εφαρμόστηκε αυτούσια μία συγκεκριμένη τυπική ontology-development methodology. Αντίθετα, οι

γενικές αρχές προσαρμόστηκαν στις ανάγκες του συγκεκριμένου healthcare monitoring proof-of-concept.

Η επιλογή αυτή είναι συμβατή με την ευρύτερη λογική της ontology engineering βιβλιογραφίας, όπου η ανάπτυξη μιας οντολογίας αντιμετωπίζεται ως επαναληπτική διαδικασία και οι σχεδιαστικές αποφάσεις καθορίζονται από το use case και το απαιτούμενο επίπεδο λεπτομέρειας [22], [28].

Το πρώτο βήμα ήταν ο **καθορισμός του πεδίου εφαρμογής και του σκοπού της οντολογίας**. Το μοντέλο περιορίστηκε σε healthcare monitoring alarm events και ειδικότερα σε ventricular tachycardia alarm events από το VTaC dataset [4]. Παράλληλα, καθορίστηκε ότι ο στόχος δεν ήταν η δημιουργία γενικής healthcare ontology ούτε η πλήρης αναπαράσταση του AI Act, αλλά η ανάπτυξη ενός εφαρμοσμένου proof-of-concept για τη σημασιολογική τεκμηρίωση πραγματικών events και τη σύνδεσή τους με επιλεγμένα AI Act-related σημεία αναφοράς.

Το δεύτερο βήμα ήταν ο **προσδιορισμός των ερωτημάτων που θα έπρεπε να μπορεί να απαντήσει η οντολογία**. Τα ερωτήματα αυτά χρησιμοποιήθηκαν ως competency questions, δηλαδή ως λειτουργικά ερωτήματα που συνδέουν τον σκοπό της οντολογίας με τις πληροφορίες που πρέπει να μπορούν να ανακτηθούν από αυτή. Η χρήση competency questions ως μέσου προσδιορισμού και ελέγχου των απαιτήσεων μιας οντολογίας αποτελεί καθιερωμένη πρακτική στην ontology engineering [22].

Στο συγκεκριμένο proof-of-concept, τα competency questions αφορούν, μεταξύ άλλων, τον εντοπισμό των πραγματικών VTaC events, τη διάκριση true και false alarm cases, τη σύνδεσή τους με evidence records, την ανάκτηση των requirement mappings και την εξέταση των risk-related σχέσεων για τα false alarm cases. Τα ερωτήματα αυτά αποτέλεσαν οδηγό τόσο για τον σχεδιασμό των εννοιών και των properties όσο και για τον μεταγενέστερο λειτουργικό έλεγχο του μοντέλου.

Το τρίτο βήμα ήταν η **αναγνώριση των βασικών εννοιών του domain**. Από την ανάλυση του use case και των ερευνητικών αναγκών προσδιορίστηκαν έννοιες όπως Waveform Record, Alert Event, Expert Review, Review Decision, Evidence Record, Compliance Requirement, Control Mechanism, Risk Indicator, Risk-Relevant Event και Potential Harm. Οι έννοιες αυτές αποτέλεσαν τη βάση για τη διαμόρφωση της ιεραρχίας των classes.

Το τέταρτο βήμα ήταν η **δημιουργία των classes και της ιεραρχίας τους στο Protégé**. Οι 23 θεματικές classes οργανώθηκαν κάτω από τέσσερις κύριες εννοιολογικές ομάδες: Operational Concept, Evidence and Assessment Concept, Normative Concept και Risk Concept. Όπως αναλύθηκε στην Ενότητα 3.9, η οργάνωση αυτή επιτρέπει τη διάκριση διαφορετικών σημασιολογικών ρόλων μέσα στο ίδιο OWL μοντέλο.

Το πέμπτο βήμα ήταν ο **ορισμός των object properties**. Δημιουργήθηκαν 12 object properties για την αναπαράσταση των σχέσεων μεταξύ των individuals. Properties όπως belongsToRecord, hasAlertEvent, hasReview, hasDecision, isDocumentedBy και supportsRequirement χρησιμοποιήθηκαν για τη βασική διαδρομή ιχνηλασιμότητας και τεκμηρίωσης, ενώ τα hasRiskIndicator, isClassifiedAs και mayLeadTo χρησιμοποιήθηκαν στο lightweight risk-related μέρος του μοντέλου.

Κατά τον σχεδιασμό των object properties καθορίστηκαν όπου απαιτούνταν οι κατάλληλες domain και range classes. Με αυτόν τον τρόπο, οι σχέσεις δεν χρησιμοποιούνται ως απλές ετικέτες μεταξύ individuals, αλλά εντάσσονται στη σημασιολογική δομή της οντολογίας.

Το έκτο βήμα ήταν ο **ορισμός των data properties**. Δημιουργήθηκαν 9 data properties για την αναπαράσταση συγκεκριμένων περιγραφικών τιμών, identifiers, labels και boolean ενδείξεων. Σε αυτά

περιλαμβάνονται τα `assessmentOutcome`, `decisionLabel`, `eventId`, `hasHumanReview`, `isLogged`, `isRiskRelevant`, `isTraceable`, `potentialHarmLabel` και `recordId`.

Η χρήση `data properties` συμπληρώνει τις `object-property` σχέσεις. Τα `object properties` αναπαριστούν κυρίως σχέσεις μεταξύ οντοτήτων, ενώ τα `data properties` επιτρέπουν τη σύνδεση `individuals` με συγκεκριμένες τιμές δεδομένων.

Το έβδομο βήμα ήταν η **δημιουργία και εισαγωγή των `individuals`**. Στο στάδιο αυτό η αφηρημένη δομή των `classes` και `properties` εφαρμόστηκε σε συγκεκριμένες περιπτώσεις. Δημιουργήθηκαν `individuals` για τα πραγματικά VTaC waveform records και alarm events, καθώς και για Expert Reviews, Review Decisions, Evidence Records, requirement `individuals`, control mechanisms και risk-related στοιχεία.

Η instantiation πραγματικών περιπτώσεων είχε ιδιαίτερη σημασία, επειδή επέτρεψε η οντολογία να ελεγχθεί ως εφαρμοσμένο proof-of-concept και όχι μόνο ως αφηρημένο εννοιολογικό σχήμα. Αντίστοιχες ontology engineering προσεγγίσεις χρησιμοποιούν την εισαγωγή συγκεκριμένων instances ως μέσο ελέγχου της εφαρμοσιμότητας του μοντέλου [28].

Το όγδοο βήμα ήταν η **προσθήκη annotations**. Χρησιμοποιήθηκαν annotations όπως `rdfs:label` και `rdfs:comment` για την ανθρώπινα κατανοητή περιγραφή των `classes`, `properties` και `individuals`. Τα annotations δεν αποτελούν `object-property` σχέσεις και δεν χρησιμοποιούνται για την αναπαράσταση των βασικών λογικών διαδρομών του μοντέλου. Ο ρόλος τους είναι τεκμηριωτικός και βοηθά στην κατανόηση του σκοπού και της σημασίας κάθε οντολογικού στοιχείου.

Το ένατο βήμα ήταν ο **έλεγχος λογικής συνέπειας της οντολογίας μέσω reasoner**. Μετά τη δημιουργία των `classes`, `properties` και `individuals`, χρησιμοποιήθηκε reasoner μέσα στο Protégé προκειμένου να εξεταστεί αν η οντολογία περιέχει λογικές ασυνέπειες και αν η inferred ιεραρχία παραμένει συμβατή με τον σχεδιασμό της. Η συγκεκριμένη διαδικασία αποτελεί τεχνικό και σημασιολογικό έλεγχο του OWL μοντέλου και όχι αξιολόγηση νομικής συμμόρφωσης.

Το δέκατο βήμα ήταν η **λειτουργική επικύρωση μέσω DL competency questions και SPARQL queries**. Τα DL Queries χρησιμοποιήθηκαν για να εξεταστεί αν η οντολογία μπορεί να επιστρέψει τα αναμενόμενα σύνολα `individuals` με βάση τις δηλωμένες και inferred σχέσεις. Τα SPARQL queries χρησιμοποιήθηκαν συμπληρωματικά για την ανάκτηση περισσότερο σύνθετων και συγκεντρωτικών διαδρομών πληροφορίας [25].

Σημαντικό στοιχείο της διαδικασίας είναι ότι τα competency questions και ο λειτουργικός έλεγχος δεν αντιμετωπίστηκαν αποκλειστικά ως τελικό στάδιο που προστέθηκε μετά την ολοκλήρωση της οντολογίας. Τα ερωτήματα που έπρεπε να μπορεί να απαντήσει το μοντέλο είχαν ήδη ληφθεί υπόψη στον σχεδιασμό των `classes`, `properties` και `individuals`. Με αυτόν τον τρόπο, υπάρχει σύνδεση μεταξύ του αρχικού ερευνητικού σκοπού, της οντολογικής δομής και της μεταγενέστερης λειτουργικής επικύρωσης.

Η συνολική διαδικασία ανάπτυξης μπορεί να συνοψιστεί στον Πίνακα 3-8.

Πίνακας 3-8 Μεθοδολογικά βήματα ανάπτυξης της οντολογίας στο Protégé.

Βήμα	Περιγραφή	Κύριος σκοπός
1	Καθορισμός σκοπού και πεδίου εφαρμογής	Περιορισμός του domain σε συγκεκριμένο και ελέγξιμο use case
2	Προσδιορισμός competency questions	Καθορισμός των ερωτημάτων που πρέπει να υποστηρίζει η οντολογία
3	Αναγνώριση βασικών εννοιών	Προσδιορισμός των κύριων concepts του μοντέλου
4	Δημιουργία classes και class hierarchy	Οργάνωση της εννοιολογικής δομής
5	Δημιουργία object properties	Αναπαράσταση σχέσεων μεταξύ οντοτήτων
6	Δημιουργία data properties	Αναπαράσταση identifiers, labels και περιγραφικών τιμών
7	Δημιουργία individuals	Εφαρμογή της οντολογίας σε πραγματικά VTaC instances
8	Προσθήκη annotations	Ανθρώπινα κατανοητή τεκμηρίωση των οντολογικών στοιχείων
9	Έλεγχος με reasoner	Έλεγχος λογικής συνέπειας και inferred δομής
10	DL Queries και SPARQL queries	Λειτουργική επικύρωση και ανάκτηση σημασιολογικών διαδρομών

Η διαδικασία αυτή δεν πρέπει να θεωρηθεί αυστηρά γραμμική. Κατά την ανάπτυξη της οντολογίας πραγματοποιήθηκαν επαναληπτικοί έλεγχοι και προσαρμογές στις classes, στις properties, στα individuals και στα queries, όταν αυτό κρίθηκε απαραίτητο. Η επαναληπτική αυτή λογική είναι συμβατή με τη γενικότερη φύση της ontology engineering, όπου η μοντελοποίηση εξελίσσεται καθώς ελέγχεται η καταλληλότητα της αναπαράστασης για το συγκεκριμένο use case [22], [28].

Το Protégé χρησιμοποιήθηκε επίσης για την οπτική επιθεώρηση και τεκμηρίωση της οντολογίας. Οι διαθέσιμες προβολές της ιεραρχίας και εργαλεία όπως το OntoGraf και το OWLViz αξιοποιήθηκαν για την παρουσίαση επιλεγμένων τμημάτων του μοντέλου. Οι απεικονίσεις αυτές λειτουργούν συμπληρωματικά προς τη λογική και λειτουργική επικύρωση και δεν αποτελούν από μόνες τους μέθοδο απόδειξης ορθότητας της οντολογίας.

Συνολικά, η μεθοδολογική πορεία επέτρεψε τη σταδιακή μετάβαση από τον ορισμό του ερευνητικού προβλήματος σε ένα υλοποιημένο και λειτουργικά ελέγξιμο OWL μοντέλο. Η οντολογία σχεδιάστηκε με συγκεκριμένο πεδίο εφαρμογής, οργανώθηκε σε σαφή εννοιολογική δομή, εμπλουτίστηκε με πραγματικά VTaC instances και ελέγχθηκε ως προς τη λογική συνέπεια και τη δυνατότητα ανάκτησης των σημασιολογικών σχέσεων που απαιτούνται από τα competency questions.

3.12 Όρια της μεθοδολογικής προσέγγισης

Η μεθοδολογική προσέγγιση της παρούσας εργασίας έχει συγκεκριμένα όρια, τα οποία είναι απαραίτητα να αποσαφηνιστούν ώστε να προσδιορίζεται με ακρίβεια τι εξετάζει και τι δεν εξετάζει η οντολογική υλοποίηση. Η σαφής οριοθέτηση είναι ιδιαίτερα σημαντική σε ένα proof-of-concept μοντέλο, καθώς επιτρέπει τα αποτελέσματα να ερμηνεύονται αποκλειστικά μέσα στο πλαίσιο του ερευνητικού σκοπού για τον οποίο σχεδιάστηκε.

Πρώτο όριο αποτελεί το πεδίο εφαρμογής της οντολογίας. Η εργασία εστιάζει σε ένα συγκεκριμένο healthcare monitoring use case και ειδικότερα σε ventricular tachycardia alarm events από το VTaC dataset [4]. Επομένως, η οντολογία δεν επιχειρεί να αναπαραστήσει όλα τα είδη ιατρικών alarms, όλα τα συστήματα παρακολούθησης υγείας ή όλες τις πιθανές εφαρμογές τεχνητής νοημοσύνης στον χώρο της υγείας. Τα αποτελέσματα αφορούν τη δυνατότητα εφαρμογής της συγκεκριμένης οντολογικής δομής στο επιλεγμένο use case και δεν πρέπει να γενικεύονται αυτομάτως σε διαφορετικά κλινικά ή τεχνολογικά περιβάλλοντα.

Δεύτερο όριο αποτελεί ο αριθμός και ο τρόπος επιλογής των events που χρησιμοποιούνται στην proof-of-concept υλοποίηση. Το μοντέλο εφαρμόζεται σε 12 επιλεγμένα VTaC alarm events, έξι true και έξι false. Η ισόρροπη αυτή κατανομή αποτελεί σκόπιμη σχεδιαστική επιλογή για τον λειτουργικό έλεγχο της οντολογίας και δεν αντανάκλα την πραγματική κατανομή των κατηγοριών στο συνολικό dataset. Τα συγκεκριμένα instances παρέχουν διαφορετικές περιπτώσεις για τον έλεγχο των σχεδιασμένων σημασιολογικών σχέσεων, αλλά δεν χρησιμοποιούνται ως στατιστικά αντιπροσωπευτικό δείγμα. Κατά συνέπεια, η εργασία δεν εξάγει συμπεράσματα για τη συχνότητα ή τη συνολική κατανομή των true και false alarms στο VTaC.

Τρίτο όριο είναι ότι η οντολογία δεν χρησιμοποιείται για πρόβλεψη ή ταξινόμηση. Η τελική απόφαση true ή false προέρχεται από τα διαθέσιμα expert labels του VTaC dataset [4] και δεν παράγεται από το οντολογικό μοντέλο. Η οντολογία οργανώνει και αναπαριστά σημασιολογικά την ήδη διαθέσιμη πληροφορία. Δεν εκπαιδεύεται πάνω στα waveform data, δεν προβλέπει την κατηγορία νέων alarm events και δεν λειτουργεί ως σύστημα αυτόματης ιατρικής διάγνωσης.

Τέταρτο όριο αφορά το διαθέσιμο επίπεδο κλινικής πληροφορίας. Η εργασία αξιοποιεί κυρίως τα identifiers των waveform records, τα alarm event identifiers και τις τελικές expert αποφάσεις που παρέχονται από το VTaC [4]. Δεν πραγματοποιείται ανάλυση των ίδιων των waveform signals και δεν επιχειρείται πλήρης κλινική ερμηνεία κάθε περιστατικού. Επιπλέον, το dataset δεν παρέχει το πλήρες κλινικό και οργανωτικό πλαίσιο που θα απαιτούνταν για την εξαγωγή συμπερασμάτων σχετικά με την πραγματική επίπτωση κάθε alarm σε ασθενείς ή επαγγελματίες υγείας. Για τον λόγο αυτό, η οντολογία εστιάζει στην οργανωμένη σημασιολογική τεκμηρίωση των διαθέσιμων πληροφοριών και όχι στην ιατρική αξιολόγηση των σημάτων.

Πέμπτο όριο αφορά το κανονιστικό εύρος της οντολογίας. Η υλοποίηση δεν επιχειρεί να μοντελοποιήσει το σύνολο του AI Act. Χρησιμοποιεί τέσσερα εφαρμοσμένα requirement individuals — req_logging, req_human_review, req_traceability και req_risk_management — τα οποία λειτουργούν ως επιλεγμένα AI Act-related σημεία αναφοράς του proof-of-concept. Όπως αναλύθηκε στην Ενότητα 3.6, οι έννοιες αυτές συνδέονται με συγκεκριμένες διαστάσεις του Κανονισμού, αλλά δεν αποτελούν πλήρη αναπαράσταση όλων των νομικών υποχρεώσεων που θα μπορούσαν να αφορούν ένα πραγματικό AI system [3].

Έκτο όριο είναι ότι το requirement-to-evidence mapping δεν αποδεικνύει πλήρη νομική συμμόρφωση. Η σύνδεση ενός Evidence Record με ένα requirement individual μέσω του

`supportsRequirement` δηλώνει τη σημασιολογική χαρτογράφηση που έχει σχεδιαστεί στο πλαίσιο του `proof-of-concept`. Δεν αποτελεί νομική αξιολόγηση, πιστοποίηση ή επίσημη δήλωση συμμόρφωσης. Η αξιολόγηση της πραγματικής συμμόρφωσης ενός συγκεκριμένου AI system με τον AI Act θα απαιτούσε πολύ ευρύτερη τεχνική, οργανωτική, νομική και διαδικαστική τεκμηρίωση [3].

Στο ίδιο πλαίσιο, `boolean data properties` όπως `isLogged`, `hasHumanReview` και `isTraceable` πρέπει να ερμηνεύονται ως στοιχεία της συγκεκριμένης οντολογικής αναπαράστασης και όχι ως ανεξάρτητες αποδείξεις ότι ένα πραγματικό εξωτερικό σύστημα ικανοποιεί τις αντίστοιχες απαιτήσεις του AI Act.

Έβδομο όριο αφορά το `lightweight risk-related` επίπεδο. Η σύνδεση των `false alarm cases` με `Risk Indicator`, `Risk-Relevant Event` και `Potential Harm` αποτυπώνει μια περιορισμένη `risk-related` σημασιολογική αναπαράσταση. Η βιβλιογραφία υποστηρίζει τη γενικότερη σύνδεση των συχνών ή μη χρήσιμων `alarms` με φαινόμενα όπως το `alarm fatigue` [6], [13], όμως η παρούσα υλοποίηση δεν αποδεικνύει ότι ένα συγκεκριμένο VTaC `false alarm` προκάλεσε πραγματική βλάβη ή `alert fatigue`. Για τον λόγο αυτό χρησιμοποιείται η σχέση `mayLeadTo`, η οποία εκφράζει πιθανή σύνδεση και όχι βέβαιη αιτιότητα.

Επιπλέον, το `lightweight risk-related` επίπεδο δεν αποτελεί πλήρες `risk management system` κατά την έννοια του άρθρου 9 του AI Act [3]. Δεν πραγματοποιείται ποσοτικοποίηση πιθανότητας ή σοβαρότητας κινδύνου, συστηματική αξιολόγηση `mitigation measures` ή παρακολούθηση ενός ολοκληρωμένου `risk management lifecycle`.

Όγδοο όριο αφορά την τεχνική κλίμακα και το περιβάλλον εφαρμογής της υλοποίησης. Η οντολογία αναπτύχθηκε ως `proof-of-concept` στο `Protégé` και όχι ως παραγωγικό σύστημα μεγάλης κλίμακας. Δεν εξετάζονται ζητήματα απόδοσης σε πολύ μεγάλο αριθμό `instances`, αυτοματοποιημένη εισαγωγής ολόκληρου του VTaC `dataset`, συνεχούς εισροής νέων δεδομένων, διασύνδεσης με πραγματικά `patient monitoring systems` ή ενσωμάτωσης σε κλινική πληροφοριακή υποδομή. Κατά συνέπεια, η παρούσα εργασία δεν αξιολογεί `scalability`, `real-world integration` ή επιχειρησιακή απόδοση ενός τέτοιου συστήματος.

Ένατο όριο αφορά το εύρος της επικύρωσης της οντολογίας. Η αξιολόγηση που πραγματοποιείται στην παρούσα εργασία είναι κυρίως λογική και λειτουργική. Περιλαμβάνει έλεγχο συνέπειας μέσω `reasoner`, εξέταση `inferred γνώσης`, `DL competency questions` και `SPARQL queries`. Με τον τρόπο αυτό ελέγχεται αν η υλοποιημένη οντολογία μπορεί να αναπαραστήσει και να ανακτήσει τις σημασιολογικές σχέσεις για τις οποίες σχεδιάστηκε.

Η συγκεκριμένη διαδικασία δεν ισοδυναμεί όμως με πλήρη εξωτερική επικύρωση της οντολογίας. Δεν πραγματοποιήθηκε ανεξάρτητη κλινική αξιολόγηση από επαγγελματίες υγείας, ανεξάρτητη νομική αξιολόγηση των `mappings` από ειδικούς του AI Act ή αξιολόγηση της χρηστικότητας της οντολογίας σε πραγματικό οργανωτικό περιβάλλον. Επίσης, η λειτουργική επικύρωση δεν πρέπει να συγχέεται με αξιολόγηση προγνωστικής ακρίβειας, καθώς δεν υπάρχει προγνωστικό μοντέλο.

Τα παραπάνω όρια καθορίζουν το σωστό πλαίσιο μέσα στο οποίο πρέπει να ερμηνευθούν τα αποτελέσματα. Η οντολογία δεν πρέπει να αξιολογηθεί ως μοντέλο μηχανικής μάθησης, ως διαγνωστικό εργαλείο, ως παραγωγικό `clinical monitoring system` ή ως πλήρες σύστημα νομικής συμμόρφωσης. Πρέπει να αξιολογηθεί ως εφαρμοσμένο `proof-of-concept` οντολογικό μοντέλο που εξετάζει αν πραγματικά `healthcare monitoring alarm events` μπορούν να οργανωθούν σημασιολογικά, να συνδεθούν με την προέλευση και την ανθρώπινη αξιολόγησή τους, να τεκμηριωθούν μέσω `Evidence Records` και να χαρτογραφηθούν σε επιλεγμένα AI Act-related σημεία αναφοράς.

Συνολικά, οι περιορισμοί της μεθοδολογικής προσέγγισης δεν αναιρούν τον σκοπό της εργασίας, αλλά καθορίζουν με σαφήνεια τα όρια της συμβολής της. Το μοντέλο εξετάζει ένα συγκεκριμένο και ελεγχόμενο ερευνητικό πρόβλημα και δημιουργεί τη βάση για μελλοντικές επεκτάσεις ως προς το πλήθος των δεδομένων, το κανονιστικό εύρος, την κλινική πληροφορία, τη risk-related μοντελοποίηση, την εξωτερική επικύρωση και τη διασύνδεση με πραγματικά συστήματα.

3.13 Επίλογος

Στο παρόν κεφάλαιο παρουσιάστηκε η μεθοδολογική προσέγγιση και ο σχεδιασμός της οντολογίας. Αποσαφηνίστηκε ότι η εργασία δεν ακολουθεί προσέγγιση μηχανικής μάθησης, αλλά αξιοποιεί πραγματικά VTaC alarm events ως instances ενός εφαρμοσμένου οντολογικού proof-of-concept. Η επιλογή αυτή συνδέεται άμεσα με τον βασικό σκοπό της διπλωματικής, ο οποίος δεν είναι η πρόβλεψη ή η αυτόματη ταξινόμηση των alarms, αλλά η σημασιολογική αναπαράσταση, η τεκμηρίωση και η ιχνηλασιμότητα πραγματικών περιστατικών.

Παρουσιάστηκαν η επιλογή των 12 πραγματικών VTaC events, τα επιλεγμένα AI Act-related σημεία αναφοράς, η βασική σημασιολογική δομή τεκμηρίωσης και το lightweight risk-related επίπεδο για τα false alarm cases. Ιδιαίτερη έμφαση δόθηκε στον ρόλο του Evidence Record ως τεκμηριωτικού κόμβου, μέσω του οποίου τα πραγματικά events μπορούν να συνδεθούν με επιλεγμένα requirement individuals χωρίς το mapping αυτό να ερμηνεύεται ως απόδειξη πλήρους νομικής συμμόρφωσης.

Η αρχιτεκτονική της οντολογίας οργανώθηκε σε τέσσερις κύριες εννοιολογικές ομάδες: Operational Concept, Evidence and Assessment Concept, Normative Concept και Risk Concept. Η διάκριση αυτή επιτρέπει να οργανώνονται διαφορετικά είδη γνώσης μέσα στο ίδιο OWL μοντέλο, διατηρώντας παράλληλα τις μεταξύ τους σημασιολογικές συνδέσεις. Το Operational μέρος περιλαμβάνει τα πραγματικά records, events, reviews και decisions, το Evidence and Assessment μέρος οργανώνει την τεκμηρίωση, το Normative μέρος περιλαμβάνει τα επιλεγμένα AI Act-related σημεία αναφοράς και το Risk μέρος παρέχει την πρόσθετη risk-related αναπαράσταση των false alarm cases.

Παράλληλα, παρουσιάστηκαν οι βασικές σχεδιαστικές επιλογές και η διαδικασία ανάπτυξης της οντολογίας στο Protégé. Ο σχεδιασμός βασίστηκε στον καθορισμό συγκεκριμένου πεδίου εφαρμογής, στη χρήση competency questions, στην ανάπτυξη των classes και properties και στη δημιουργία πραγματικών individuals. Η διαδικασία σχεδιάστηκε εξ αρχής ώστε η οντολογία να μπορεί στη συνέχεια να ελεγχθεί λειτουργικά μέσω reasoner, DL Queries και SPARQL queries. Με αυτόν τον τρόπο, υπάρχει άμεση σύνδεση μεταξύ των ερευνητικών ερωτημάτων, του σχεδιασμού της οντολογίας και της μεθόδου αξιολόγησής της.

Ιδιαίτερη σημασία έχει επίσης η σαφής οριοθέτηση της προτεινόμενης προσέγγισης. Η οντολογία δεν αποτελεί σύστημα πρόβλεψης, διαγνωστικό εργαλείο, πλήρες risk management system ή μηχανισμό απόδειξης νομικής συμμόρφωσης με τον AI Act. Πρόκειται για ένα περιορισμένο proof-of-concept, το οποίο εξετάζει αν πραγματικά healthcare monitoring events μπορούν να οργανωθούν σε ένα συνεκτικό και αναζητήσιμο σημασιολογικό μοντέλο και να συνδεθούν με την προέλευσή τους, την ανθρώπινη αξιολόγηση, τα διαθέσιμα evidence records και επιλεγμένα κανονιστικά σημεία αναφοράς.

Η μεθοδολογική αυτή βάση αποτελεί το απαραίτητο υπόβαθρο για τα επόμενα στάδια της εργασίας. Στο επόμενο κεφάλαιο παρουσιάζεται αναλυτικά η πραγματική υλοποίηση της οντολογίας στο Protégé, με τις classes, τα object properties, τα data properties και τα individuals που συγκροτούν το μοντέλο. Στη συνέχεια, η λειτουργικότητα της υλοποίησης εξετάζεται μέσω των μηχανισμών λογικού και σημασιολογικού ελέγχου που έχουν ήδη προσδιοριστεί στο παρόν κεφάλαιο.

Κεφάλαιο 4ο: Υλοποίηση της οντολογίας στο Protégé

4.1 Εισαγωγή

Στο παρόν κεφάλαιο παρουσιάζεται η πρακτική υλοποίηση της οντολογίας στο περιβάλλον Protégé [26]. Μετά τη μεθοδολογική ανάλυση και τον σχεδιασμό που παρουσιάστηκαν στο προηγούμενο κεφάλαιο, η παρούσα ενότητα εστιάζει στον τρόπο με τον οποίο το οντολογικό μοντέλο αποτυπώθηκε τεχνικά μέσα από classes, object properties, data properties, individuals και annotations.

Η υλοποίηση της οντολογίας έχει ως βασικό στόχο τη σημασιολογική αναπαράσταση πραγματικών VTaC healthcare monitoring alarm events και τη σύνδεσή τους με evidence records και επιλεγμένα AI Act-related σημεία αναφοράς. Για τον σκοπό αυτό, η οντολογία δεν περιορίζεται σε μια απλή ιεραρχία όρων. Αντίθετα, συγκροτεί ένα λειτουργικό σημασιολογικό μοντέλο, στο οποίο κάθε έννοια, κάθε σχέση και κάθε individual έχει συγκεκριμένο ρόλο.

Η τελική οντολογία περιλαμβάνει:

- **23 classes**
- **12 object properties**
- **9 data properties**
- **83 individuals**

Τα στοιχεία αυτά δεν δημιουργήθηκαν ως ανεξάρτητα ή αποσπασματικά αντικείμενα. Κάθε class, property και individual αντιστοιχεί σε συγκεκριμένη ανάγκη της μοντελοποίησης. Άλλες έννοιες χρησιμοποιούνται για την αναπαράσταση των πραγματικών VTaC events, άλλες για την expert αξιολόγηση, άλλες για τα evidence records, άλλες για τα επιλεγμένα AI Act-related σημεία αναφοράς και άλλες για το lightweight risk layer.

Το κεφάλαιο αυτό έχει διπλό ρόλο. Από τη μία πλευρά, παρουσιάζει αναλυτικά τη δομή της οντολογίας όπως υλοποιήθηκε στο Protégé. Από την άλλη πλευρά, τεκμηριώνει ότι η υλοποίηση δεν αποτελεί θεωρητική περιγραφή, αλλά ένα συγκεκριμένο οντολογικό μοντέλο που μπορεί να ελεγχθεί, να αναζητηθεί και να παρουσιαστεί με οργανωμένο τρόπο.

Η ανάλυση που ακολουθεί οργανώνεται με βάση τα βασικά δομικά στοιχεία της οντολογίας. Αρχικά παρουσιάζεται το περιβάλλον υλοποίησης και η γενική δομή του μοντέλου. Στη συνέχεια αναλύεται η χρήση annotations και ακολουθεί η αναλυτική παρουσίαση των classes, των object properties, των data properties και των individuals. Τέλος, παρουσιάζονται τα βασικά διαγράμματα της οντολογίας και γίνεται συνολική αποτίμηση της υλοποίησης.

4.2 Περιβάλλον υλοποίησης

Η οντολογία υλοποιήθηκε στο περιβάλλον Protégé, το οποίο χρησιμοποιήθηκε ως βασικό εργαλείο ανάπτυξης, οργάνωσης και τεκμηρίωσης του μοντέλου. Η επιλογή του Protégé συνδέεται με τη δυνατότητα που παρέχει για δημιουργία ιεραρχίας κλάσεων, ορισμό σχέσεων, εισαγωγή individuals, χρήση annotations, λογικό έλεγχο και οπτική παρουσίαση της οντολογικής δομής.

Στην παρούσα υλοποίηση, το Protégé χρησιμοποιείται για την ενιαία αναπαράσταση των διαφορετικών εννοιολογικών ομάδων που παρουσιάστηκαν στο προηγούμενο κεφάλαιο. Οι ομάδες αυτές δεν

αποτελούν ανεξάρτητα τεχνικά modules, αλλά τμήματα της ίδιας OWL οντολογίας και συνδέονται μεταξύ τους μέσω object και data properties [24].

Παράλληλα, το Protégé χρησιμοποιήθηκε για τον ορισμό των object properties, μέσω των οποίων συνδέονται τα individuals μεταξύ τους. Οι σχέσεις αυτές επιτρέπουν στην οντολογία να αναπαραστήσει τη διαδρομή από το waveform record μέχρι το alert event, το expert review, τη review decision, το evidence record και το compliance requirement. Επίσης, επιτρέπουν την αναπαράσταση της risk-related διαδρομής για τα false alarm cases.

Εκτός από τα object properties, στο Protégé ορίστηκαν και data properties. Τα data properties χρησιμοποιούνται για την καταγραφή τιμών που περιγράφουν συγκεκριμένα individuals, όπως identifiers, labels και λογικές ενδείξεις. Με αυτόν τον τρόπο, η οντολογία συνδυάζει δομικές σχέσεις με περιγραφικά στοιχεία, τα οποία βοηθούν στην αναζήτηση και την τεκμηρίωση.

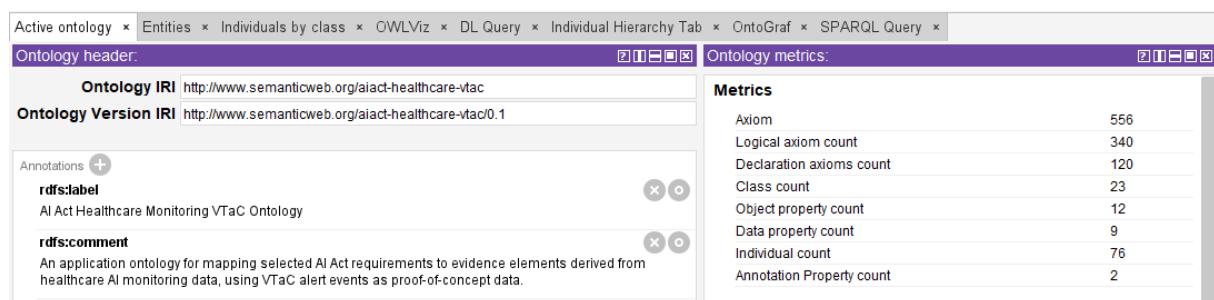
Ιδιαίτερη σημασία είχε η εισαγωγή πραγματικών individuals από το VTaC dataset [4]. Τα selected VTaC waveform records και alert events εισήχθησαν ως συγκεκριμένα instances της οντολογίας και συνδέθηκαν με expert reviews, review decisions, evidence records και selected requirements. Με αυτόν τον τρόπο, η οντολογία αποκτά εφαρμοσμένο περιεχόμενο και δεν παραμένει απλώς θεωρητική ταξινόμηση εννοιών.

Το Protégé χρησιμοποιήθηκε επίσης για την προσθήκη annotations. Τα annotations, όπως `rdfs:label` και `rdfs:comment`, βοηθούν στην κατανόηση του ρόλου κάθε στοιχείου της οντολογίας. Δεν μεταβάλλουν τη λογική λειτουργία του μοντέλου, αλλά προσφέρουν επεξηγηματικό επίπεδο που διευκολύνει την ανάγνωση, την παρουσίαση και την τεκμηρίωση της υλοποίησης.

Επιπλέον, το περιβάλλον Protégé αξιοποιήθηκε για την οπτική παρουσίαση της οντολογίας. Μέσω εργαλείων όπως το OntoGraf και το OWLViz δημιουργήθηκαν διαγράμματα που απεικονίζουν την ιεραρχία των classes, τη διάταξη των βασικών εννοιολογικών ομάδων και τη σύνδεση επιμέρους individuals. Τα διαγράμματα αυτά χρησιμοποιούνται στο κεφάλαιο ως υποστηρικτικά στοιχεία τεκμηρίωσης.

Το σχήμα παρουσιάζει το περιβάλλον υλοποίησης της οντολογίας και λειτουργεί ως εισαγωγική αποτύπωση του μοντέλου. Μέσω του Protégé οργανώθηκαν οι classes, τα object properties, τα data properties και τα individuals που χρησιμοποιούνται για τη σημασιολογική σύνδεση πραγματικών VTaC healthcare monitoring events με evidence records και επιλεγμένα AI Act-related σημεία αναφοράς.

Συνολικά, το Protégé λειτούργησε ως το βασικό περιβάλλον μέσα στο οποίο η θεωρητική και μεθοδολογική προσέγγιση της εργασίας μετατράπηκε σε υλοποιημένο οντολογικό μοντέλο. Η χρήση του επέτρεψε τη συγκροτημένη οργάνωση της γνώσης, την αναπαράσταση πραγματικών περιστατικών, τη σύνδεση με επιλεγμένες απαιτήσεις και την προετοιμασία της οντολογίας για λειτουργικό έλεγχο μέσω reasoner, DL Queries και SPARQL Queries.



The screenshot shows the Protégé interface with the 'Ontology header' and 'Metrics' panels. The 'Ontology header' panel displays the 'Ontology IRI' as <http://www.semanticweb.org/aiact-healthcare-vtac> and the 'Ontology Version IRI' as <http://www.semanticweb.org/aiact-healthcare-vtac/0.1>. Below this, there are two annotations: an `rdfs:label` 'AI Act Healthcare Monitoring VTaC Ontology' and an `rdfs:comment` 'An application ontology for mapping selected AI Act requirements to evidence elements derived from healthcare AI monitoring data, using VTaC alert events as proof-of-concept data.' The 'Metrics' panel shows a table with various counts for the ontology.

Metrics	
Axiom	556
Logical axiom count	340
Declaration axioms count	120
Class count	23
Object property count	12
Data property count	9
Individual count	76
Annotation Property count	2

Σχήμα 4.1 Γενική προβολή της οντολογίας στο Protégé.

4.3 Γενική δομή της οντολογίας

Η οντολογία που υλοποιήθηκε στο Protégé οργανώνεται γύρω από τέσσερα βασικά είδη οντολογικών στοιχείων: classes, object properties, data properties και individuals. Τα στοιχεία αυτά λειτουργούν συμπληρωματικά μεταξύ τους και επιτρέπουν τη δημιουργία ενός ενιαίου σημασιολογικού μοντέλου.

Οι classes αποτελούν το εννοιολογικό υπόβαθρο της οντολογίας. Μέσω αυτών ορίζονται οι βασικές κατηγορίες εννοιών που χρησιμοποιούνται για την αναπαράσταση των healthcare monitoring events, της αξιολόγησής τους, της τεκμηρίωσης, των επιλεγμένων απαιτήσεων και των risk-related στοιχείων.

Τα object properties χρησιμοποιούνται για να αποτυπώσουν τις σχέσεις ανάμεσα στα individuals. Μέσω αυτών δημιουργείται η βασική σημασιολογική διαδρομή από το waveform record προς το alert event, το expert review, τη review decision, το evidence record και το compliance requirement. Επίσης, μέσω object properties αναπαρίσταται η πρόσθετη risk-related διαδρομή για τα false alarm cases.

Τα data properties χρησιμοποιούνται για την αποθήκευση συγκεκριμένων τιμών που περιγράφουν τα individuals. Τέτοιες τιμές μπορεί να είναι identifiers, labels ή λογικές ενδείξεις, όπως αν ένα event είναι logged, traceable ή risk-relevant. Με αυτόν τον τρόπο, η οντολογία δεν περιέχει μόνο σχέσεις μεταξύ εννοιών, αλλά και περιγραφικά δεδομένα που υποστηρίζουν την τεκμηρίωση.

Τα individuals αποτελούν τις συγκεκριμένες περιπτώσεις που εισάγονται στην οντολογία. Σε αυτά περιλαμβάνονται τα πραγματικά VTaC waveform records, τα VTaC alert events, τα expert reviews, οι review decisions, τα evidence records, τα requirement individuals, τα control mechanisms και τα risk-related individuals. Η ύπαρξη των individuals μετατρέπει την οντολογία από απλό θεωρητικό σχήμα σε εφαρμοσμένο μοντέλο με πραγματικά παραδείγματα.

Η τελική υλοποίηση της οντολογίας περιλαμβάνει συνολικά 23 classes, 12 object properties, 9 data properties και 83 individuals. Η κατανομή αυτή παρουσιάζεται συνοπτικά στον Πίνακα 4.1.

Στο επίπεδο των classes, η οντολογία οργανώνεται κάτω από τη γενική ρίζα owl:Thing, η οποία αποτελεί την ανώτερη προκαθορισμένη κλάση κάθε OWL οντολογίας. Η owl:Thing δεν αποτελεί θεματική κλάση της παρούσας υλοποίησης. Λειτουργεί ως τεχνικό σημείο εκκίνησης της ιεραρχίας, κάτω από το οποίο τοποθετούνται οι θεματικές κλάσεις του μοντέλου.

Οι 23 θεματικές classes οργανώνονται σε τέσσερις κύριες εννοιολογικές ομάδες:

- Evidence and Assessment Concept
- Normative Concept
- Operational Concept
- Risk Concept

Πίνακας 4-1 Συνοπτική δομή της υλοποιημένης οντολογίας.

Δομικό στοιχείο	Πλήθος	Ρόλος στην οντολογία
Classes	23	Ορίζουν τις βασικές κατηγορίες εννοιών του μοντέλου.
Object properties	12	Συνδέουν τα individuals μεταξύ τους και δημιουργούν τις σημασιολογικές διαδρομές.
Data properties	9	Αποθηκεύουν περιγραφικές τιμές, labels και λογικές ενδείξεις.
Individuals	76	Αναπαριστούν τα πραγματικά VTaC instances και τα στοιχεία τεκμηρίωσης, απαιτήσεων και κινδύνου.

Η οργάνωση αυτή αντιστοιχεί στις τέσσερις εννοιολογικές ομάδες που παρουσιάστηκαν στο προηγούμενο κεφάλαιο. Το Operational Concept συγκεντρώνει τις έννοιες που σχετίζονται με τα πραγματικά healthcare monitoring events. Το Evidence and Assessment Concept περιλαμβάνει τις έννοιες τεκμηρίωσης και αξιολόγησης. Το Normative Concept περιλαμβάνει τις απαιτήσεις αναφοράς. Το Risk Concept περιλαμβάνει τις έννοιες που σχετίζονται με την αναπαράσταση κινδύνου.

Η γενική δομή της οντολογίας δεν βασίζεται μόνο στην ιεραρχία των classes. Η λειτουργικότητα του μοντέλου προκύπτει κυρίως από τη σύνδεση των individuals μέσω object properties. Για παράδειγμα, ένα VTaC alert event δεν αποκτά σημασιολογική αξία μόνο από την ταξινόμησή του ως instance της class Ventricular Tachycardia Alarm. Συνδέεται με το waveform record προέλευσής του, με την αντίστοιχη Expert Review και Review Decision και, παράλληλα, με το Evidence Record που το τεκμηριώνει. Μέσω του Evidence Record πραγματοποιείται στη συνέχεια το mapping προς τα επιλεγμένα AI Act-related requirement individuals.

Το Evidence Record λειτουργεί ως τεκμηριωτικός κόμβος μεταξύ του operational μέρους του μοντέλου και των επιλεγμένων AI Act-related requirement individuals. Αυτή η σχεδιαστική επιλογή είναι κεντρική για το requirement-to-evidence mapping της εργασίας.

Η γενική δομή της οντολογίας μπορεί να αποδοθεί συνοπτικά ως εξής:

Classes → ορίζουν τις έννοιες

Object properties → ορίζουν τις σχέσεις

Data properties → προσθέτουν τιμές και περιγραφές

Individuals → υλοποιούν πραγματικές περιπτώσεις

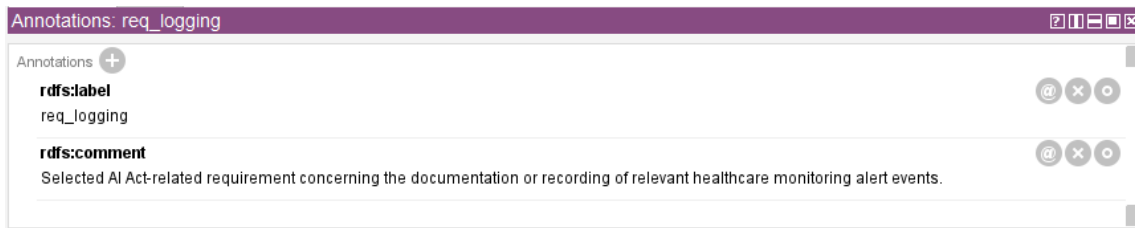
Με αυτή τη δομή, η οντολογία μπορεί να υποστηρίξει τόσο την αναπαράσταση γνώσης όσο και τη λειτουργική αναζήτηση. Οι classes παρέχουν το εννοιολογικό πλαίσιο, τα properties παρέχουν τη συνδεσιμότητα και τα individuals παρέχουν το εφαρμοσμένο περιεχόμενο. Έτσι, το μοντέλο μπορεί να ελεγχθεί στη συνέχεια μέσω reasoner, DL Queries και SPARQL Queries.

Η γενική δομή της οντολογίας δείχνει ότι η υλοποίηση δεν αποτελεί απλή καταγραφή όρων στο Protégé. Αντίθετα, αποτελεί οργανωμένο σημασιολογικό μοντέλο, στο οποίο οι έννοιες, οι σχέσεις, οι τιμές και τα πραγματικά instances συνεργάζονται για να υποστηρίξουν ιχνηλασιμότητα, evidence-based τεκμηρίωση και σύνδεση με επιλεγμένα AI Act-related σημεία αναφοράς.

4.4 Χρήση annotations στην οντολογία

Εκτός από τις classes, τα object properties, τα data properties και τα individuals, στην οντολογία χρησιμοποιήθηκαν και annotations [24]. Τα annotations δεν αποτελούν λογικές σχέσεις με την ίδια έννοια που αποτελούν τα object properties ή τα data properties. Δεν χρησιμοποιούνται για να συνδέσουν δύο individuals μεταξύ τους, ούτε για να ορίσουν περιορισμούς που αξιοποιεί ο reasoner. Ο ρόλος τους είναι κυρίως επεξηγηματικός και τεκμηριωτικός.

Στο Protégé, τα annotations επιτρέπουν την προσθήκη ανθρώπινα αναγνώσιμων πληροφοριών σε στοιχεία της οντολογίας, όπως classes, properties και individuals. Με αυτόν τον τρόπο, κάθε στοιχείο μπορεί να συνοδεύεται από μια σύντομη περιγραφή, ετικέτα ή σχόλιο που εξηγεί τον ρόλο του μέσα στο μοντέλο.



Σχήμα 4.2 Παράδειγμα χρήσης annotations σε requirement individual.

Στην παρούσα υλοποίηση χρησιμοποιήθηκαν κυρίως τα annotations `rdfs:label` και `rdfs:comment`. Το `rdfs:label` χρησιμοποιείται για να δώσει μια καθαρή και αναγνώσιμη ονομασία στο στοιχείο της οντολογίας. Το `rdfs:comment` χρησιμοποιείται για να εξηγήσει σύντομα τι σημαίνει το συγκεκριμένο στοιχείο και ποιος είναι ο ρόλος του στο μοντέλο.

Για παράδειγμα, ένα requirement individual όπως το `req_logging` συνοδεύεται από annotation που διευκρινίζει ότι σχετίζεται στο proof-of-concept με το AI Act-related σημείο αναφοράς για record-keeping/logging και με τη διαθέσιμη τεκμηρίωση των healthcare monitoring alert events. Αντίστοιχα, annotations χρησιμοποιούνται σε classes και properties, ώστε ο αναγνώστης να μπορεί να κατανοήσει ευκολότερα τη λειτουργία κάθε στοιχείου χωρίς να χρειάζεται να ερμηνεύσει μόνο το τεχνικό του όνομα.

Η χρήση annotations είναι σημαντική για τρεις βασικούς λόγους. Πρώτον, βελτιώνει την αναγνωσιμότητα της οντολογίας μέσα στο Protégé. Ο αναγνώστης δεν βλέπει μόνο τεχνικά ονόματα, όπως `req_traceability` ή `risk_management_control`, αλλά μπορεί να διαβάσει και μια σύντομη περιγραφή της σημασίας τους.

Δεύτερον, τα annotations βοηθούν στην τεκμηρίωση των σχεδιαστικών επιλογών. Μέσω αυτών μπορεί να εξηγηθεί γιατί υπάρχει ένα συγκεκριμένο στοιχείο στην οντολογία και πώς χρησιμοποιείται στο πλαίσιο του μοντέλου. Αυτό είναι ιδιαίτερα χρήσιμο σε μια διπλωματική εργασία, όπου δεν αρκεί να παρουσιαστεί μόνο το τελικό τεχνικό αποτέλεσμα, αλλά πρέπει να είναι σαφής και η λογική πίσω από αυτό.

Τρίτον, τα annotations ενισχύουν τη δυνατότητα παρουσίασης και ελέγχου της οντολογίας. Λειτουργούν ως συνοδευτικό επεξηγηματικό επίπεδο, το οποίο διευκολύνει τόσο την ανάγνωση της οντολογίας όσο και την αξιολόγησή της από τρίτους. Έτσι, η οντολογία γίνεται πιο κατανοητή όχι μόνο για τον δημιουργό της, αλλά και για έναν αναγνώστη που την εξετάζει για πρώτη φορά.

Ιδιαίτερη σημασία έχουν τα annotations στα requirement individuals. Οι απαιτήσεις `req_logging`, `req_human_review`, `req_traceability` και `req_risk_management` δεν είναι απλές ονομασίες. Κάθε μία αντιστοιχεί σε επιλεγμένη AI Act-related απαίτηση που χρησιμοποιείται στο μοντέλο για requirement-to-evidence mapping. Τα annotations βοηθούν ώστε να γίνεται σαφές τι αντιπροσωπεύει κάθε απαίτηση και γιατί συνδέεται με συγκεκριμένα evidence records ή control mechanisms.

Πρέπει επίσης να τονιστεί ότι τα annotations δεν αντικαθιστούν τη λογική δομή της οντολογίας. Η σημασιολογική λειτουργία του μοντέλου βασίζεται στις classes, στα object properties, στα data properties και στα individuals. Τα annotations λειτουργούν συμπληρωματικά, ως επίπεδο τεκμηρίωσης και ερμηνείας. Με άλλα λόγια, δεν αλλάζουν τα λογικά συμπεράσματα του reasoner, αλλά κάνουν την οντολογία πιο κατανοητή, ελέγξιμη και παρουσιάσιμη.

Τα annotations χρησιμοποιούνται αποκλειστικά ως επεξηγηματικό και τεκμηριωτικό επίπεδο της υλοποίησης. Δεν αποτελούν λογικά axioms, δεν δημιουργούν από μόνα τους inferred γνώση και δεν

χρησιμοποιούνται ως evidence για την υποστήριξη των AI Act-related mappings. Οι λειτουργικές σημασιολογικές συνδέσεις του μοντέλου υλοποιούνται μέσω των classes, object properties, data properties και individuals που παρουσιάζονται στις επόμενες ενότητες.

Συνεπώς, η χρήση annotations συμβάλλει στη συνολική ποιότητα της υλοποίησης. Η οντολογία δεν παρουσιάζεται μόνο ως τεχνικό αρχείο OWL, αλλά ως τεκμηριωμένο μοντέλο γνώσης, στο οποίο τα βασικά στοιχεία συνοδεύονται από επεξηγήσεις που διευκολύνουν την κατανόηση από τεχνικούς και μη τεχνικούς αναγνώστες.

4.5 Classes της οντολογίας

Οι classes αποτελούν το βασικό εννοιολογικό υπόβαθρο της οντολογίας. Μέσω αυτών ορίζονται οι βασικές κατηγορίες εννοιών που χρειάζονται για την αναπαράσταση των healthcare monitoring alarm events, της αξιολόγησής τους, της τεκμηρίωσης και της σύνδεσής τους με επιλεγμένες απαιτήσεις συμμόρφωσης.

Στο Protégé, όλες οι classes εμφανίζονται κάτω από τη γενική ρίζα owl:Thing, η οποία αποτελεί την ανώτερη προκαθορισμένη κλάση κάθε OWL οντολογίας. Η owl:Thing δεν αποτελεί θεματική κλάση της υλοποίησης, αλλά το τεχνικό σημείο εκκίνησης της ιεραρχίας. Κάτω από αυτήν οργανώνονται οι θεματικές classes που συγκροτούν το μοντέλο.

Η οντολογία περιλαμβάνει 23 θεματικές classes, οργανωμένες σε τέσσερις κύριες εννοιολογικές ομάδες:

- Evidence and Assessment Concept
- Normative Concept
- Operational Concept
- Risk Concept

Ο διαχωρισμός αυτός βοηθά ώστε κάθε έννοια να τοποθετείται στο κατάλληλο επίπεδο. Άλλες classes περιγράφουν πραγματικά συμβάντα και waveform records, άλλες αναπαριστούν διαδικασίες αξιολόγησης, άλλες οργανώνουν τεκμήρια και απαιτήσεις, ενώ άλλες χρησιμοποιούνται για τη risk-related αναπαράσταση των false alarm cases.

Η ιεραρχία των classes δεν λειτουργεί ως απλή ταξινόμηση όρων. Αντίθετα, αποτελεί τη βάση πάνω στην οποία στηρίζονται οι object properties, τα data properties και τα individuals. Οι classes ορίζουν τις βασικές κατηγορίες του μοντέλου, ενώ οι σχέσεις και τα individuals δίνουν λειτουργικό περιεχόμενο σε αυτές τις κατηγορίες.

Για παράδειγμα, η class Alert Event ορίζει τη γενική έννοια ενός alarm event. Όμως η πλήρης σημασία ενός συγκεκριμένου event προκύπτει όταν αυτό εισαχθεί ως individual και συνδεθεί με Waveform Record, Expert Review, Review Decision, Evidence Record και Compliance Requirement. Επομένως, οι classes είναι το πρώτο επίπεδο της οντολογίας, αλλά αποκτούν πλήρη λειτουργικότητα μέσα από τις σχέσεις και τα instances.

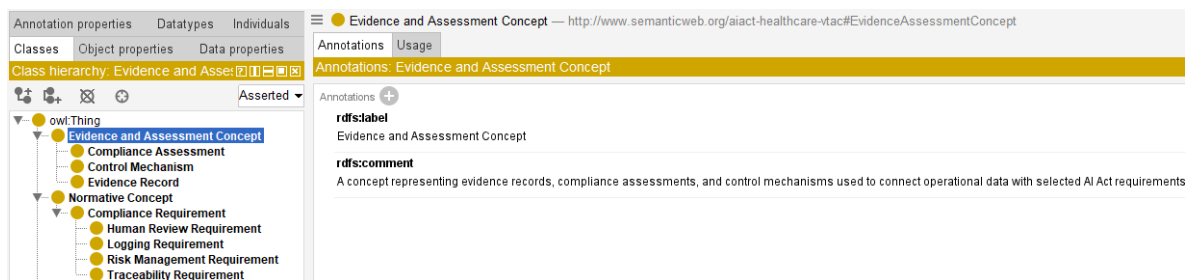
Στις επόμενες υποενότητες παρουσιάζονται οι classes της οντολογίας με βάση την οργάνωσή τους στο Protégé. Η ανάλυση ξεκινά από το Evidence and Assessment Concept, επειδή το επίπεδο αυτό περιλαμβάνει τον βασικό μηχανισμό τεκμηρίωσης της οντολογίας, και συνεχίζει με τις υπόλοιπες εννοιολογικές ομάδες.

4.5.1 Evidence and Assessment Concept

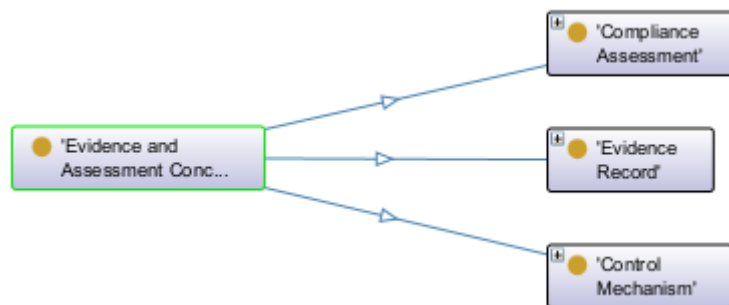
Η class Evidence and Assessment Concept λειτουργεί ως γενική κατηγορία για τις έννοιες που σχετίζονται με την τεκμηρίωση και την αξιολόγηση. Δεν αναπαριστά ένα συγκεκριμένο τεκμήριο ή μία συγκεκριμένη αξιολόγηση. Αντίθετα, χρησιμοποιείται ως ανώτερη οργανωτική class κάτω από την οποία τοποθετούνται έννοιες όπως το Evidence Record, το Compliance Assessment και το Control Mechanism.

Ο ρόλος της είναι να διαχωρίζει το επίπεδο της τεκμηρίωσης από το operational επίπεδο των πραγματικών alarm events. Με αυτόν τον τρόπο, η οντολογία δείχνει ότι άλλο πράγμα είναι το ίδιο το συμβάν και άλλο πράγμα η τεκμηρίωση ή η αξιολόγηση που συνδέεται με αυτό μέσα στο μοντέλο.

Η ύπαρξη της class αυτής είναι σημαντική, επειδή η παρούσα οντολογία δεν περιορίζεται στην απλή αναπαράσταση των δεδομένων του VTaC. Δημιουργεί επιπλέον ένα επίπεδο evidence-based documentation, το οποίο επιτρέπει τη σύνδεση πραγματικών events με evidence records και selected requirements. Έτσι, το Evidence and Assessment Concept λειτουργεί ως η γενική ομπρέλα του μέρους της οντολογίας που σχετίζεται με την απόδειξη, την αξιολόγηση και την υποστήριξη requirement mappings.



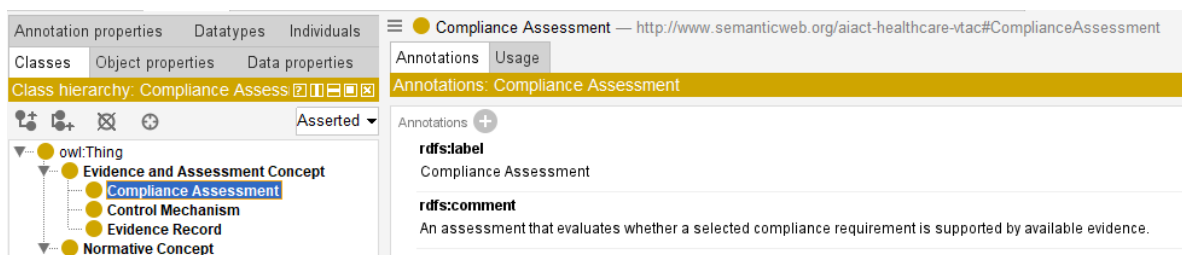
Σχήμα 4.3 H class Evidence and Assessment Concept και τα annotations της στο Protégé.



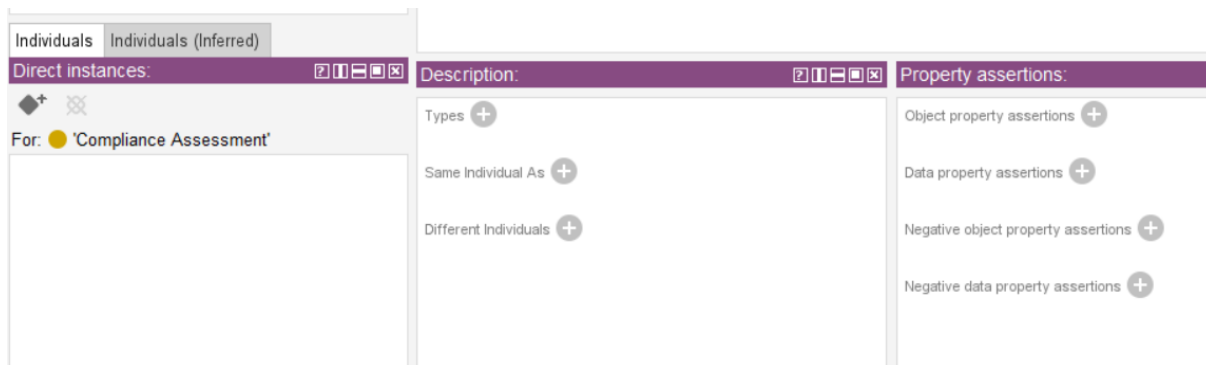
Σχήμα 4.4 H class Evidence and Assessment Concept και οι βασικές υποκλάσεις της.

4.5.2 Compliance Assessment

Η class Compliance Assessment αναπαριστά σε εννοιολογικό επίπεδο την αξιολόγηση διαθέσιμης τεκμηρίωσης σε σχέση με ένα event και ένα επιλεγμένο AI Act-related σημείο αναφοράς. Δεν χρησιμοποιείται για να δηλώσει πλήρη νομική συμμόρφωση, αλλά λειτουργεί ως δομικό σημείο της οντολογίας για πιθανή αναπαράσταση διαδικασιών αξιολόγησης evidence.



Σχήμα 4.5 Η class Compliance Assessment και τα annotations της στο Protégé.



Σχήμα 4.6 Η class Compliance Assessment χωρίς direct instances στην παρούσα proof-of-concept υλοποίηση.

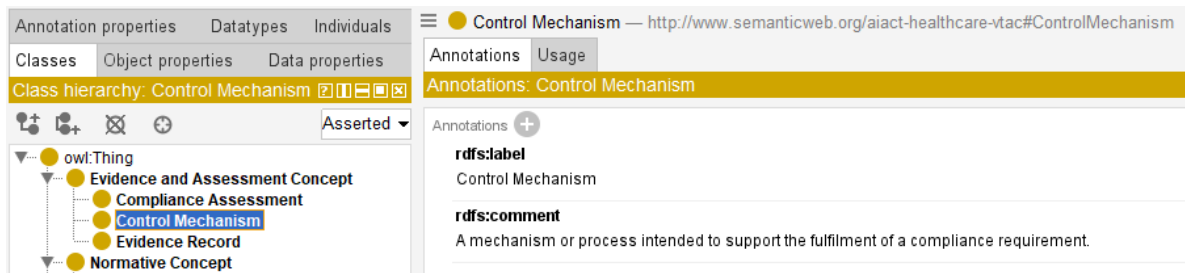
Στην παρούσα proof-of-concept έκδοση δεν έχουν δημιουργηθεί direct individuals της class Compliance Assessment. Η class διατηρείται στο μοντέλο ώστε να υποστηρίζει τη σχεδιαστική λογική των properties `assessesEvent`, `assessesRequirement`, `usesEvidence` και `assessmentOutcome` και να αποτελεί σημείο επέκτασης για μελλοντικές υλοποιήσεις στις οποίες θα αναπαρίστανται ρητά assessment instances.

Με αυτόν τον τρόπο, η Compliance Assessment έχει κυρίως εννοιολογικό και επεκτάσιμο ρόλο στην τρέχουσα ontology και δεν συμμετέχει άμεσα στα instance-level mappings που χρησιμοποιούνται στα 12 VTaC events.

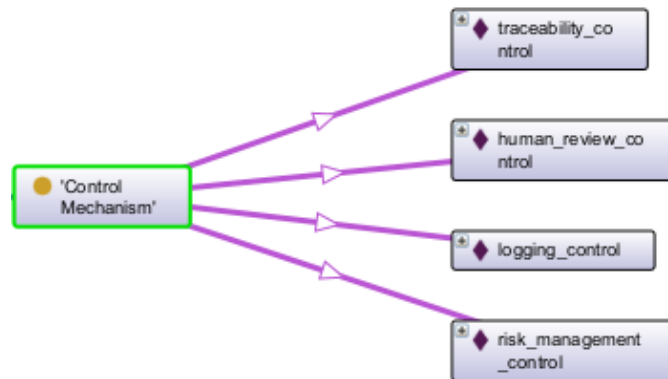
4.5.3 Control Mechanism

Η class Control Mechanism αναπαριστά μηχανισμούς ελέγχου ή οργανωτικά/τεχνικά μέτρα που μπορούν να συνδεθούν με την τεκμηρίωση και τις επιλεγμένες απαιτήσεις. Σε ένα σύστημα συμμόρφωσης, ένα control mechanism μπορεί να αντιστοιχεί σε μηχανισμό logging, διαδικασία ανθρώπινης αξιολόγησης, μηχανισμό traceability ή διαδικασία risk monitoring.

Στο συγκεκριμένο μοντέλο, η class αυτή λειτουργεί ως έννοια που υποστηρίζει τη μετάβαση από την απαίτηση στην πρακτική εφαρμογή της. Για παράδειγμα, η απαίτηση για human review δεν είναι μόνο μια θεωρητική απαίτηση. Χρειάζεται κάποιος μηχανισμός ή διαδικασία που να δείχνει πώς αυτή η απαίτηση υποστηρίζεται στην πράξη.



Σχήμα 4.7 Η class Control Mechanism και τα annotations της στο Protégé.



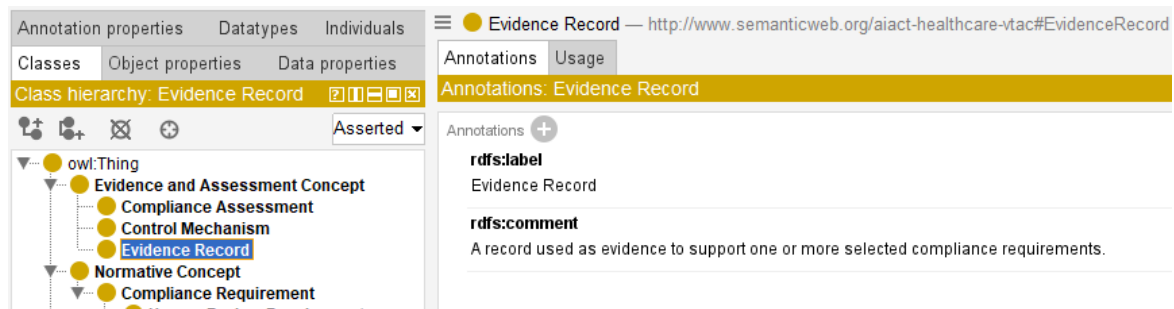
Σχήμα 4.8 Η class Control Mechanism με τα individuals της.

Η χρήση της class Control Mechanism βοηθά να φανεί ότι η οντολογία δεν καταγράφει μόνο requirements, αλλά μπορεί να οργανώσει και τα μέσα ή τις διαδικασίες που τα υποστηρίζουν. Με αυτόν τον τρόπο, το μοντέλο δεν περιορίζεται στο να δηλώνει ότι υπάρχει μια απαίτηση, αλλά μπορεί να συνδέει την απαίτηση με στοιχεία που δείχνουν πώς αυτή μπορεί να υποστηριχθεί.

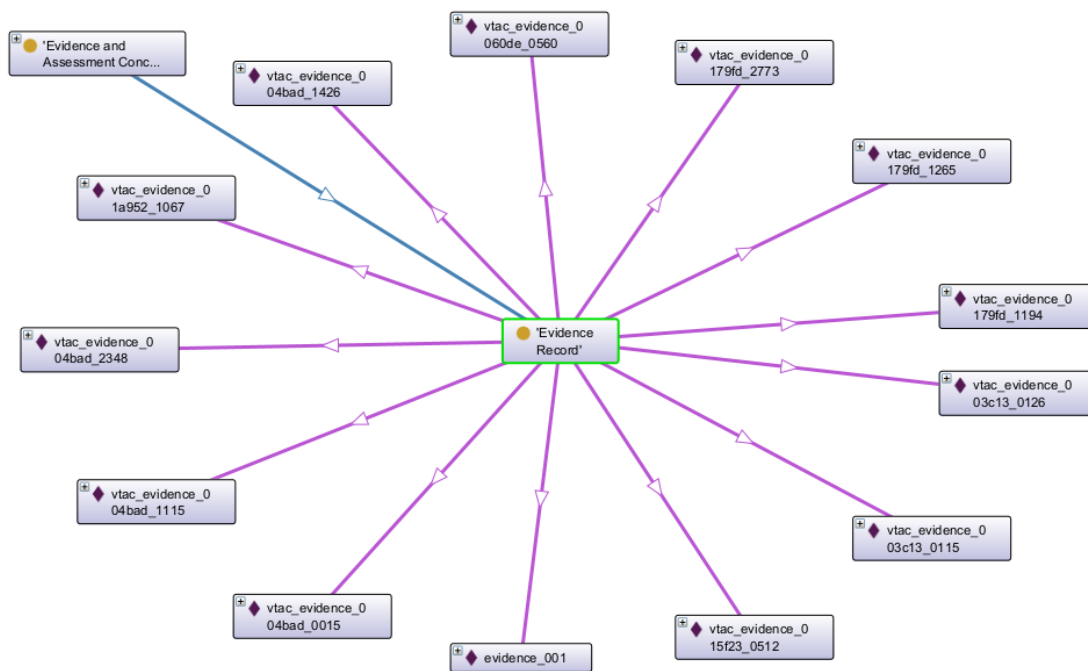
Στην παρούσα proof-of-concept υλοποίηση, τα control mechanism individuals χρησιμοποιούνται σε σχέση με τις επιλεγμένες απαιτήσεις της οντολογίας. Επομένως, λειτουργούν ως βοηθητικά στοιχεία που ενισχύουν το requirement-to-evidence mapping και υποστηρίζουν τη συνολική τεκμηριωτική λογική του μοντέλου.

4.5.4 Evidence Record

Η class Evidence Record αποτελεί μία από τις κεντρικότερες classes της οντολογίας. Αναπαριστά τον τεκμηριωτικό κόμβο που συνδέεται με ένα συγκεκριμένο alarm event και οργανώνει τη διαθέσιμη πληροφορία που χρησιμοποιείται στο requirement-to-evidence mapping. Το Evidence Record λειτουργεί ως γέφυρα μεταξύ του operational μέρους της οντολογίας και των επιλεγμένων AI Act-related requirement individuals.



Σχήμα 4.9 Η class Evidence Record και τα annotations της στο Protégé.



Σχήμα 4.10 Η class Evidence Record με τα individuals τεκμηρίωσης.

Ο ρόλος της class είναι να αποφεύγεται η απευθείας σύνδεση ενός alarm event με ένα κανονιστικό σημείο αναφοράς. Ένα event αποτελεί το πραγματικό περιστατικό του healthcare monitoring use case, ενώ το Evidence Record αποτελεί το διακριτό στοιχείο τεκμηρίωσης μέσω του οποίου πραγματοποιείται το mapping προς τα requirement individuals.

Στην πραγματική υλοποίηση, η βασική σχέση αποτυπώνεται ως εξής:

Alert Event --isDocumentedBy→ Evidence Record --supportsRequirement→ Requirement individual

Παράλληλα, η πληροφορία για την ανθρώπινη αξιολόγηση του ίδιου event αναπαρίσταται μέσω διαφορετικής σημασιολογικής διαδρομής:

Alert Event --hasReview→ Expert Review --hasDecision→ Review Decision

Οι δύο αυτές διαδρομές συνδυάζονται σημασιολογικά γύρω από το ίδιο alarm event, χωρίς να απαιτείται άμεση object-property σχέση μεταξύ Review Decision και Evidence Record.

Με αυτόν τον τρόπο, το Evidence Record επιτρέπει τη σαφή διάκριση μεταξύ πραγματικού περιστατικού, διαθέσιμης αξιολόγησης, τεκμηρίωσης και AI Act-related mapping. Η ύπαρξη του mapping δεν αποτελεί απόδειξη πλήρους νομικής συμμόρφωσης, αλλά στοιχείο της proof-of-concept σημασιολογικής αναπαράστασης.

4.5.5 Normative Concept

Η class Normative Concept λειτουργεί ως γενική κατηγορία για τις έννοιες που σχετίζονται με απαιτήσεις, κανόνες ή κανονιστικά σημεία αναφοράς. Στην ιεραρχία της οντολογίας, κάτω από αυτήν τοποθετείται η class Compliance Requirement.

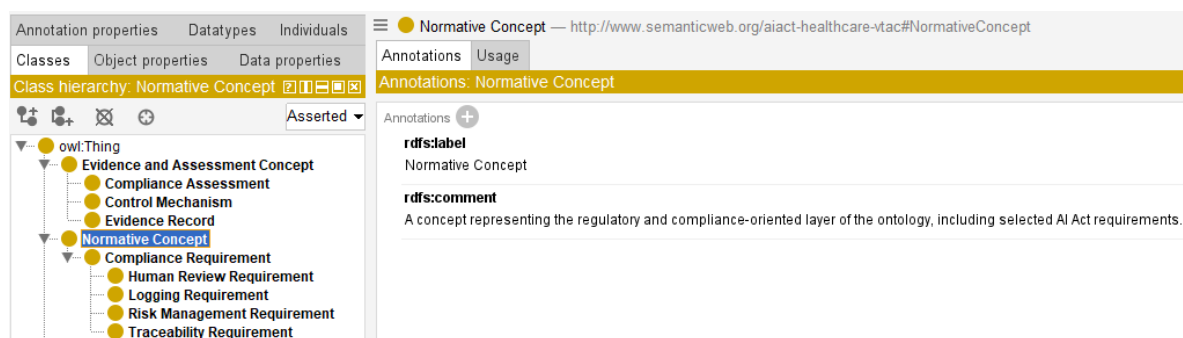
Ο ρόλος της είναι να διαχωρίζει τις κανονιστικές έννοιες από τα πραγματικά δεδομένα και τα evidence records. Με αυτόν τον τρόπο, γίνεται σαφές ότι τα requirements δεν είναι operational events, αλλά ανήκουν σε διαφορετικό επίπεδο νοήματος. Ένα alarm event αποτελεί πραγματικό περιστατικό του healthcare monitoring περιβάλλοντος, ενώ ένα requirement αποτελεί επιλεγμένο σημείο αναφοράς με το οποίο μπορεί να συσχετιστεί η τεκμηρίωση.

Η ύπαρξη της class Normative Concept είναι σημαντική επειδή επιτρέπει στην οντολογία να διαθέτει ξεχωριστό επίπεδο για τις απαιτήσεις. Έτσι, το μοντέλο μπορεί να υποστηρίξει τη χαρτογράφηση από evidence records προς επιλεγμένες AI Act-related έννοιες, χωρίς να συγχέει τα πραγματικά events με τις απαιτήσεις που αυτά ενδέχεται να υποστηρίζουν μέσω τεκμηρίωσης.

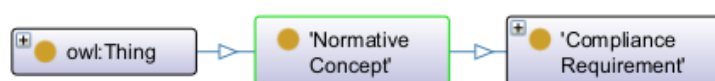
4.5.6 Compliance Requirement

Η class Compliance Requirement αναπαριστά τη γενική έννοια της απαίτησης συμμόρφωσης. Κάτω από αυτήν βρίσκονται οι ειδικότερες απαιτήσεις που χρησιμοποιούνται στο μοντέλο: Human Review Requirement, Logging Requirement, Risk Management Requirement και Traceability Requirement.

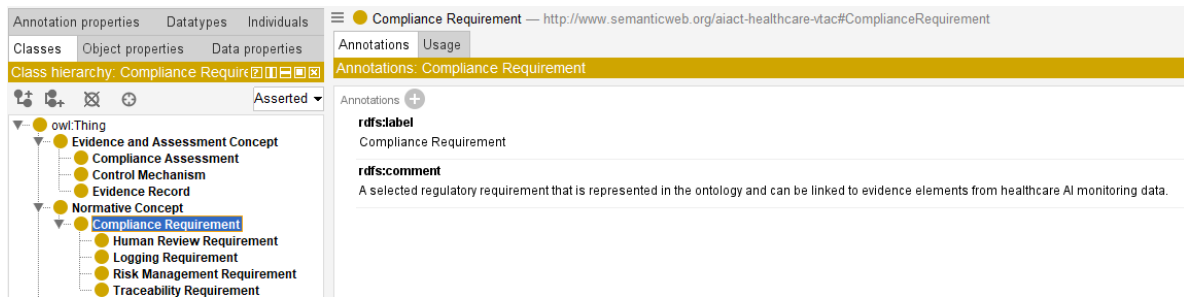
Ο ρόλος της είναι να παρέχει μια κοινή υπερκλάση για όλες τις επιλεγμένες απαιτήσεις. Αυτό επιτρέπει στην οντολογία να αναζητά είτε όλες τις απαιτήσεις συνολικά είτε συγκεκριμένους τύπους απαιτήσεων. Για παράδειγμα, ένα query μπορεί να ζητήσει όλα τα evidence records που συνδέονται με οποιαδήποτε Compliance Requirement, χωρίς να χρειάζεται να αναφέρει ξεχωριστά κάθε ειδικότερη απαίτηση.



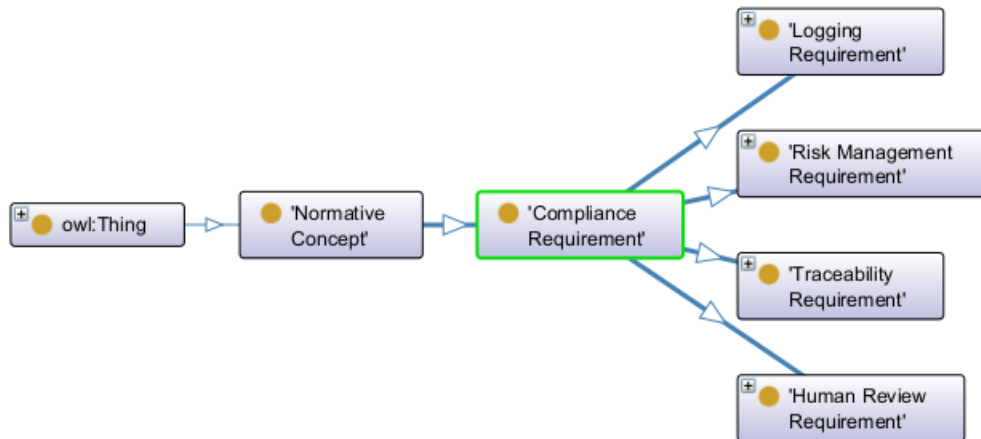
Σχήμα 4.11 Η class Normative Concept και τα annotations της στο Protégé.



Σχήμα 4.12 Θέση της class Normative Concept στην ιεραρχία της οντολογίας.



Σχήμα 4.13 Η class Compliance Requirement και τα annotations της στο Protégé.



Σχήμα 4.14 Υποκλάσεις της class Compliance Requirement.

Η class αυτή είναι κρίσιμη για το requirement-to-evidence mapping της οντολογίας. Τα evidence records δεν συνδέονται με απαιτήσεις με αόριστο τρόπο, αλλά με συγκεκριμένα requirement individuals που ανήκουν σε ειδικότερες classes. Έτσι, το μοντέλο μπορεί να διακρίνει αν ένα evidence record υποστηρίζει logging, human review, traceability ή risk management.

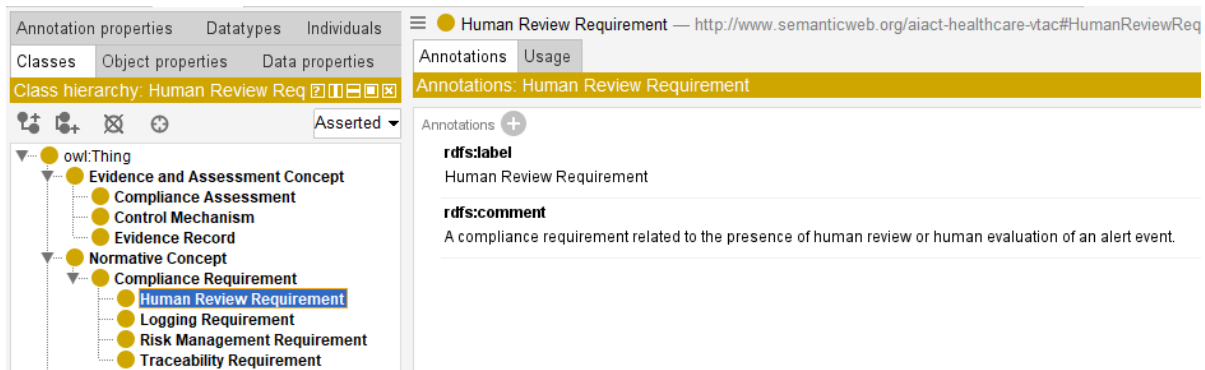
Πρέπει να τονιστεί ότι η class Compliance Requirement δεν χρησιμοποιείται για να δηλώσει πλήρη νομική συμμόρφωση. Χρησιμοποιείται ως οντολογική αναπαράσταση επιλεγμένων απαιτήσεων που σχετίζονται με τον AI Act και μπορούν να συνδεθούν με τα διαθέσιμα στοιχεία του συγκεκριμένου use case.

Στο πλαίσιο της παρούσας εφαρμογής, οι ειδικότερες classes χρησιμοποιούνται για την οργάνωση των τεσσάρων εφαρμοσμένων AI Act-related σημείων αναφοράς και δεν αποτελούν πλήρη αναπαράσταση του συνόλου των νομικών απαιτήσεων του Κανονισμού.

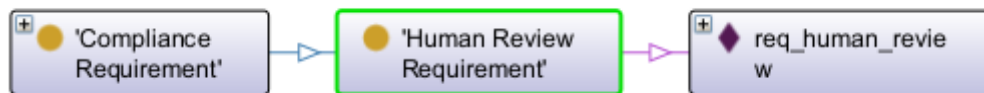
4.5.7 Human Review Requirement

Η class Human Review Requirement αναπαριστά στο proof-of-concept το AI Act-related σημείο αναφοράς που αφορά την ανθρώπινη διάσταση της αξιολόγησης. Στην παρούσα εφαρμογή συνδέεται με το requirement individual `req_human_review`.

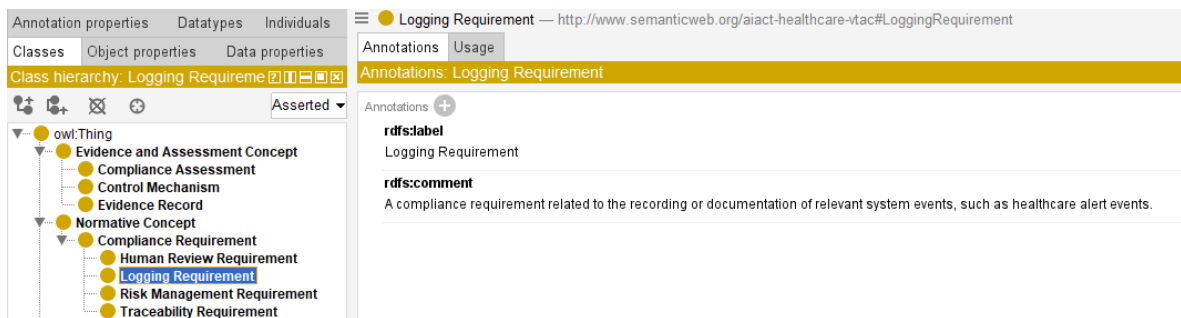
Κάθε επιλεγμένο VTaC event συνδέεται με αντίστοιχο Expert Review και Review Decision, καθώς οι τελικές αποφάσεις true ή false προέρχονται από την ανθρώπινη διαδικασία αξιολόγησης του dataset [4]. Μέσω των Evidence Records, η πληροφορία αυτή μπορεί να χαρτογραφηθεί στο `req_human_review`.



Σχήμα 4.15 H class Human Review Requirement και τα annotations της στο Protégé.



Σχήμα 4.16 H class Human Review Requirement και το αντίστοιχο requirement individual.



Σχήμα 4.17 H class Logging Requirement και τα annotations της στο Protégé.

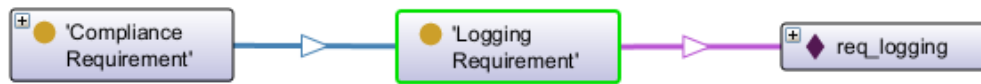
Η συγκεκριμένη αναπαράσταση δεν πρέπει να ταυτίζεται με πλήρη εφαρμογή του human oversight που προβλέπεται στο άρθρο 14 του AI Act [3]. Το VTaC Expert Review αποτελεί ανθρώπινη αξιολόγηση των alarm events, ενώ η ανθρώπινη εποπτεία ενός πραγματικού high-risk AI system αποτελεί ευρύτερη τεχνική και οργανωτική απαίτηση.

4.5.8 Logging Requirement

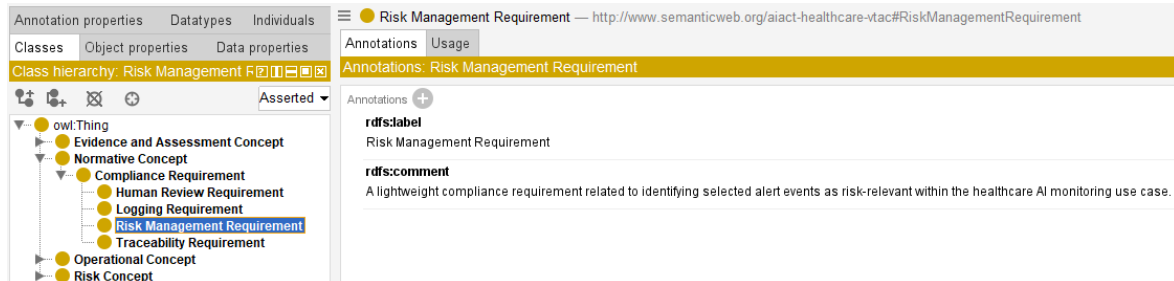
Η class Logging Requirement αναπαριστά στο proof-of-concept το AI Act-related σημείο αναφοράς που σχετίζεται με record-keeping και logging. Στην παρούσα υλοποίηση συνδέεται με το requirement individual req_logging.

Η επιλογή της σχετίζεται με το γεγονός ότι τα VTaC alarm events και τα waveform records διαθέτουν αναγνωρίσιμα identifiers [4], τα οποία μπορούν να αναπαρασταθούν και να ανακτηθούν μέσα στην οντολογία. Η διαθέσιμη αυτή πληροφορία χρησιμοποιείται ως στοιχείο τεκμηρίωσης σχετικό με τη λογική της καταγραφής.

Η class Logging Requirement λειτουργεί επομένως ως κανονιστικό σημείο αναφοράς για τη σημασιολογική οργάνωση διαθέσιμης καταγεγραμμένης πληροφορίας. Η ύπαρξη event identifiers και σχετικών evidence records στο proof-of-concept δεν αποδεικνύει από μόνη της την πλήρη automatic logging capability που προβλέπει το άρθρο 12 για ένα πραγματικό high-risk AI system [3].



Σχήμα 4.18 H class Logging Requirement και το αντίστοιχο requirement individual.



Σχήμα 4.19 H class Risk Management Requirement και τα annotations της στο Protégé.



Σχήμα 4.20 H class Risk Management Requirement και το αντίστοιχο requirement individual.

4.5.9 Risk Management Requirement

Η class Risk Management Requirement αναπαριστά στο proof-of-concept το AI Act-related σημείο αναφοράς που σχετίζεται με τη διαχείριση κινδύνων. Στην παρούσα υλοποίηση συνδέεται κυρίως με τα false alarm cases και με το lightweight risk-related μέρος της οντολογίας.

Το αντίστοιχο requirement individual `req_risk_management` μπορεί να υποστηρίζεται από τα evidence records των false alarm cases. Παράλληλα, τα συγκεκριμένα events διαθέτουν πρόσθετη risk-related αναπαράσταση μέσω Risk Indicator, Risk-Relevant Event και Potential Harm.

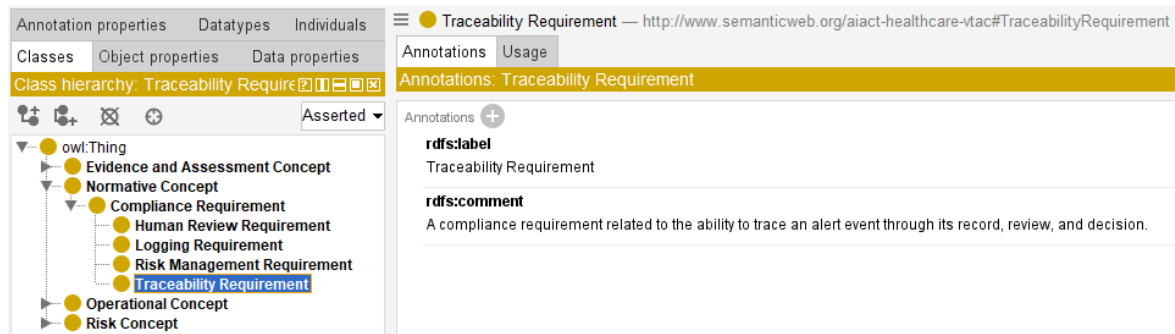
Η επιλογή αυτή επιτρέπει να αναπαρασταθεί σημασιολογικά η πρόσθετη risk-related σημασία των false alarms, χωρίς να υποστηρίζεται ότι κάθε false alarm προκαλεί πραγματική βλάβη ή ότι η ontology υλοποιεί πλήρες risk management system.

Το άρθρο 9 του AI Act προβλέπει ένα πολύ ευρύτερο πλαίσιο διαχείρισης κινδύνων για high-risk AI systems [3]. Επομένως, η Risk Management Requirement της παρούσας οντολογίας αποτελεί περιορισμένο AI Act-related σημείο αναφοράς του proof-of-concept και όχι πλήρη αναπαράσταση του άρθρου 9.

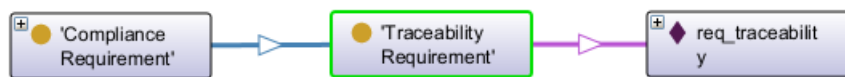
4.5.10 Traceability Requirement

Η class Traceability Requirement αναπαριστά στο proof-of-concept το AI Act-related σημείο αναφοράς που σχετίζεται με την ιχνηλασιμότητα. Στην παρούσα υλοποίηση συνδέεται με το requirement individual `req_traceability`.

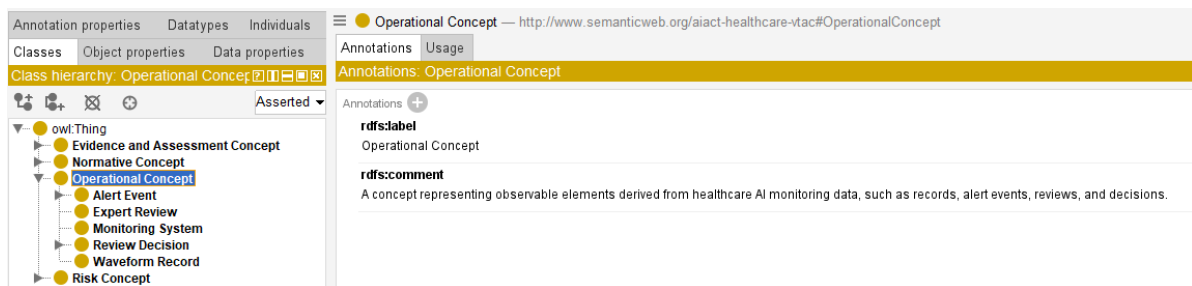
Η ιχνηλασιμότητα αποτυπώνεται μέσα από τη δυνατότητα ανάκτησης των διαφορετικών σημασιολογικών σχέσεων που αφορούν ένα alarm event. Για ένα συγκεκριμένο event μπορεί να εντοπιστεί το waveform record προέλευσής του, η Expert Review, η αντίστοιχη Review Decision και το Evidence Record που το τεκμηριώνει.



Σχήμα 4.21 Η class Traceability Requirement και τα annotations της στο Protégé.



Σχήμα 4.22 Η class Traceability Requirement και το αντίστοιχο requirement individual.



Σχήμα 4.23 Η class Operational Concept και τα annotations της στο Protégé.

Οι βασικές σχέσεις είναι:

Alert Event --belongsToRecord→ Waveform Record

Alert Event --hasReview→ Expert Review --hasDecision→ Review Decision

Alert Event --isDocumentedBy→ Evidence Record

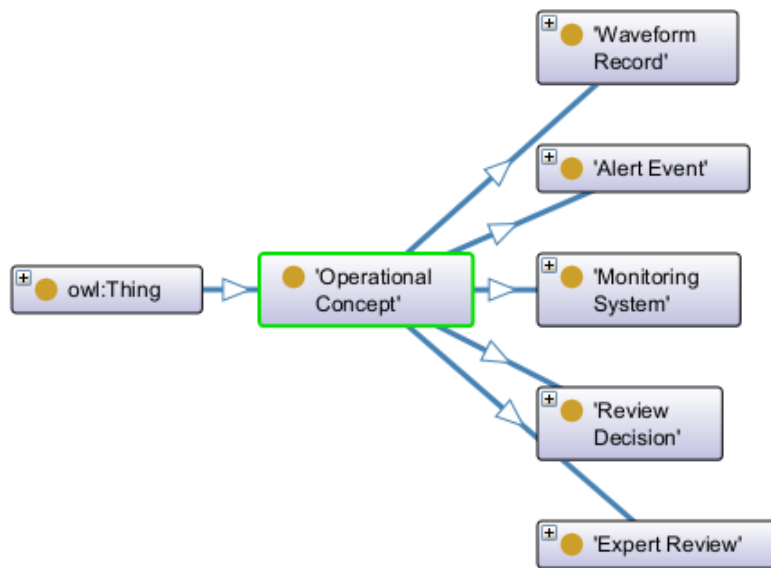
Η σύνθεση των σχέσεων αυτών δημιουργεί τη διαδρομή ιχνηλασιμότητας της παρούσας οντολογίας. Δεν πρόκειται για μία μοναδική γραμμική object-property αλυσίδα, αλλά για συνδεδεμένες σημασιολογικές διαδρομές γύρω από το ίδιο alarm event.

Το άρθρο 12 του AI Act συνδέει τις logging capabilities με την εξασφάλιση κατάλληλου επιπέδου traceability [3]. Στην παρούσα εργασία το req_traceability χρησιμοποιείται ως εφαρμοσμένο σημείο αναφοράς αυτής της διάστασης και όχι ως απόδειξη πλήρους συμμόρφωσης με το άρθρο 12.

4.5.11 Operational Concept

Η class Operational Concept αποτελεί τη γενική κατηγορία για τις έννοιες που σχετίζονται με τη λειτουργία του healthcare monitoring σεναρίου. Κάτω από αυτήν οργανώνονται οι βασικές λειτουργικές classes της οντολογίας, όπως το Waveform Record, το Alert Event, το Ventricular Tachycardia Alarm, το Expert Review, το Review Decision και το Monitoring System.

Ο ρόλος της class αυτής είναι να ομαδοποιεί τα στοιχεία που ανήκουν στο πραγματικό ή λειτουργικό επίπεδο της υλοποίησης. Οι έννοιες αυτές δεν αποτελούν κανονιστικές απαιτήσεις ούτε risk-related στοιχεία. Αντίθετα, αναπαριστούν τα βασικά αντικείμενα του healthcare monitoring περιβάλλοντος, δηλαδή τα records, τα alarm events, τις αξιολογήσεις και τις αποφάσεις.



Σχήμα 4.24 Υποκλάσεις της class Operational Concept.

The screenshot shows the Protégé interface. On the left, the 'Class hierarchy' pane displays a tree structure starting from 'owl:Thing', with 'Operational Concept' expanded to show its subclasses: 'Evidence and Assessment Concept', 'Normative Concept', 'Alert Event', 'Ventricular Tachycardia Alarm', 'Expert Review', 'Monitoring System', 'Review Decision', 'Waveform Record', and 'Risk Concept'. The 'Alert Event' class is selected. On the right, the 'Annotations' pane for 'Alert Event' shows two annotations: 'rdfs:label' with the value 'Alert Event' and 'rdfs:comment' with the value 'An observable alarm event derived from healthcare monitoring data.' The top bar shows the URI for 'Alert Event' as 'http://www.semanticweb.org/aiact-healthcare-vtac#AlertEvent'.

Σχήμα 4.25 Η class Alert Event και τα annotations της στο Protégé.

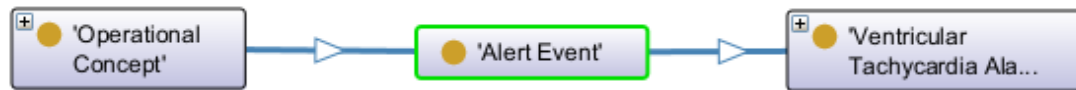
Η ύπαρξη του Operational Concept είναι σημαντική, επειδή διαχωρίζει το επίπεδο των πραγματικών περιστατικών από τα επίπεδα τεκμηρίωσης, απαιτήσεων και κινδύνου. Με αυτόν τον τρόπο, η οντολογία διατηρεί σαφή διάκριση ανάμεσα στο τι συνέβη στο σύστημα παρακολούθησης και στο πώς αυτό τεκμηριώνεται ή συσχετίζεται με selected requirements.

4.5.12 Alert Event

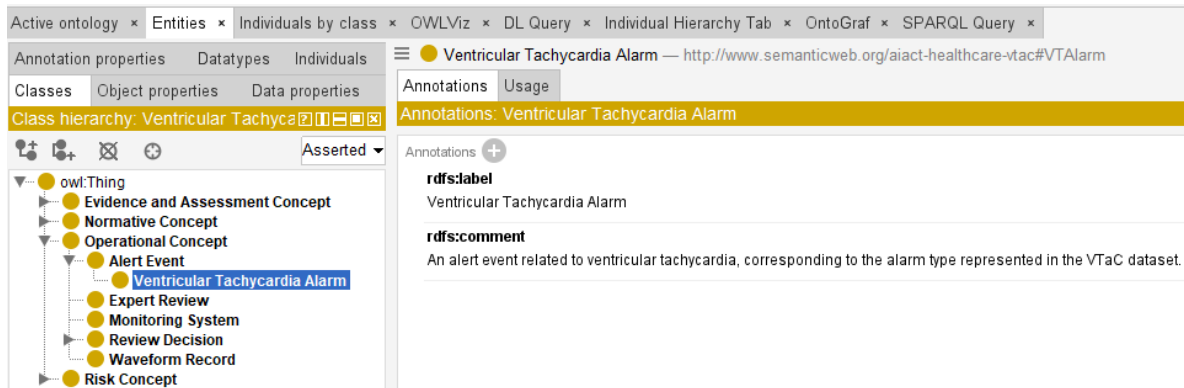
Η class Alert Event αναπαριστά τη γενική έννοια ενός alarm event. Είναι η βασική class για τα συμβάντα που προέρχονται από το VTaC dataset. Κάθε συγκεκριμένο alarm που επιλέχθηκε για την υλοποίηση εισάγεται στην οντολογία ως individual αυτής της class ή ειδικότερης υποκλάσης της.

Ο ρόλος της class Alert Event είναι κεντρικός, επειδή το alert event αποτελεί την αφετηρία της βασικής σημασιολογικής διαδρομής. Από αυτό ξεκινούν οι συνδέσεις προς το waveform record, το expert review, τη review decision, το evidence record και τις επιλεγμένες απαιτήσεις.

Η class αυτή επιτρέπει στην οντολογία να μη διαχειρίζεται τα alarms απλώς ως γραμμές ενός dataset, αλλά ως οντολογικά αντικείμενα που μπορούν να συνδεθούν με άλλα στοιχεία. Ένα alert event αποκτά πλήρη σημασία όταν συνδεθεί με το record προέλευσής του, την αξιολόγηση, την απόφαση και το evidence record που το τεκμηριώνει.



Σχήμα 4.26 Θέση της class Alert Event στο λειτουργικό επίπεδο της οντολογίας.



Σχήμα 4.27 Η class Ventricular Tachycardia Alarm και τα annotations της στο Protégé.

Στην παρούσα υλοποίηση, τα πραγματικά VTaC events αποτελούν instances της ειδικότερης class Ventricular Tachycardia Alarm, η οποία ανήκει κάτω από την Alert Event. Με αυτόν τον τρόπο, το μοντέλο διατηρεί τόσο τη γενική έννοια του alarm event όσο και την ειδική πληροφορία ότι πρόκειται για ventricular tachycardia alarms.

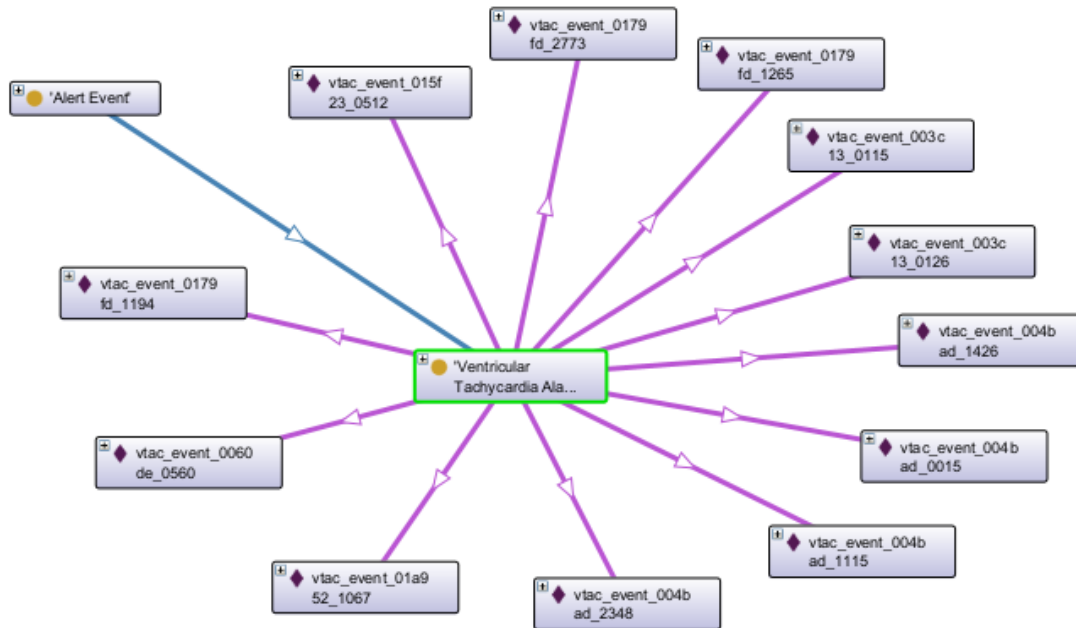
4.5.13 Ventricular Tachycardia Alarm

Η class Ventricular Tachycardia Alarm είναι υποκλάση της Alert Event και εξειδικεύει το είδος του alarm που εξετάζεται στην παρούσα εργασία. Το VTaC dataset αφορά ventricular tachycardia alarms, επομένως η ύπαρξη αυτής της υποκλάσης επιτρέπει στο μοντέλο να αποτυπώσει με μεγαλύτερη ακρίβεια το domain των events [4].

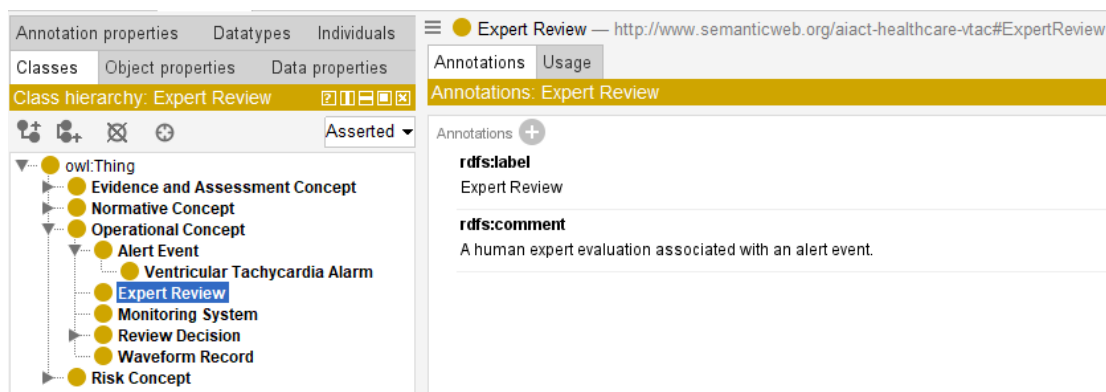
Ο ρόλος της class αυτής είναι να δηλώσει ότι τα συγκεκριμένα alert events δεν είναι γενικά alarms οποιουδήποτε τύπου, αλλά alarms που σχετίζονται με κοιλιακή ταχυκαρδία. Έτσι, η οντολογία παραμένει συνδεδεμένη με το περιεχόμενο του dataset και δεν γενικεύει υπερβολικά το use case.

Τα 12 επιλεγμένα VTaC alert events εισάγονται ως individuals της class Ventricular Tachycardia Alarm. Με αυτόν τον τρόπο, η οντολογία διατηρεί την πληροφορία ότι τα events προέρχονται από το συγκεκριμένο medical monitoring domain και μπορούν να ανακτηθούν μέσω queries που αναζητούν instances αυτής της class.

Η ύπαρξη της class είναι επίσης χρήσιμη για τον έλεγχο της οντολογίας. Μέσω DL Queries μπορεί να ανακτηθεί το σύνολο των Ventricular Tachycardia Alarm individuals, ώστε να επιβεβαιωθεί ότι όλα τα επιλεγμένα events έχουν εισαχθεί και ταξινομηθεί σωστά.



Σχήμα 4.28 Η class Ventricular Tachycardia Alarm με τα 12 πραγματικά VTaC event individuals.



Σχήμα 4.29 Η class Expert Review και τα annotations της στο Protégé.

4.5.14 Expert Review

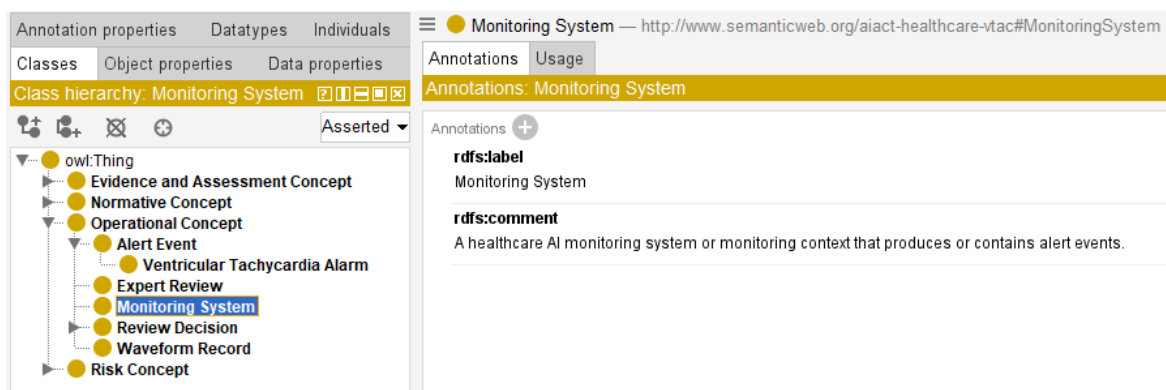
Η class Expert Review αναπαριστά την αξιολόγηση που πραγματοποιείται για ένα alert event. Στο VTaC dataset, τα alarms έχουν αξιολογηθεί από ανθρώπινους annotators με σχετική εμπειρία, και η πληροφορία αυτή είναι κρίσιμη για τη λογική του μοντέλου [4].

Ο ρόλος της class είναι να αποτυπώσει το στάδιο αξιολόγησης πριν από την τελική απόφαση. Το alarm event δεν οδηγεί απευθείας σε true ή false label μέσα στην οντολογία. Παρεμβάλλεται το Expert Review, ώστε να φαίνεται ότι η απόφαση προκύπτει από διαδικασία αξιολόγησης.

Η διάκριση αυτή είναι σημαντική, επειδή διαχωρίζει το ίδιο το event από την αξιολόγησή του. Το event αποτελεί το πραγματικό συμβάν, ενώ το expert review αποτελεί τη διαδικασία μέσω της οποίας το event εξετάζεται. Έτσι, η οντολογία μπορεί να υποστηρίξει πιο καθαρή αναπαράσταση της ανθρώπινης ή expert συμμετοχής στη διαδικασία τεκμηρίωσης.



Σχήμα 4.30 Η class Expert Review με τα 12 review individuals των VTaC events.



Σχήμα 4.31 Η class Monitoring System και τα annotations της στο Protégé.

Στην υλοποίηση, κάθε επιλεγμένο VTaC event συνδέεται με αντίστοιχο Expert Review individual. Αυτή η σύνδεση επιτρέπει στο μοντέλο να απαντά σε ερωτήματα σχετικά με το ποια events έχουν αξιολογηθεί και ποια review decisions προέκυψαν από αυτή τη διαδικασία.

4.5.15 Monitoring System

Η class Monitoring System αναπαριστά το σύστημα παρακολούθησης από το οποίο προέρχονται τα alarms. Στο πραγματικό περιβάλλον του VTaC, τα alarms παράγονται από ICU bedside monitors. Στην οντολογία, η class αυτή επιτρέπει να αποτυπωθεί η πηγή παραγωγής ή παρακολούθησης των events σε πιο γενικό επίπεδο.

Ο ρόλος της είναι να συνδέσει τα alert events με το τεχνολογικό περιβάλλον στο οποίο δημιουργούνται. Ένα alarm event δεν εμφανίζεται ανεξάρτητα, αλλά παράγεται ή καταγράφεται στο πλαίσιο ενός monitoring system. Η ύπαρξη αυτής της class βοηθά να γίνει σαφές ότι τα events ανήκουν σε συγκεκριμένο healthcare monitoring context.

Στην παρούσα υλοποίηση, η class Monitoring System χρησιμοποιείται κυρίως ως μέρος του operational layer. Περιλαμβάνει το individual vtac_icu_monitoring_system, το οποίο αναπαριστά το monitoring περιβάλλον του VTaC use case. Παρότι η υλοποίηση περιλαμβάνει ένα μόνο τέτοιο individual, η class μπορεί να υποστηρίξει μελλοντική επέκταση με περισσότερα monitoring systems ή διαφορετικά healthcare monitoring περιβάλλοντα.

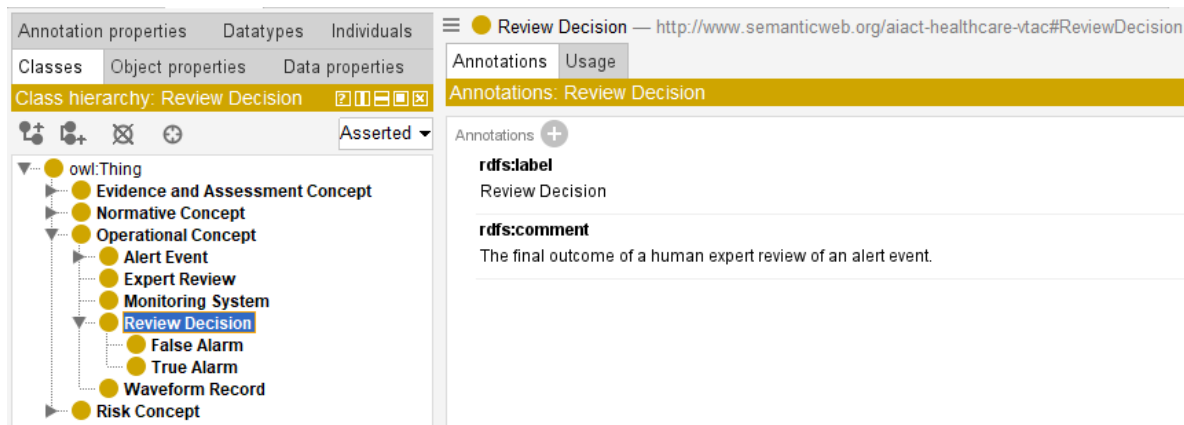


Σχήμα 4.32 Η class Monitoring System και το αντίστοιχο individual της.

Η ύπαρξη της class είναι χρήσιμη για τη γενίκευση του μοντέλου. Εάν μελλοντικά η οντολογία επεκταθεί σε άλλα datasets ή σε διαφορετικούς τύπους monitoring systems, η class Monitoring System μπορεί να λειτουργήσει ως σημείο ένταξης αυτών των νέων στοιχείων.

4.5.16 Review Decision

Η class Review Decision αναπαριστά το τελικό αποτέλεσμα της αξιολόγησης ενός alert event. Αν η class Expert Review εκφράζει τη διαδικασία της ανθρώπινης ή expert αξιολόγησης, τότε η Review Decision εκφράζει το αποτέλεσμα αυτής της διαδικασίας. Με άλλα λόγια, το review είναι η πράξη της αξιολόγησης, ενώ το decision είναι το συμπέρασμα στο οποίο καταλήγει η αξιολόγηση.



Σχήμα 4.33 Η class Review Decision και τα annotations της στο Protégé.

Η διάκριση αυτή είναι σημαντική για τη λογική της οντολογίας. Ένα alarm event δεν χαρακτηρίζεται απευθείας ως true ή false χωρίς να υπάρχει ενδιάμεσο στάδιο αξιολόγησης. Η οντολογία αποτυπώνει ότι πρώτα υπάρχει ένα Alert Event, στη συνέχεια πραγματοποιείται Expert Review και έπειτα προκύπτει μία Review Decision.

Η Review Decision αποτελεί το αποτέλεσμα της διαδρομής ανθρώπινης αξιολόγησης και συνδέεται με το Expert Review μέσω του hasDecision. Δεν υπάρχει άμεση object-property σύνδεση μεταξύ Review Decision και Evidence Record. Η τεκμηρίωση οργανώνεται σε ξεχωριστή διαδρομή από το ίδιο Alert Event μέσω του isDocumentedBy.



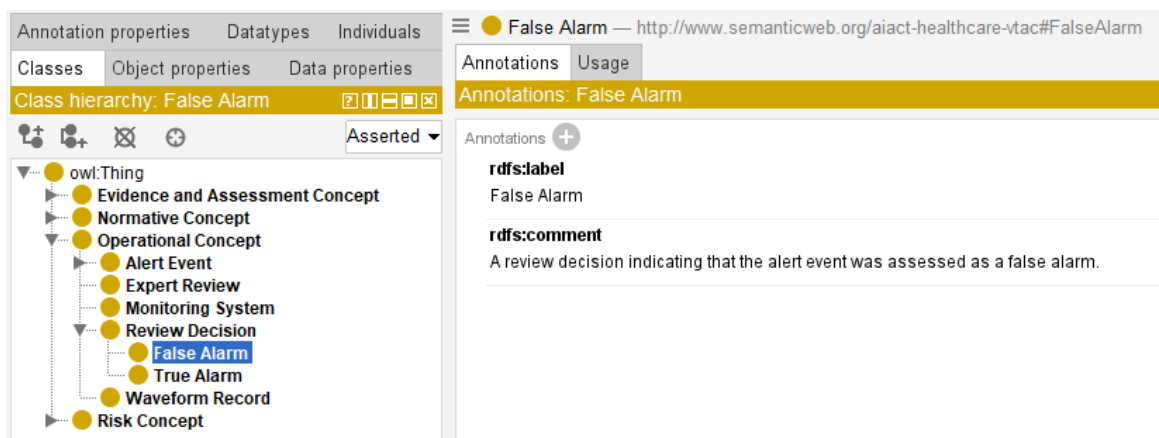
Σχήμα 4.34 Υποκλάσεις της class Review Decision.

Η class αυτή είναι ιδιαίτερα σημαντική επειδή συνδέει το operational επίπεδο με τα επόμενα επίπεδα τεκμηρίωσης. Από τη στιγμή που ένα alarm event έχει review decision, μπορεί να δημιουργηθεί evidence record και να γίνει mapping προς επιλεγμένες απαιτήσεις.

4.5.17 False Alarm

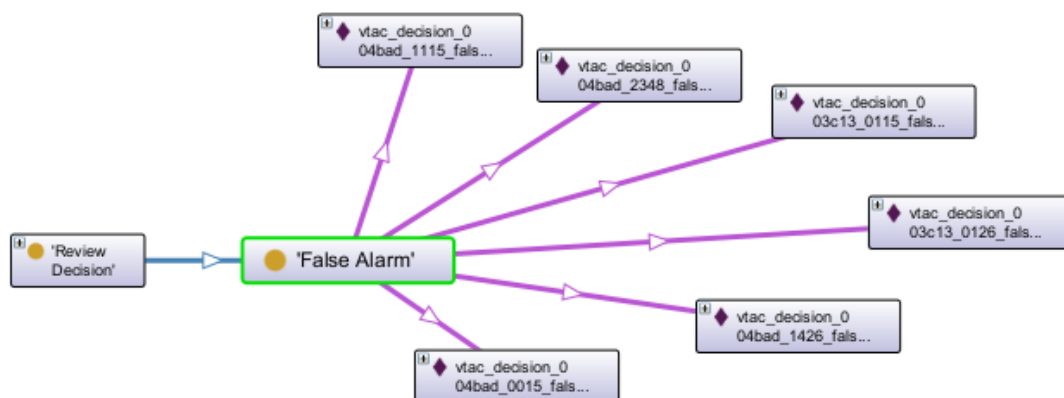
Η class False Alarm είναι υποκλάση της Review Decision και αναπαριστά την απόφαση ότι ένα alarm ήταν λανθασμένο. Δεν σημαίνει ότι το event δεν υπήρξε ως καταγεγραμμένο alarm. Σημαίνει ότι, μετά την expert αξιολόγηση, το alarm χαρακτηρίστηκε ως μη πραγματικό ή μη έγκυρο ως VT συμβάν.

Ο ρόλος της class είναι ιδιαίτερα σημαντικός, επειδή τα false alarms ενεργοποιούν το risk-related μέρος της οντολογίας. Ένα false alarm μπορεί να συνδεθεί με risk indicator, risk-relevant event και potential harm, όπως το alert fatigue.



Σχήμα 4.35 Η class False Alarm και τα annotations της στο Protégé.

Στην παρούσα υλοποίηση, τα false alarm decision individuals αντιστοιχούν στα έξι VTaC events που έχουν αξιολογηθεί ως false alarms. Η σύνδεση αυτή επιτρέπει στην οντολογία να διακρίνει τα false alarm cases από τα true alarm cases και να τα αξιοποιεί διαφορετικά στο requirement mapping.



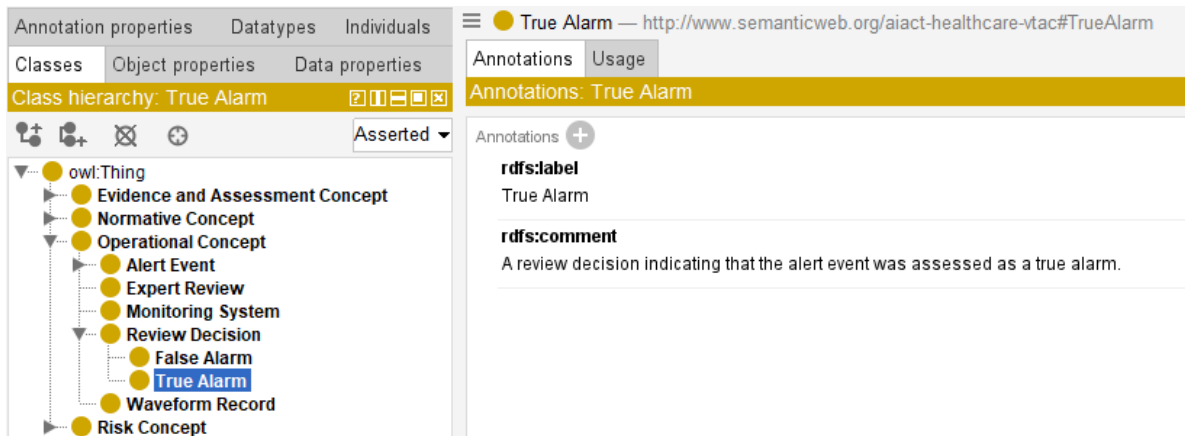
Σχήμα 4.36 Η class False Alarm με τα αντίστοιχα decision individuals.

Η class False Alarm είναι επίσης σημαντική για τα queries. Μέσω αυτής μπορεί να αναζητηθεί ποια events έχουν οδηγήσει σε false alarm decision και ποια από αυτά συνδέονται με risk management. Έτσι, η οντολογία μπορεί να αναπαραστήσει όχι μόνο την απόφαση, αλλά και την πρόσθετη σημασία των false alarms στο lightweight risk layer.

4.5.18 True Alarm

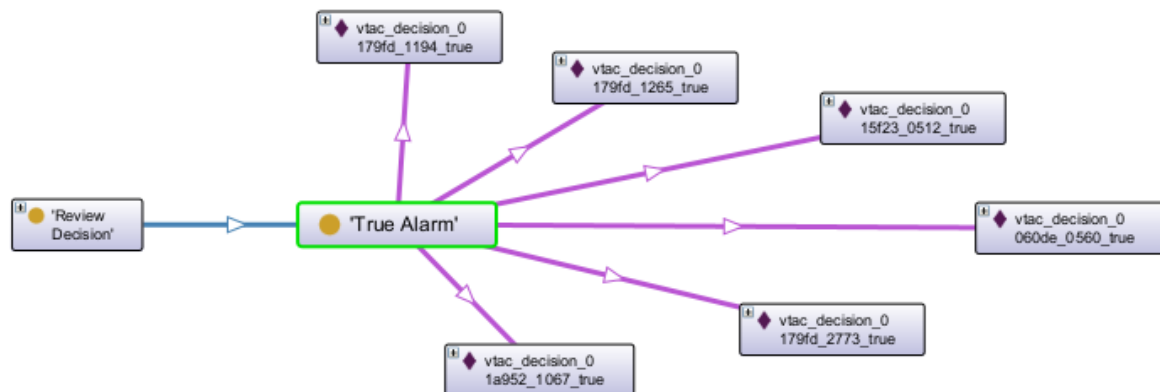
Η class True Alarm είναι υποκλάση της Review Decision και αναπαριστά την απόφαση ότι ένα alarm επιβεβαιώθηκε ως πραγματικό VT alarm. Δηλώνει ότι, μετά την expert αξιολόγηση, το alarm θεωρήθηκε έγκυρο.

Ο ρόλος της είναι να διαχωρίζει τις επιβεβαιωμένες περιπτώσεις από τα false alarms. Τα true alarm cases συμμετέχουν στην ίδια event-centered σημασιολογική δομή με τα υπόλοιπα events: το Alert Event συνδέεται με το Waveform Record προέλευσής του, με το Expert Review και τη Review Decision και, παράλληλα, με το αντίστοιχο Evidence Record. Σε αντίθεση με τα επιλεγμένα false alarm cases, δεν εντάσσονται στο lightweight risk-related επίπεδο της παρούσας υλοποίησης.



Σχήμα 4.37 Η class True Alarm και τα annotations της στο Protégé.

Στην οντολογία, τα true alarm decision individuals αντιστοιχούν στα έξι VTaC events που έχουν αξιολογηθεί ως true alarms. Η ύπαρξη της class True Alarm επιτρέπει στο μοντέλο να ανακτά ξεχωριστά τις επιβεβαιωμένες περιπτώσεις και να τις διακρίνει από τις false alarm decisions.



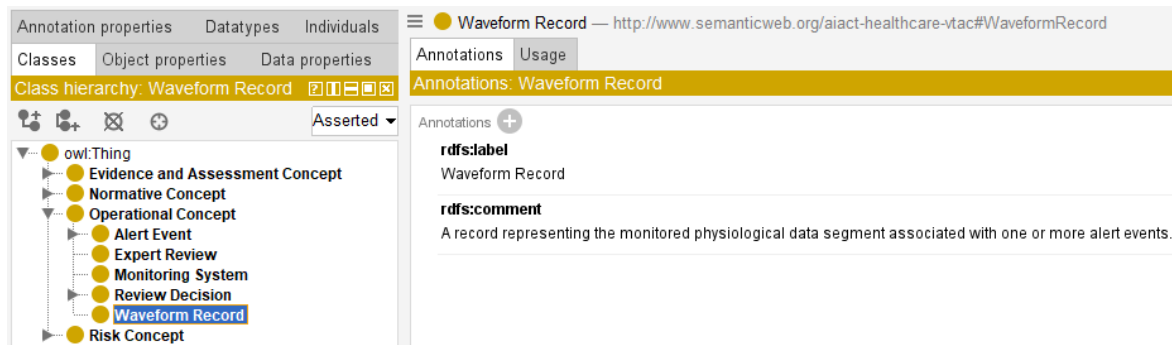
Σχήμα 4.38 Η class True Alarm με τα αντίστοιχα decision individuals.

Η class αυτή είναι χρήσιμη τόσο για την παρουσίαση των selected events όσο και για τα DL Queries και SPARQL Queries που συγκρίνουν ή ανακτούν true και false alarm cases. Έτσι, η οντολογία μπορεί να απαντήσει σε ερωτήματα σχετικά με το ποια events έχουν αξιολογηθεί ως true alarms και ποια evidence records σχετίζονται με αυτά.

4.5.19 Waveform Record

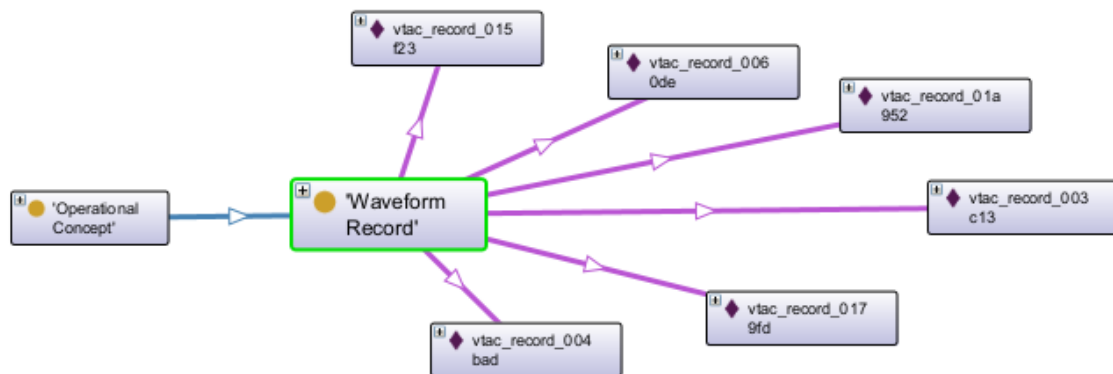
Η class Waveform Record αναπαριστά το record από το οποίο προέρχεται ένα alarm event. Στο VTaC dataset, το πεδίο record δηλώνει το patient waveform record, ενώ κάθε record μπορεί να συνδέεται με ένα ή περισσότερα alarm events [4].

Ο ρόλος της class είναι να υποστηρίζει την ιχνηλασιμότητα. Αν δεν υπήρχε ξεχωριστή class για τα waveform records, τα events θα εμφανίζονταν αποκομμένα από την πηγή τους. Με την class αυτή, μπορεί να φανεί ποιο event ανήκει σε ποιο record.



Σχήμα 4.39 Η class Waveform Record και τα annotations της στο Protégé.

Η class Waveform Record είναι ιδιαίτερα σημαντική στις περιπτώσεις όπου ένα record συνδέεται με περισσότερα από ένα events. Στην παρούσα υλοποίηση υπάρχουν waveform records που συνδέονται με ένα event και άλλα που συνδέονται με περισσότερα. Αυτό επιτρέπει στην οντολογία να αναπαραστήσει σχέσεις one-to-many και να υποστηρίξει queries που ξεκινούν από ένα record και αναζητούν τα events που συνδέονται με αυτό.



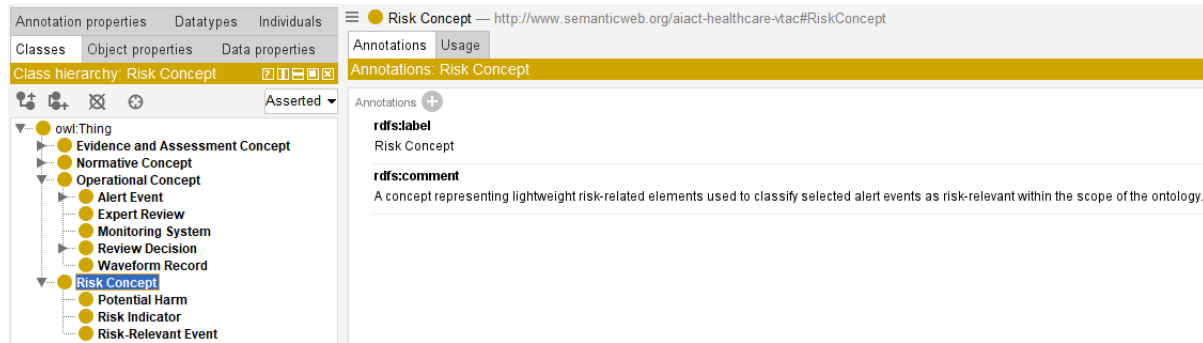
Σχήμα 4.40 Η class Waveform Record με τα αντίστοιχα record individuals.

Η ύπαρξη της class αυτής ενισχύει τη βασική αλυσίδα ιχνηλασιμότητας της οντολογίας. Η διαδρομή δεν ξεκινά από το alarm ως απομονωμένο περιστατικό, αλλά από το record προέλευσής του. Έτσι, κάθε event μπορεί να τοποθετηθεί μέσα στο πραγματικό πλαίσιο από το οποίο προήλθε.

4.5.20 Risk Concept

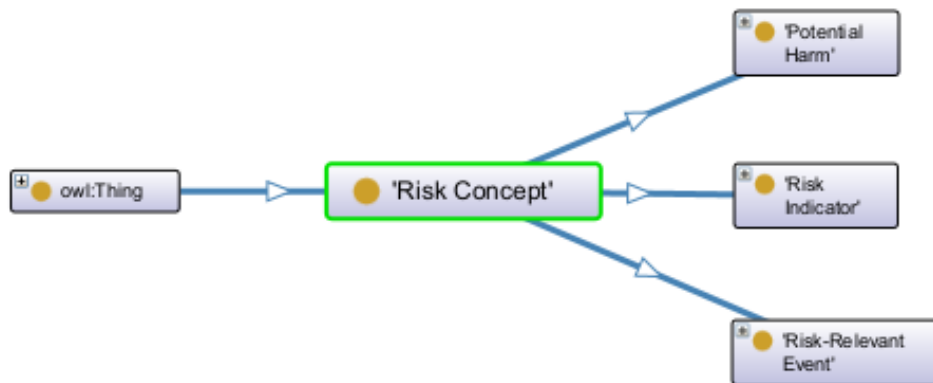
Η class Risk Concept αποτελεί τη γενική κατηγορία για τις έννοιες που σχετίζονται με το lightweight risk layer της οντολογίας. Κάτω από αυτήν οργανώνονται οι classes Potential Harm, Risk Indicator και Risk-Relevant Event.

Ο ρόλος της class αυτής είναι να διαχωρίζει το risk-related μέρος της οντολογίας από τα operational, evidence και normative στοιχεία. Με αυτόν τον τρόπο, οι έννοιες κινδύνου δεν αναμειγνύονται άμεσα με τα alarm events, τα evidence records ή τις απαιτήσεις, αλλά οργανώνονται σε δικό τους επίπεδο.



Σχήμα 4.41 Η class Risk Concept και τα annotations της στο Protégé.

Η ύπαρξη της class Risk Concept είναι σημαντική, επειδή η οντολογία δεν περιορίζεται μόνο στην αναπαράσταση true και false alarm decisions. Για τα false alarm cases προστίθεται ένα επιπλέον επίπεδο ερμηνείας, το οποίο επιτρέπει να δηλωθεί ότι ένα false alarm μπορεί να έχει σημασία από πλευράς κινδύνου. Αυτό δεν σημαίνει ότι αποδεικνύεται πραγματική βλάβη, αλλά ότι το event μπορεί να οργανωθεί ως risk-related περίπτωση μέσα στο μοντέλο.



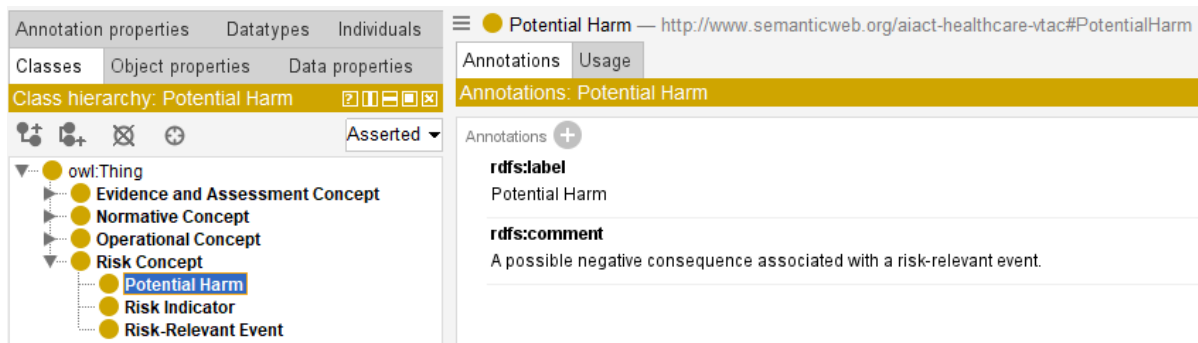
Σχήμα 4.42 Υποκλάσεις της class Risk Concept.

Στην παρούσα υλοποίηση, το Risk Concept χρησιμοποιείται κυρίως για τα false alarm cases που συνδέονται με risk indicators, risk-relevant events και potential harm. Έτσι, η οντολογία μπορεί να υποστηρίξει ερωτήματα που αφορούν τη σχέση ανάμεσα σε false alarms και risk management.

4.5.21 Potential Harm

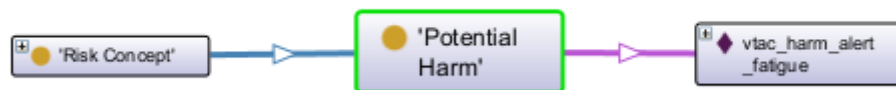
Η class Potential Harm αναπαριστά πιθανή αρνητική επίπτωση που μπορεί να σχετίζεται με ένα risk-relevant event. Στο healthcare monitoring σενάριο της παρούσας εργασίας, η πιο χαρακτηριστική πιθανή επίπτωση που χρησιμοποιείται είναι το alert fatigue [6].

Ο ρόλος της class είναι να επιτρέπει στην οντολογία να περάσει από το επίπεδο του false alarm στο επίπεδο της πιθανής συνέπειας. Ένα false alarm δεν είναι μόνο μια λανθασμένη απόφαση αξιολόγησης. Σε περιβάλλοντα όπως η μονάδα εντατικής θεραπείας, τα συχνά ή μη έγκυρα alarms μπορούν να συνδεθούν με ευρύτερα προβλήματα, όπως η κόπωση του προσωπικού απέναντι στους συναγερμούς.



Σχήμα 4.43 H class Potential Harm και τα annotations της στο Protégé.

Στην παρούσα υλοποίηση, η class Potential Harm περιλαμβάνει το individual vtac_harm_alert_fatigue. Το individual αυτό χρησιμοποιείται για να αναπαραστήσει την πιθανή επίπτωση του alert fatigue. Η σύνδεσή του με false alarm cases γίνεται μέσω του risk-related layer και ειδικότερα μέσω των risk-relevant events.



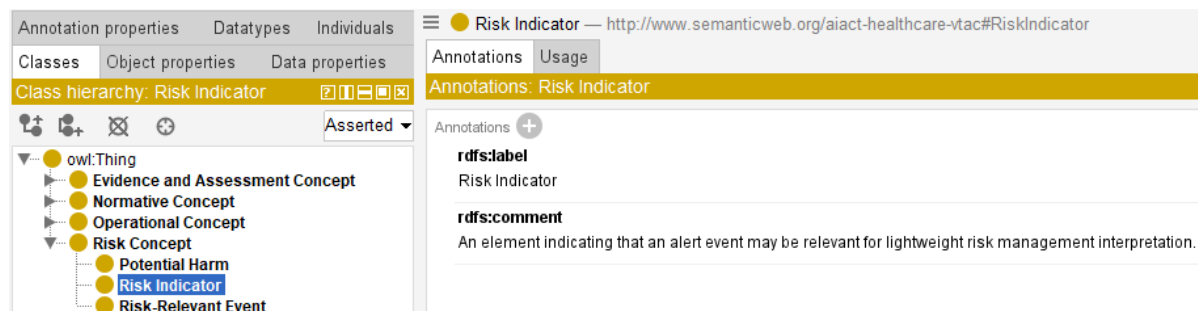
Σχήμα 4.44 H class Potential Harm και το αντίστοιχο individual vtac_harm_alert_fatigue.

Είναι σημαντικό να τονιστεί ότι η class Potential Harm δεν χρησιμοποιείται για να δηλώσει ότι μια πραγματική βλάβη έχει συμβεί. Η οντολογία δεν κάνει κλινική ή οργανωτική αξιολόγηση της επίπτωσης. Αντίθετα, οργανώνει σημασιολογικά τη δυνατότητα σύνδεσης ενός false alarm case με μια πιθανή αρνητική επίπτωση.

4.5.22 Risk Indicator

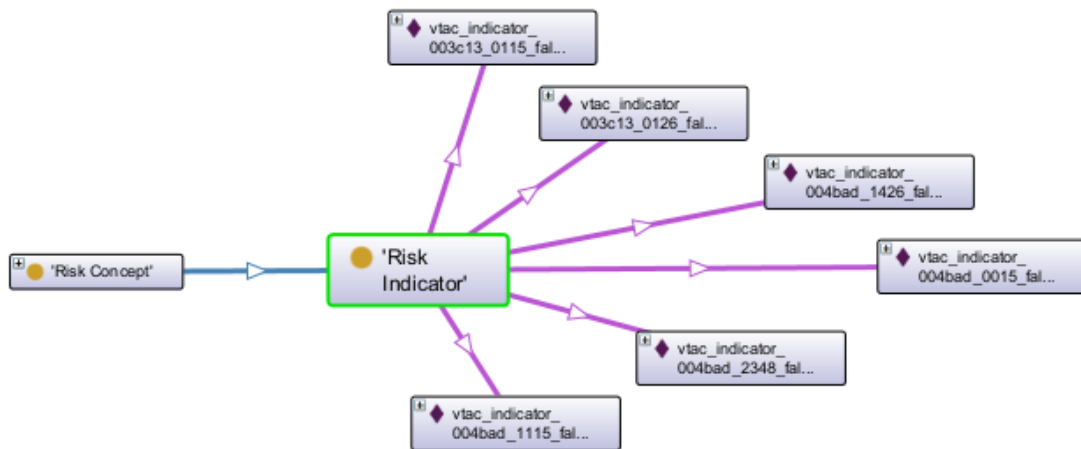
Η class Risk Indicator αναπαριστά ένδειξη κινδύνου. Στο μοντέλο, ένα false alarm μπορεί να λειτουργήσει ως risk indicator επειδή μπορεί να σχετίζεται με πιθανές επιπτώσεις στη λειτουργία ενός healthcare monitoring περιβάλλοντος.

Ο ρόλος της class είναι να δηλώσει ότι ένα στοιχείο δεν αποτελεί απαραίτητα βλάβη από μόνο του, αλλά μπορεί να υποδεικνύει πιθανότητα ή ένδειξη κινδύνου. Αυτός ο διαχωρισμός είναι σημαντικός. Το Risk Indicator δεν είναι ίδιο με το Potential Harm. Το πρώτο δηλώνει την ένδειξη ή το σήμα κινδύνου, ενώ το δεύτερο δηλώνει την πιθανή αρνητική επίπτωση.



Σχήμα 4.45 H class Risk Indicator και τα annotations της στο Protégé.

Στην παρούσα υλοποίηση, risk indicator individuals δημιουργούνται για τα false alarm cases. Με αυτόν τον τρόπο, η οντολογία μπορεί να διακρίνει τα false alarms που συνδέονται με risk-related ερμηνεία και να τα αναζητήσει μέσω queries.



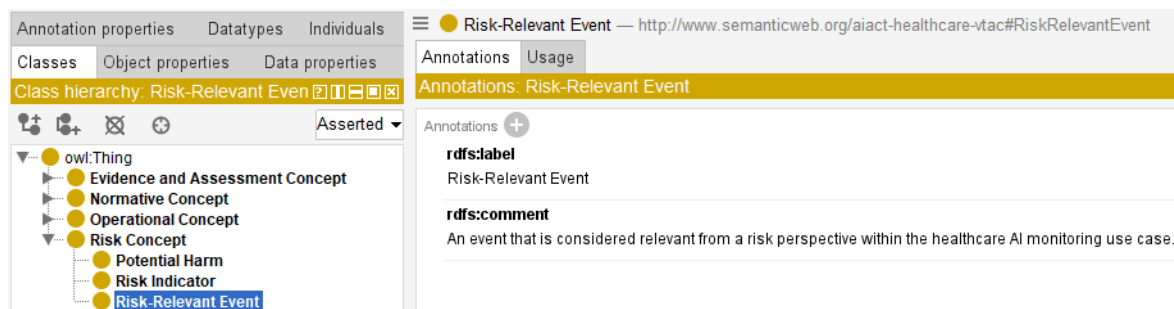
Σχήμα 4.46 Η class Risk Indicator με τα αντίστοιχα indicator individuals.

Η class Risk Indicator είναι χρήσιμη επειδή επιτρέπει την ενδιάμεση σύνδεση ανάμεσα στο operational event και στο risk layer. Το alert event παραμένει μέρος του operational επιπέδου, αλλά μέσω του risk indicator μπορεί να αποκτήσει πρόσθετη σημασιολογική ερμηνεία ως περίπτωση που σχετίζεται με risk management.

4.5.23 Risk-Relevant Event

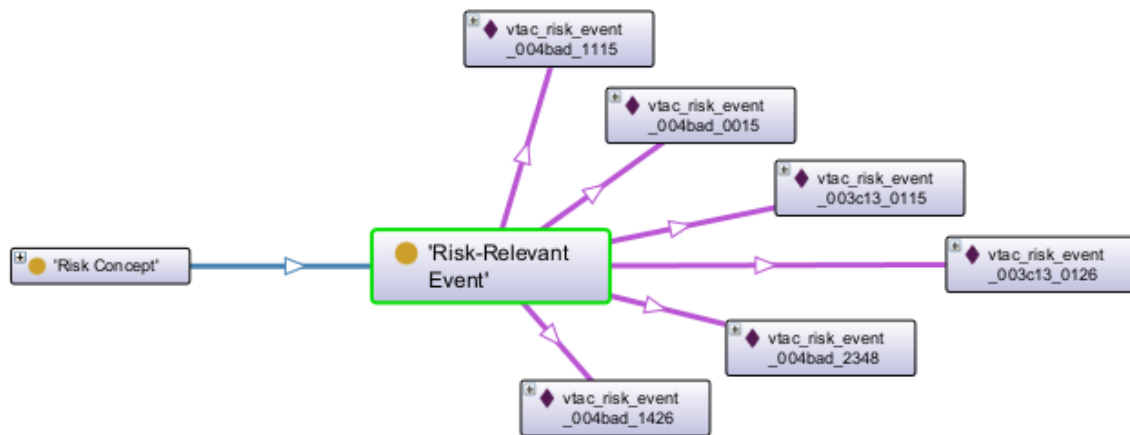
Η class Risk-Relevant Event αναπαριστά ένα event που έχει σημασία από πλευράς κινδύνου. Δεν αποτελεί απαραίτητα διαφορετικό φυσικό συμβάν από το αρχικό alert event. Αντίθετα, λειτουργεί ως risk-oriented αναπαράσταση ή ερμηνεία ενός false alarm case.

Ο ρόλος της class είναι να επιτρέπει στην οντολογία να ξεχωρίζει τα events που έχουν risk significance από τα απλά operational events. Έτσι, ένα false alarm μπορεί να είναι ταυτόχρονα alert event στο operational επίπεδο και risk-relevant event στο risk layer, εφόσον συνδέεται με πιθανή επίπτωση.



Σχήμα 4.47 Η class Risk-Relevant Event και τα annotations της στο Protégé.

Στην παρούσα υλοποίηση, τα risk-relevant event individuals αντιστοιχούν στα false alarm cases. Η σχέση ανάμεσα σε ένα alert event και στο αντίστοιχο risk-relevant event αποτυπώνεται με το object property isClassifiedAs. Με αυτόν τον τρόπο, ένα συγκεκριμένο false alarm event μπορεί να ταξινομηθεί ως risk-relevant μέσα στο lightweight risk layer.



Σχήμα 4.48 Η class Risk-Relevant Event με τα αντίστοιχα risk event individuals.

Η class αυτή είναι ιδιαίτερα σημαντική για τη σύνδεση με το req_risk_management. Τα false alarm cases που συνδέονται με risk-relevant events μπορούν να τεκμηριώσουν, μέσω evidence records, την απαίτηση που αφορά risk management. Έτσι, η οντολογία μπορεί να απαντήσει σε ερωτήματα όπως ποια false alarms δεν είναι απλώς λανθασμένες αποφάσεις, αλλά έχουν και risk-related σημασία.

4.5.24 Συνολική αποτίμηση της ιεραρχίας των classes

Η ιεραρχία των 23 classes δείχνει ότι η οντολογία έχει σχεδιαστεί με σαφή εννοιολογική δομή. Οι classes δεν αποτελούν απλή λίστα όρων, αλλά οργανώνονται σε τέσσερα βασικά επίπεδα: operational, evidence and assessment, normative και risk.

Το operational επίπεδο συνδέει την οντολογία με τα πραγματικά VTaC events. Μέσα από classes όπως Waveform Record, Alert Event, Ventricular Tachycardia Alarm, Expert Review και Review Decision, το μοντέλο μπορεί να αναπαραστήσει το πραγματικό συμβάν, την πηγή του, την αξιολόγησή του και την τελική απόφαση.

Το evidence and assessment επίπεδο οργανώνει την τεκμηρίωση. Η class Evidence Record είναι κεντρική σε αυτό το επίπεδο, επειδή λειτουργεί ως ενδιάμεσος κόμβος ανάμεσα στα πραγματικά events και στις επιλεγμένες απαιτήσεις. Οι classes Compliance Assessment και Control Mechanism υποστηρίζουν επιπλέον τη λογική της αξιολόγησης και των μηχανισμών τεκμηρίωσης.

Το normative επίπεδο περιλαμβάνει τις επιλεγμένες AI Act-related απαιτήσεις. Μέσα από τη γενική class Compliance Requirement και τις ειδικότερες classes Logging Requirement, Human Review Requirement, Traceability Requirement και Risk Management Requirement, η οντολογία οργανώνει τα κανονιστικά σημεία αναφοράς στα οποία χαρτογραφούνται τα evidence records.

Το risk επίπεδο επιτρέπει την ειδική αναπαράσταση των false alarm cases ως risk-related περιπτώσεων. Οι classes Risk Indicator, Risk-Relevant Event και Potential Harm επιτρέπουν στο μοντέλο να συνδέσει false alarms με πιθανές επιπτώσεις, όπως το alert fatigue, χωρίς να ισχυρίζεται ότι έχει αποδειχθεί πραγματική βλάβη.

Η πλήρης ανάλυση των classes είναι απαραίτητη, επειδή όλες οι επόμενες λειτουργίες της οντολογίας βασίζονται σε αυτές. Τα object properties δείχνουν πώς συνδέονται οι classes μεταξύ τους μέσω των individuals. Τα data properties προσθέτουν συγκεκριμένες τιμές στα individuals. Τα individuals δείχνουν τα πραγματικά παραδείγματα της υλοποίησης. Τέλος, τα DL Queries και τα SPARQL Queries

αξιοποιούν την ταξινόμηση των individuals στις σωστές classes για να επιστρέψουν ελέγξιμα αποτελέσματα.

Συνολικά, η ιεραρχία των classes προσφέρει το εννοιολογικό θεμέλιο της οντολογίας. Πάνω σε αυτή τη βάση χτίζονται οι σχέσεις, τα instances, η τεκμηρίωση και οι μηχανισμοί ελέγχου που παρουσιάζονται στις επόμενες ενότητες.

4.6 Object properties της οντολογίας

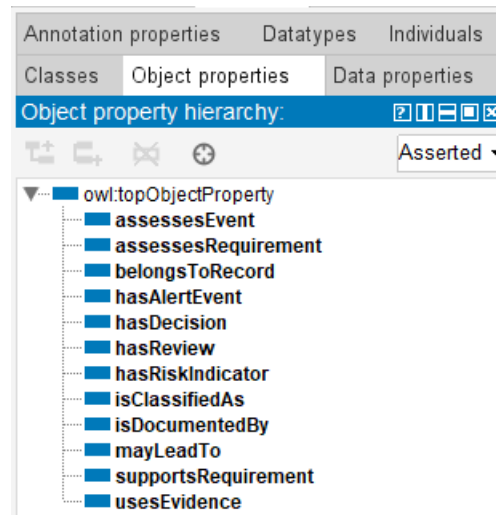
Τα object properties αποτελούν το βασικό μέσο σύνδεσης των individuals μέσα στην οντολογία [24]. Ενώ οι classes ορίζουν ποιες έννοιες υπάρχουν στο μοντέλο, τα object properties δείχνουν πώς αυτές οι έννοιες συνδέονται μεταξύ τους σε επίπεδο πραγματικών instances.

Στην παρούσα υλοποίηση, τα object properties έχουν ιδιαίτερα σημαντικό ρόλο, επειδή επιτρέπουν την αναπαράσταση της βασικής σημασιολογικής αλυσίδας της οντολογίας. Ένα alert event δεν αντιμετωπίζεται ως μεμονωμένη εγγραφή. Συνδέεται με το waveform record από το οποίο προέρχεται, με το expert review που το αξιολογεί, με τη review decision που προκύπτει, με το evidence record που το τεκμηριώνει και με τις επιλεγμένες απαιτήσεις που υποστηρίζονται από αυτό το evidence.

Επιπλέον, στις περιπτώσεις false alarms, τα object properties επιτρέπουν τη σύνδεση του event με risk indicators, risk-relevant events και πιθανές επιπτώσεις. Με αυτόν τον τρόπο, η οντολογία μπορεί να αποτυπώσει όχι μόνο τη βασική διαδρομή τεκμηρίωσης, αλλά και την πρόσθετη risk-related διάσταση των false alarm cases.

Συνολικά, στην οντολογία χρησιμοποιήθηκαν δώδεκα object properties:

- assessesEvent
- assessesRequirement
- belongsToRecord
- hasAlertEvent
- hasDecision
- hasReview
- hasRiskIndicator
- isClassifiedAs
- isDocumentedBy
- mayLeadTo
- supportsRequirement
- usesEvidence



Σχήμα 4.49 Ιεραρχία των object properties της οντολογίας στο Protégé.

Κάθε object property έχει συγκεκριμένο ρόλο και δεν δημιουργήθηκε ως απλή τεχνική προσθήκη. Άλλα object properties χρησιμοποιούνται για να αναπαραστήσουν τη λειτουργική διαδρομή του event, άλλα για να συνδέσουν την αξιολόγηση με την απόφαση, άλλα για να οργανώσουν την τεκμηρίωση, άλλα για να υποστηρίξουν τη σύνδεση με απαιτήσεις συμμόρφωσης και άλλα για να εκφράσουν τη risk-related διάσταση των false alarms.

Η ανάλυση που ακολουθεί παρουσιάζει κάθε object property ξεχωριστά, με βάση τον ρόλο του μέσα στο μοντέλο.

4.6.1 assessesEvent

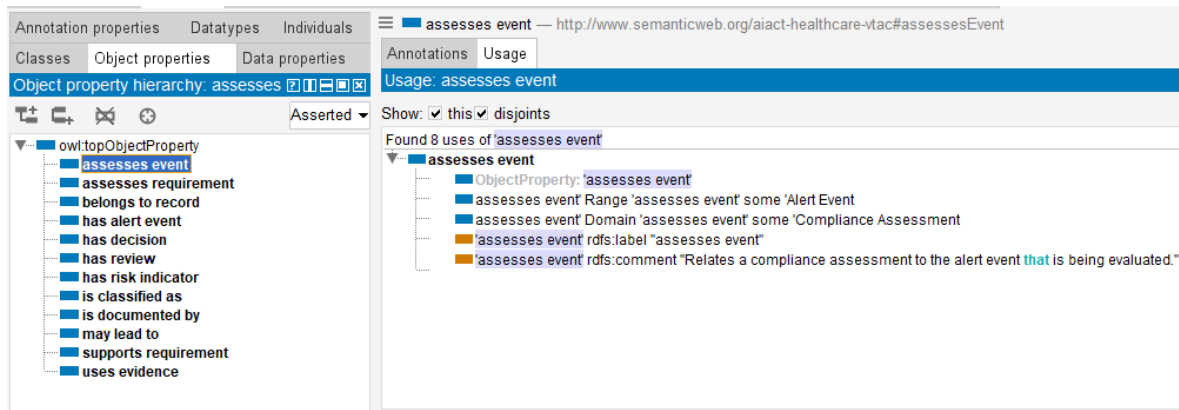
Το object property `assessesEvent` έχει σχεδιαστεί για να συνδέει ένα Compliance Assessment με το Alert Event στο οποίο αναφέρεται.

Το domain του property είναι η class `Compliance Assessment` και το range η class `Alert Event`. Η σχεδιαστική λογική του μπορεί να αποδοθεί ως εξής:

Compliance Assessment → assessesEvent → Alert Event

Στην παρούσα proof-of-concept έκδοση δεν υπάρχουν direct individuals της class `Compliance Assessment` και, κατά συνέπεια, το `assessesEvent` δεν χρησιμοποιείται σε instance-level assertions. Διατηρείται στο schema της οντολογίας ως μέρος μιας επεκτάσιμης δομής, ώστε σε μελλοντική έκδοση να μπορεί ένα ρητά αναπαριστώμενο assessment να συνδέεται με το συγκεκριμένο event που αξιολογεί.

Η παρουσία του property επομένως εκφράζει σχεδιαστική δυνατότητα του μοντέλου και όχι ενεργή σχέση μεταξύ των 76 individuals της τελικής proof-of-concept υλοποίησης.



Σχήμα 4.50 Object property `assessesEvent` στο Protégé μαζί με τα annotations, το domain και το range του.

Το `assessesEvent` επομένως ενισχύει τη διαδρομή από τα πραγματικά δεδομένα προς την αξιολόγηση. Δεν συνδέει απλώς δύο τεχνικά στοιχεία, αλλά δείχνει ότι η αξιολόγηση βασίζεται σε συγκεκριμένο alarm event και όχι σε γενική ή αόριστη παραδοχή.

4.6.2 `assessesRequirement`

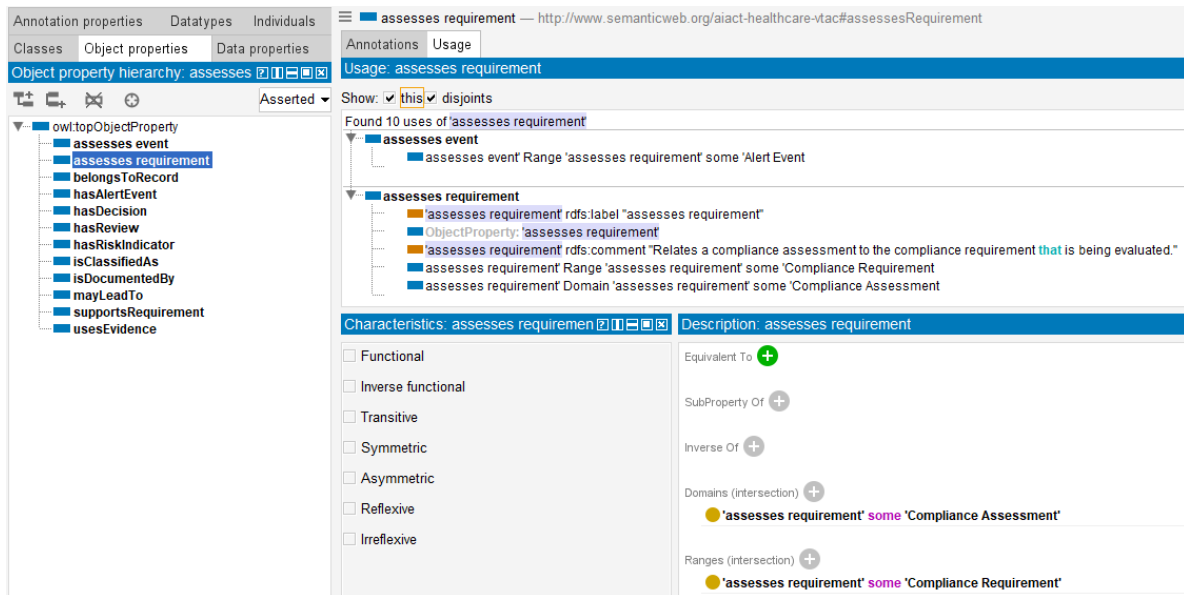
Το object property `assessesRequirement` έχει σχεδιαστεί για να συνδέει ένα `Compliance Assessment` με το `Compliance Requirement` που αποτελεί αντικείμενο της αντίστοιχης αξιολόγησης.

Το domain του property είναι η class `Compliance Assessment` και το range η class `Compliance Requirement`:

`Compliance Assessment` → `assessesRequirement` → `Compliance Requirement`

Στην τελική proof-of-concept έκδοση δεν έχουν δημιουργηθεί direct `Compliance Assessment` individuals και επομένως το property δεν συμμετέχει σε instance-level assertions. Παραμένει στο schema ως σημείο μελλοντικής επέκτασης, ώστε μια περισσότερο αναλυτική υλοποίηση να μπορεί να αναπαριστά ρητά ποιο AI Act-related σημείο αναφοράς εξετάζεται από ένα συγκεκριμένο assessment.

Η ύπαρξη του `assessesRequirement` δεν δηλώνει ότι πραγματοποιήθηκε πραγματική αξιολόγηση νομικής συμμόρφωσης. Αντίθετα, προσδιορίζει τη σημασιολογική δομή που θα μπορούσε να χρησιμοποιηθεί σε μελλοντική επέκταση για την οργανωμένη αναπαράσταση μιας τέτοιας διαδικασίας.



Σχήμα 4.51 Object property `assessesRequirement` στο Protégé μαζί με τα annotations, το usage, το domain και το range του.

4.6.3 belongsToRecord

Το object property `belongsToRecord` χρησιμοποιείται για να συνδέσει ένα Alert Event με το Waveform Record από το οποίο προέρχεται. Η σχέση αυτή είναι βασική για την ιχνηλασιμότητα της οντολογίας, επειδή επιτρέπει να γνωρίζουμε σε ποιο καταγεγραμμένο τμήμα φυσιολογικού σήματος ανήκει κάθε alarm event.

Στο VTaC dataset, τα alarm events δεν υπάρχουν αποκομμένα από το αρχικό record. Κάθε event σχετίζεται με ένα συγκεκριμένο waveform record, το οποίο λειτουργεί ως πηγή ή πλαίσιο προέλευσης. Για τον λόγο αυτό, η οντολογία χρειάζεται μια ρητή σχέση που να αποτυπώνει αυτή τη σύνδεση.

Το `belongsToRecord` απαντά στο ερώτημα:

Από ποιο waveform record προέρχεται το συγκεκριμένο alert event;

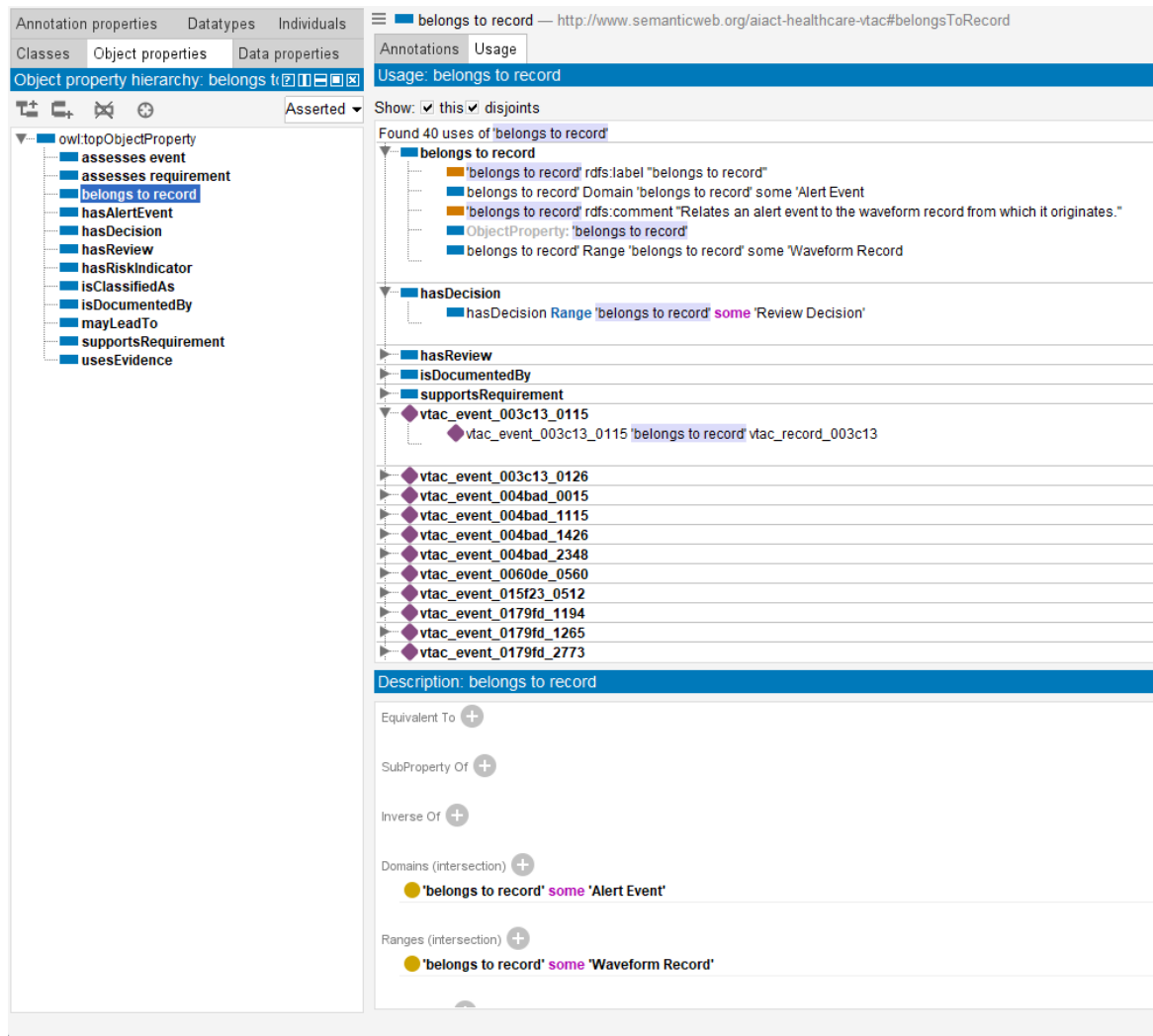
Η απάντηση σε αυτό το ερώτημα είναι σημαντική, γιατί χωρίς τη σύνδεση event–record η οντολογία θα μπορούσε να δείξει ότι υπάρχει ένα alarm event, αλλά όχι από πού προήλθε. Με τη χρήση του `belongsToRecord`, κάθε event τοποθετείται μέσα στο πραγματικό πλαίσιο του dataset.

Η σχέση αυτή έχει ιδιαίτερη σημασία στις περιπτώσεις όπου ένα waveform record περιλαμβάνει περισσότερα από ένα alert events. Για παράδειγμα, records όπως το `vtac_record_004bad` και το `vtac_record_0179fd` συνδέονται με περισσότερα από ένα VTaC events. Το `belongsToRecord` επιτρέπει να αναπαρασταθεί αυτή η πολλαπλή σύνδεση με σαφή και αναζητήσιμο τρόπο.

Σε επίπεδο οντολογικής αλυσίδας, η σχέση μπορεί να αποδοθεί ως εξής:

Alert Event → belongsToRecord → Waveform Record

Με αυτόν τον τρόπο, η οντολογία δεν ξεκινά την τεκμηρίωση από το event ως απομονωμένο στοιχείο, αλλά το συνδέει με το πραγματικό record προέλευσής του. Αυτό ενισχύει τη δυνατότητα traceability, επειδή κάθε event μπορεί να εντοπιστεί πίσω στο αντίστοιχο waveform record.



Σχήμα 4.52 Object property belongsToRecord στο Protégé μαζί με τα annotations, το usage, το domain και το range του.

4.6.4 hasAlertEvent

Το object property hasAlertEvent χρησιμοποιείται για να συνδέσει ένα Waveform Record με ένα Alert Event που περιλαμβάνεται ή σχετίζεται με αυτό το record. Η σχέση αυτή λειτουργεί από την πλευρά του record προς το event.

Σε σχέση με το belongsToRecord, το hasAlertEvent εκφράζει την ίδια βασική σύνδεση από την αντίθετη οπτική. Το belongsToRecord δείχνει σε ποιο record ανήκει ένα event. Το hasAlertEvent δείχνει ποια events περιλαμβάνει ή συνδέει ένα waveform record.

Η σχέση αυτή μπορεί να αποδοθεί ως εξής:

Waveform Record → hasAlertEvent → Alert Event

Το hasAlertEvent είναι σημαντικό επειδή επιτρέπει στην οντολογία να περιγράψει όχι μόνο από ποιο record προέρχεται ένα event, αλλά και ποια events συνδέονται με ένα συγκεκριμένο record. Αυτό είναι ιδιαίτερα χρήσιμο στις περιπτώσεις όπου ένα waveform record έχει περισσότερα από ένα alert events.

Ο ρόλος του `hasAlertEvent` βρίσκεται στο `operational layer` της οντολογίας. Δεν αφορά ακόμη την αξιολόγηση, την απόφαση, το `evidence` ή την απαίτηση συμμόρφωσης. Αφορά το αρχικό επίπεδο των πραγματικών δεδομένων: το `record` και τα `alert events` που προέρχονται από αυτό.

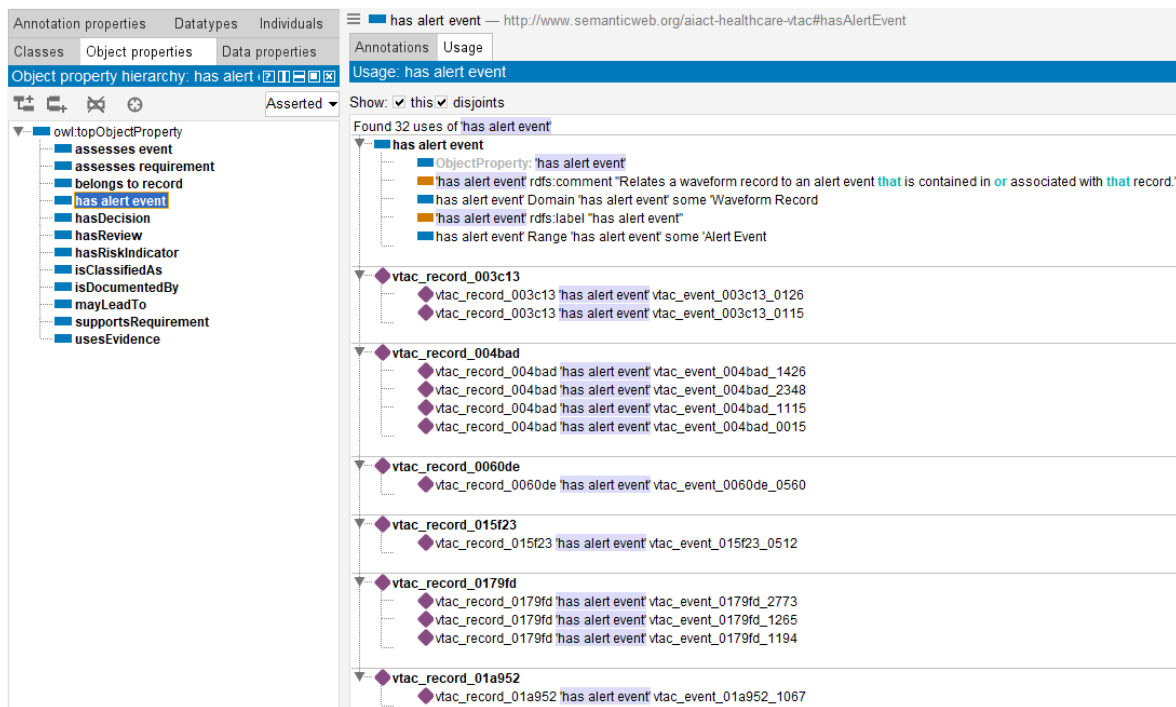
Η σχέση αυτή μπορεί να απαντήσει σε ερωτήματα όπως:

Ποια `alert events` περιλαμβάνει ένα συγκεκριμένο `waveform record`;

ή:

Ποια `records` συνδέονται με περισσότερα από ένα `events`;

Τέτοια ερωτήματα είναι χρήσιμα, επειδή δείχνουν ότι η οντολογία δεν περιορίζεται στην απλή καταγραφή μεμονωμένων `alarms`, αλλά μπορεί να αναπαραστήσει και τη δομή του `dataset`.



Σχήμα 4.53 Object property `hasAlertEvent` στο Protégé μαζί με τα annotations, το usage, το domain και το range του.

Το `hasAlertEvent` αποτελεί τη σχέση που επιτρέπει την ανάκτηση των `alarm events` από το αντίστοιχο `Waveform Record`. Από τη σκοπιά του `Alert Event`, η αντίστοιχη σχέση προέλευσης αποτυπώνεται μέσω του `belongsToRecord`. Οι σχέσεις αυτές αποτελούν το `operational` μέρος της οντολογίας και λειτουργούν παράλληλα με τις ξεχωριστές διαδρομές αξιολόγησης και τεκμηρίωσης.

4.6.5 `hasDecision`

Το object property `hasDecision` χρησιμοποιείται για να συνδέσει ένα `Expert Review` με την απόφαση που προκύπτει από αυτό το `review`. Δηλαδή, δηλώνει ότι μια διαδικασία αξιολόγησης καταλήγει σε ένα συγκεκριμένο αποτέλεσμα.

Η σχέση αυτή είναι σημαντική επειδή στην οντολογία διαχωρίζεται η πράξη της αξιολόγησης από το αποτέλεσμα της αξιολόγησης. Η class `Expert Review` αναπαριστά το στάδιο όπου ένα `alert event`

εξετάζεται από άνθρωπο ή ειδικό αξιολογητή. Η class Review Decision αναπαριστά το συμπέρασμα αυτής της αξιολόγησης. Το hasDecision είναι η σχέση που ενώνει αυτά τα δύο στοιχεία.

Το hasDecision απαντά στο ερώτημα:

Ποια απόφαση προέκυψε από αυτό το expert review;

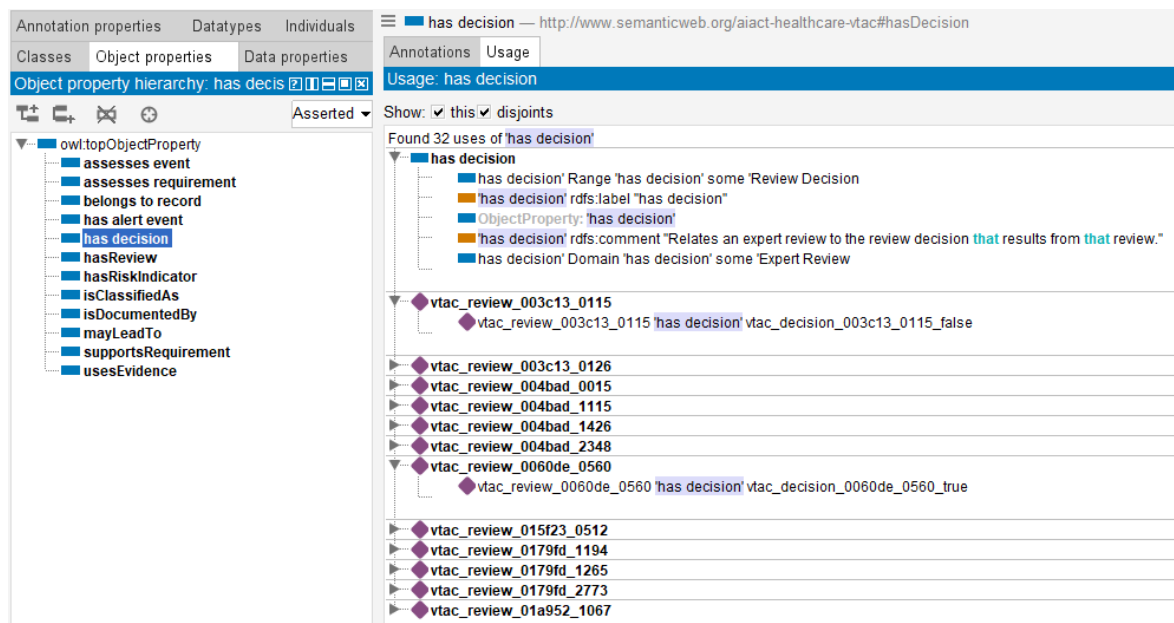
Η απάντηση μπορεί να οδηγεί σε ένα individual που ανήκει είτε στην class True Alarm είτε στην class False Alarm. Έτσι, η οντολογία μπορεί να δείξει αν ένα συγκεκριμένο review κατέληξε σε απόφαση true alarm ή false alarm.

Η ύπαρξη αυτής της σχέσης είναι απαραίτητη για να μη φαίνεται η απόφαση ως ανεξάρτητο ή αποκομμένο στοιχείο. Ένα vtac_decision ..._true ή vtac_decision ..._false individual δεν εμφανίζεται αυθαίρετα μέσα στην οντολογία. Συνδέεται με το αντίστοιχο review, το οποίο με τη σειρά του συνδέεται με το alert event που αξιολογήθηκε.

Σε επίπεδο βασικής αλυσίδας, το hasDecision υποστηρίζει το τμήμα:

Alert Event → Expert Review → Review Decision

Η σχέση αυτή είναι ιδιαίτερα χρήσιμη για τη διάκριση true και false alarms. Τα false alarms μπορούν στη συνέχεια να συνδεθούν επιπλέον με risk-related στοιχεία, ενώ τα true alarms ακολουθούν τη βασική διαδρομή τεκμηρίωσης.



Σχήμα 4.54 Object property hasDecision στο Protégé μαζί με τα annotations, το usage, το domain και το range του.

Το hasDecision συμβάλλει επίσης στην απαίτηση human review. Δεν αρκεί να υπάρχει μια τελική απόφαση. Πρέπει να φαίνεται ότι η απόφαση αυτή προήλθε από διαδικασία αξιολόγησης. Η σχέση hasDecision αποτυπώνει ακριβώς αυτή τη σύνδεση.

4.6.6 hasReview

Το object property `hasReview` χρησιμοποιείται για να συνδέσει ένα `Alert Event` με το `Expert Review` που σχετίζεται με αυτό. Δηλαδή, δηλώνει ότι ένα συγκεκριμένο `alarm event` έχει αξιολογηθεί μέσω μιας συγκεκριμένης διαδικασίας ανθρώπινης ή `expert` αξιολόγησης.

Η σχέση αυτή είναι βασική για τη λογική της οντολογίας, επειδή το `alert event` δεν πρέπει να εμφανίζεται ως ένα απλό τεχνικό συμβάν χωρίς αξιολόγηση. Για να μπορεί να τεκμηριωθεί η πορεία ενός `alarm`, πρέπει να φαίνεται ότι το event εξετάστηκε και ότι υπάρχει συγκεκριμένο `review` που συνδέεται με αυτό.

Το `hasReview` απαντά στο ερώτημα:

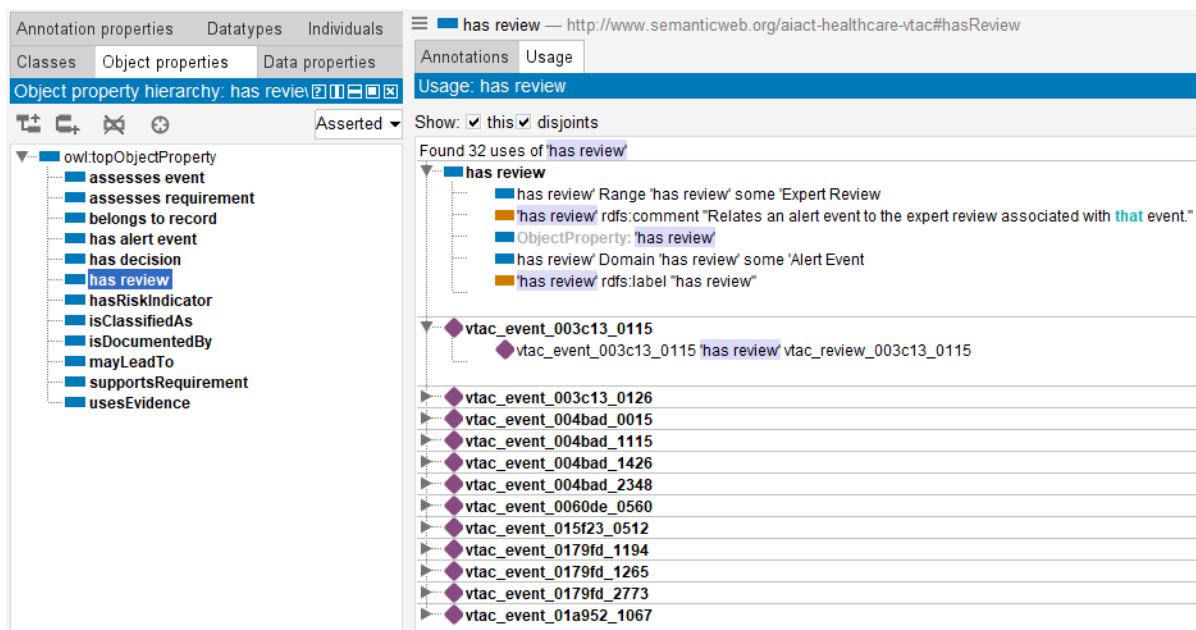
Ποιο expert review αντιστοιχεί σε αυτό το alert event;

Με αυτή τη σχέση, κάθε `vtac_event ... individual` μπορεί να συνδεθεί με το αντίστοιχο `vtac_review ... individual`. Έτσι, η οντολογία δεν αποθηκεύει απλώς το αποτέλεσμα της αξιολόγησης, αλλά αποτυπώνει και το ενδιάμεσο στάδιο της αξιολόγησης.

Η σχέση αυτή βρίσκεται στο `operational layer` και υποστηρίζει το τμήμα της βασικής οντολογικής αλυσίδας:

Alert Event → Expert Review → Review Decision

Το `hasReview` καλύπτει το πρώτο μέρος αυτής της διαδρομής, δηλαδή τη σύνδεση του `alarm event` με τη διαδικασία αξιολόγησης. Το `hasDecision` συνδέει στη συνέχεια το `review` με την τελική απόφαση. Επομένως, τα δύο properties λειτουργούν συμπληρωματικά.



Σχήμα 4.55 Object property `hasReview` στο Protégé μαζί με τα annotations, το usage, το domain και το range του.

Η ύπαρξη του `hasReview` είναι σημαντική και για τη σύνδεση με την απαίτηση `human review`. Η οντολογία δεν δηλώνει απλώς ότι υπάρχει ανθρώπινη αξιολόγηση σε γενικό επίπεδο. Αντίθετα, δείχνει για κάθε συγκεκριμένο VTaC event ποιο review το αφορά. Αυτό ενισχύει την τεκμηρίωση, επειδή η ανθρώπινη αξιολόγηση συνδέεται με πραγματικά `dataset-based events` και όχι με αφηρημένη περιγραφή.

Σε επίπεδο queries, η σχέση hasReview επιτρέπει την ανάκτηση όλων των alert events που έχουν συνδεθεί με expert review. Επίσης, επιτρέπει να εντοπιστεί ποιο review αντιστοιχεί σε κάθε event και να ακολουθηθεί η αλυσίδα μέχρι την τελική απόφαση.

4.6.7 hasRiskIndicator

Το object property hasRiskIndicator χρησιμοποιείται για να συνδέσει ένα Alert Event με ένα Risk Indicator. Η σχέση αυτή χρησιμοποιείται κυρίως για τα false alarm cases, τα οποία στην παρούσα οντολογία αποκτούν πρόσθετη risk-related σημασία.

Ο ρόλος του hasRiskIndicator είναι να δηλώσει ότι ένα συγκεκριμένο alert event μπορεί να συσχετιστεί με μια ένδειξη κινδύνου. Αυτό δεν σημαίνει ότι το event προκάλεσε πραγματική βλάβη. Σημαίνει ότι, στο πλαίσιο της οντολογικής αναπαράστασης, το event μπορεί να θεωρηθεί σχετικό με πιθανό κίνδυνο ή με διαδικασία risk management.

Το property αυτό απαντά στο ερώτημα:

Ποια ένδειξη κινδύνου συνδέεται με ένα συγκεκριμένο alert event;

Η σχέση αυτή είναι ιδιαίτερα χρήσιμη για τα false alarms. Ένα false alarm δεν είναι μόνο μια λανθασμένη ή μη επιβεβαιωμένη ειδοποίηση. Σε ένα healthcare monitoring περιβάλλον μπορεί να έχει σημασία, επειδή τα επαναλαμβανόμενα ή μη έγκυρα alarms μπορούν να σχετίζονται με φαινόμενα όπως το alert fatigue.

Στην οντολογία, το hasRiskIndicator λειτουργεί ως το πρώτο βήμα της risk-related αλυσίδας:

Alert Event → Risk Indicator → Risk-Relevant Event → Potential Harm

Με αυτόν τον τρόπο, η οντολογία μπορεί να δείξει ποια false alarm events συνδέονται με risk indicators και μπορούν στη συνέχεια να συσχετιστούν με risk-relevant events και potential harm.

The screenshot displays the Protégé interface with the 'Object property hierarchy' tab selected. The hierarchy for 'has risk indicator' is shown on the left, including properties like 'assesses event', 'assesses requirement', 'belongs to record', 'has alert event', 'has decision', 'has review', 'has risk indicator', 'isClassifiedAs', 'isDocumentedBy', 'mayLeadTo', 'supportsRequirement', and 'usesEvidence'. The 'has risk indicator' property is highlighted. On the right, the 'Usage' tab is active, showing 20 uses of the property. The uses are listed as follows:

- has risk indicator** (rdf:type: 'has risk indicator')
- ObjectProperty: 'has risk indicator'**
- has risk indicator** Range 'has risk indicator' some 'Risk Indicator'
- has risk indicator** rdfs:comment "Relates an alert event to a risk indicator associated with *that* event."
- has risk indicator** Domain 'has risk indicator' some 'Alert Event'

Below these, several instances are listed, each with a diamond icon indicating a domain restriction:

- vtac_event_003c13_0115** (diamond icon) **vtac_event_003c13_0115** 'has risk indicator' vtac_indicator_003c13_0115_false_alarm
- vtac_event_003c13_0126** (diamond icon) **vtac_event_003c13_0126** 'has risk indicator' vtac_indicator_003c13_0126_false_alarm
- vtac_event_004bad_0015** (diamond icon) **vtac_event_004bad_0015** 'has risk indicator' vtac_indicator_004bad_0015_false_alarm
- vtac_event_004bad_1115** (diamond icon) **vtac_event_004bad_1115** 'has risk indicator' vtac_indicator_004bad_1115_false_alarm
- vtac_event_004bad_1426** (diamond icon) **vtac_event_004bad_1426** 'has risk indicator' vtac_indicator_004bad_1426_false_alarm
- vtac_event_004bad_2348** (diamond icon) **vtac_event_004bad_2348** 'has risk indicator' vtac_indicator_004bad_2348_false_alarm

Σχήμα 4.56 Object property hasRiskIndicator στο Protégé μαζί με τα annotations, το usage, το domain και το range του.

Η χρήση του hasRiskIndicator συμβάλλει επίσης στη σύνδεση με το requirement individual req_risk_management. Μέσω αυτής της σχέσης, τα false alarm cases μπορούν να υποστηρίξουν την ανάγκη τεκμηρίωσης και διαχείρισης κινδύνων, χωρίς η οντολογία να ισχυρίζεται ότι εκτελεί πλήρη risk assessment.

4.6.8 isClassifiedAs

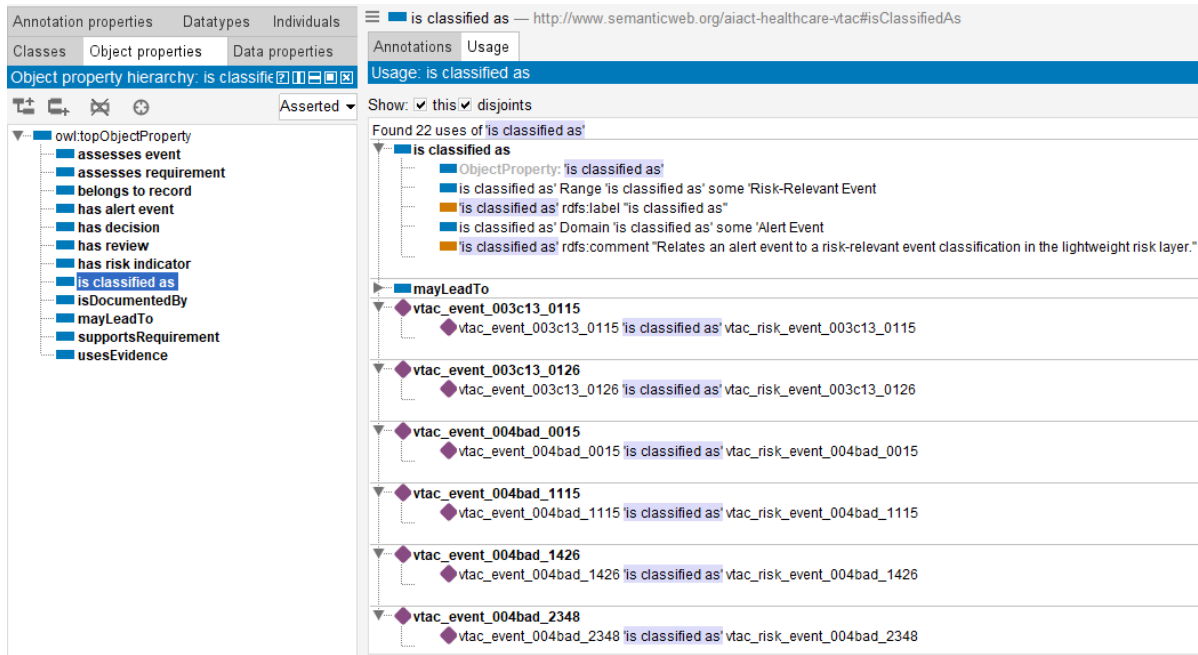
Το object property isClassifiedAs χρησιμοποιείται για να συνδέσει ένα Alert Event με ένα Risk-Related Event. Μέσω αυτής της σχέσης, ένα alarm event μπορεί να χαρακτηριστεί ως event που έχει σημασία από πλευράς κινδύνου.

Η σχέση αυτή είναι σημαντική, επειδή επιτρέπει τη μετάβαση από το operational επίπεδο στο risk-related επίπεδο. Το alert event παραμένει το πραγματικό alarm που προέρχεται από το VTaC dataset. Το risk-relevant event αποτελεί την ερμηνεία του ίδιου περιστατικού ως στοιχείου που μπορεί να σχετίζεται με risk management.

Το property αυτό απαντά στο ερώτημα:

Ως τι ταξινομείται ένα συγκεκριμένο alert event στο risk layer;

Στην παρούσα υλοποίηση, η σχέση isClassifiedAs χρησιμοποιείται κυρίως για τα false alarm cases. Ένα false alarm μπορεί να ταξινομηθεί ως Risk-Related Event, επειδή έχει πιθανή σημασία για τη διαχείριση κινδύνων στο healthcare monitoring περιβάλλον.



Σχήμα 4.57 Object property isClassifiedAs στο Protégé μαζί με τα annotations, το usage, το domain και το range του.

Η χρήση του isClassifiedAs βοηθά να μη συγχέεται το αρχικό alarm event με την risk-related ερμηνεία του. Το event παραμένει instance της operational class Ventricular Tachycardia Alarm, ενώ παράλληλα μπορεί να συνδεθεί με risk-relevant αναπαράσταση μέσω της σχέσης αυτής.

Σε επίπεδο risk-related αλυσίδας, το isClassifiedAs υποστηρίζει τη μετάβαση:

Alert Event → Risk-Relevant Event

Έτσι, το μοντέλο μπορεί να αναζητήσει ποια VTaC alert events έχουν ταξινομηθεί ως risk-relevant και ποια από αυτά συνδέονται με potential harm ή με το requirement req_risk_management.

4.6.9 isDocumentedBy

Το object property isDocumentedBy συνδέει ένα Alert Event με το Evidence Record που χρησιμοποιείται για την τεκμηρίωσή του. Η σχέση αυτή είναι ιδιαίτερα σημαντική, επειδή αποφεύγει την απευθείας σύνδεση του πραγματικού event με ένα requirement individual.

Η βασική διαδρομή τεκμηρίωσης αποτυπώνεται ως εξής:

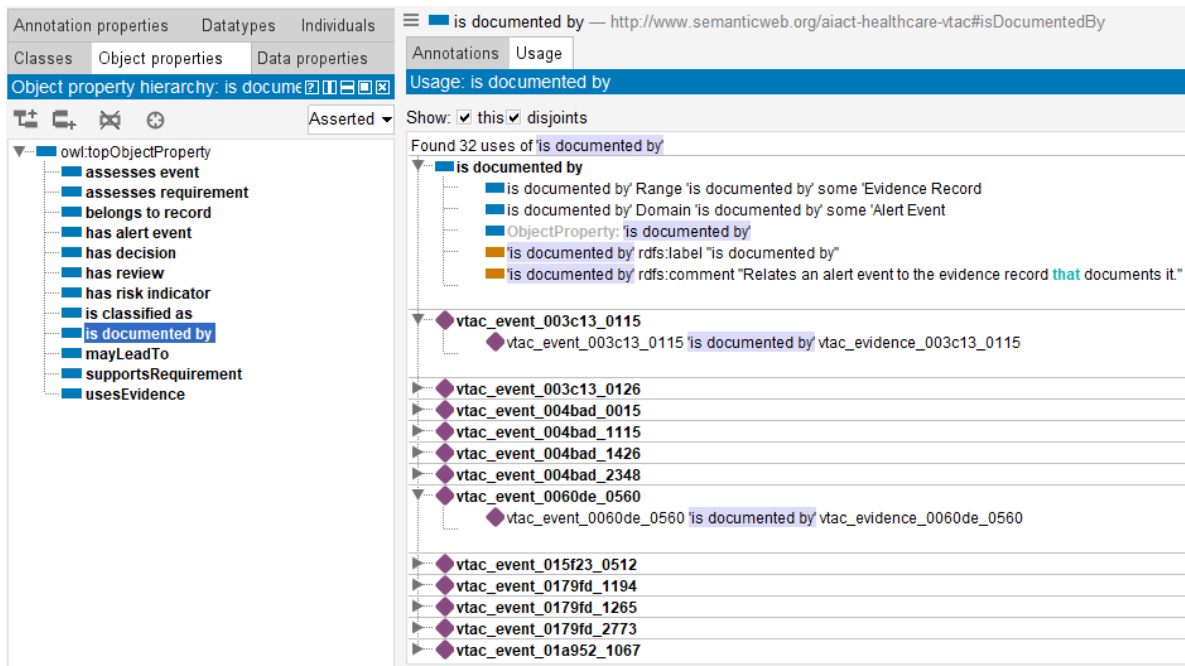
Alert Event --isDocumentedBy→ Evidence Record --supportsRequirement→ Requirement individual

Παράλληλα, η ανθρώπινη αξιολόγηση του ίδιου event αναπαρίσταται μέσω της ξεχωριστής διαδρομής:

Alert Event --hasReview→ Expert Review --hasDecision→ Review Decision

Οι δύο διαδρομές συνδυάζονται γύρω από το ίδιο Alert Event. Δεν υπάρχει άμεση object-property σχέση μεταξύ Review Decision και Evidence Record.

Με αυτόν τον τρόπο, το isDocumentedBy λειτουργεί ως κρίσιμος σύνδεσμος ανάμεσα στο operational μέρος και στο evidence and assessment μέρος της οντολογίας και υποστηρίζει το requirement-to-evidence mapping χωρίς να δηλώνει από μόνο του νομική συμμόρφωση.



Σχήμα 4.58 Object property isDocumentedBy στο Protégé μαζί με τα annotations, το usage, το domain και το range του.

Η χρήση του property αυτού ενισχύει τη διαφάνεια και την ιχνηλασιμότητα του μοντέλου. Κάθε evidence record έχει σαφές event αναφοράς, ενώ κάθε event μπορεί να εντοπιστεί μέσα από το τεκμήριο που το συνοδεύει. Αυτό είναι απαραίτητο για την τεκμηριωμένη σύνδεση με τις selected AI Act-related απαιτήσεις.

4.6.10 mayLeadTo

Το object property mayLeadTo χρησιμοποιείται για να συνδέσει ένα Risk-Relevant Event με ένα Potential Harm. Η σχέση αυτή εκφράζει ότι ένα risk-relevant event μπορεί να σχετίζεται με πιθανή αρνητική επίπτωση.

Ο ρόλος του mayLeadTo είναι να ολοκληρώσει τη risk-related διαδρομή της οντολογίας. Μέσω αυτής της σχέσης, ένα false alarm case που έχει ταξινομηθεί ως risk-relevant μπορεί να συνδεθεί με πιθανή επίπτωση, όπως το alert fatigue.

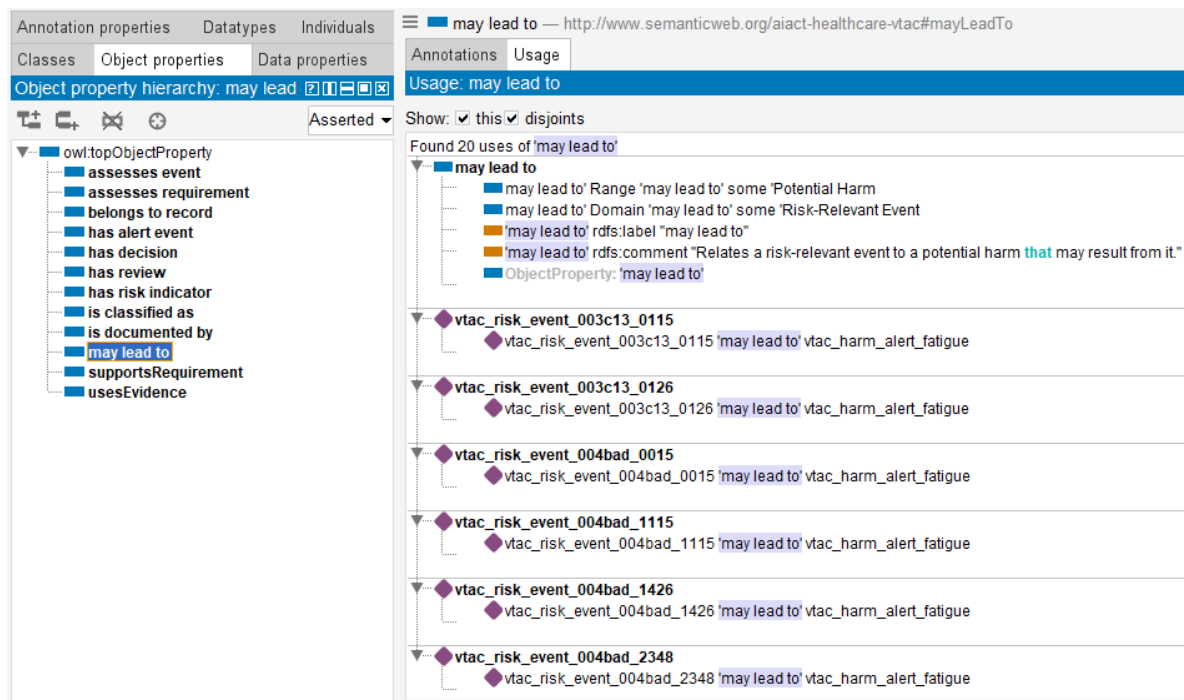
Το property αυτό απαντά στο ερώτημα:

Σε ποια πιθανή επίπτωση μπορεί να οδηγήσει ή με ποια πιθανή επίπτωση μπορεί να συνδεθεί ένα risk-relevant event;

Η διατύπωση mayLeadTo είναι σκόπιμα προσεκτική. Δεν δηλώνει ότι μια βλάβη έχει πράγματι συμβεί. Δηλώνει πιθανή σύνδεση ή πιθανή συνέπεια. Αυτό είναι σημαντικό, επειδή η παρούσα οντολογία δεν περιλαμβάνει κλινική αξιολόγηση πραγματικής βλάβης ούτε ποσοτική εκτίμηση κινδύνου.

Σε επίπεδο risk-related αλυσίδας, η σχέση μπορεί να αποδοθεί ως εξής:

Risk-Relevant Event → mayLeadTo → Potential Harm



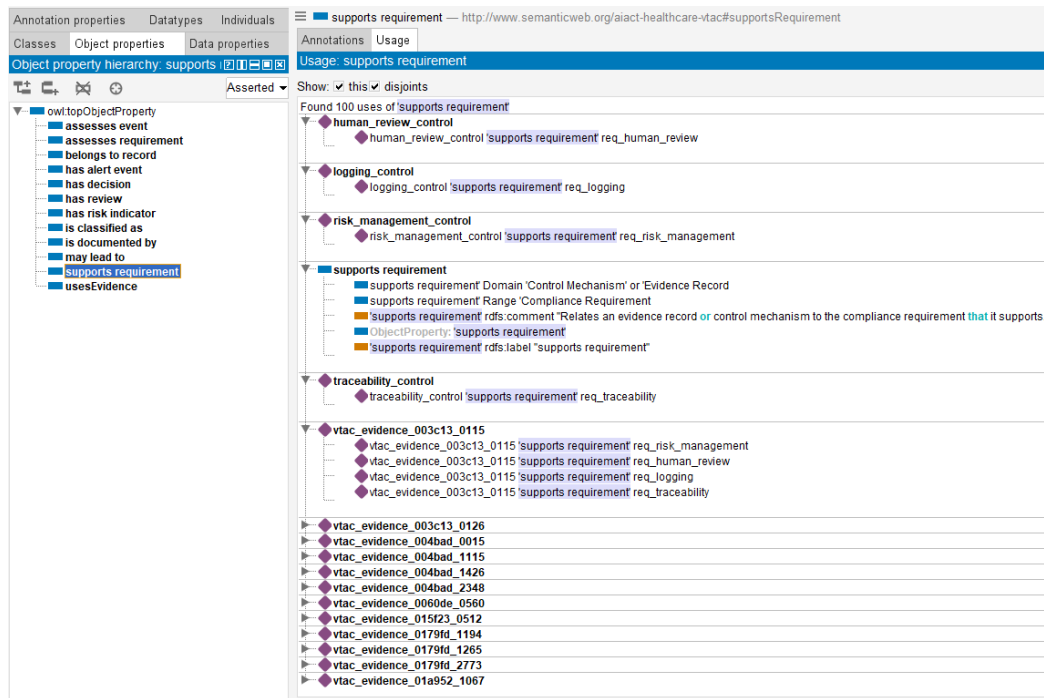
Σχήμα 4.59 Object property mayLeadTo στο Protégé μαζί με τα annotations, το usage, το domain και το range του.

Στην υλοποίηση, το mayLeadTo επιτρέπει τη σύνδεση των risk-relevant false alarm cases με το individual vtac_harm_alert_fatigue. Με αυτόν τον τρόπο, η οντολογία μπορεί να δείξει ότι τα false alarm cases έχουν potential harm-related σημασία, χωρίς να υπερβαίνει τα όρια του proof-of-concept.

Η σχέση αυτή είναι επίσης χρήσιμη για τη σύνδεση με το req_risk_management. Τα Risk-Relevant Event individuals που συνδέονται με Potential Harm μέσω του mayLeadTo παρέχουν την πρόσθετη risk-related αναπαράσταση που χρησιμοποιείται στο proof-of-concept για τη σύνδεση με το req_risk_management. Η σχέση δεν δηλώνει ότι έχει αποδειχθεί πραγματική βλάβη ή αιτιώδης σχέση.

4.6.11 supportsRequirement

Το object property supportsRequirement αποτελεί ένα από τα σημαντικότερα properties της οντολογίας, καθώς υλοποιεί το requirement-to-evidence mapping της παρούσας εργασίας. Ο ρόλος του είναι να συνδέει ένα Evidence Record με ένα Compliance Requirement individual, επιτρέποντας έτσι τη σημασιολογική μετάβαση από την τεκμηρίωση ενός πραγματικού healthcare monitoring event προς ένα επιλεγμένο AI Act-related σημείο αναφοράς.



Σχήμα 4.60 Object property supportsRequirement στο Protégé μαζί με τα annotations, το usage, το domain και το range του.

Το domain του supportsRequirement είναι η class Evidence Record, ενώ το range του είναι η class Compliance Requirement. Η βασική σχέση αποτυπώνεται ως εξής:

Evidence Record --supportsRequirement→ Compliance Requirement

Στην πραγματική εφαρμογή της οντολογίας, η σχέση αυτή αποτελεί συνέχεια της διαδρομής τεκμηρίωσης που ξεκινά από το Alert Event:

Alert Event --isDocumentedBy→ Evidence Record --supportsRequirement→ Requirement individual

Με αυτή τη δομή, το πραγματικό alarm event δεν συνδέεται απευθείας με ένα requirement individual. Παρεμβάλλεται το Evidence Record, το οποίο λειτουργεί ως τεκμηριωτικός κόμβος και διαχωρίζει το operational περιστατικό από το normative μέρος του μοντέλου. Η σχεδιαστική αυτή επιλογή επιτρέπει το requirement mapping να παραμένει ρητό, αναζητήσιμο και διακριτό από το ίδιο το event.

Στην παρούσα υλοποίηση, το range του supportsRequirement περιλαμβάνει individuals που ανήκουν στις ειδικότερες classes της Compliance Requirement. Τα τέσσερα requirement individuals που χρησιμοποιούνται είναι:

- req_logging
- req_human_review
- req_traceability
- req_risk_management

Τα req_logging, req_human_review και req_traceability χρησιμοποιούνται για τη σημασιολογική χαρτογράφηση των evidence records που σχετίζονται με τα βασικά στοιχεία τεκμηρίωσης των επιλεγμένων VTaC events. Τα evidence records των false alarm cases μπορούν επιπλέον, στο πλαίσιο του lightweight risk-related σχεδιασμού της οντολογίας, να συνδέονται με το req_risk_management.

Η χρήση του `supportsRequirement` δεν σημαίνει ότι το συγκεκριμένο `requirement` έχει νομικά ικανοποιηθεί ούτε ότι το `Evidence Record` αποτελεί από μόνο του επαρκή απόδειξη συμμόρφωσης με τον AI Act. Τα τέσσερα `requirement individuals` αποτελούν εφαρμοσμένα AI Act-related σημεία αναφοράς του συγκεκριμένου `proof-of-concept`. Κατά συνέπεια, η σχέση δηλώνει ότι το διαθέσιμο `evidence` έχει χαρτογραφηθεί σημασιολογικά προς ένα συγκεκριμένο σημείο αναφοράς και όχι ότι έχει πραγματοποιηθεί πλήρης αξιολόγηση νομικής συμμόρφωσης [3].

Για παράδειγμα, ένα `Evidence Record` που τεκμηριώνει ένα συγκεκριμένο VTaC alarm event μπορεί να συνδέεται μέσω του `supportsRequirement` με το `req_logging`, το `req_human_review` και το `req_traceability`. Η πληροφορία αυτή επιτρέπει να ανακτηθούν μέσω της οντολογίας τα `evidence records` που έχουν συσχετιστεί με κάθε επιλεγμένο `requirement individual` και, αντίστροφα, να εξεταστεί ποια `requirement mappings` έχουν αποδοθεί στο `evidence` ενός συγκεκριμένου event.

Στα false alarm cases, το corresponding `Evidence Record` μπορεί να συνδέεται επιπλέον με το `req_risk_management`. Η σύνδεση αυτή λειτουργεί συμπληρωματικά προς το lightweight risk-related μέρος της οντολογίας, όπου το ίδιο alarm event μπορεί να συνδέεται με Risk Indicator και Risk-Relevant Event και, μέσω του τελευταίου, με Potential Harm. Η σύνδεση του `Evidence Record` με το `req_risk_management` αποτελεί σχεδιαστική επιλογή του `proof-of-concept` και δεν πρέπει να ερμηνεύεται ως κανονιστικός κανόνας σύμφωνα με τον οποίο κάθε false alarm συνεπάγεται αυτομάτως συγκεκριμένη υποχρέωση διαχείρισης κινδύνου.

Το `supportsRequirement` είναι επίσης ιδιαίτερα σημαντικό για τη λειτουργική επικύρωση της οντολογίας. Μέσω DL Queries μπορεί να εξεταστεί ποια `Evidence Record individuals` υποστηρίζουν συγκεκριμένα `requirement individuals`, ενώ μέσω SPARQL μπορεί να ανακτηθεί συγκεντρωτικά η διαδρομή:

Alert Event → Evidence Record → Requirement individual

Η δυνατότητα αυτή αποτελεί βασικό στοιχείο της απαίτησης για αναζητήσιμη και ελέγξιμη σημασιολογική τεκμηρίωση που επιδιώκει η παρούσα εργασία.

Συνολικά, το `supportsRequirement` λειτουργεί ως ο βασικός σύνδεσμος μεταξύ του `evidence and assessment` μέρους και του `normative / AI Act-related` μέρους της οντολογίας. Μέσω αυτού υλοποιείται το `requirement-to-evidence mapping`, διατηρώντας σαφή διάκριση μεταξύ πραγματικού event, διαθέσιμου `evidence` και κανονιστικού σημείου αναφοράς.

.

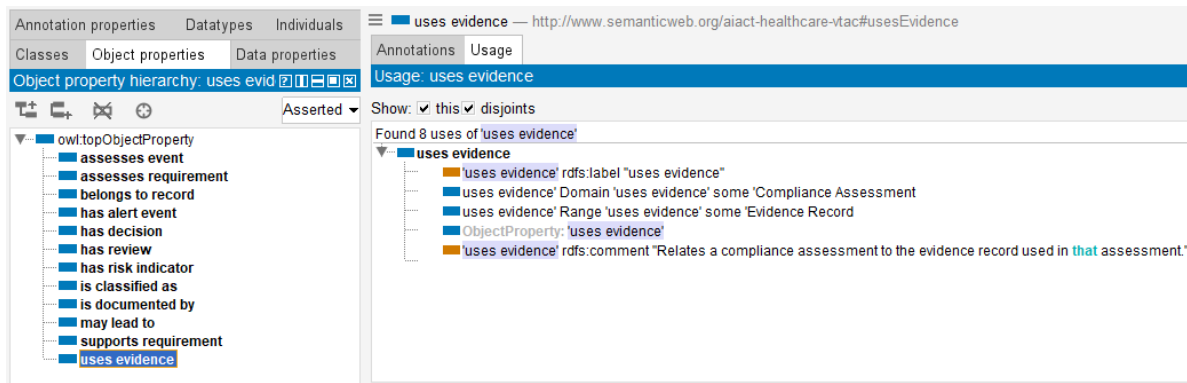
4.6.12 usesEvidence

Το object property `usesEvidence` έχει σχεδιαστεί για να συνδέει ένα `Compliance Assessment` με το `Evidence Record` που θα χρησιμοποιείται ως τεκμηριωτική βάση της αντίστοιχης αξιολόγησης.

Το domain του property είναι η class `Compliance Assessment` και το range η class `Evidence Record`:

`Compliance Assessment → usesEvidence → Evidence Record`

Το property συμπληρώνει σε επίπεδο schema τα `assessesEvent` και `assessesRequirement`. Σε μια μελλοντική επέκταση, τα τρία properties θα μπορούσαν να χρησιμοποιηθούν από κοινού ώστε ένα `Compliance Assessment individual` να δηλώνει ποιο event αξιολογεί, ποιο `requirement concept` εξετάζει και ποιο `Evidence Record` χρησιμοποιεί.



Σχήμα 4.61 Object property usesEvidence στο Protégé μαζί με τα annotations, το usage, το domain και το range του.

Στην παρούσα proof-of-concept έκδοση δεν υπάρχουν direct Compliance Assessment individuals και επομένως το usesEvidence δεν χρησιμοποιείται σε instance-level assertions. Η πραγματικά υλοποιημένη σύνδεση τεκμηρίωσης πραγματοποιείται μέσω των properties:

Alert Event → isDocumentedBy → Evidence Record

Evidence Record → supportsRequirement → Compliance Requirement

Το usesEvidence διατηρείται επομένως ως στοιχείο του επεκτάσιμου schema και δεν πρέπει να συγχέεται με τις instance-level διαδρομές που χρησιμοποιούνται για τα 12 VTaC events.

4.6.13 Συνολική αποτίμηση των object properties

Τα object properties αποτελούν τον βασικό μηχανισμό μέσω του οποίου οι classes και τα individuals της οντολογίας συνδέονται σε λειτουργικές σημασιολογικές διαδρομές.

Η πρώτη ομάδα σχέσεων υποστηρίζει τη σύνδεση των VTaC events με τα waveform records και την ανθρώπινη αξιολόγηση. Τα belongsToRecord, hasAlertEvent, hasReview και hasDecision επιτρέπουν την ανάκτηση της προέλευσης ενός event και της αντίστοιχης διαδικασίας αξιολόγησής του.

Η δεύτερη ομάδα αφορά την τεκμηρίωση και την αξιολόγηση. Τα isDocumentedBy και supportsRequirement συμμετέχουν άμεσα στις instance-level διαδρομές τεκμηρίωσης και requirement mapping των 12 VTaC events. Αντίθετα, τα assessesEvent, assessesRequirement και usesEvidence αποτελούν schema-level properties που έχουν προβλεφθεί για πιθανή μελλοντική αναπαράσταση ρητών Compliance Assessment individuals.

Η τρίτη ομάδα υποστηρίζει το requirement-to-evidence mapping. Το supportsRequirement συνδέει τα Evidence Records με τα επιλεγμένα requirement individuals req_logging, req_human_review, req_traceability και req_risk_management. Η σύνδεση αυτή αποτελεί σημασιολογικό mapping προς επιλεγμένα AI Act-related σημεία αναφοράς και όχι απόδειξη πλήρους νομικής συμμόρφωσης.

Η τέταρτη ομάδα υποστηρίζει το lightweight risk-related μέρος της οντολογίας. Τα hasRiskIndicator, isClassifiedAs και mayLeadTo επιτρέπουν την πρόσθετη αναπαράσταση των false alarm cases μέσω Risk Indicator, Risk-Relevant Event και Potential Harm.

Συνολικά, τα δώδεκα object properties επιτρέπουν στο μοντέλο να αναπαριστά όχι μία μοναδική γραμμική αλυσίδα, αλλά ένα σύνολο συνδεδεμένων σημασιολογικών διαδρομών. Μέσω αυτών

υποστηρίζονται η προέλευση των πραγματικών VTaC events, η ανθρώπινη αξιολόγηση, η τεκμηρίωση, το requirement-to-evidence mapping και η πρόσθετη risk-related αναπαράσταση.

4.7 Data properties της οντολογίας

Τα data properties χρησιμοποιούνται στην οντολογία για να συνδέσουν individuals με συγκεκριμένες τιμές δεδομένων. Σε αντίθεση με τα object properties, τα οποία συνδέουν ένα individual με ένα άλλο individual, τα data properties συνδέουν ένα individual με μια απλή τιμή, όπως κείμενο, identifier, label ή boolean ένδειξη.

Στην παρούσα υλοποίηση, τα data properties έχουν κυρίως περιγραφικό και τεκμηριωτικό ρόλο. Δεν δημιουργούν τη βασική σημασιολογική αλυσίδα της οντολογίας, καθώς αυτή υλοποιείται κυρίως μέσω των object properties. Ωστόσο, συμπληρώνουν το μοντέλο προσθέτοντας συγκεκριμένες τιμές στα individuals που συμμετέχουν στην αλυσίδα.

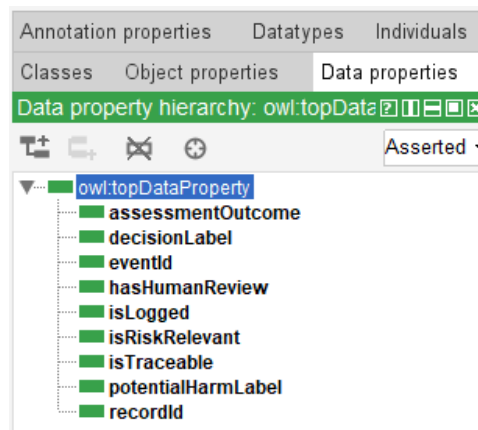
Η χρήση τους είναι σημαντική, επειδή επιτρέπει στην οντολογία να μην περιορίζεται μόνο στη δομή των σχέσεων. Μέσω των data properties, τα πραγματικά VTaC events μπορούν να εμφανίζονται με σαφή αναγνωριστικά, οι review decisions με λεκτικές τιμές απόφασης, τα compliance assessments με αποτέλεσμα αξιολόγησης και τα false alarm cases με risk-related ενδείξεις.

Για παράδειγμα, ένα object property όπως το hasReview δείχνει ότι ένα alert event συνδέεται με expert review. Το object property hasDecision δείχνει ότι το review οδηγεί σε συγκεκριμένη review decision. Το data property decisionLabel, όμως, προσθέτει την τιμή της απόφασης, όπως true alarm ή false alarm. Επομένως, τα data properties λειτουργούν ως υποστηρικτικό επίπεδο πληροφορίας πάνω στη βασική δομή των σχέσεων.

Συνολικά, στην οντολογία χρησιμοποιήθηκαν εννέα data properties:

- assessmentOutcome
- decisionLabel
- eventId
- hasHumanReview
- isLogged
- isRiskRelevant
- isTraceable
- potentialHarmLabel
- recordId

Στις επόμενες υποενότητες παρουσιάζονται τα εννέα data properties της οντολογίας με τη σειρά που εμφανίζονται στο Protégé.

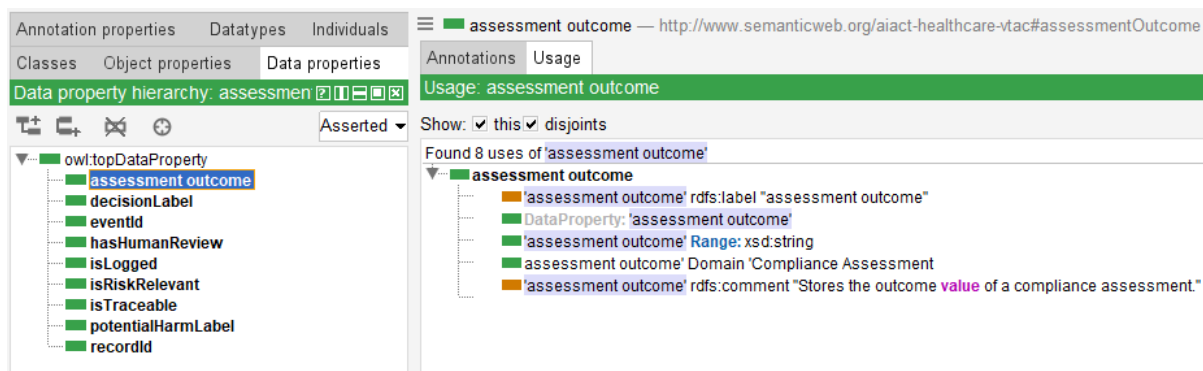


Σχήμα 4.62 Ιεραρχία των data properties της οντολογίας στο Protégé.

4.7.1 assessmentOutcome

Το data property `assessmentOutcome` έχει σχεδιαστεί για την αναπαράσταση περιγραφικού αποτελέσματος μιας Compliance Assessment. Το domain του συνδέεται με τη class `Compliance Assessment` και επιτρέπει, σε μελλοντική επέκταση της οντολογίας, ένα `assessment individual` να συνοδεύεται από τιμή που περιγράφει το αποτέλεσμα της αξιολόγησης.

Στην παρούσα proof-of-concept έκδοση δεν έχουν δημιουργηθεί direct `Compliance Assessment` individuals και, κατά συνέπεια, το `assessmentOutcome` δεν χρησιμοποιείται σε instance-level assertions. Διατηρείται στο schema της ontology ως στοιχείο της σχεδιασμένης δομής για πιθανή μελλοντική αναπαράσταση ρητών διαδικασιών αξιολόγησης.



Σχήμα 4.63 Data property assessmentOutcome στο Protégé.

Η σχεδιαστική του σχέση με τα object properties `assessesEvent`, `assessesRequirement` και `usesEvidence` παραμένει η εξής:

Compliance Assessment → `assessesEvent` → **Alert Event**

Compliance Assessment → `assessesRequirement` → **Compliance Requirement**

Compliance Assessment → `usesEvidence` → **Evidence Record**

Compliance Assessment → `assessmentOutcome` → **περιγραφική τιμή αποτελέσματος**

Η δομή αυτή δεν αποτελεί απόδειξη κανονιστικής συμμόρφωσης, αλλά προδιαγράφει έναν πιθανό μηχανισμό επέκτασης του μοντέλου για τη ρητή αναπαράσταση `assessment instances`.

4.7.2 decisionLabel

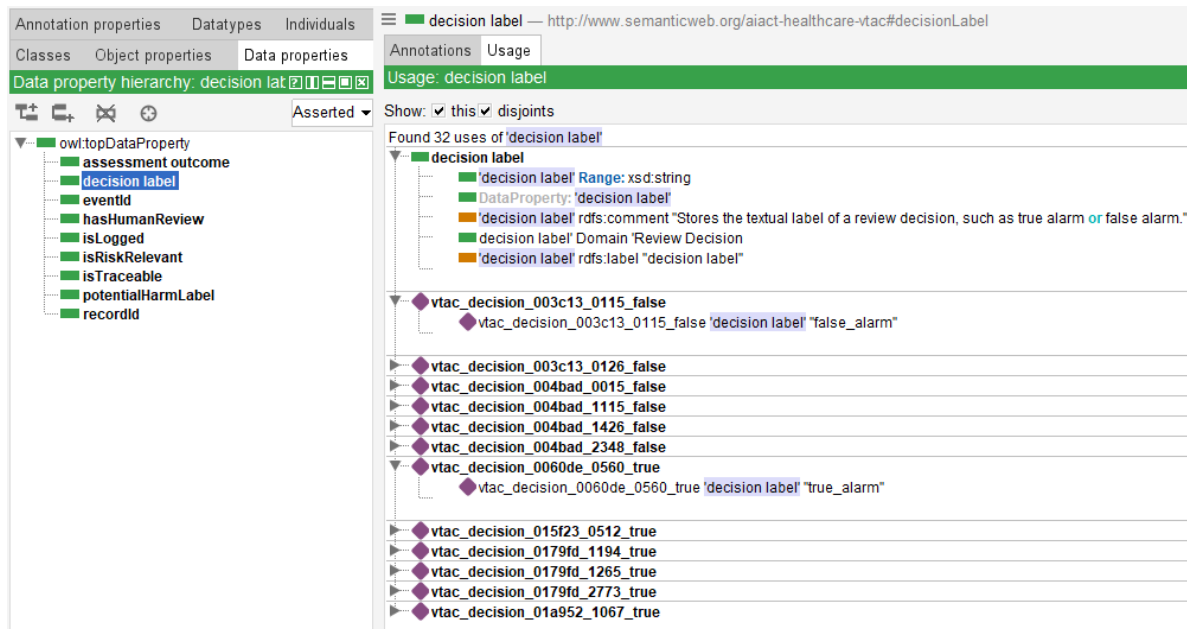
Το data property decisionLabel χρησιμοποιείται για να αποδώσει λεκτική τιμή σε ένα Review Decision individual. Στην παρούσα υλοποίηση, οι αποφάσεις των VTaC events διακρίνονται σε true alarm και false alarm, σύμφωνα με τα expert labels του dataset.

Ο ρόλος του decisionLabel είναι να αποθηκεύει με απλό και αναγνώσιμο τρόπο το αποτέλεσμα της αξιολόγησης. Ενώ οι classes True Alarm και False Alarm οργανώνουν οντολογικά τις αποφάσεις, το decisionLabel προσθέτει την αντίστοιχη τιμή ως περιγραφικό δεδομένο.

Το property αυτό απαντά στο ερώτημα:

Ποια είναι η τιμή της απόφασης για ένα συγκεκριμένο review decision;

Για παράδειγμα, ένα decision individual που αντιστοιχεί σε true alarm μπορεί να έχει decision label που δηλώνει ότι η απόφαση είναι true alarm. Αντίστοιχα, ένα decision individual που αντιστοιχεί σε false alarm μπορεί να έχει decision label που δηλώνει ότι η απόφαση είναι false alarm.



Σχήμα 4.64 Data property decisionLabel στο Protégé.

Η ύπαρξη του decisionLabel είναι χρήσιμη επειδή προσθέτει ένα άμεσα αναγνώσιμο επίπεδο πληροφορίας. Ο αναγνώστης μπορεί να δει όχι μόνο σε ποια class ανήκει ένα decision individual, αλλά και ποια λεκτική τιμή decision έχει αποδοθεί σε αυτό.

Επιπλέον, το decisionLabel βοηθά στη σύνδεση της οντολογίας με τα αρχικά labels του VTaC dataset. Η οντολογία δεν παράγει μόνη της την απόφαση. Η απόφαση προέρχεται από τα expert labels του dataset και εισάγεται στο μοντέλο ως δεδομένο [4].

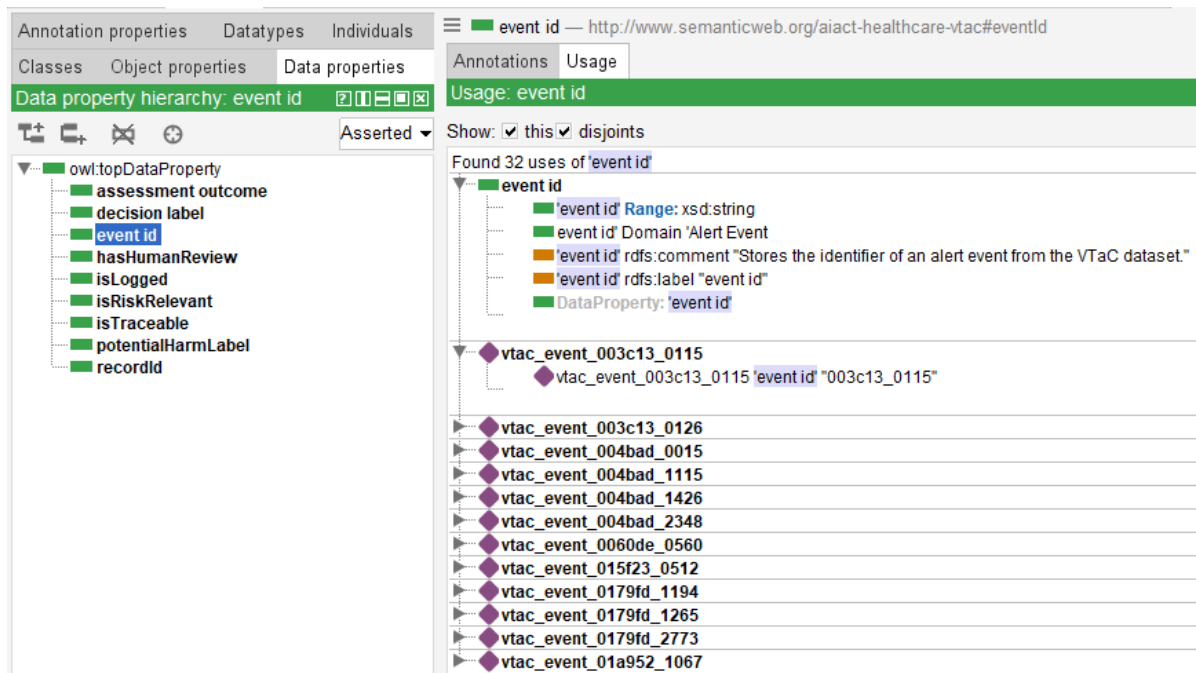
4.7.3 eventId

Το data property eventId χρησιμοποιείται για να αποδώσει σε ένα Alert Event το αναγνωριστικό του event. Στο VTaC dataset, κάθε event διαθέτει συγκεκριμένο identifier, το οποίο επιτρέπει την αναφορά σε αυτό με σαφή και μοναδικό τρόπο.

Ο ρόλος του `eventId` είναι σημαντικός για την ιχνηλασιμότητα. Ένα event δεν πρέπει να εμφανίζεται μόνο ως οντολογικό individual, αλλά πρέπει να μπορεί να συνδεθεί και με το αντίστοιχο identifier του αρχικού dataset. Με αυτόν τον τρόπο, η οντολογία διατηρεί σύνδεση με την πραγματική πηγή των δεδομένων.

Το property αυτό απαντά στο ερώτημα:

Ποιο είναι το αναγνωριστικό ενός συγκεκριμένου alert event;



Σχήμα 4.65 Data property `eventId` στο Protégé.

Για παράδειγμα, ένα individual όπως το `vtac_event_0060de_0560` μπορεί να διαθέτει `eventId` που αντιστοιχεί στο συγκεκριμένο VTaC event [4]. Αυτό βοηθά τον αναγνώστη να καταλάβει ότι το individual δεν είναι υποθετικό, αλλά βασίζεται σε πραγματική εγγραφή του dataset.

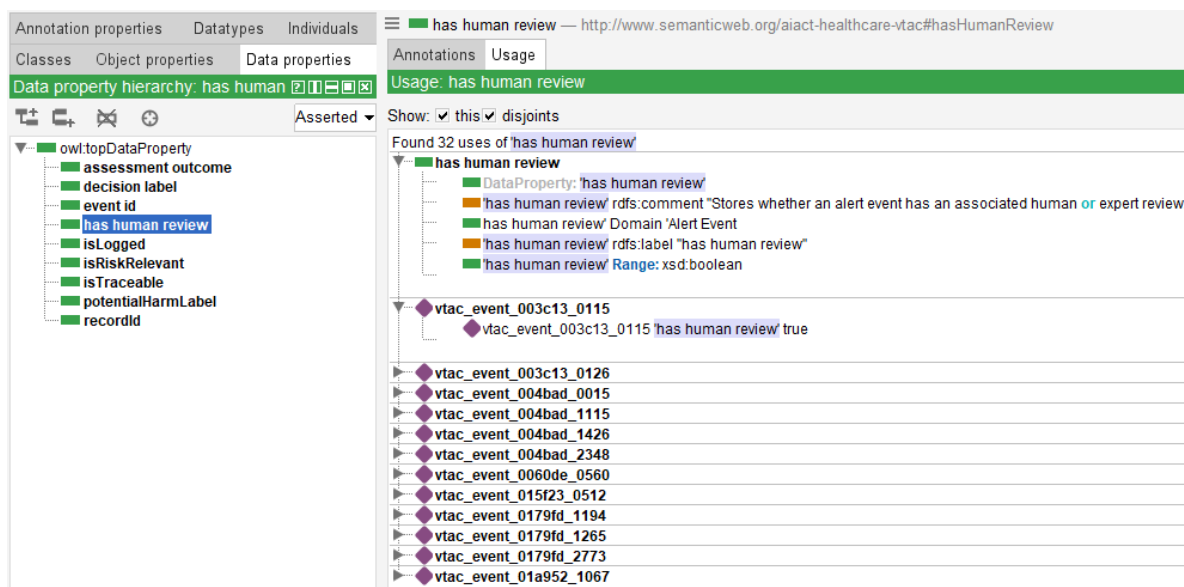
Η ύπαρξη του `eventId` υποστηρίζει επίσης τη μελλοντική αναζήτηση ή αντιστοίχιση με εξωτερικά αρχεία. Αν κάποιος θέλει να εντοπίσει ένα event στο αρχικό dataset ή να το συσχετίσει με άλλο τεκμήριο, το id λειτουργεί ως σημείο αναφοράς.

Σε επίπεδο τεκμηρίωσης, το `eventId` ενισχύει τη διαφάνεια της υλοποίησης. Δείχνει ότι το alert event της οντολογίας αντιστοιχεί σε συγκεκριμένο dataset-based event και όχι σε γενική ή τεχνητά κατασκευασμένη περίπτωση.

4.7.4 `hasHumanReview`

Το data property `hasHumanReview` χρησιμοποιείται για να δηλώσει, μέσω boolean τιμής, αν ένα Alert Event έχει ενταχθεί στην οντολογική αναπαράσταση ως event που διαθέτει ανθρώπινη expert αξιολόγηση. Στο πλαίσιο της παρούσας υλοποίησης, η ουσιαστική σημασιολογική αναπαράσταση της ανθρώπινης αξιολόγησης πραγματοποιείται μέσω της σύνδεσης ενός Alert Event με Expert Review και στη συνέχεια με Review Decision.

Ο ρόλος του `hasHumanReview` είναι επομένως συμπληρωματικός. Παρέχει μια άμεσα αναγνώσιμη boolean ένδειξη, χωρίς να αντικαθιστά τις object-property σχέσεις που αναπαριστούν αναλυτικότερα τη διαδικασία αξιολόγησης.



Σχήμα 4.66 Data property hasHumanReview στο Protégé.

Η βασική δομή είναι:

Alert Event --hasReview→ Expert Review --hasDecision→ Review Decision

και συμπληρωματικά:

Alert Event --hasHumanReview→ true/false

Το property αυτό απαντά στο ερώτημα:

Υπάρχει καταγεγραμμένη ένδειξη ανθρώπινης expert αξιολόγησης για το συγκεκριμένο event;

Στο επίπεδο του requirement mapping, τα σχετικά Evidence Records μπορούν να συνδέονται με το requirement individual req_human_review. Το hasHumanReview λειτουργεί συμπληρωματικά προς αυτή τη χαρτογράφηση και δεν την αντικαθιστά.

Ιδιαίτερα σημαντικό είναι ότι η τιμή hasHumanReview = true δεν πρέπει να ερμηνεύεται ως απόδειξη ότι ικανοποιείται πλήρως η απαίτηση human oversight του άρθρου 14 του AI Act. Στην παρούσα εφαρμογή δηλώνει ότι το συγκεκριμένο VTaC event διαθέτει την ανθρώπινη expert αξιολόγηση που έχει αναπαρασταθεί στο proof-of-concept [4].

4.7.5 isLogged

Το data property isLogged χρησιμοποιείται ως boolean ένδειξη ότι ένα Alert Event έχει αναπαρασταθεί στο proof-of-concept ως καταγεγραμμένο στοιχείο για το οποίο διατίθενται συγκεκριμένες πληροφορίες τεκμηρίωσης.

Ο ρόλος του isLogged είναι περιγραφικός και συμπληρωματικός. Η βασική σημασιολογική σύνδεση με το AI Act-related σημείο αναφοράς για logging πραγματοποιείται μέσω του Evidence Record και του object property supportsRequirement:

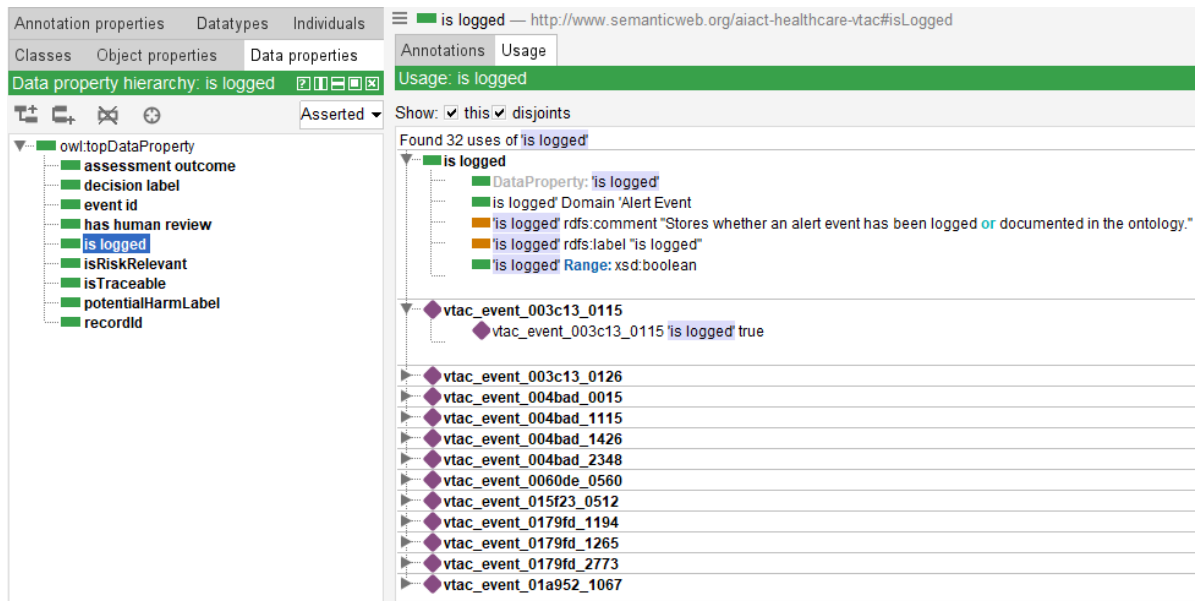
Alert Event --isDocumentedBy→ Evidence Record --supportsRequirement→ req_logging

Παράλληλα, το data property μπορεί να αποδώσει:

Alert Event --isLogged→ true/false

Το property αυτό απαντά στο ερώτημα:

Έχει το συγκεκριμένο event χαρακτηριστεί στο proof-of-concept ως καταγεγραμμένο στοιχείο;



Σχήμα 4.67 Data property isLogged στο Protégé.

Στην παρούσα εφαρμογή, η ένδειξη αυτή χρησιμοποιείται σε συνδυασμό με διαθέσιμες πληροφορίες όπως το eventId, τη σύνδεση με το αντίστοιχο Waveform Record και το Evidence Record που τεκμηριώνει το event.

Το isLogged δεν αντικαθιστά το requirement-to-evidence mapping και δεν αποτελεί αυτόνομη απόδειξη ότι ένα πραγματικό AI system διαθέτει τις automatic logging capabilities που προβλέπονται στο άρθρο 12 του AI Act [3]. Αποτελεί boolean χαρακτηριστικό της συγκεκριμένης proof-of-concept αναπαράστασης.

4.7.6 isRiskRelevant

Το data property isRiskRelevant χρησιμοποιείται για να δηλώσει, μέσω boolean τιμής, αν ένα Alert Event έχει risk-related σημασία. Στην παρούσα υλοποίηση, το property αυτό χρησιμοποιείται κυρίως για τα false alarm cases που συνδέονται με το lightweight risk layer.

Ο ρόλος του isRiskRelevant είναι να αποτυπώνει ως απλή ένδειξη ότι ένα event συνδέεται με risk-related ερμηνεία. Η κύρια σημασιολογική αναπαράσταση του κινδύνου γίνεται μέσω των object properties hasRiskIndicator, isClassifiedAs και mayLeadTo. Το data property λειτουργεί συμπληρωματικά, ως ένδειξη risk relevance.

Το property αυτό απαντά στο ερώτημα:

Ποια alert events θεωρούνται risk-relevant;

ή:

Ποια false alarms συμμετέχουν στο lightweight risk layer;

Η χρήση του είναι ιδιαίτερα χρήσιμη για τη διάκριση των false alarm cases από τα true alarm cases. Στο συγκεκριμένο proof-of-concept, τα έξι επιλεγμένα false alarm cases έχουν χαρακτηριστεί ως risk-relevant και συνδέονται με risk indicators, risk-relevant events και potential harm. Τα true alarms, αντίθετα, δεν χρησιμοποιούνται για την ενεργοποίηση του lightweight risk layer.

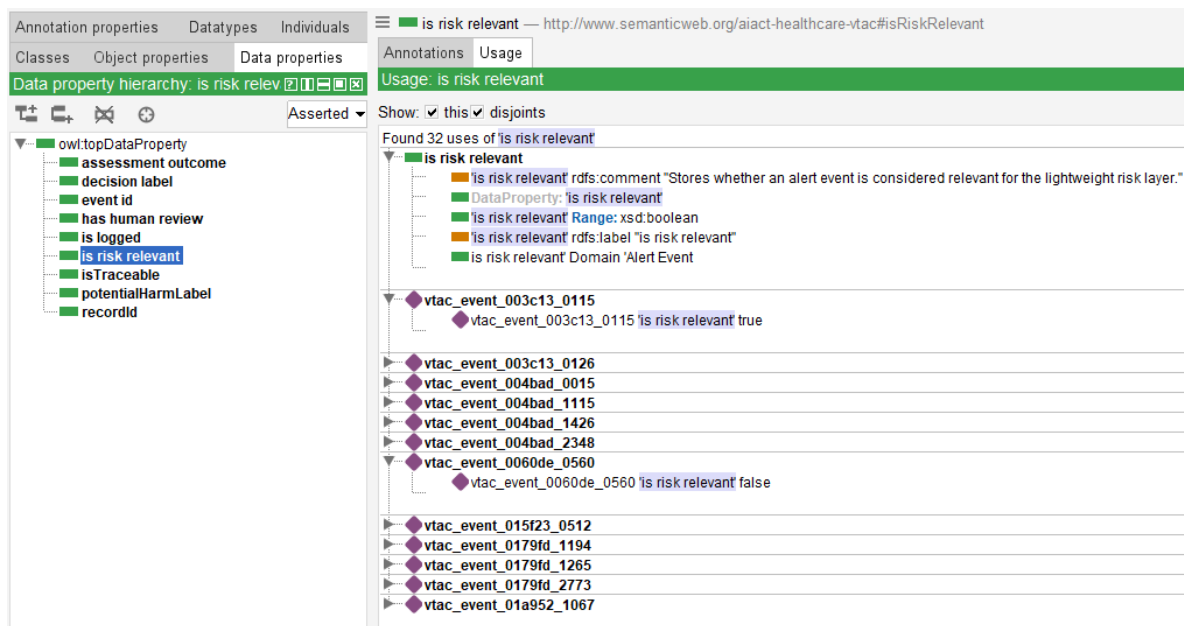
Η λογική μπορεί να αποδοθεί ως εξής:

Alert Event → **isRiskRelevant** → **true/false**

Alert Event → **hasRiskIndicator** → **Risk Indicator**

Alert Event → **isClassifiedAs** → **Risk-Relevant Event**

Risk-Relevant Event → **mayLeadTo** → **Potential Harm**



Σχήμα 4.68 Data property isRiskRelevant στο Protégé.

Το isRiskRelevant δεν αποδεικνύει ότι έχει συμβεί πραγματικός κίνδυνος ή πραγματική βλάβη. Δηλώνει μόνο ότι το συγκεκριμένο event έχει χαρακτηριστεί ως σχετικό με τη risk-related διάσταση της οντολογίας.

4.7.7 isTraceable

Το data property isTraceable χρησιμοποιείται ως boolean ένδειξη ότι ένα Alert Event συμμετέχει στη διαδρομή ιχνηλασιμότητας που έχει σχεδιαστεί στο proof-of-concept.

Ο ρόλος του είναι συμπληρωματικός. Το isTraceable δεν δημιουργεί από μόνο του την ιχνηλασιμότητα και δεν αντικαθιστά τις object-property σχέσεις από τις οποίες προκύπτει η σχετική σημασιολογική διαδρομή.

Το property αυτό μπορεί να αποδοθεί ως:

Alert Event --isTraceable→ **true/false**

Η ουσιαστική αναπαράσταση της ιχνηλασιμότητας προκύπτει από τη σύνθεση των σχέσεων:

Alert Event --belongsToRecord→ **Waveform Record**

Alert Event --hasReview→ Expert Review --hasDecision→ Review Decision

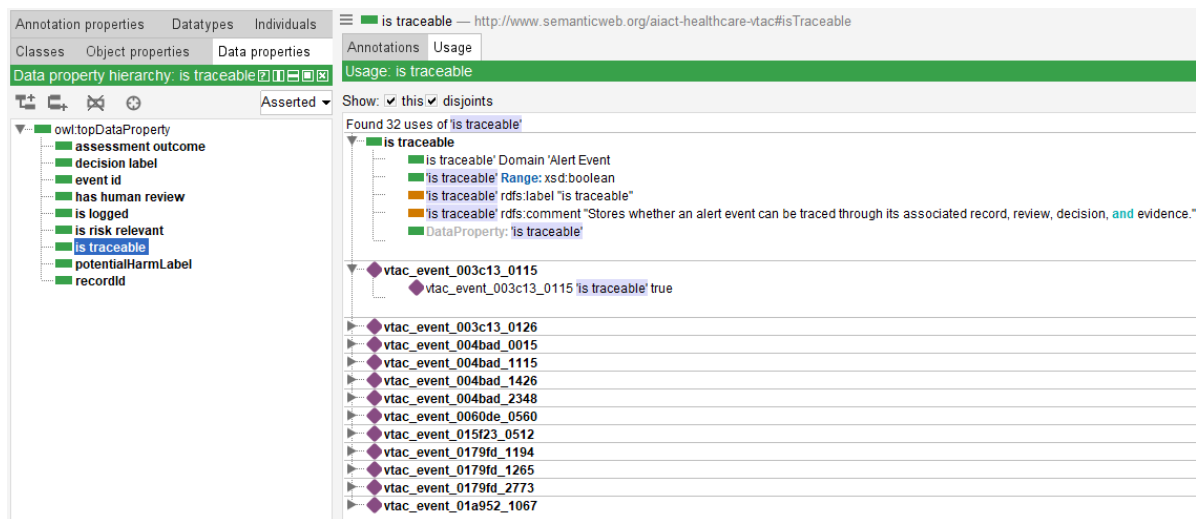
Alert Event --isDocumentedBy→ Evidence Record

και, στο επίπεδο του requirement mapping:

Evidence Record --supportsRequirement→ req_traceability

Με αυτόν τον τρόπο, για ένα συγκεκριμένο event μπορεί να ανακτηθεί το record προέλευσής του, η ανθρώπινη αξιολόγηση, η τελική απόφαση και το Evidence Record που το τεκμηριώνει.

Το isTraceable λειτουργεί επομένως ως βοηθητική boolean ένδειξη ότι το event συμμετέχει στην παραπάνω δομή. Δεν αποτελεί από μόνο του απόδειξη πλήρους ιχνηλασιμότητας ενός πραγματικού AI system ούτε απόδειξη συμμόρφωσης με το άρθρο 12 του AI Act [3].



Σχήμα 4.69 Data property isTraceable στο Protégé.

4.7.8 potentialHarmLabel

Το data property potentialHarmLabel χρησιμοποιείται για να αποδώσει λεκτική τιμή σε ένα Potential Harm individual. Στην παρούσα υλοποίηση, το βασικό potential harm που χρησιμοποιείται είναι το alert fatigue [13].

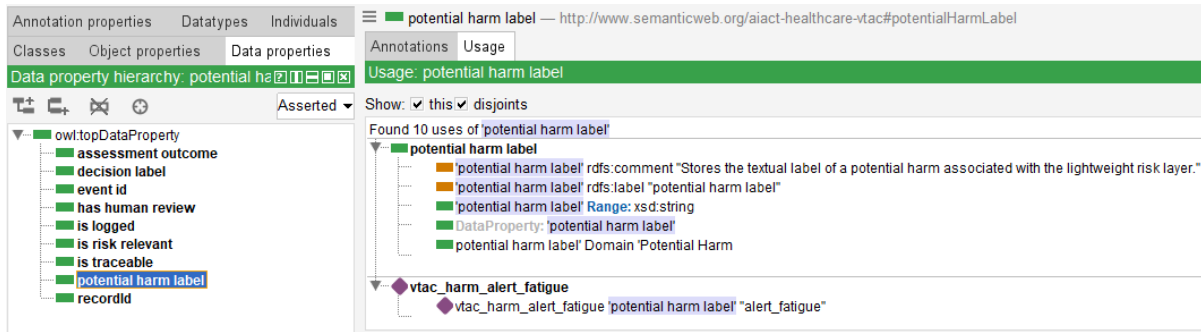
Ο ρόλος του potentialHarmLabel είναι να κάνει πιο αναγνώσιμη την πιθανή επίπτωση που συνδέεται με τα risk-relevant false alarm cases. Ενώ η class Potential Harm οργανώνει οντολογικά την έννοια της πιθανής επίπτωσης, το potentialHarmLabel προσθέτει μια ανθρώπινα αναγνώσιμη περιγραφή.

Το property αυτό απαντά στο ερώτημα:

Ποια είναι η περιγραφική ετικέτα μιας πιθανής επίπτωσης;

Για παράδειγμα, το individual vtac_harm_alert_fatigue μπορεί να διαθέτει potential harm label που δηλώνει το alert fatigue. Με αυτόν τον τρόπο, ο αναγνώστης μπορεί να κατανοήσει ευκολότερα τι αντιπροσωπεύει το συγκεκριμένο potential harm individual.

Η ύπαρξη του potentialHarmLabel είναι χρήσιμη επειδή το risk layer περιλαμβάνει έννοιες που χρειάζονται σαφή ερμηνεία. Η πιθανή επίπτωση δεν πρέπει να εμφανίζεται μόνο ως τεχνικό όνομα individual, αλλά και ως περιγραφικό label που εξηγεί το νόημά της.



Σχήμα 4.70 Data property potentialHarmLabel στο Protégé.

Πρέπει να σημειωθεί ότι το potentialHarmLabel δεν δηλώνει πραγματική βλάβη. Περιγράφει μια πιθανή επίπτωση που χρησιμοποιείται στο πλαίσιο της risk-related αναπαράστασης.

4.7.9 recordId

Το data property recordId χρησιμοποιείται για να αποδώσει σε ένα Waveform Record το αναγνωριστικό του record. Στο VTaC dataset, το record id δηλώνει το waveform record από το οποίο προέρχονται ένα ή περισσότερα alarm events [4].

Ο ρόλος του recordId είναι να ενισχύει τη σύνδεση ανάμεσα στην οντολογία και το αρχικό dataset. Όπως το eventId βοηθά στην αναγνώριση ενός event, έτσι και το recordId βοηθά στην αναγνώριση του record από το οποίο προέρχεται το event.

Το property αυτό απαντά στο ερώτημα:

Ποιο είναι το αναγνωριστικό ενός συγκεκριμένου waveform record;

Η ύπαρξη του recordId είναι ιδιαίτερα σημαντική σε περιπτώσεις όπου ένα waveform record συνδέεται με περισσότερα από ένα alert events. Για παράδειγμα, αν ένα record περιλαμβάνει πολλαπλά events, το record id επιτρέπει να διατηρηθεί η κοινή πηγή αυτών των events μέσα στην οντολογία.

Σε επίπεδο traceability, το recordId συμπληρώνει τη σχέση belongsToRecord. Το object property δείχνει ότι ένα event ανήκει σε ένα record, ενώ το data property δίνει την αναγνωρίσιμη τιμή του record. Έτσι, η οντολογία συνδυάζει σημασιολογική σύνδεση και πρακτική αναγνωρισιμότητα.

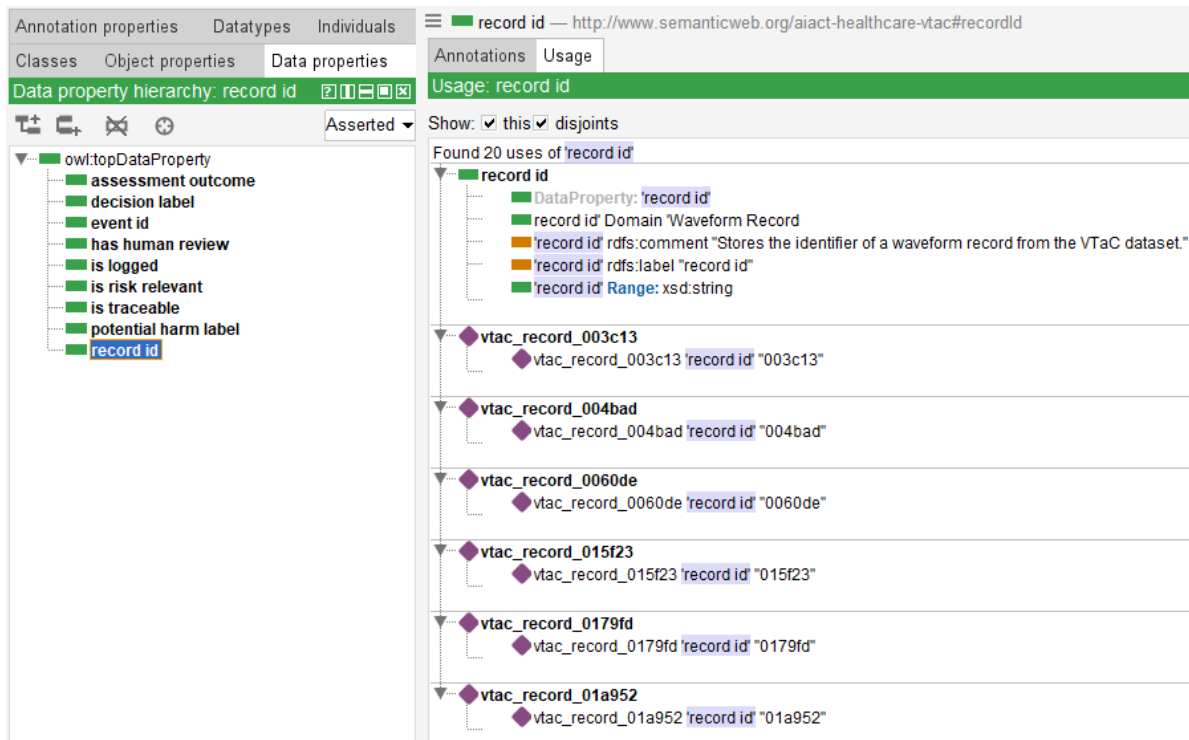
Η λογική μπορεί να αποδοθεί ως εξής:

Alert Event → belongsToRecord → Waveform Record

Waveform Record → recordId → τιμή record id

Με αυτόν τον τρόπο, το recordId υποστηρίζει τη δυνατότητα αναδρομής από ένα alert event προς το αρχικό waveform record του VTaC dataset.

Είναι ένα απλό αλλά απαραίτητο data property. Επιτρέπει στην οντολογία να διατηρεί καθαρή σύνδεση με τα πραγματικά VTaC waveform records και υποστηρίζει τη βασική απαίτηση ιχνηλασιμότητας της εργασίας.



Σχήμα 4.71 Data property recordId στο Protégé.

4.7.10 Συνολική αποτίμηση των data properties

Τα data properties συμπληρώνουν τη δομή της οντολογίας προσθέτοντας συγκεκριμένες τιμές στα individuals. Δεν δημιουργούν τις κύριες σημασιολογικές σχέσεις μεταξύ των στοιχείων του μοντέλου, οι οποίες αναπαρίστανται κυρίως μέσω των object properties, αλλά ενισχύουν την αναγνωσιμότητα, την τεκμηρίωση και την πρακτική αξιοποίηση της οντολογίας.

Τα `eventId` και `recordId` διατηρούν τη σύνδεση των οντολογικών individuals με τα αντίστοιχα πραγματικά στοιχεία του VTaC dataset [4]. Το `decisionLabel` αναπαριστά περιγραφικά την τελική expert απόφαση, ενώ το `assessmentOutcome` έχει προβλεφθεί για την απόδοση αποτελέσματος σε μελλοντικά Compliance Assessment individuals, χωρίς να χρησιμοποιείται σε instance-level assertions στην παρούσα proof-of-concept έκδοση.

Τα `hasHumanReview`, `isLogged`, `isRiskRelevant` και `isTraceable` λειτουργούν ως boolean ενδείξεις. Οι τιμές αυτές πρέπει να ερμηνεύονται αποκλειστικά μέσα στο πλαίσιο του συγκεκριμένου proof-of-concept. Δεν αντικαθιστούν τις object-property σχέσεις, τα Evidence Records ή το requirement-to-evidence mapping και δεν αποτελούν από μόνες τους αποδείξεις ότι ένα πραγματικό εξωτερικό σύστημα ικανοποιεί κάποια τεχνική ή νομική απαίτηση του AI Act.

Ειδικότερα, μια τιμή `hasHumanReview = true` δηλώνει ότι στο μοντέλο έχει καταγραφεί ανθρώπινη expert αξιολόγηση για το συγκεκριμένο event, χωρίς να αποδεικνύει πλήρη εφαρμογή του Article 14 human oversight. Αντίστοιχα, οι τιμές `isLogged = true` και `isTraceable = true` δηλώνουν χαρακτηριστικά της proof-of-concept αναπαράστασης και δεν αποτελούν αυτόνομη απόδειξη πλήρους εκπλήρωσης των απαιτήσεων του άρθρου 12 [3].

Το `isRiskRelevant` δηλώνει ότι ένα event έχει ενταχθεί στο lightweight risk-related μέρος της συγκεκριμένης οντολογίας. Δεν δηλώνει ότι έχει αποδειχθεί πραγματικός κίνδυνος ή βλάβη. Η

ουσιαστική risk-related αναπαράσταση πραγματοποιείται μέσω των object properties `hasRiskIndicator`, `isClassifiedAs` και `mayLeadTo`.

Τέλος, το `potentialHarmLabel` προσθέτει ανθρώπινα αναγνώσιμη περιγραφή στο `Potential Harm individual`, επιτρέποντας στην πιθανή επίπτωση που χρησιμοποιείται στο μοντέλο να παρουσιάζεται με σαφή τρόπο.

Συνολικά, τα εννέα data properties λειτουργούν ως περιγραφικό επίπεδο πάνω στο συνδεδεμένο οντολογικό μοντέλο. Τα object properties αναπαριστούν τις κύριες σχέσεις μεταξύ των individuals, ενώ τα data properties προσθέτουν identifiers, labels, outcomes και boolean ενδείξεις που διευκολύνουν την κατανόηση, τα queries και την τεκμηρίωση της υλοποίησης.

4.8 Individuals της οντολογίας

Μετά την παρουσίαση των classes, των object properties και των data properties, η παρούσα ενότητα επικεντρώνεται στα individuals της οντολογίας. Τα individuals αποτελούν τα συγκεκριμένα παραδείγματα ή στιγμιότυπα των classes και χρησιμοποιούνται για να δείξουν πώς το οντολογικό μοντέλο εφαρμόζεται στην πράξη.

Οι classes ορίζουν τις βασικές κατηγορίες εννοιών, όπως `Alert Event`, `Waveform Record`, `Expert Review`, `Review Decision`, `Evidence Record`, `Compliance Requirement` και `Risk-Related Event`. Τα individuals, αντίθετα, αναπαριστούν συγκεκριμένες περιπτώσεις αυτών των εννοιών. Για παράδειγμα, ένα πραγματικό VTaC alarm event αναπαρίσταται ως individual της class `Ventricular Tachycardia Alarm`, ενώ ένα συγκεκριμένο evidence record αναπαρίσταται ως individual της class `Evidence Record`.

Η χρήση individuals είναι απαραίτητη για την παρούσα εργασία, επειδή η οντολογία δεν αποτελεί μόνο ένα θεωρητικό σχήμα εννοιών. Αντίθετα, χρησιμοποιείται για να οργανώσει πραγματικά παραδείγματα από το VTaC dataset και να τα συνδέσει με αξιολογήσεις, αποφάσεις, τεκμήρια και επιλεγμένες απαιτήσεις συμμόρφωσης.

Τα individuals της οντολογίας οργανώνονται σε διαφορετικές ομάδες, ανάλογα με τον ρόλο τους. Υπάρχουν individuals που αναπαριστούν επιλεγμένες απαιτήσεις συμμόρφωσης, individuals που αναπαριστούν control mechanisms, individuals που αναπαριστούν waveform records, πραγματικά VTaC alert events, expert reviews, review decisions, evidence records και risk-related στοιχεία.

Η ανάλυση των individuals είναι σημαντική επειδή δείχνει ότι οι βασικές σημασιολογικές διαδρομές της οντολογίας δεν παραμένουν αφηρημένες, αλλά υλοποιούνται με συγκεκριμένα instances. Οι κύριες διαδρομές οργανώνονται γύρω από το ίδιο `Alert Event`:

Alert Event → belongsToRecord → Waveform Record

Alert Event → hasReview → Expert Review → hasDecision → Review Decision

Alert Event → isDocumentedBy → Evidence Record → supportsRequirement → Compliance Requirement

Για τα false alarm cases χρησιμοποιούνται επιπλέον risk-related individuals, τα οποία οργανώνονται μέσω δύο παράλληλων διαδρομών:

Alert Event → hasRiskIndicator → Risk Indicator

Alert Event → isClassifiedAs → Risk-Related Event → mayLeadTo → Potential Harm

Με αυτόν τον τρόπο, τα individuals μετατρέπουν τη θεωρητική δομή της οντολογίας σε λειτουργικό proof-of-concept μοντέλο, χωρίς να υπονοείται μία μοναδική γραμμική object-property αλυσίδα.

Επειδή το πλήθος των individuals δεν είναι πρακτικό να παρουσιαστεί αποκλειστικά μέσα από μία ενιαία εικόνα του Protégé, η παρούσα ενότητα χρησιμοποιεί αρχικά έναν συγκεντρωτικό πίνακα. Ο πίνακας ομαδοποιεί τα βασικά individuals ανάλογα με τον ρόλο τους στην οντολογία, ώστε να φαίνεται καθαρά πώς τα πραγματικά VTaC events, τα waveform records, τα reviews, οι decisions, τα evidence records, τα requirements και τα risk-related στοιχεία συμμετέχουν στη συνολική σημασιολογική δομή.

Πίνακας 4-2 Συγκεντρωτική παρουσίαση των βασικών individuals της οντολογίας.

Ομάδα individuals	Κλάση / ρόλος	Individuals
Compliance requirement individuals	Compliance Requirement	req_logging, req_human_review, req_traceability, req_risk_management
Control mechanism individuals	Control Mechanism	logging_control, human_review_control, traceability_control, risk_management_control
Monitoring system individual	Monitoring System	vtac_icu_monitoring_system
Waveform record individuals	Waveform Record	vtac_record_003c13, vtac_record_004bad, vtac_record_0060de, vtac_record_015f23, vtac_record_0179fd, vtac_record_01a952
VTaC alert event individuals	Ventricular Tachycardia Alarm / Alert Event	vtac_event_003c13_0115, vtac_event_003c13_0126, vtac_event_004bad_0015, vtac_event_004bad_1115, vtac_event_004bad_1426, vtac_event_004bad_2348, vtac_event_0060de_0560, vtac_event_015f23_0512, vtac_event_0179fd_1194, vtac_event_0179fd_1265, vtac_event_0179fd_2773, vtac_event_01a952_1067
Expert review individuals	Expert Review	vtac_review_003c13_0115, vtac_review_003c13_0126, vtac_review_004bad_0015, vtac_review_004bad_1115, vtac_review_004bad_1426, vtac_review_004bad_2348, vtac_review_0060de_0560, vtac_review_015f23_0512, vtac_review_0179fd_1194, vtac_review_0179fd_1265, vtac_review_0179fd_2773, vtac_review_01a952_1067
Review decision individuals — false alarms	False Alarm	vtac_decision_003c13_0115_false, vtac_decision_003c13_0126_false, vtac_decision_004bad_0015_false, vtac_decision_004bad_1115_false, vtac_decision_004bad_1426_false, vtac_decision_004bad_2348_false
Review decision individuals — true alarms	True Alarm	vtac_decision_0060de_0560_true, vtac_decision_015f23_0512_true, vtac_decision_0179fd_1194_true, vtac_decision_0179fd_1265_true, vtac_decision_0179fd_2773_true, vtac_decision_01a952_1067_true

Evidence record individuals	Evidence Record	vtac_evidence_003c13_0115, vtac_evidence_003c13_0126, vtac_evidence_004bad_0015, vtac_evidence_004bad_1115, vtac_evidence_004bad_1426, vtac_evidence_004bad_2348, vtac_evidence_0060de_0560, vtac_evidence_015f23_0512, vtac_evidence_0179fd_1194, vtac_evidence_0179fd_1265, vtac_evidence_0179fd_2773, vtac_evidence_01a952_1067
Risk indicator individuals	Risk Indicator	vtac_indicator_003c13_0115_false_alarm, vtac_indicator_003c13_0126_false_alarm, vtac_indicator_004bad_0015_false_alarm, vtac_indicator_004bad_1115_false_alarm, vtac_indicator_004bad_1426_false_alarm, vtac_indicator_004bad_2348_false_alarm
Risk-relevant event individuals	Risk-Relevant Event	vtac_risk_event_003c13_0115, vtac_risk_event_003c13_0126, vtac_risk_event_004bad_0015, vtac_risk_event_004bad_1115, vtac_risk_event_004bad_1426, vtac_risk_event_004bad_2348
Potential harm individual	Potential Harm	vtac_harm_alert_fatigue

Ο Πίνακας 4.2 δείχνει ότι τα individuals της οντολογίας δεν έχουν δημιουργηθεί τυχαία, αλλά ακολουθούν τη βασική λογική του μοντέλου. Τα waveform records και τα VTaC alert events αναπαριστούν το λειτουργικό επίπεδο των πραγματικών δεδομένων. Τα expert reviews και οι review decisions αναπαριστούν την ανθρώπινη αξιολόγηση και το αποτέλεσμα της. Τα evidence records λειτουργούν ως ενδιάμεσο επίπεδο τεκμηρίωσης, ενώ τα compliance requirements και τα control mechanisms αναπαριστούν το κανονιστικό και ελεγκτικό επίπεδο της οντολογίας.

Επιπλέον, τα risk indicator, risk-relevant event και potential harm individuals χρησιμοποιούνται για την αναπαράσταση του lightweight risk layer. Η ομάδα αυτή αφορά κυρίως τα false alarm cases, τα οποία δεν αντιμετωπίζονται μόνο ως λανθασμένες ειδοποιήσεις, αλλά και ως περιστατικά που μπορούν να αποκτήσουν σημασία από πλευράς risk management.

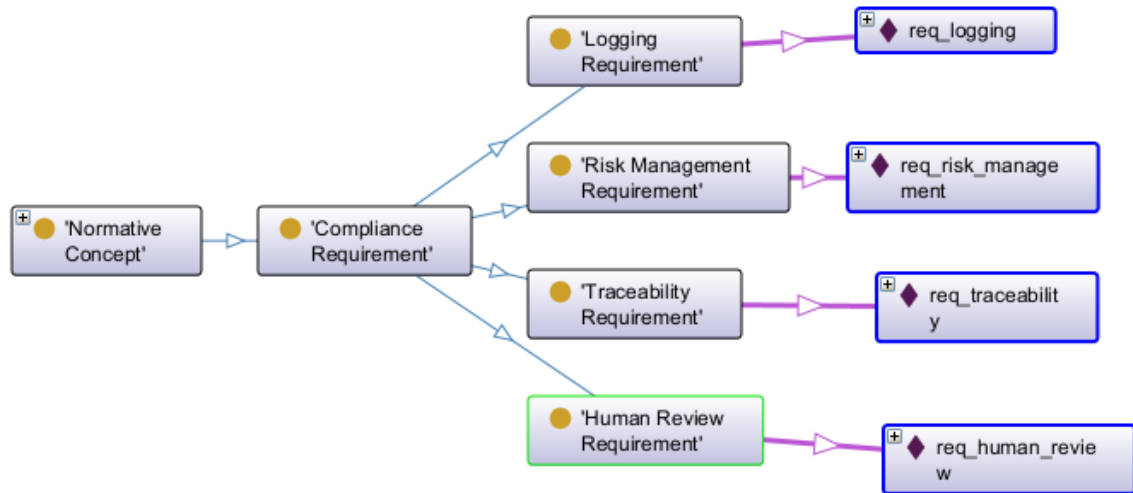
Με αυτόν τον τρόπο, ο πίνακας τεκμηριώνει ότι η οντολογία έχει εφαρμοστεί σε συγκεκριμένα instances και όχι μόνο σε επίπεδο αφηρημένων classes. Τα individuals υλοποιούν στην πράξη τη διαδρομή από τα πραγματικά VTaC records και alert events προς τις αξιολογήσεις, τις αποφάσεις, τα evidence records και τις επιλεγμένες AI Act-related requirements.

4.8.1 Requirement individuals

Τα requirement individuals αναπαριστούν τα τέσσερα επιλεγμένα AI Act-related σημεία αναφοράς που χρησιμοποιούνται στο proof-of-concept μοντέλο. Ανήκουν στις αντίστοιχες ειδικότερες classes της Compliance Requirement και είναι:

- req_logging

- req_human_review
- req_traceability
- req_risk_management



Σχήμα 4.72 Requirement individuals που αναπαριστούν τα επιλεγμένα AI Act-related σημεία αναφοράς της οντολογίας.

Το req_logging χρησιμοποιείται ως εφαρμοσμένο σημείο αναφοράς για τη διάσταση record-keeping/logging, το req_human_review για την ανθρώπινη expert αξιολόγηση, το req_traceability για τη δυνατότητα ιχνηλάτησης των σχετικών πληροφοριών και το req_risk_management για το lightweight risk-related μέρος της υλοποίησης.

Τα requirement individuals συνδέονται με τα Evidence Record individuals μέσω του object property supportsRequirement. Με τον τρόπο αυτό υλοποιείται το requirement-to-evidence mapping:

Evidence Record --supportsRequirement→ Requirement individual

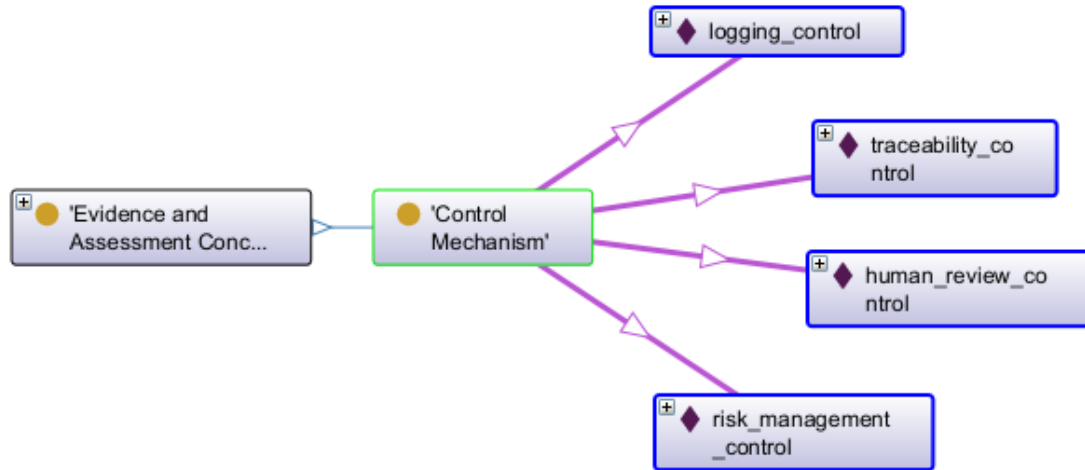
Τα συγκεκριμένα individuals δεν δηλώνουν ότι έχει αποδειχθεί πλήρης νομική συμμόρφωση με τον AI Act. Αποτελούν εφαρμοσμένες έννοιες της παρούσας οντολογίας που χρησιμοποιούνται για να οργανωθεί και να ελεγχθεί η σύνδεση διαθέσιμων evidence records με επιλεγμένες διαστάσεις του Κανονισμού [3].

4.8.2 Control mechanism individuals

Η οντολογία περιλαμβάνει τέσσερα individuals της class Control Mechanism:

- logging_control
- human_review_control
- traceability_control
- risk_management_control

Τα individuals αυτά χρησιμοποιούνται για να αναπαραστήσουν εννοιολογικά τέσσερις βασικές κατηγορίες μηχανισμών ελέγχου που σχετίζονται με το πεδίο της εργασίας: logging, ανθρώπινη αξιολόγηση, traceability και risk management.



Σχήμα 4.73 Individuals της class Control Mechanism στο Protégé.

Το `logging_control` αναπαριστά την έννοια ενός μηχανισμού σχετικού με την καταγραφή και τεκμηρίωση. Το `human_review_control` αναπαριστά την έννοια ενός μηχανισμού σχετικού με ανθρώπινη ή expert αξιολόγηση. Το `traceability_control` αφορά την έννοια μηχανισμού που υποστηρίζει την ιχνηλασιμότητα, ενώ το `risk_management_control` αφορά την έννοια μηχανισμού σχετικού με τη διαχείριση κινδύνου.

Στην τρέχουσα proof-of-concept υλοποίηση, τα συγκεκριμένα individuals έχουν κυρίως εννοιολογικό και οργανωτικό ρόλο. Δεν αποτελούν ενεργούς τεχνικούς μηχανισμούς και δεν εκτελούν αυτόματα logging, human oversight, traceability ή risk management. Επίσης, οι βασικές διαδρομές event-review-decision και event-evidence-requirement υλοποιούνται μέσω των object properties που παρουσιάστηκαν στην Ενότητα 4.6 και όχι μέσω των Control Mechanism individuals.

Η παρουσία τους στο μοντέλο παρέχει ένα σημείο μελλοντικής επέκτασης, στο οποίο θα μπορούσαν να προστεθούν επιπλέον σχέσεις με πραγματικά τεχνικά ή οργανωτικά controls. Στην παρούσα εργασία, όμως, δεν πρέπει να ερμηνεύονται ως απόδειξη ότι τέτοιοι μηχανισμοί έχουν πραγματικά υλοποιηθεί ή λειτουργούν σε εξωτερικό AI system.

4.8.3 Monitoring system individual

Για την class Monitoring System δημιουργήθηκε το individual:

```
vtac_icu_monitoring_system
```

Το individual αυτό χρησιμοποιείται για να αναπαραστήσει το γενικό healthcare monitoring περιβάλλον στο οποίο εντάσσονται τα ventricular tachycardia alarm events του VTaC. Το dataset βασίζεται σε alarms που προέρχονται από ICU bedside monitoring περιβάλλον [4].

Στην παρούσα proof-of-concept υλοποίηση, το `vtac_icu_monitoring_system` λειτουργεί κυρίως ως contextual individual, δηλαδή ως σημασιολογική αναπαράσταση του γενικού τεχνολογικού περιβάλλοντος προέλευσης των events.

Δεν χρησιμοποιείται στην τρέχουσα έκδοση object property που να συνδέει ρητά το συγκεκριμένο Monitoring System individual με κάθε Alert Event ή Waveform Record. Επομένως, ο ρόλος του είναι να αποτυπώνει το monitoring context και όχι να δημιουργεί ενεργή διαδρομή ιχνηλασιμότητας μέσα στην υλοποίηση.

Η επιλογή αυτή διατηρεί τη δυνατότητα μελλοντικής επέκτασης του μοντέλου, όπου θα μπορούσαν να εισαχθούν επιπλέον properties για τη ρητή σύνδεση monitoring systems, records και events.

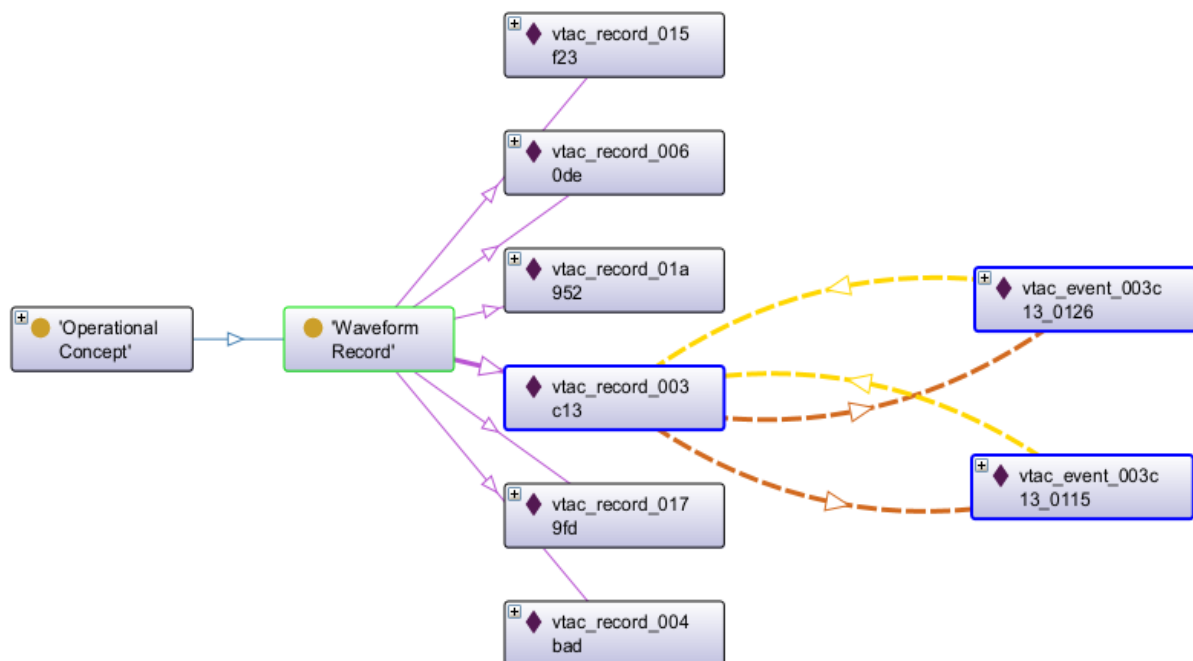
4.8.4 Waveform record individuals

Τα waveform record individuals αποτελούν βασικό μέρος του operational layer της οντολογίας, καθώς αναπαριστούν τα records φυσιολογικών δεδομένων από τα οποία προέρχονται τα alert events του VTaC dataset. Ο ρόλος τους είναι ιδιαίτερα σημαντικός, επειδή επιτρέπουν τη σύνδεση κάθε alert event με το αρχικό record στο οποίο ανήκει, ενισχύοντας έτσι την ιχνηλασιμότητα της οντολογίας.

Στην παρούσα υλοποίηση έχουν εισαχθεί έξι waveform record individuals:

- vtac_record_003c13
- vtac_record_004bad
- vtac_record_0060de
- vtac_record_015f23
- vtac_record_0179fd
- vtac_record_01a952

Κάθε ένα από αυτά λειτουργεί ως σημείο αναφοράς για ένα ή περισσότερα alert events. Με αυτόν τον τρόπο, η οντολογία δεν αντιμετωπίζει τα alarm events ως ανεξάρτητες και αποκομμένες εγγραφές, αλλά τα τοποθετεί στο record προέλευσής τους.



Σχήμα 4.74 Ενδεικτική απεικόνιση της σχέσης μεταξύ waveform record και alert events.

Στο Σχήμα 4.74 παρουσιάζεται ενδεικτικά το παράδειγμα του vtac_record_003c13. Το συγκεκριμένο waveform record συνδέεται με δύο alert events:

- vtac_event_003c13_0115
- vtac_event_003c13_0126

Η σχέση αυτή αποτυπώνει ότι περισσότερα από ένα alert events μπορούν να προέρχονται από το ίδιο waveform record. Αυτό είναι συμβατό με τη λογική του VTaC dataset και με τον τρόπο που οργανώνονται τα πραγματικά alarm περιστατικά στην οντολογία.

Η αναπαράσταση αυτή έχει ουσιαστική σημασία, επειδή δείχνει ότι ένα alert event δεν εμφανίζεται μόνο του μέσα στο μοντέλο. Αντίθετα, εντάσσεται σε ένα συγκεκριμένο πλαίσιο προέλευσης. Στο συγκεκριμένο παράδειγμα, το vtac_record_003c13 λειτουργεί ως κοινή πηγή δεδομένων για τα δύο αντίστοιχα alert events. Στη συνέχεια, κάθε ένα από αυτά τα events μπορεί να συνδεθεί με expert review, review decision, evidence record και compliance requirement.

Η σύνδεση ανάμεσα στα waveform records και στα alert events πραγματοποιείται μέσω των object properties hasAlertEvent και belongsToRecord. Το hasAlertEvent χρησιμοποιείται από την πλευρά του waveform record και δηλώνει ποια alert events συνδέονται με αυτό. Αντίθετα, το belongsToRecord χρησιμοποιείται από την πλευρά του alert event και δηλώνει σε ποιο waveform record ανήκει.

Η σχέση αυτή μπορεί να αποδοθεί ως εξής:

Waveform Record → hasAlertEvent → Alert Event

και αντίστροφα:

Alert Event → belongsToRecord → Waveform Record

Η ύπαρξη αυτών των δύο σχέσεων ενισχύει την ιχνηλασιμότητα. Ο αναγνώστης μπορεί να ξεκινήσει είτε από το waveform record και να δει ποια events συνδέονται με αυτό, είτε από ένα alert event και να εντοπίσει το record προέλευσής του. Αυτό είναι σημαντικό για τη λογική της εργασίας, επειδή η ιχνηλασιμότητα δεν αφορά μόνο το τελικό evidence ή την απαίτηση, αλλά ξεκινά ήδη από το αρχικό record των δεδομένων.

Επιπλέον, τα waveform records συνδέονται με το data property recordId, το οποίο αποθηκεύει το αναγνωριστικό του record σε μορφή κειμένου. Για παράδειγμα, το individual vtac_record_003c13 αντιστοιχεί στο recordId "003c13". Με αυτόν τον τρόπο, η οντολογία διατηρεί σαφή σύνδεση με την αρχική κωδικοποίηση του dataset.

Τα waveform record individuals αποτελούν κρίσιμο σημείο της οντολογικής μοντελοποίησης, επειδή παρέχουν το αρχικό επίπεδο τεκμηρίωσης της operational πληροφορίας. Το παράδειγμα του vtac_record_003c13 δείχνει με κατανοητό τρόπο πώς από το record εντοπίζεται το event και στη συνέχεια γύρω από το event αναπτύσσονται οι διαδρομές αξιολόγησης και τεκμηρίωσης.

4.8.5 VTaC alert event individuals

Τα VTaC alert event individuals αποτελούν τον κεντρικό πυρήνα της εφαρμοσμένης υλοποίησης της οντολογίας. Σε αντίθεση με τα requirement individuals ή τα control mechanism individuals, τα οποία αναπαριστούν κανονιστικές και οργανωτικές έννοιες, τα alert event individuals αντιστοιχούν σε πραγματικά περιστατικά alarm από το VTaC dataset.

Στην παρούσα οντολογία χρησιμοποιούνται δώδεκα VTaC alert event individuals. Όλα τα συγκεκριμένα events έχουν μοντελοποιηθεί ως instances της class Ventricular Tachycardia Alarm, η

οποία αποτελεί ειδικότερη μορφή της class Alert Event. Η επιλογή αυτή έγινε επειδή τα events που χρησιμοποιούνται στην υλοποίηση αντιστοιχούν σε ventricular tachycardia alarms.

Τα 12 alert event individuals είναι τα εξής:

- vtac_event_003c13_0115
- vtac_event_003c13_0126
- vtac_event_004bad_0015
- vtac_event_004bad_1115
- vtac_event_004bad_1426
- vtac_event_004bad_2348
- vtac_event_0060de_0560
- vtac_event_015f23_0512
- vtac_event_0179fd_1194
- vtac_event_0179fd_1265
- vtac_event_0179fd_2773
- vtac_event_01a952_1067

Η ύπαρξη αυτών των individuals είναι κρίσιμη, επειδή η οντολογία δεν λειτουργεί μόνο ως αφηρημένο μοντέλο εννοιών. Αντίθετα, χρησιμοποιεί συγκεκριμένα πραγματικά events για να δείξει πώς μπορεί να οργανωθεί η πληροφορία γύρω από ένα healthcare alarm event.

Κάθε alert event individual συνδέεται με το αντίστοιχο waveform record προέλευσης μέσω του object property belongsToRecord. Με αυτόν τον τρόπο, κάθε event μπορεί να ιχνηλατηθεί πίσω στο αρχικό record από το οποίο προέρχεται. Για παράδειγμα, τα vtac_event_003c13_0115 και vtac_event_003c13_0126 συνδέονται με το waveform record vtac_record_003c13.

Εκτός από τη σύνδεση με το record προέλευσης, κάθε alert event συνδέεται και με ένα expert review μέσω του object property hasReview. Αυτό σημαίνει ότι το event δεν παραμένει μια απλή τεχνική εγγραφή, αλλά εντάσσεται σε διαδικασία ανθρώπινης ή expert αξιολόγησης. Η ανθρώπινη αξιολόγηση αποτελεί βασικό σημείο της οντολογίας, επειδή συνδέεται με τη λογική της human oversight τεκμηρίωσης.

Στη συνέχεια, το expert review οδηγεί σε ένα review decision. Η απόφαση αυτή μπορεί να είναι true alarm ή false alarm. Είναι σημαντικό να σημειωθεί ότι η οντολογία δεν προσπαθεί να προβλέψει αν ένα alarm είναι true ή false. Δεν χρησιμοποιείται μοντέλο μηχανικής μάθησης. Η τιμή true ή false αξιοποιείται ως δεδομένη πληροφορία από την ανθρώπινη αξιολόγηση του VTaC dataset και ενσωματώνεται στην οντολογία ως στοιχείο τεκμηρίωσης.

Τα alert event individuals αποτελούν το κοινό σημείο γύρω από το οποίο οργανώνονται οι βασικές συνδεδεμένες σημασιολογικές διαδρομές της υλοποίησης:

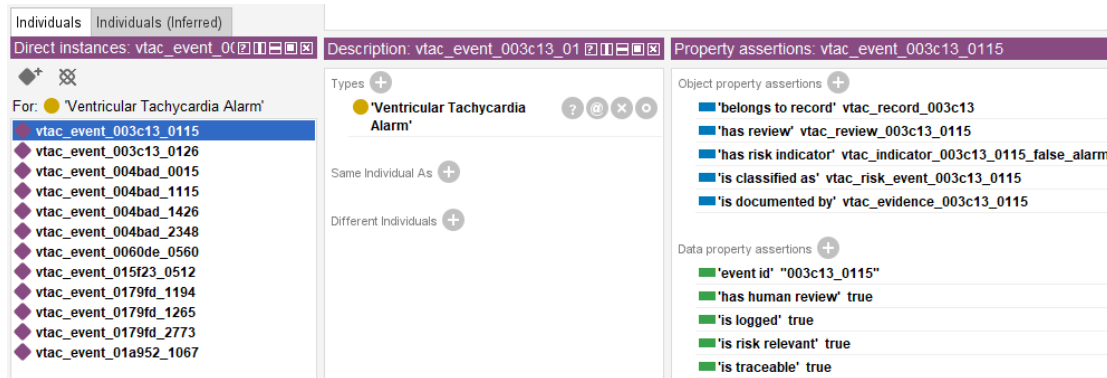
Alert Event → belongsToRecord → Waveform Record

Alert Event → hasReview → Expert Review → hasDecision → Review Decision

Alert Event → isDocumentedBy → Evidence Record → supportsRequirement → Compliance Requirement

Οι διαδρομές αυτές παραμένουν διακριτές αλλά συνδέονται μέσω του ίδιου Alert Event. Με αυτόν τον τρόπο, για κάθε event μπορούν να ανακτηθούν η προέλευση, η ανθρώπινη αξιολόγηση, η τελική απόφαση και το αντίστοιχο Evidence Record, καθώς και τα requirement mappings του Evidence Record, χωρίς να υπονοείται μία μοναδική γραμμική object-property αλυσίδα.

Τα alert event individuals φέρουν επίσης data properties, τα οποία αποθηκεύουν συγκεκριμένες τιμές. Για παράδειγμα, το eventId αποθηκεύει το αναγνωριστικό του event, το hasHumanReview δηλώνει αν υπάρχει ανθρώπινη αξιολόγηση, το isLogged δηλώνει αν το event έχει καταγραφεί, το isTraceable δηλώνει αν το event είναι ιχνηλατήσιμο και το isRiskRelevant δηλώνει αν το event σχετίζεται με το lightweight risk layer.



Σχήμα 4.75 VTaC alert event individuals και ιδιότητες επιλεγμένου event στο Protégé.

Το σχήμα 4.75 παρουσιάζει τα 12 πραγματικά VTaC alert event individuals που έχουν εισαχθεί στην οντολογία ως direct instances της class Ventricular Tachycardia Alarm. Παράλληλα, παρουσιάζονται ενδεικτικά οι object property και data property assertions ενός επιλεγμένου event, ώστε να φανεί ότι τα VTaC alert events δεν εισάγονται ως απλές εγγραφές, αλλά ως πλήρως συνδεδεμένα και τεκμηριωμένα individuals.

Ιδιαίτερη σημασία έχει η διάκριση ανάμεσα σε false alarm και true alarm cases. Στην παρούσα υλοποίηση, έξι από τα δώδεκα events έχουν χαρακτηριστεί ως false alarms και έξι ως true alarms.

Τα false alarm event individuals είναι:

- vtac_event_003c13_0115
- vtac_event_003c13_0126
- vtac_event_004bad_0015
- vtac_event_004bad_1115
- vtac_event_004bad_1426
- vtac_event_004bad_2348

Τα true alarm event individuals είναι:

- vtac_event_0060de_0560
- vtac_event_015f23_0512
- vtac_event_0179fd_1194
- vtac_event_0179fd_1265

- vtac_event_0179fd_2773
- vtac_event_01a952_1067

Η διάκριση αυτή δεν χρησιμοποιείται για στατιστική αξιολόγηση ή για εκπαίδευση μοντέλου. Χρησιμοποιείται για να ελεγχθεί αν η οντολογία μπορεί να αναπαραστήσει και τις δύο βασικές περιπτώσεις απόφασης. Τα true alarms ακολουθούν τη βασική διαδρομή τεκμηρίωσης, ενώ τα false alarms αποκτούν επιπλέον risk-related αναπαράσταση.

Συγκεκριμένα, τα false alarm events συνδέονται με risk indicators, risk-relevant events και potential harm. Με αυτόν τον τρόπο, η οντολογία δείχνει ότι ένα false alarm δεν είναι μόνο μια λανθασμένη ειδοποίηση, αλλά μπορεί να αποκτήσει σημασία από πλευράς risk management, ιδίως σε σχέση με πιθανές επιπτώσεις όπως το alert fatigue.

Η πρόσθετη risk-related αναπαράσταση για τα false alarm cases οργανώνεται μέσω δύο συνδεδεμένων διαδρομών που ξεκινούν από το ίδιο Alert Event:

Alert Event → hasRiskIndicator → Risk Indicator

Alert Event → isClassifiedAs → Risk-Relevant Event → mayLeadTo → Potential Harm

Οι διαδρομές αυτές λειτουργούν παράλληλα και δεν αποτελούν μία μοναδική γραμμική object-property αλυσίδα.

Αυτό σημαίνει ότι τα alert event individuals αποτελούν το σημείο σύνδεσης ανάμεσα στα πραγματικά δεδομένα του VTaC dataset και στη συνολική λογική της οντολογίας. Από αυτά ξεκινά η διαδρομή προς την ανθρώπινη αξιολόγηση, την απόφαση, την τεκμηρίωση, τις selected AI Act-related requirements και, όπου χρειάζεται, το lightweight risk layer.

Τα VTaC alert event individuals αποτελούν το πιο σημαντικό εφαρμοσμένο επίπεδο της οντολογίας. Μέσω αυτών, η εργασία δείχνει ότι η οντολογική δομή δεν παραμένει θεωρητική, αλλά εφαρμόζεται σε πραγματικά healthcare monitoring events. Κάθε event αποκτά record προέλευσης, expert review, review decision, evidence record, data property values και requirement mappings. Έτσι, τα VTaC alert event individuals λειτουργούν ως ο βασικός κόμβος γύρω από τον οποίο οργανώνεται ολόκληρη η proof-of-concept υλοποίηση.

4.8.6 Expert review individuals

Μετά την παρουσίαση των VTaC alert event individuals, η οντολογία προχωρά στο επίπεδο της ανθρώπινης ή expert αξιολόγησης. Το επίπεδο αυτό είναι ιδιαίτερα σημαντικό, επειδή συνδέει τα πραγματικά alarm events με τη διαδικασία αξιολόγησης που οδηγεί στην τελική απόφαση.

Στην παρούσα οντολογία, κάθε alert event συνδέεται με ένα αντίστοιχο expert review individual. Τα expert review individuals ανήκουν στην class Expert Review και αναπαριστούν την ανθρώπινη αξιολόγηση που έχει συσχετιστεί με κάθε VTaC alarm event.

Τα expert review individuals που χρησιμοποιούνται είναι τα εξής:

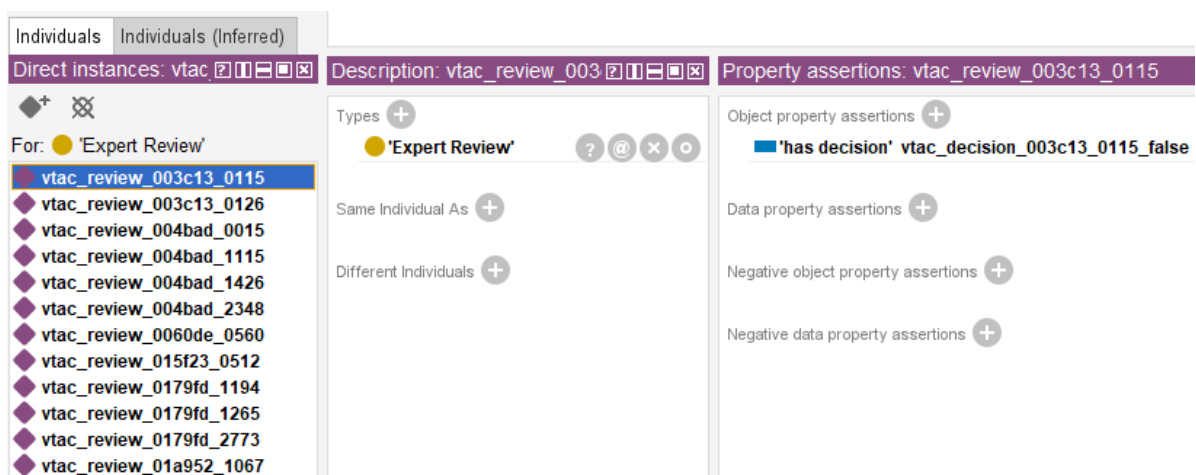
- vtac_review_003c13_0115
- vtac_review_003c13_0126
- vtac_review_004bad_0015

- vtac_review_004bad_1115
- vtac_review_004bad_1426
- vtac_review_004bad_2348
- vtac_review_0060de_0560
- vtac_review_015f23_0512
- vtac_review_0179fd_1194
- vtac_review_0179fd_1265
- vtac_review_0179fd_2773
- vtac_review_01a952_1067

Η σχέση ανάμεσα σε ένα alert event και στο αντίστοιχο expert review αποτυπώνεται με το object property hasReview. Για παράδειγμα, το event vtac_event_003c13_0115 συνδέεται με το review vtac_review_003c13_0115. Με αυτόν τον τρόπο, η οντολογία δείχνει ότι το συγκεκριμένο alarm event έχει αξιολογηθεί και δεν αποτελεί απλή τεχνική εγγραφή χωρίς ανθρώπινο ή expert έλεγχο.

Η σχέση αυτή μπορεί να αποδοθεί ως εξής:

Alert Event → hasReview → Expert Review



Σχήμα 4.76 Individuals της class Expert Review στο Protégé.

Το σχήμα 4.76 παρουσιάζει τα individuals που έχουν εισαχθεί ως instances της class Expert Review στην οντολογία. Τα individuals αυτά εκφράζουν το στάδιο της ανθρώπινης ή expert αξιολόγησης των VTaC alert events και συνδέονται με τα αντίστοιχα review decision individuals μέσω του object property hasDecision. Ενδεικτικά, το individual vtac_review_003c13_0115 οδηγεί στο vtac_decision_003c13_0115_false.

Το expert review λειτουργεί ως ενδιάμεσο στάδιο ανάμεσα στο alert event και στην τελική απόφαση. Δηλαδή, πρώτα υπάρχει το alarm event, στη συνέχεια υπάρχει η ανθρώπινη ή expert αξιολόγηση και έπειτα προκύπτει η review decision.

Η διαδρομή μπορεί να αποδοθεί ως εξής:

Alert Event → Expert Review → Review Decision

ή, με βάση τα object properties της οντολογίας:

Alert Event → hasReview → Expert Review

Expert Review → hasDecision → Review Decision

Η διάκριση αυτή είναι σημαντική για τη λογική της οντολογίας. Αν η απόφαση true ή false συνδεόταν απευθείας με το alert event, θα χανόταν το ενδιάμεσο στάδιο της αξιολόγησης. Με την ύπαρξη των expert review individuals, η οντολογία δείχνει ότι η τελική απόφαση δεν εμφανίζεται αυθαίρετα, αλλά προκύπτει από συγκεκριμένη διαδικασία αξιολόγησης.

Τα expert review individuals συνδέονται επίσης με τη λογική της απαίτησης req_human_review. Η ύπαρξη review για κάθε VTaC event υποστηρίζει την οντολογική τεκμηρίωση της ανθρώπινης αξιολόγησης μέσα στο μοντέλο. Αυτό δεν σημαίνει ότι αποδεικνύεται πλήρης νομική συμμόρφωση, αλλά ότι η οντολογία οργανώνει ρητά την πληροφορία που σχετίζεται με human review.

Επιπλέον, τα expert review individuals συμβάλλουν στην ιχνηλασιμότητα της υλοποίησης. Μέσω αυτών, ο αναγνώστης μπορεί να ακολουθήσει τη διαδρομή από το αρχικό alert event μέχρι την απόφαση. Για παράδειγμα, το vtac_event_003c13_0115 συνδέεται με το vtac_review_003c13_0115, το οποίο στη συνέχεια οδηγεί στο αντίστοιχο decision individual. Έτσι, η οντολογία δεν παρουσιάζει μόνο το αποτέλεσμα, αλλά και το στάδιο που προηγείται του αποτελέσματος.

Συνολικά, τα expert review individuals αποτελούν κρίσιμο ενδιάμεσο επίπεδο στην εφαρμοσμένη υλοποίηση της οντολογίας. Συνδέουν τα πραγματικά VTaC alert events με τις review decisions, υποστηρίζουν τη λογική της human review τεκμηρίωσης και ενισχύουν τη δυνατότητα traceability από το event μέχρι το evidence record και τις selected AI Act-related requirements.

4.8.7 Review decision individuals

Τα review decision individuals αναπαριστούν το αποτέλεσμα της ανθρώπινης ή expert αξιολόγησης κάθε VTaC alert event. Αν τα expert review individuals εκφράζουν τη διαδικασία της αξιολόγησης, τότε τα review decision individuals εκφράζουν το συμπέρασμα αυτής της διαδικασίας.

Στην παρούσα οντολογία, τα review decision individuals ανήκουν στην class Review Decision, η οποία εξειδικεύεται σε δύο υποκλάσεις:

- False Alarm
- True Alarm

Με αυτόν τον τρόπο, η οντολογία μπορεί να αναπαραστήσει αν ένα alert event αξιολογήθηκε ως πραγματικό alarm ή ως false alarm. Η διάκριση αυτή είναι βασική για τη λειτουργία του μοντέλου, επειδή επηρεάζει τη συνέχεια της σημασιολογικής διαδρομής. Τα true alarm cases ακολουθούν τη βασική διαδρομή τεκμηρίωσης, ενώ τα false alarm cases συνδέονται επιπλέον με το lightweight risk layer.

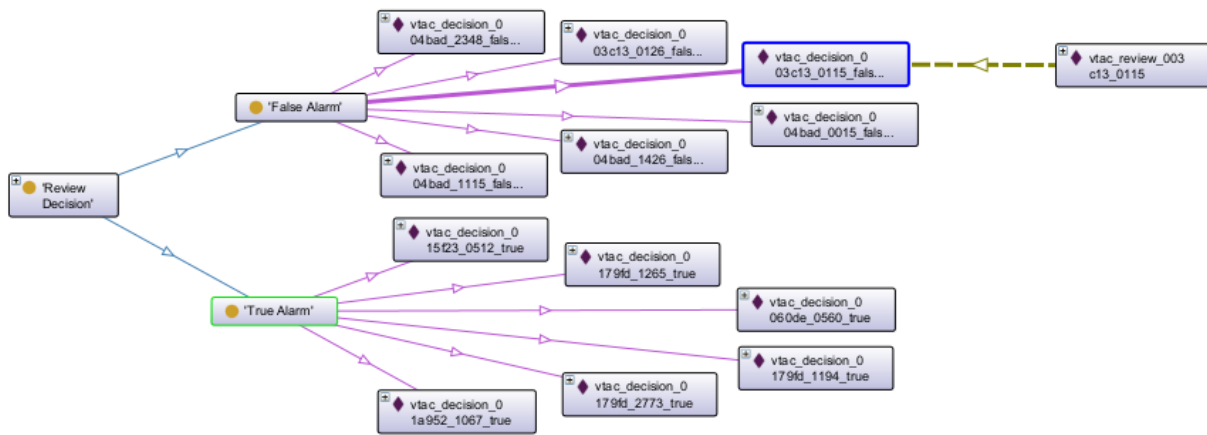
Τα false alarm decision individuals είναι:

- vtac_decision_003c13_0115_false
- vtac_decision_003c13_0126_false
- vtac_decision_004bad_0015_false
- vtac_decision_004bad_1115_false

- vtac_decision_004bad_1426_false
- vtac_decision_004bad_2348_false

Τα true alarm decision individuals είναι:

- vtac_decision_0060de_0560_true
- vtac_decision_015f23_0512_true
- vtac_decision_0179fd_1194_true
- vtac_decision_0179fd_1265_true
- vtac_decision_0179fd_2773_true
- vtac_decision_01a952_1067_true



Σχήμα 4.77 Αναπαράσταση των individuals της class Review Decision.

Το σχήμα 4.77 παρουσιάζει την class Review Decision και τις δύο υποκλάσεις της, False Alarm και True Alarm, στις οποίες ταξινομούνται τα individuals των τελικών αποφάσεων. Με αυτόν τον τρόπο, η οντολογία αποτυπώνει τη διάκριση ανάμεσα στα false και true alarm cases που προέρχονται από την ανθρώπινη αξιολόγηση των VTaC events [4].

Η σχέση ανάμεσα στο expert review και στο αντίστοιχο review decision αποτυπώνεται μέσω του object property hasDecision. Για παράδειγμα, το individual vtac_review_003c13_0115 συνδέεται με το decision individual vtac_decision_003c13_0115_false. Αντίστοιχα, ένα review που αφορά true alarm case συνδέεται με το αντίστοιχο true alarm decision individual.

Η σχέση αυτή μπορεί να αποδοθεί ως εξής:

Expert Review → hasDecision → Review Decision

Η ύπαρξη των review decision individuals είναι σημαντική επειδή η οντολογία δεν αποθηκεύει απλώς μια τιμή true ή false πάνω στο alert event. Αντίθετα, δημιουργεί ξεχωριστό individual για την απόφαση, το οποίο συνδέεται με το expert review και ταξινομείται στην κατάλληλη υποκλάση. Έτσι, η απόφαση αποκτά δική της θέση μέσα στο οντολογικό μοντέλο.

Η επιλογή αυτή κάνει τη σημασιολογική αλυσίδα πιο καθαρή. Το alert event συνδέεται πρώτα με το expert review και στη συνέχεια το expert review συνδέεται με τη review decision. Άρα η απόφαση δεν εμφανίζεται ως αυθαίρετη ή απομονωμένη τιμή, αλλά ως αποτέλεσμα συγκεκριμένης αξιολόγησης.

Η Review Decision αποτελεί το τελικό στοιχείο της διαδρομής ανθρώπινης αξιολόγησης του alarm event:

Alert Event → hasReview → Expert Review → hasDecision → Review Decision

Στα true alarm cases, το Review Decision ταξινομείται ως True Alarm, ενώ στα false alarm cases ταξινομείται ως False Alarm. Παράλληλα και ανεξάρτητα από τη διαδρομή της ανθρώπινης αξιολόγησης, το ίδιο Alert Event συνδέεται με το αντίστοιχο Evidence Record μέσω του isDocumentedBy, ενώ το Evidence Record χαρτογραφείται προς τα επιλεγμένα AI Act-related σημεία αναφοράς μέσω του supportsRequirement.

Για τα false alarm cases προστίθεται επιπλέον η lightweight risk-related αναπαράσταση, η οποία οργανώνεται μέσω δύο διαδρομών:

Alert Event → hasRiskIndicator → Risk Indicator

Alert Event → isClassifiedAs → Risk-Relevant Event → mayLeadTo → Potential Harm

Με αυτόν τον τρόπο, η διάκριση true/false παραμένει μέρος της ανθρώπινης αξιολόγησης, ενώ η τεκμηρίωση και η πρόσθετη risk-related πληροφορία αναπαρίστανται ως ξεχωριστές αλλά συνδεδεμένες σημασιολογικές όψεις του ίδιου Alert Event.

Αυτός ο διπλός τρόπος αναπαράστασης ενισχύει την αναγνωσιμότητα και τη δυνατότητα αναζήτησης. Ο χρήστης ή ο ερευνητής μπορεί να δει σε ποια class ανήκει ένα decision individual, αλλά μπορεί επίσης να αξιοποιήσει την τιμή του decisionLabel σε queries ή σε αναφορές.

Η ύπαρξη των review decision individuals είναι επίσης σημαντική για τα competency questions και τα SPARQL queries. Μέσω αυτών μπορεί να αναζητηθεί ποια events χαρακτηρίστηκαν ως false alarms, ποια ως true alarms και ποια false alarm cases συνδέονται επιπλέον με risk-related στοιχεία.

Πρέπει να σημειωθεί ότι η οντολογία δεν παράγει η ίδια την απόφαση true ή false. Δεν λειτουργεί ως μηχανισμός πρόβλεψης ούτε ως μοντέλο μηχανικής μάθησης. Η απόφαση αξιοποιείται ως υπάρχουσα πληροφορία από το VTaC dataset και εισάγεται στην οντολογία ως στοιχείο τεκμηρίωσης.

Τα review decision individuals αποτελούν κρίσιμο σημείο της εφαρμοσμένης οντολογικής υλοποίησης. Μέσω αυτών, τα VTaC alert events αποκτούν αξιολογική ερμηνεία, διακρίνονται σε true και false alarm cases και μπορούν να ενταχθούν στη συνέχεια σε evidence records, requirement mappings και, όπου χρειάζεται, στο lightweight risk layer.

4.8.8 Evidence record individuals

Τα evidence record individuals αποτελούν έναν από τους πιο σημαντικούς μηχανισμούς της οντολογίας, επειδή λειτουργούν ως γέφυρα ανάμεσα στα πραγματικά VTaC alert events και στις επιλεγμένες AI Act-related requirements. Αν τα alert event individuals αναπαριστούν τα πραγματικά περιστατικά και τα review decision individuals αναπαριστούν το αποτέλεσμα της αξιολόγησης, τότε τα evidence record individuals οργανώνουν την πληροφορία αυτή ως τεκμήριο.

Στην παρούσα υλοποίηση, για κάθε VTaC alert event δημιουργήθηκε ένα αντίστοιχο evidence record individual. Τα evidence record individuals ανήκουν στην class Evidence Record και χρησιμοποιούνται για να τεκμηριώσουν το αντίστοιχο event μέσα στο μοντέλο.

Τα evidence record individuals που χρησιμοποιούνται είναι τα εξής:

- vtac_evidence_003c13_0115
- vtac_evidence_003c13_0126
- vtac_evidence_004bad_0015
- vtac_evidence_004bad_1115
- vtac_evidence_004bad_1426
- vtac_evidence_004bad_2348
- vtac_evidence_0060de_0560
- vtac_evidence_015f23_0512
- vtac_evidence_0179fd_1194
- vtac_evidence_0179fd_1265
- vtac_evidence_0179fd_2773
- vtac_evidence_01a952_1067

Η σύνδεση ενός alert event με το αντίστοιχο evidence record γίνεται μέσω του object property `isDocumentedBy`. Για παράδειγμα, το event `vtac_event_003c13_0115` συνδέεται με το evidence record `vtac_evidence_003c13_0115`. Αυτό σημαίνει ότι το συγκεκριμένο event δεν υπάρχει απλώς ως alarm μέσα στην οντολογία, αλλά διαθέτει αντίστοιχο τεκμήριο που το υποστηρίζει.

Η σχέση αυτή μπορεί να αποδοθεί ως εξής:

Alert Event → isDocumentedBy → Evidence Record

Το evidence record είναι ιδιαίτερα σημαντικό επειδή αποφεύγει την απευθείας και αδύναμη σύνδεση ενός event με μια απαίτηση συμμόρφωσης. Η οντολογία δεν δηλώνει απλώς ότι ένα event σχετίζεται με μια απαίτηση. Αντίθετα, εισάγει ένα ενδιάμεσο τεκμήριο. Το event τεκμηριώνεται από evidence record και το evidence record υποστηρίζει συγκεκριμένες απαιτήσεις.

Η δεύτερη βασική σχέση των evidence records είναι το object property `supportsRequirement`. Μέσω αυτής της σχέσης, κάθε evidence record συνδέεται με ένα ή περισσότερα επιλεγμένα AI Act-related σημεία αναφοράς.

Η σχέση μπορεί να αποδοθεί ως εξής:

Evidence Record → supportsRequirement → Compliance Requirement

Στην παρούσα υλοποίηση, όλα τα πραγματικά VTaC alert events συνδέονται με evidence records που υποστηρίζουν τις απαιτήσεις:

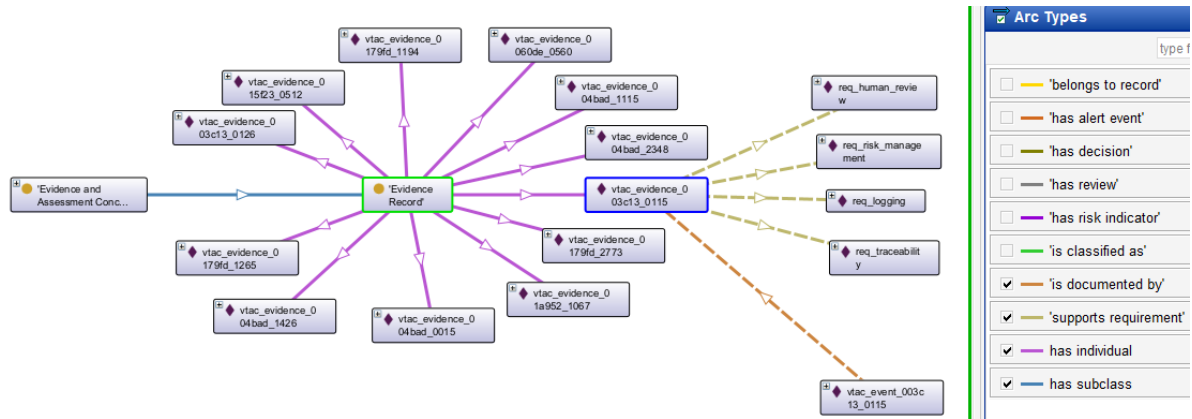
- req_logging
- req_human_review
- req_traceability

Αυτό συμβαίνει επειδή όλα τα επιλεγμένα events έχουν καταγραφεί, διαθέτουν ανθρώπινη ή expert αξιολόγηση και μπορούν να ιχνηλατηθούν μέσα από τη σημασιολογική αλυσίδα της οντολογίας.

Για τα false alarm cases, τα evidence records συνδέονται επιπλέον και με το requirement individual:

- req_risk_management

Αυτό συμβαίνει επειδή τα false alarms στην παρούσα οντολογία έχουν συνδεθεί με το lightweight risk layer μέσω risk indicators, risk-relevant events και potential harm. Επομένως, τα evidence records των false alarm cases δεν υποστηρίζουν μόνο logging, human review και traceability, αλλά και risk management.



Σχήμα 4.78 Evidence Record individuals και mappings προς τα επιλεγμένα AI Act-related σημεία αναφοράς στο Protégé.

Ένα ενδεικτικό παράδειγμα είναι το vtac_evidence_003c13_0115. Το evidence record αυτό τεκμηριώνει το event vtac_event_003c13_0115, το οποίο έχει χαρακτηριστεί ως false alarm. Για τον λόγο αυτό, το συγκεκριμένο evidence record μπορεί να συνδεθεί με τις απαιτήσεις logging, human review, traceability και risk management.

Η πλήρης λογική για το συγκεκριμένο παράδειγμα μπορεί να αποδοθεί ως εξής:

vtac_event_003c13_0115 → isDocumentedBy → vtac_evidence_003c13_0115

και στη συνέχεια:

vtac_evidence_003c13_0115 → supportsRequirement → req_logging

vtac_evidence_003c13_0115 → supportsRequirement → req_human_review

vtac_evidence_003c13_0115 → supportsRequirement → req_traceability

vtac_evidence_003c13_0115 → supportsRequirement → req_risk_management

Με αυτόν τον τρόπο, η οντολογία υλοποιεί στην πράξη το requirement-to-evidence mapping. Δεν παρουσιάζει τις απαιτήσεις ως αφηρημένες έννοιες, αλλά τις συνδέει με συγκεκριμένα evidence records που προκύπτουν από πραγματικά VTaC events.

Η διαφοροποίηση ανάμεσα σε true και false alarm cases φαίνεται καθαρά στο επίπεδο των evidence records. Τα evidence records που αντιστοιχούν σε true alarm cases υποστηρίζουν τα τρία βασικά AI Act-related σημεία αναφοράς: logging, human review και traceability. Αντίθετα, τα evidence records που αντιστοιχούν σε false alarm cases στο συγκεκριμένο proof-of-concept έχουν χαρτογραφηθεί επιπλέον και προς το req_risk_management λόγω της πρόσθετης risk-related αναπαράστασης των false alarm cases.

Η διάκριση αυτή είναι σημαντική, γιατί δείχνει ότι η οντολογία δεν αντιμετωπίζει όλα τα events με τον ίδιο τρόπο. Όλα τα events διαθέτουν βασική τεκμηρίωση, αλλά τα false alarms αποκτούν πρόσθετη σημασία από πλευράς risk management. Έτσι, το evidence layer λειτουργεί όχι μόνο ως γενικός

μηχανισμός τεκμηρίωσης, αλλά και ως σημείο διαφοροποίησης ανάμεσα σε διαφορετικούς τύπους περιστατικών.

Τα evidence record individuals είναι επίσης σημαντικά για τα queries. Μέσω αυτών μπορούν να αναζητηθούν ποια evidence records υποστηρίζουν κάθε απαίτηση, ποια events τεκμηριώνονται από συγκεκριμένα evidence records και ποιες απαιτήσεις συνδέονται με false ή true alarm cases. Αυτό καθιστά τα evidence records βασικό στοιχείο για την αναζήτηση, την αναφορά αποτελεσμάτων και την αξιολόγηση της λειτουργίας της οντολογίας.

Συνολικά, τα evidence record individuals αποτελούν τον βασικό μηχανισμό μετάβασης από τα operational δεδομένα προς το normative layer. Μέσω αυτών, τα πραγματικά VTaC alert events συνδέονται με οργανωμένα Evidence Records, τα οποία μπορούν στη συνέχεια να χαρτογραφούνται προς επιλεγμένα AI Act-related σημεία αναφοράς. Έτσι, η οντολογία αποκτά πρακτική αξία ως proof-of-concept μοντέλο για requirement-to-evidence mapping.

4.8.9 Risk-related individuals

Τα risk-related individuals αποτελούν το πρόσθετο επίπεδο της οντολογίας που σχετίζεται με το lightweight risk layer. Το επίπεδο αυτό δημιουργήθηκε ώστε να αποτυπώσει ότι ορισμένα alert events, κυρίως τα false alarm cases, μπορούν να έχουν πρόσθετη σημασία από πλευράς risk management.

Στην παρούσα οντολογία, τα risk-related individuals δεν χρησιμοποιούνται για να δηλώσουν ότι έχει συμβεί πραγματική βλάβη. Αντίθετα, χρησιμοποιούνται για να οργανώσουν σημασιολογικά την πιθανή risk-related ερμηνεία συγκεκριμένων false alarm events. Με αυτόν τον τρόπο, η οντολογία μπορεί να δείξει ότι ένα false alarm δεν είναι απλώς μια λανθασμένη ειδοποίηση, αλλά μπορεί να συνδεθεί με πιθανές επιπτώσεις, όπως το alert fatigue [6].

Το lightweight risk layer αποτελείται από τρεις βασικές ομάδες individuals:

- Risk Indicator individuals
- Risk-Relevant Event individuals
- Potential Harm individual

Η πρώτη ομάδα αφορά τα Risk Indicator individuals. Τα individuals αυτά χρησιμοποιούνται για να δηλώσουν ότι ένα alert event έχει χαρακτηριστικό που το καθιστά σχετικό με risk-related ερμηνεία. Στην παρούσα υλοποίηση, ο βασικός risk indicator είναι η περίπτωση false alarm. Για κάθε false alarm event δημιουργήθηκε αντίστοιχο risk indicator individual.

Τα risk indicator individuals που χρησιμοποιούνται είναι τα εξής:

- vtac_indicator_003c13_0115_false_alarm
- vtac_indicator_003c13_0126_false_alarm
- vtac_indicator_004bad_0015_false_alarm
- vtac_indicator_004bad_1115_false_alarm
- vtac_indicator_004bad_1426_false_alarm
- vtac_indicator_004bad_2348_false_alarm

Η σχέση ανάμεσα σε ένα alert event και στο αντίστοιχο risk indicator αποτυπώνεται με το object property `hasRiskIndicator`.

Η σχέση αυτή μπορεί να αποδοθεί ως εξής:

Alert Event → `hasRiskIndicator` → Risk Indicator

Για παράδειγμα, το `vtac_event_003c13_0115` συνδέεται με το `vtac_indicator_003c13_0115_false_alarm`. Η σύνδεση αυτή δείχνει ότι το συγκεκριμένο false alarm event διαθέτει risk-related ένδειξη μέσα στο μοντέλο.

Η δεύτερη ομάδα αφορά τα Risk-Relevant Event individuals. Τα individuals αυτά χρησιμοποιούνται για να αναπαραστήσουν τη risk-related ερμηνεία ενός alert event. Δηλαδή, ενώ το αρχικό alert event ανήκει στο operational layer, το αντίστοιχο risk-relevant event ανήκει στο risk layer.

Τα risk-relevant event individuals είναι τα εξής:

- `vtac_risk_event_003c13_0115`
- `vtac_risk_event_003c13_0126`
- `vtac_risk_event_004bad_0015`
- `vtac_risk_event_004bad_1115`
- `vtac_risk_event_004bad_1426`
- `vtac_risk_event_004bad_2348`

Η σχέση ανάμεσα σε ένα alert event και στο αντίστοιχο risk-relevant event αποτυπώνεται με το object property `isClassifiedAs`.

Η σχέση αυτή μπορεί να αποδοθεί ως εξής:

Alert Event → `isClassifiedAs` → Risk-Relevant Event

Για παράδειγμα, το `vtac_event_003c13_0115` συνδέεται με το `vtac_risk_event_003c13_0115`. Αυτό σημαίνει ότι το συγκεκριμένο alert event έχει ταξινομηθεί ως risk-relevant μέσα στο lightweight risk layer της οντολογίας.

Η τρίτη ομάδα αφορά το Potential Harm individual. Στην παρούσα υλοποίηση χρησιμοποιείται ένα βασικό potential harm individual:

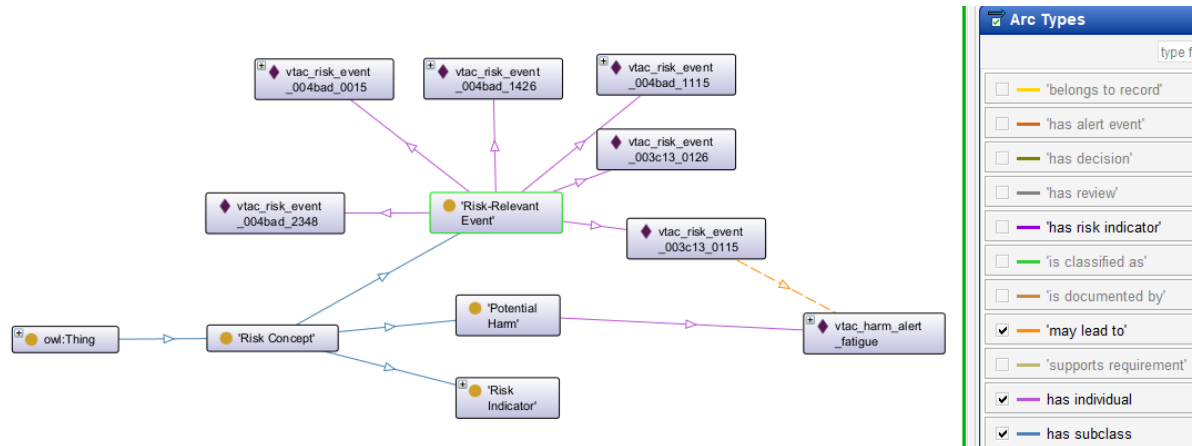
- `vtac_harm_alert_fatigue`

Το individual αυτό αναπαριστά την πιθανή επίπτωση του alert fatigue. Η επιλογή αυτή συνδέεται με το γεγονός ότι τα false alarms σε περιβάλλοντα παρακολούθησης υγείας μπορούν να συμβάλουν σε κόπωση συναγερμών. Η οντολογία, όμως, δεν δηλώνει ότι το alert fatigue συνέβη πραγματικά σε κάποιο συγκεκριμένο περιστατικό. Δηλώνει μόνο ότι μπορεί να αποτελέσει πιθανή risk-related επίπτωση που σχετίζεται με false alarm cases.

Η σύνδεση ανάμεσα σε ένα risk-relevant event και στο potential harm γίνεται μέσω του object property `mayLeadTo`.

Η σχέση αυτή μπορεί να αποδοθεί ως εξής:

Risk-Relevant Event → `mayLeadTo` → Potential Harm



Σχήμα 4.79 Risk-related individuals και σύνδεση με potential harm.

Το σχήμα 4.79 παρουσιάζει το lightweight risk layer της οντολογίας, το οποίο οργανώνεται κάτω από την class Risk Concept. Στο διάγραμμα εμφανίζονται οι βασικές risk-related classes Risk-Relevant Event, Potential Harm και Risk Indicator, καθώς και τα risk-relevant event individuals που αντιστοιχούν στα false alarm cases. Ενδεικτικά, το individual vtac_risk_event_003c13_0115 συνδέεται με το potential harm individual vtac_harm_alert_fatigue μέσω της σχέσης mayLeadTo. Η απεικόνιση δείχνει πώς ένα false alarm case μπορεί να αποκτήσει πρόσθετη risk-related ερμηνεία μέσα στην οντολογία, χωρίς να δηλώνεται ότι η πιθανή επίπτωση έχει πράγματι συμβεί.

Η συνολική risk-related αλυσίδα μπορεί να αποδοθεί ως εξής:

Alert Event → hasRiskIndicator → Risk Indicator

Alert Event → isClassifiedAs → Risk-Relevant Event

Risk-Relevant Event → mayLeadTo → Potential Harm

Η αλυσίδα αυτή επεκτείνει τη βασική διαδρομή τεκμηρίωσης της οντολογίας. Η βασική διαδρομή συνδέει το alert event με expert review, review decision, evidence record και compliance requirement. Το risk layer προσθέτει μια επιπλέον διάσταση μόνο για τα false alarm cases, ώστε να αποτυπωθεί η πιθανή σχέση τους με risk management.

Η σύνδεση των risk-related individuals με την απαίτηση req_risk_management γίνεται έμμεσα μέσω των evidence records. Τα false alarm events τεκμηριώνονται από αντίστοιχα evidence records και τα evidence records μπορούν να υποστηρίξουν το requirement individual req_risk_management μέσω του object property supportsRequirement.

Η λογική αυτή μπορεί να αποδοθεί ως εξής:

False Alarm Event → isDocumentedBy → Evidence Record

Evidence Record → supportsRequirement → req_risk_management

Με αυτόν τον τρόπο, η οντολογία δεν ισχυρίζεται ότι πραγματοποιεί πλήρη διαχείριση κινδύνου. Αντίθετα, δείχνει πώς μπορεί να μοντελοποιηθεί σημασιολογικά μια απλή risk-related διάσταση, ώστε να υποστηριχθεί η τεκμηρίωση και η ιχνηλασιμότητα.

Επιπλέον, τα false alarm alert events μπορούν να φέρουν και το data property isRiskRelevant με τιμή true. Το data property αυτό δεν αντικαθιστά τις object property συνδέσεις με Risk Indicator, Risk-Relevant Event και Potential Harm. Λειτουργεί συμπληρωματικά, ως απλή boolean ένδειξη ότι το συγκεκριμένο event συμμετέχει στο lightweight risk layer.

Τα true alarm cases δεν εντάσσονται στο lightweight risk-related επίπεδο της παρούσας υλοποίησης. Παραμένουν συνδεδεμένα με τις βασικές πληροφορίες προέλευσης, ανθρώπινης αξιολόγησης και τεκμηρίωσης. Αντίθετα, τα έξι επιλεγμένα false alarm cases αποκτούν επιπλέον risk-related αναπαράσταση μέσω Risk Indicator και Risk-Relevant Event.

Με αυτόν τον τρόπο, τα risk-related individuals επιτρέπουν να αναπαρασταθεί ότι τα false alarm cases μπορούν να εξεταστούν και από πρόσθετη risk-related οπτική. Η σύνδεση με το req_risk_management πραγματοποιείται μέσω του αντίστοιχου Evidence Record ως εφαρμοσμένο AI Act-related mapping και δεν αποτελεί απόδειξη πλήρους υλοποίησης του risk management system του άρθρου 9.

Συνολικά, το lightweight risk-related επίπεδο λειτουργεί ως συμπληρωματική επέκταση της βασικής event-centered αναπαράστασης. Η πραγματική τεχνική δομή περιλαμβάνει τις παράλληλες σχέσεις Alert Event → Risk Indicator και Alert Event → Risk-Relevant Event → Potential Harm. Η σύνδεση με Potential Harm εκφράζει πιθανή σημασιολογική συσχέτιση και όχι τεκμηριωμένη πραγματική πρόκληση βλάβης.

4.8.10 Παράδειγμα συνδεδεμένων σημασιολογικών διαδρομών για επιλεγμένο individual

Για να παρουσιαστεί συγκεντρωτικά ο τρόπος με τον οποίο συνεργάζονται τα διαφορετικά στοιχεία της οντολογίας, χρησιμοποιείται ως παράδειγμα το πραγματικό VTaC alert event:

```
vtac_event_003c13_0115
```

Το συγκεκριμένο individual αποτελεί instance της class Ventricular Tachycardia Alarm και αντιστοιχεί σε πραγματικό alarm event του VTaC dataset, το οποίο έχει τελική expert απόφαση false alarm [4]. Η επιλογή του είναι ιδιαίτερα χρήσιμη, επειδή επιτρέπει να παρουσιαστούν τόσο η βασική διαδρομή προέλευσης, αξιολόγησης και τεκμηρίωσης όσο και η πρόσθετη lightweight risk-related αναπαράσταση.

Πρέπει να διευκρινιστεί ότι η συνολική αναπαράσταση δεν αποτελεί μία μοναδική γραμμική object-property αλυσίδα. Αποτελεί συνδυασμό διαφορετικών σημασιολογικών διαδρομών που έχουν ως κοινό κεντρικό σημείο το ίδιο Alert Event.

Αρχικά, το event συνδέεται με το waveform record προέλευσής του:

```
vtac_event_003c13_0115 --belongsToRecord→ vtac_record_003c13
```

Αντίστροφα, από την πλευρά του waveform record:

```
vtac_record_003c13 --hasAlertEvent→ vtac_event_003c13_0115
```

Η σχέση αυτή επιτρέπει στο πραγματικό event να ιχνηλατηθεί προς το record από το οποίο προέρχεται.

Η ανθρώπινη αξιολόγηση αναπαρίσταται μέσω διαφορετικής διαδρομής. Το event συνδέεται με το αντίστοιχο Expert Review individual:

```
vtac_event_003c13_0115 --hasReview→ vtac_review_003c13_0115
```

και το Expert Review συνδέεται με την τελική Review Decision:

```
vtac_review_003c13_0115 --hasDecision→ vtac_decision_003c13_0115_false
```

Το decision individual ανήκει στην class False Alarm. Η συγκεκριμένη απόφαση δεν παράγεται από την οντολογία. Αναπαριστά την τελική expert απόφαση που είναι διαθέσιμη στο VTaC dataset [4].

Παράλληλα, το ίδιο Alert Event συνδέεται με το Evidence Record που χρησιμοποιείται για την τεκμηρίωσή του:

```
vtac_event_003c13_0115 --isDocumentedBy→ vtac_evidence_003c13_0115
```

Το Evidence Record αποτελεί τον τεκμηριωτικό κόμβο μέσω του οποίου πραγματοποιείται το requirement-to-evidence mapping. Στο συγκεκριμένο false alarm case έχουν υλοποιηθεί οι σχέσεις:

```
vtac_evidence_003c13_0115 --supportsRequirement→ req_logging
```

```
vtac_evidence_003c13_0115 --supportsRequirement→ req_human_review
```

```
vtac_evidence_003c13_0115 --supportsRequirement→ req_traceability
```

```
vtac_evidence_003c13_0115 --supportsRequirement→ req_risk_management
```

Οι σχέσεις αυτές δηλώνουν ότι το συγκεκριμένο evidence record έχει χαρτογραφηθεί προς τα τέσσερα εφαρμοσμένα AI Act-related σημεία αναφοράς του proof-of-concept. Δεν δηλώνουν ότι έχει αποδειχθεί πλήρης νομική συμμόρφωση με τις αντίστοιχες διατάξεις του AI Act [3].

Επειδή το vtac_event_003c13_0115 αποτελεί ένα από τα έξι επιλεγμένα false alarm cases, συμμετέχει επιπλέον στο lightweight risk-related μέρος της οντολογίας.

Η πρώτη σχετική σύνδεση είναι:

```
vtac_event_003c13_0115 --hasRiskIndicator→ vtac_indicator_003c13_0115_false_alarm
```

Η δεύτερη είναι:

```
vtac_event_003c13_0115 --isClassifiedAs→ vtac_risk_event_003c13_0115
```

και στη συνέχεια:

```
vtac_risk_event_003c13_0115 --mayLeadTo→ vtac_harm_alert_fatigue
```

Η τελευταία σχέση αποτυπώνει πιθανή risk-related σύνδεση με το alert fatigue και όχι αποδεδειγμένη αιτιώδη σχέση ή πραγματική βλάβη. Το vtac_harm_alert_fatigue λειτουργεί ως Potential Harm individual της συγκεκριμένης proof-of-concept αναπαράστασης.

Παράλληλα με τις object-property assertions, το Alert Event διαθέτει data properties όπως eventId, hasHumanReview, isLogged, isRiskRelevant και isTraceable. Οι boolean τιμές των properties αυτών λειτουργούν ως συμπληρωματικές ενδείξεις της συγκεκριμένης οντολογικής αναπαράστασης. Δεν αντικαθιστούν τις αναλυτικές object-property σχέσεις ούτε αποτελούν αυτοτελείς αποδείξεις συμμόρφωσης.

The screenshot displays the Protégé interface for the individual `vtac_event_003c13_0115`. The left pane shows the 'Types' section with 'Ventricular Tachycardia Alarm' selected. The right pane shows the 'Property assertions' section, which is divided into 'Object property assertions' and 'Data property assertions'.

Object property assertions:

- 'belongs to record' `vtac_record_003c13`
- 'has review' `vtac_review_003c13_0115`
- 'has risk indicator' `vtac_indicator_003c13_0115_false_alarm`
- 'is classified as' `vtac_risk_event_003c13_0115`
- 'is documented by' `vtac_evidence_003c13_0115`

Data property assertions:

- 'event id' `"003c13_0115"`
- 'has human review' `true`
- 'is logged' `true`
- 'is risk relevant' `true`
- 'is traceable' `true`

Σχήμα 4.80 Άμεσες object και data property assertions του `vtac_event_003c13_0115`.

Το Σχήμα 4.80 παρουσιάζει συγκεντρωτικά τις άμεσες assertions του συγκεκριμένου event στο Protégé. Η εικόνα είναι ιδιαίτερα χρήσιμη επειδή δείχνει ότι ένα πραγματικό VTaC alarm event δεν παραμένει απομονωμένη εγγραφή. Γύρω από αυτό οργανώνονται διαφορετικές αλλά συνδεδεμένες πληροφορίες που αφορούν την προέλευση, την ανθρώπινη αξιολόγηση, την απόφαση, την τεκμηρίωση, το requirement mapping και την πρόσθετη risk-related αναπαράσταση.

Το παράδειγμα του `vtac_event_003c13_0115` συνοψίζει επομένως τον βασικό σκοπό της υλοποίησης. Η οντολογία δεν προβλέπει αν το alarm είναι true ή false, δεν πραγματοποιεί ιατρική διάγνωση και δεν πιστοποιεί νομική συμμόρφωση. Παρέχει ένα σημασιολογικό πλαίσιο μέσα στο οποίο ένα πραγματικό healthcare monitoring event μπορεί να συνδεθεί με τις διαθέσιμες πληροφορίες και τα τεκμήρια που το αφορούν και να ανακτηθεί μέσα από ελέγξιμες σημασιολογικές διαδρομές.

4.8.11 Συνολική αποτίμηση των individuals

Η εισαγωγή των individuals αποτελεί το σημείο στο οποίο η αφηρημένη δομή των classes και properties εφαρμόζεται σε συγκεκριμένο περιεχόμενο. Μέσω των individuals, η οντολογία μετατρέπεται από ένα γενικό εννοιολογικό σχήμα σε εφαρμοσμένο proof-of-concept που μπορεί να ελεγχθεί μέσω reasoner, DL Queries και SPARQL queries.

Είναι σημαντικό, ωστόσο, να γίνεται διάκριση μεταξύ των στοιχείων που προέρχονται από το πραγματικό VTaC dataset και των οντολογικών individuals που δημιουργήθηκαν στο πλαίσιο της παρούσας εργασίας. Τα waveform record identifiers, τα alarm event identifiers και οι τελικές expert αποφάσεις true ή false βασίζονται στα πραγματικά δεδομένα του VTaC [4]. Αντίθετα, τα Expert Review, Evidence Record, requirement, Control Mechanism και risk-related individuals αποτελούν σημασιολογικές αναπαραστάσεις που δημιουργήθηκαν στην οντολογία για να οργανώσουν και να συνδέσουν τις διαθέσιμες πληροφορίες.

Η διάκριση αυτή δεν μειώνει τον πραγματολογικό χαρακτήρα του proof-of-concept. Αντίθετα, αποσαφηνίζει ποιο μέρος της γνώσης παρέχεται από το dataset και ποιο μέρος αποτελεί τη μοντελοποίηση που προτείνεται από την εργασία.

Τα waveform record και alert event individuals διατηρούν τη σύνδεση με τις πραγματικές εγγραφές του VTaC. Τα Expert Review και Review Decision individuals οργανώνουν σημασιολογικά την πληροφορία της ανθρώπινης αξιολόγησης και των τελικών labels. Τα Evidence Record individuals

λειτουργούν ως τεκμηριωτικοί κόμβοι, ενώ τα requirement individuals παρέχουν τα επιλεγμένα AI Act-related σημεία αναφοράς για το requirement-to-evidence mapping.

Για τα επιλεγμένα false alarm cases, τα Risk Indicator, Risk-Relevant Event και Potential Harm individuals προσθέτουν ένα περιορισμένο risk-related επίπεδο. Η αναπαράσταση αυτή δεν εισάγει νέα κλινικά γεγονότα και δεν αποδεικνύει πραγματική βλάβη. Αποτελεί οντολογική επέκταση που επιτρέπει να εξεταστεί η δυνατότητα σημασιολογικής οργάνωσης πιθανής risk-related πληροφορίας.

Τα Control Mechanism individuals και το Monitoring System individual έχουν περισσότερο εννοιολογικό και contextual ρόλο στην τρέχουσα proof-of-concept έκδοση. Παρέχουν δομικά σημεία τα οποία μπορούν να αξιοποιηθούν σε μελλοντική επέκταση της οντολογίας με επιπλέον object properties και πραγματικά τεχνικά ή οργανωτικά δεδομένα.

Συνολικά, τα individuals δείχνουν ότι η υλοποίηση δεν περιορίζεται σε μια ταξινομική ιεραρχία classes. Το μοντέλο περιλαμβάνει συγκεκριμένα instances και assertions που επιτρέπουν να ανακτηθούν πραγματικά VTaC events, η προέλευσή τους, οι διαθέσιμες expert αποφάσεις, τα evidence records, τα requirement mappings και οι risk-related σχέσεις που έχουν σχεδιαστεί στο proof-of-concept.

Η λειτουργική ορθότητα αυτών των συνδέσεων δεν αξιολογείται μόνο οπτικά μέσα από τις καρτέλες του Protégé. Στο επόμενο κεφάλαιο εξετάζεται συστηματικά μέσω reasoner, DL competency questions και SPARQL queries, ώστε να διαπιστωθεί αν η γνώση που έχει εισαχθεί μπορεί να ανακτηθεί σύμφωνα με τον σχεδιασμό της οντολογίας.

4.9 Συνολική αποτίμηση της υλοποίησης

Η υλοποίηση της οντολογίας στο Protégé μετέτρεψε τον μεθοδολογικό σχεδιασμό του προηγούμενου κεφαλαίου σε ένα συγκεκριμένο OWL μοντέλο, αποτελούμενο από classes, object properties, data properties και individuals. Η δομή δεν περιορίζεται σε μια απλή ταξινομική ιεραρχία, αλλά περιλαμβάνει ρητές σημασιολογικές σχέσεις και συγκεκριμένα instances που επιτρέπουν την εφαρμογή του μοντέλου σε πραγματικά VTaC healthcare monitoring alarm events [4].

Οι classes οργανώνουν τις βασικές έννοιες της οντολογίας στις τέσσερις κύριες εννοιολογικές ομάδες: Operational Concept, Evidence and Assessment Concept, Normative Concept και Risk Concept. Η οργάνωση αυτή επιτρέπει τη διάκριση μεταξύ των πραγματικών healthcare monitoring στοιχείων, της τεκμηρίωσης και αξιολόγησης, των επιλεγμένων AI Act-related σημείων αναφοράς και της πρόσθετης risk-related πληροφορίας.

Τα object properties αποτελούν τον κύριο μηχανισμό σύνδεσης των διαφορετικών στοιχείων του μοντέλου. Μέσω των belongsToRecord και hasAlertEvent διατηρείται η σχέση μεταξύ waveform records και alarm events. Τα hasReview και hasDecision οργανώνουν τη διαδρομή ανθρώπινης αξιολόγησης, ενώ τα isDocumentedBy και supportsRequirement υλοποιούν τη διαδρομή τεκμηρίωσης και το requirement-to-evidence mapping.

Η συνολική σημασιολογική δομή δεν αποτελεί μία μοναδική γραμμική αλυσίδα. Αντίθετα, γύρω από κάθε Alert Event σχηματίζονται συνδεδεμένες διαδρομές:

Alert Event --belongsToRecord→ Waveform Record

Alert Event --hasReview→ Expert Review --hasDecision→ Review Decision

Alert Event --isDocumentedBy→ Evidence Record --supportsRequirement→ Requirement individual

Η σύνθεση των σχέσεων αυτών επιτρέπει την ιχνηλάτηση της προέλευσης, της ανθρώπινης αξιολόγησης, της απόφασης, της τεκμηρίωσης και του αντίστοιχου requirement mapping.

Τα data properties συμπληρώνουν τις παραπάνω σχέσεις με identifiers, labels, outcomes και boolean ενδείξεις. Ιδιαίτερα τα hasHumanReview, isLogged, isTraceable και isRiskRelevant πρέπει να ερμηνεύονται ως χαρακτηριστικά της proof-of-concept αναπαράστασης και όχι ως αυτοτελείς αποδείξεις ότι ένα πραγματικό εξωτερικό σύστημα ικανοποιεί αντίστοιχες νομικές ή τεχνικές απαιτήσεις του AI Act.

Ιδιαίτερη σημασία έχει η χρήση των Evidence Record individuals. Μέσω αυτών διαχωρίζεται το πραγματικό alarm event από το κανονιστικό επίπεδο της οντολογίας. Ένα event τεκμηριώνεται μέσω συγκεκριμένου Evidence Record και το evidence χαρτογραφείται στη συνέχεια προς τα req_logging, req_human_review, req_traceability και, όπου έχει προβλεφθεί από τον σχεδιασμό, req_risk_management. Τα mappings αυτά αποτελούν εφαρμοσμένες συνδέσεις προς επιλεγμένα AI Act-related σημεία αναφοράς και όχι απόδειξη πλήρους νομικής συμμόρφωσης [3].

Για τα έξι επιλεγμένα false alarm cases, η βασική αναπαράσταση επεκτείνεται με το lightweight risk-related μέρος:

Alert Event --hasRiskIndicator→ Risk Indicator

Alert Event --isClassifiedAs→ Risk-Relevant Event

Risk-Relevant Event --mayLeadTo→ Potential Harm

Με τον τρόπο αυτό, τα false alarm cases μπορούν να αποκτήσουν πρόσθετη risk-related σημασιολογική πληροφορία χωρίς να δηλώνεται ότι ένα συγκεκριμένο false alarm προκάλεσε πραγματική βλάβη.

Τα individuals επιτρέπουν την εφαρμογή της αφηρημένης δομής σε συγκεκριμένες περιπτώσεις. Είναι όμως σημαντικό να διακρίνονται τα στοιχεία που έχουν πραγματική βάση στο VTaC — όπως waveform record identifiers, alarm event identifiers και expert decisions — από τα οντολογικά individuals που δημιουργήθηκαν στο πλαίσιο της εργασίας για τη σημασιολογική οργάνωση αυτών των πληροφοριών. Evidence Records, requirement individuals, Control Mechanisms και risk-related individuals αποτελούν μέρος της προτεινόμενης μοντελοποίησης.

Παράλληλα, ορισμένα στοιχεία της οντολογίας, όπως τα Control Mechanism individuals και το vtac_icu_monitoring_system, έχουν στην τρέχουσα έκδοση κυρίως εννοιολογικό ή contextual ρόλο. Δεν συμμετέχουν όλα στις βασικές object-property διαδρομές και δεν πρέπει να ερμηνεύονται ως ενεργοί τεχνικοί μηχανισμοί. Η παρουσία τους παρέχει, ωστόσο, σημεία στα οποία το μοντέλο μπορεί να επεκταθεί μελλοντικά.

Συνολικά, η υλοποίηση δείχνει ότι η προτεινόμενη αρχιτεκτονική μπορεί να εφαρμοστεί σε συγκεκριμένα πραγματικά healthcare monitoring events και να οργανώσει σε ένα ενιαίο σημασιολογικό μοντέλο την προέλευση, την αξιολόγηση, την απόφαση, την τεκμηρίωση, τα επιλεγμένα requirement mappings και την πρόσθετη risk-related πληροφορία.

Το Κεφάλαιο 4 παρουσιάζει επομένως **τι έχει υλοποιηθεί** στην οντολογία. Το επόμενο στάδιο είναι να εξεταστεί αν οι δηλωμένες σχέσεις και τα instances λειτουργούν όπως έχει σχεδιαστεί. Για τον λόγο αυτό, στο Κεφάλαιο 5 πραγματοποιείται η λογική και λειτουργική επικύρωση της οντολογίας μέσω reasoner, DL competency questions και SPARQL queries.

4.10 Επίλογος

Στο παρόν κεφάλαιο παρουσιάστηκε αναλυτικά η υλοποίηση της οντολογίας στο Protégé. Αναλύθηκαν η ιεραρχία των classes, τα object properties, τα data properties, τα annotations και τα individuals που συγκροτούν το τελικό proof-of-concept μοντέλο.

Η υλοποίηση επιτρέπει την αναπαράσταση πραγματικών VTaC alarm events και τη σύνδεσή τους με waveform records, Expert Reviews, Review Decisions και Evidence Records. Μέσω των Evidence Records πραγματοποιείται το requirement-to-evidence mapping προς τα επιλεγμένα AI Act-related σημεία αναφοράς, ενώ για τα false alarm cases έχει ενσωματωθεί πρόσθετη lightweight risk-related αναπαράσταση.

Ιδιαίτερη σημασία έχει ότι οι διαφορετικές πληροφορίες δεν οργανώνονται ως μία μοναδική γραμμική αλυσίδα, αλλά μέσω συνδεδεμένων σημασιολογικών διαδρομών γύρω από το Alert Event. Η δομή αυτή διατηρεί διακριτά το πραγματικό περιστατικό, την ανθρώπινη αξιολόγηση, την τεκμηρίωση και το κανονιστικό mapping, ενώ παράλληλα επιτρέπει την αναζήτηση και την ιχνηλάτησή τους.

Η υλοποίηση δεν αποτελεί σύστημα πρόβλεψης, κλινικό διαγνωστικό εργαλείο ή μηχανισμό πιστοποίησης συμμόρφωσης με τον AI Act. Αποτελεί ένα εφαρμοσμένο οντολογικό proof-of-concept, μέσα στο οποίο πραγματικά healthcare monitoring events οργανώνονται σε σαφή και ελέγξιμη σημασιολογική μορφή.

Με την ολοκλήρωση της υλοποίησης, το επόμενο βήμα είναι η εξέταση της λειτουργικής συμπεριφοράς της οντολογίας. Στο Κεφάλαιο 5 παρουσιάζεται ο έλεγχος λογικής συνέπειας με reasoner και η λειτουργική επικύρωση μέσω DL competency questions και SPARQL queries, προκειμένου να ελεγχθεί αν οι σημασιολογικές σχέσεις που σχεδιάστηκαν και υλοποιήθηκαν μπορούν πράγματι να ανακτηθούν όπως αναμένεται.

Κεφάλαιο 5ο: Έλεγχος και λειτουργική επικύρωση της οντολογίας

5.1 Εισαγωγή

Μετά την παρουσίαση της υλοποίησης της οντολογίας στο Protégé, το παρόν κεφάλαιο επικεντρώνεται στον έλεγχο της λειτουργίας της. Στο προηγούμενο κεφάλαιο αναλύθηκαν οι classes, τα object properties, τα data properties και τα individuals που συγκροτούν το οντολογικό μοντέλο. Ωστόσο, η ύπαρξη αυτών των στοιχείων δεν αρκεί από μόνη της. Είναι απαραίτητο να εξεταστεί αν η οντολογία είναι λογικά συνεπής και αν μπορεί να απαντήσει σε συγκεκριμένα ερωτήματα που σχετίζονται με τη δομή και τη χρήση της.

Η λειτουργική επικύρωση της οντολογίας πραγματοποιείται με τρεις βασικούς τρόπους. Πρώτον, χρησιμοποιείται reasoner, ώστε να ελεγχθεί η λογική συνέπεια του μοντέλου. Δεύτερον, αξιοποιούνται DL competency questions, μέσα από τα οποία ελέγχεται αν η οντολογία μπορεί να επιστρέψει τα αναμενόμενα αποτελέσματα με βάση τη σημασιολογική της δομή. Τρίτον, χρησιμοποιούνται SPARQL queries, ώστε να εξεταστεί πιο δομημένα η δυνατότητα αναζήτησης και ανάκτησης πληροφορίας από την οντολογία.

Ο έλεγχος αυτός είναι ιδιαίτερα σημαντικός, επειδή η εργασία δεν περιορίζεται στη θεωρητική σχεδίαση μιας οντολογίας. Αντίθετα, προτείνει ένα proof-of-concept μοντέλο που εφαρμόζεται σε πραγματικά VTaC healthcare monitoring events και τα συνδέει με expert reviews, review decisions, evidence records και επιλεγμένα AI Act-related σημεία αναφοράς. Επομένως, πρέπει να ελεγχθεί αν οι βασικές σημασιολογικές διαδρομές που σχεδιάστηκαν μπορούν πράγματι να χρησιμοποιηθούν στην πράξη.

Το κεφάλαιο αυτό δεν εξετάζει την απόδοση κάποιου αλγορίθμου μηχανικής μάθησης και δεν επιχειρεί να αποδείξει πλήρη νομική συμμόρφωση με τον AI Act. Αντίθετα, εστιάζει στη λειτουργική αξιολόγηση της οντολογίας ως σημασιολογικού μοντέλου τεκμηρίωσης, ιχνηλασιμότητας και requirement-to-evidence mapping.

Στις επόμενες ενότητες παρουσιάζεται πρώτα ο σκοπός και η λογική της λειτουργικής επικύρωσης. Στη συνέχεια ακολουθεί ο έλεγχος λογικής συνέπειας με reasoner, η αξιολόγηση μέσω DL competency questions και η αξιοποίηση SPARQL queries για πιο σύνθετη αναζήτηση και συγκεντρωτική αποτίμηση των αποτελεσμάτων.

5.2 Σκοπός και λογική της λειτουργικής επικύρωσης

Η λειτουργική επικύρωση της οντολογίας έχει ως σκοπό να εξετάσει αν το μοντέλο που υλοποιήθηκε στο Protégé μπορεί να χρησιμοποιηθεί με συνεπή, ελέγξιμο και κατανοητό τρόπο. Η επικύρωση αυτή δεν αφορά μόνο την ύπαρξη των δομικών στοιχείων της οντολογίας, αλλά κυρίως τον τρόπο με τον οποίο αυτά συνεργάζονται μεταξύ τους.

Στο πλαίσιο της παρούσας εργασίας, το βασικό ερώτημα δεν είναι αν η οντολογία μπορεί να προβλέψει σωστά ένα αποτέλεσμα, όπως θα συνέβαινε σε ένα σύστημα μηχανικής μάθησης. Το βασικό ερώτημα είναι αν η οντολογία μπορεί να αναπαραστήσει σωστά τις σχέσεις ανάμεσα στα πραγματικά VTaC alarm events, στις expert reviews, στις review decisions, στα evidence records και στις επιλεγμένες AI Act-related requirements.

Για τον λόγο αυτό, η λειτουργική επικύρωση δεν βασίζεται σε μετρικές όπως accuracy, precision, recall ή F1-score. Τέτοιες μετρικές θα ήταν κατάλληλες για την αξιολόγηση ενός αλγορίθμου ταξινόμησης ή πρόβλεψης. Στην παρούσα εργασία, όμως, δεν εκπαιδεύεται ούτε αξιολογείται κάποιο μοντέλο πρόβλεψης. Τα true και false labels των VTaC events χρησιμοποιούνται ως υπάρχουσα πληροφορία του dataset και όχι ως αποτέλεσμα ενός νέου υπολογιστικού μοντέλου.

Η επικύρωση εστιάζει στη σημασιολογική λειτουργικότητα της οντολογίας. Με άλλα λόγια, εξετάζεται αν το μοντέλο μπορεί να υποστηρίξει τις βασικές συνδεδεμένες σημασιολογικές διαδρομές που έχουν σχεδιαστεί γύρω από το ίδιο Alert Event:

Alert Event → belongsToRecord → Waveform Record

Alert Event → hasReview → Expert Review → hasDecision → Review Decision

Alert Event → isDocumentedBy → Evidence Record → supportsRequirement → Compliance Requirement

Μέσω των διαδρομών αυτών ελέγχεται αν ένα πραγματικό alert event μπορεί να συνδεθεί με το waveform record προέλευσής του, με την expert αξιολόγηση και την τελική απόφαση, καθώς και με το Evidence Record και τα αντίστοιχα requirement mappings. Οι σχέσεις αυτές δεν αποτελούν μία μοναδική γραμμική object-property αλυσίδα, αλλά διακριτές και συνδεδεμένες σημασιολογικές διαδρομές.

Παράλληλα, η λειτουργική επικύρωση εξετάζει και την πρόσθετη lightweight risk-related αναπαράσταση που εφαρμόζεται στα false alarm cases μέσω δύο παράλληλων διαδρομών:

Alert Event → hasRiskIndicator → Risk Indicator

Alert Event → isClassifiedAs → Risk-Relevant Event → mayLeadTo → Potential Harm

Ο έλεγχος των σχέσεων αυτών εξετάζει αν τα επιλεγμένα false alarm events συνδέονται σωστά με τα αντίστοιχα risk-related individuals και με το Potential Harm, χωρίς να υπονοείται πραγματική βλάβη σε συγκεκριμένο περιστατικό.

Η λειτουργική επικύρωση πραγματοποιείται σε τρία επίπεδα. Το πρώτο επίπεδο αφορά τον έλεγχο λογικής συνέπειας μέσω reasoner. Σε αυτό το στάδιο εξετάζεται αν η οντολογία περιέχει λογικές αντιφάσεις, ασυνεπείς κλάσεις ή προβληματικές σχέσεις που θα μπορούσαν να επηρεάσουν τη λειτουργία της.

Το δεύτερο επίπεδο αφορά τη χρήση DL competency questions. Τα ερωτήματα αυτά χρησιμοποιούνται για να ελεγχθεί αν η οντολογία μπορεί να επιστρέψει τις αναμενόμενες κατηγορίες ή τα αναμενόμενα individuals με βάση τη δομή της. Για παράδειγμα, μπορεί να εξεταστεί ποια events έχουν human review, ποια είναι false alarms, ποια Evidence Records έχουν χαρτογραφηθεί προς συγκεκριμένα requirement individuals και ποια events έχουν ενταχθεί στο lightweight risk-related επίπεδο.

Το τρίτο επίπεδο αφορά τη χρήση SPARQL queries. Τα SPARQL queries επιτρέπουν πιο δομημένη ανάκτηση πληροφορίας από την οντολογία. Σε αντίθεση με τα απλούστερα DL queries, τα SPARQL queries μπορούν να συνδυάσουν περισσότερες σχέσεις και να επιστρέψουν συγκεντρωτικά αποτελέσματα, όπως event identifiers, decisions, Evidence Records και requirement mappings.

Η επιλογή αυτών των τριών μηχανισμών ελέγχου συνδέεται άμεσα με τον χαρακτήρα της εργασίας. Εφόσον η εργασία προτείνει μια οντολογική αναπαράσταση και όχι ένα προγνωστικό μοντέλο, η αξιολόγηση πρέπει να εξετάζει τη λογική συνέπεια, την ανακτησιμότητα της πληροφορίας και την ορθότητα των σημασιολογικών διαδρομών. Με αυτόν τον τρόπο, η επικύρωση είναι συμβατή με τον σκοπό της οντολογίας.

Ο έλεγχος που πραγματοποιείται δείχνει ότι η οντολογία μπορεί να οργανώσει και να τεκμηριώσει επιλεγμένες AI Act-related requirements σε επίπεδο proof-of-concept. Δεν αντικαθιστά νομική αξιολόγηση, πλήρη τεχνικό έλεγχο, οργανωτική αξιολόγηση ή επίσημη διαδικασία πιστοποίησης.

Συνολικά, η λειτουργική επικύρωση έχει διπλό ρόλο. Από τη μία πλευρά, ελέγχει ότι η οντολογία είναι τεχνικά συνεπής και μπορεί να χρησιμοποιηθεί χωρίς λογικές ασυνέπειες. Από την άλλη πλευρά, δείχνει ότι το μοντέλο μπορεί να απαντήσει σε ουσιαστικά ερωτήματα σχετικά με τα VTaC events, την τεκμηρίωση, την ανθρώπινη αξιολόγηση, την ιχνηλασιμότητα και τη σύνδεση με επιλεγμένες απαιτήσεις.

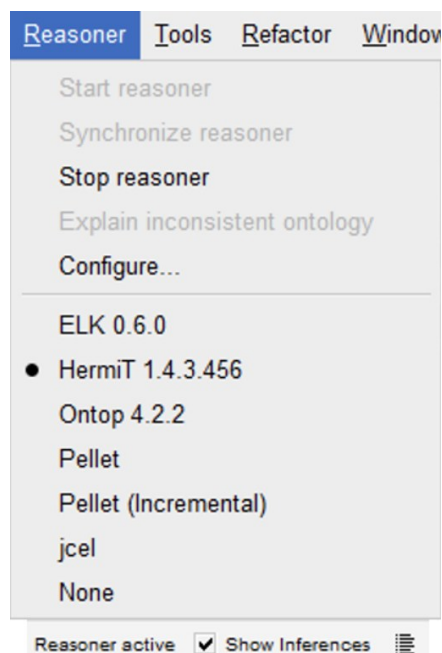
Η λειτουργική επικύρωση που πραγματοποιείται στην παρούσα εργασία πρέπει να ερμηνεύεται μέσα στα όρια του proof-of-concept. Ο έλεγχος με reasoner, DL competency questions και SPARQL queries εξετάζει τη λογική συνέπεια και τη δυνατότητα ανάκτησης των σημασιολογικών σχέσεων που έχουν υλοποιηθεί. Δεν αποτελεί πλήρη εξωτερική επικύρωση της οντολογίας, κλινική αξιολόγηση από επαγγελματίες υγείας, νομική αξιολόγηση των AI Act-related mappings ή επίσημη διαδικασία conformity assessment.

5.3 Έλεγχος συνέπειας με reasoner

Το πρώτο στάδιο της λειτουργικής επικύρωσης αφορά τον έλεγχο λογικής συνέπειας της οντολογίας μέσω reasoner. Στην παρούσα εργασία χρησιμοποιήθηκε ο HermiT reasoner, ένας OWL 2 reasoner που υποστηρίζει λογική ταξινόμηση και reasoning πάνω σε OWL ontologies [30].

Ο έλεγχος με reasoner έχει ως στόχο να εξεταστεί αν το σύνολο των axioms της οντολογίας μπορεί να ερμηνευθεί χωρίς λογική αντίφαση και αν υπάρχουν classes που καθίστανται μη ικανοποιήσιμες εξαιτίας του συνδυασμού των δηλωμένων περιορισμών.

Στην πράξη, ο HermiT ενεργοποιήθηκε μέσα από το Protégé και χρησιμοποιήθηκε για την ταξινόμηση της οντολογίας και την παραγωγή inferred γνώσης.



Σχήμα 5.1 Ενεργοποίηση του HermiT reasoner στο Protégé.

Η εκτέλεση του reasoner ολοκληρώθηκε χωρίς να εμφανιστεί inconsistency στην ontology και χωρίς να εντοπιστούν unsatisfiable classes στην εξεταζόμενη έκδοση του μοντέλου.

Το αποτέλεσμα αυτό σημαίνει ότι, με βάση τα axioms που έχουν δηλωθεί στην παρούσα υλοποίηση, ο reasoner δεν εντόπισε λογική αντίφαση που να καθιστά την ontology inconsistent. Δεν σημαίνει ότι κάθε σχεδιαστική επιλογή είναι κατ' ανάγκη βέλτιστη, ότι όλες οι πληροφορίες του domain έχουν μοντελοποιηθεί πλήρως ή ότι η ontology είναι νομικά ή κλινικά ορθή.

Ο Πίνακας 5.1 συνοψίζει τον έλεγχο λογικής συνέπειας που πραγματοποιήθηκε.

Πίνακας 5-1 Σύνοψη ελέγχου λογικής συνέπειας με reasoner.

Στοιχείο ελέγχου	Περιγραφή
Περιβάλλον υλοποίησης	Protégé
Reasoner	HermiT
Αντικείμενο ελέγχου	Λογική συνέπεια της οντολογίας
Εξεταζόμενα στοιχεία	Classes, object properties, data properties, individuals
Βασικός στόχος	Εντοπισμός πιθανών ασυνεπειών ή μη ικανοποιησιμων classes
Αποτέλεσμα	Η οντολογία ταξινομήθηκε χωρίς εμφανείς λογικές ασυνέπειες
Ρόλος στον συνολικό έλεγχο	Προκαταρκτική επιβεβαίωση πριν από DL queries και SPARQL queries

Η επιτυχής ολοκλήρωση του ελέγχου αποτελεί απαραίτητη βάση για τα επόμενα στάδια λειτουργικής επικύρωσης. Παρέχει ένδειξη ότι η λογική δομή του μοντέλου είναι συνεπής με τα δηλωμένα axioms και επιτρέπει να συνεχιστεί ο έλεγχος της inferred γνώσης και των competency questions.

5.4 Έλεγχος inferred γνώσης στην ιεραρχία κλάσεων

Μετά την ενεργοποίηση του HermiT reasoner εξετάστηκε η inferred ιεραρχία των classes. Η inferred προβολή επιτρέπει να ελεγχθεί πώς ταξινομούνται οι classes όταν ληφθούν υπόψη όχι μόνο οι άμεσες δηλώσεις της οντολογίας αλλά και οι λογικές συνέπειες που μπορούν να εξαχθούν από τα axioms της [30].

Ο έλεγχος επικεντρώθηκε στις βασικές εννοιολογικές ομάδες της οντολογίας. Συγκεκριμένα, εξετάστηκε αν οι classes που σχετίζονται με την τεκμηρίωση και την αξιολόγηση παραμένουν κάτω από το Evidence and Assessment Concept, αν οι απαιτήσεις οργανώνονται κάτω από το Normative Concept, αν οι λειτουργικές έννοιες των VTaC events παραμένουν κάτω από το Operational Concept και αν οι risk-related classes οργανώνονται κάτω από το Risk Concept.

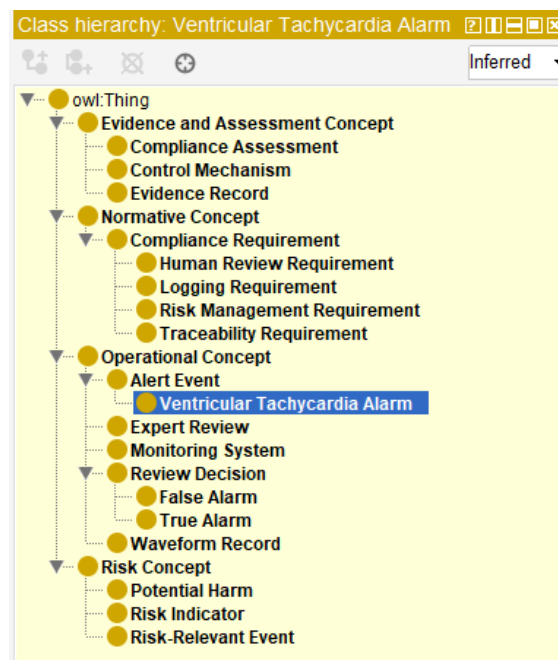
Στο operational επίπεδο, ελέγχθηκε ότι η class Ventricular Tachycardia Alarm διατηρείται ως ειδικότερη μορφή της class Alert Event. Η ταξινόμηση αυτή είναι σημαντική, επειδή τα 12 πραγματικά VTaC alarm events της υλοποίησης ανήκουν σε αυτή την κατηγορία και αποτελούν το βασικό σημείο εκκίνησης της σημασιολογικής αλυσίδας.

Επίσης, εξετάστηκε η θέση των classes Expert Review, Review Decision, False Alarm και True Alarm. Η διάκριση ανάμεσα σε False Alarm και True Alarm κάτω από τη γενικότερη class Review Decision είναι σημαντική, επειδή επιτρέπει στην οντολογία να αναπαριστά το αποτέλεσμα της expert

αξιολόγησης χωρίς να το αντιμετωπίζει ως απλή λεκτική τιμή. Με αυτόν τον τρόπο, η απόφαση μπορεί να χρησιμοποιηθεί ως κανονικό οντολογικό στοιχείο και να συμμετέχει σε queries.

Στο normative επίπεδο, ελέγχθηκε ότι οι classes Logging Requirement, Human Review Requirement, Traceability Requirement και Risk Management Requirement παραμένουν κάτω από τη γενικότερη class Compliance Requirement. Η οργάνωση αυτή είναι απαραίτητη, επειδή οι συγκεκριμένες requirements αποτελούν τα επιλεγμένα AI Act-related σημεία αναφοράς της εργασίας.

Στο risk-related επίπεδο, εξετάστηκε ότι οι classes Risk Indicator, Risk-Relevant Event και Potential Harm διατηρούνται ως μέρος της risk-related διάστασης της οντολογίας. Η ορθή θέση τους στην ιεραρχία είναι σημαντική, επειδή επιτρέπει τη διάκριση ανάμεσα στη βασική αλυσίδα τεκμηρίωσης και στην πρόσθετη risk-related αλυσίδα που εφαρμόζεται στα false alarm cases.



Σχήμα 5.2 Προβολή inferred ιεραρχίας των classes μετά την εκτέλεση του HermiT reasoner.

Το αποτέλεσμα του ελέγχου έδειξε ότι η inferred ιεραρχία των classes είναι συμβατή με τη σχεδίαση της οντολογίας. Δεν εμφανίστηκαν classes ως μη ικανοποιήσιμες και δεν παρατηρήθηκαν μη αναμενόμενες ταξινομήσεις που να αλλοιώνουν τη βασική δομή του μοντέλου.

Η επιβεβαίωση αυτή είναι σημαντική, επειδή δείχνει ότι ο reasoner δεν εντοπίζει μόνο απουσία αντιφάσεων, αλλά ταξινομεί τη δομή της οντολογίας με τρόπο που συμφωνεί με τη μεθοδολογική της λογική. Η οντολογία διατηρεί τις τέσσερις βασικές εννοιολογικές περιοχές της: operational, evidence and assessment, normative και lightweight risk layer.

Ο έλεγχος inferred γνώσης δεν αποδεικνύει από μόνος του ότι όλα τα competency questions θα επιστρέψουν τα αναμενόμενα αποτελέσματα. Αποτελεί όμως αναγκαίο ενδιάμεσο βήμα ανάμεσα στον έλεγχο συνέπειας και στα queries. Αφού επιβεβαιωθεί ότι η inferred ιεραρχία είναι σωστή, μπορεί να ακολουθήσει ο έλεγχος της ανάκτησης πληροφορίας μέσω DL competency questions.

5.5 Έλεγχος της οντολογίας μέσω DL competency questions

Μετά τον έλεγχο της οντολογίας με reasoner και την εξέταση της inferred ιεραρχίας των classes, το επόμενο στάδιο λειτουργικής επικύρωσης πραγματοποιήθηκε μέσω DL competency questions. Τα competency questions αποτελούν ερωτήματα ελέγχου, τα οποία χρησιμοποιούνται για να εξεταστεί αν η οντολογία μπορεί να απαντήσει σε βασικές ερωτήσεις που σχετίζονται με τον σκοπό της.

Στην παρούσα εργασία, τα DL competency questions χρησιμοποιήθηκαν για να ελεγχθεί αν η οντολογία μπορεί να ανακτήσει σωστά τις βασικές συνδεδεμένες σημασιολογικές διαδρομές που οργανώνονται γύρω από το ίδιο Alert Event:

Alert Event → belongsToRecord → Waveform Record

Alert Event → hasReview → Expert Review → hasDecision → Review Decision

Alert Event → isDocumentedBy → Evidence Record → supportsRequirement → Compliance Requirement

Επιπλέον, για τα false alarm cases, τα competency questions ελέγχουν και την πρόσθετη lightweight risk-related αναπαράσταση μέσω δύο παράλληλων διαδρομών:

Alert Event → hasRiskIndicator → Risk Indicator

Alert Event → isClassifiedAs → Risk-Relevant Event → mayLeadTo → Potential Harm

Με αυτόν τον τρόπο, τα competency questions ελέγχουν τη συνύπαρξη και την ανακτησιμότητα των βασικών σχέσεων της οντολογίας, χωρίς να θεωρείται ότι αυτές αποτελούν μία μοναδική γραμμική object-property αλυσίδα.

Η χρήση των DL queries στο Protégé είναι σημαντική επειδή επιτρέπει τον άμεσο έλεγχο της οντολογίας με βάση τη λογική της δομή. Με άλλα λόγια, τα ερωτήματα δεν εξετάζουν μόνο αν υπάρχουν συγκεκριμένα individuals, αλλά αν αυτά μπορούν να ανακτηθούν με βάση τις classes και τις σχέσεις που έχουν οριστεί.

Στην παρούσα υλοποίηση, τα competency questions σχεδιάστηκαν ώστε να καλύπτουν τις βασικές λειτουργικές απαιτήσεις της οντολογίας. Συγκεκριμένα, εξετάζουν αν μπορούν να εντοπιστούν τα alert events, τα events που έχουν αξιολογηθεί ως true ή false alarms, τα events που διαθέτουν evidence για logging, human review και traceability, τα evidence records που υποστηρίζουν απαιτήσεις, καθώς και τα false alarm cases που συνδέονται με το lightweight risk layer.

Η επιτυχής εκτέλεση των DL queries δείχνει ότι τα βασικά στοιχεία της οντολογίας έχουν συνδεθεί σωστά. Αν ένα query επιστρέφει τα αναμενόμενα individuals, τότε επιβεβαιώνεται ότι οι αντίστοιχες classes, σχέσεις και property assertions έχουν υλοποιηθεί με λειτουργικό τρόπο.

Στο Protégé, κάθε DL query εκτελείται από την καρτέλα DL Query, με ενεργό reasoner.

Πίνακας 5-2 DL competency questions της οντολογίας.

DL Query	Ερώτημα ελέγχου	Αναμενόμενο αποτέλεσμα
DL Query 1	Ποια individuals έχουν εισαχθεί στην οντολογία ως VTaC alert events τύπου ventricular tachycardia alarm;	12 VTaC alert event individuals
DL Query 2	Ποια VT alarm events έχουν αξιολογηθεί ως true alarms;	6 true alarm events

DL Query 3	Ποια VT alarm events έχουν αξιολογηθεί ως false alarms;	6 false alarm events
DL Query 4	Ποια VT alarm events διαθέτουν evidence για logging;	12 VT alarm events
DL Query 5	Ποια VT alarm events διαθέτουν evidence για human review;	12 VT alarm events
DL Query 6	Ποια VT alarm events διαθέτουν evidence για traceability;	12 VT alarm events
DL Query 7	Ποια evidence records υποστηρίζουν ταυτόχρονα τρία βασικά AI Act-related σημεία αναφοράς;	12 evidence records
DL Query 8	Ποια false alarm events έχουν ενταχθεί στο risk layer;	6 false alarm events
DL Query 9	Ποια risk-relevant events συνδέονται με potential harm;	6 risk-relevant event individuals
DL Query 10	Ποια Evidence Records έχουν mapping προς το req_risk_management;	6 evidence records
DL Query 11	Ποια events διαθέτουν πλήρη αλυσίδα traceability;	12 VT alarm events
DL Query 12	Ποια waveform records συνδέονται με VT alarm events;	6 waveform records

Η παραπάνω σειρά δεν είναι τυχαία. Τα πρώτα DL queries ελέγχουν το operational επίπεδο της οντολογίας, δηλαδή αν τα πραγματικά VTaC alarm events έχουν εισαχθεί σωστά και αν μπορούν να διακριθούν σε true και false alarm cases μέσω της διαδρομής Alert Event → Expert Review → Review Decision.

Στη συνέχεια, τα DL queries ελέγχουν το evidence layer. Συγκεκριμένα, εξετάζουν αν τα VT alarm events διαθέτουν evidence για logging, human review και traceability, καθώς και αν τα evidence records υποστηρίζουν ταυτόχρονα τα τρία βασικά AI Act-related σημεία αναφοράς. Με αυτόν τον τρόπο επιβεβαιώνεται ότι το μοντέλο δεν συνδέει απευθείας τα events με τις απαιτήσεις, αλλά χρησιμοποιεί τα evidence records ως ενδιάμεσο επίπεδο τεκμηρίωσης.

Ακολουθούν τα queries που αφορούν το lightweight risk layer. Τα ερωτήματα αυτά εξετάζουν αν τα false alarm events έχουν συνδεθεί με risk indicators και risk-relevant events, αν τα risk-relevant events συνδέονται με potential harm και αν τα αντίστοιχα evidence records υποστηρίζουν την απαίτηση req_risk_management.

Το req_risk_management αποτελεί περιορισμένο AI Act-related σημείο αναφοράς της παρούσας proof-of-concept υλοποίησης και το συγκεκριμένο mapping δεν ισοδυναμεί με πλήρη υλοποίηση ή αξιολόγηση του risk management system του άρθρου 9.

Τέλος, τα δύο τελευταία queries ελέγχουν πιο συνολικά τη λειτουργία της οντολογίας. Το DL Query 11 εξετάζει αν τα VT alarm events διαθέτουν πλήρη αλυσίδα traceability, ενώ το DL Query 12 επιβεβαιώνει ότι τα waveform records συνδέονται με τα αντίστοιχα VT alarm events.

Με αυτόν τον τρόπο, τα DL competency questions καλύπτουν σταδιακά την πλήρη λογική της οντολογίας: από την ύπαρξη των πραγματικών VTaC events, στη διάκριση true και false alarm cases, στη σύνδεση με evidence records, στη χαρτογράφηση προς απαιτήσεις και τέλος στην πρόσθετη risk-related αναπαράσταση των false alarm cases.

5.5.1 DL Query 1 – Εντοπισμός των VTaC alert events

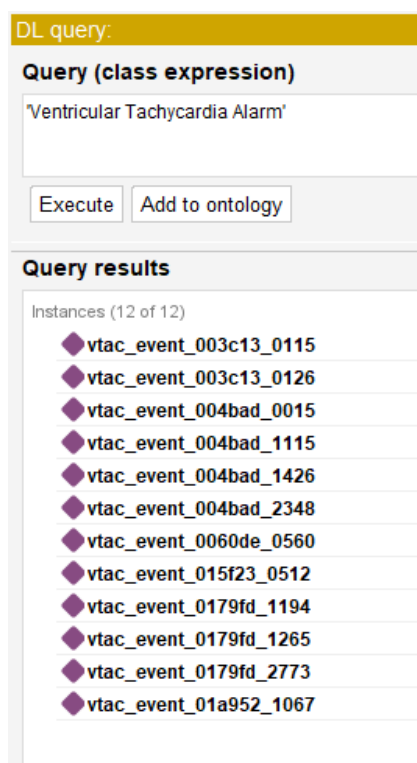
Το πρώτο DL query χρησιμοποιήθηκε για να ελεγχθεί αν τα πραγματικά VTaC alarm events έχουν εισαχθεί σωστά στην οντολογία ως individuals της class `Ventricular Tachycardia Alarm`. Η class αυτή αποτελεί εξειδίκευση της class `Alert Event` και αναπαριστά τα alarm events κοιλιακής ταχυκαρδίας που χρησιμοποιούνται στο proof-of-concept μοντέλο.

Το ερώτημα που εξετάζεται είναι το εξής:

Ποια individuals έχουν εισαχθεί στην οντολογία ως VTaC alert events τύπου ventricular tachycardia alarm;

Ο αντίστοιχος κώδικας DL Query που εκτελέστηκε στο Protégé είναι:

'Ventricular Tachycardia Alarm'



DL query:

Query (class expression)

'Ventricular Tachycardia Alarm'

Execute Add to ontology

Query results

Instances (12 of 12)

- vtac_event_003c13_0115
- vtac_event_003c13_0126
- vtac_event_004bad_0015
- vtac_event_004bad_1115
- vtac_event_004bad_1426
- vtac_event_004bad_2348
- vtac_event_0060de_0560
- vtac_event_015f23_0512
- vtac_event_0179fd_1194
- vtac_event_0179fd_1265
- vtac_event_0179fd_2773
- vtac_event_01a952_1067

Σχήμα 5.3 DL Query 1 — Εντοπισμός των VTaC alert event individuals.

Όπως φαίνεται στο Σχήμα 5.3, το query επιστρέφει 12 instances. Τα αποτελέσματα αντιστοιχούν στα 12 VTaC alert event individuals που έχουν εισαχθεί στην οντολογία.

Το αποτέλεσμα αυτό επιβεβαιώνει ότι τα 12 επιλεγμένα VTaC event individuals έχουν ανακτηθεί ως instances της class `Ventricular Tachycardia Alarm`, σύμφωνα με τον σχεδιασμό της οντολογίας. Επομένως, το βασικό operational επίπεδο της οντολογίας είναι διαθέσιμο για τους επόμενους ελέγχους που αφορούν τη σύνδεση των events με expert reviews, review decisions, evidence records, traceability και risk-related mappings.

Η επιτυχής εκτέλεση του πρώτου query αποτελεί το αρχικό σημείο της λειτουργικής επικύρωσης. Αν τα VTaC alert events δεν ανακτώνταν σωστά σε αυτό το στάδιο, δεν θα ήταν δυνατή η αξιόπιστη εκτέλεση των επόμενων DL queries, τα οποία βασίζονται στις σχέσεις αυτών των events με τα υπόλοιπα στοιχεία της οντολογίας.

5.5.2 DL Query 2 – Ποια VT alarm events έχουν αξιολογηθεί ως true alarms;

Το δεύτερο DL query χρησιμοποιήθηκε για να ελεγχθεί ποια από τα VTaC alarm events της οντολογίας έχουν αξιολογηθεί ως true alarms. Σε αντίθεση με το πρώτο query, το οποίο εντόπιζε απλώς τα individuals της class `Ventricular Tachycardia Alarm`, το παρόν query ελέγχει μια πιο σύνθετη σημασιολογική διαδρομή.

Το ερώτημα που εξετάζεται είναι το εξής:

Ποια από τα events της class `Ventricular Tachycardia Alarm` συνδέονται, μέσω της διαδικασίας review, με τελική απόφαση `True Alarm`;

Το query δεν ελέγχει μόνο αν ένα event υπάρχει στην οντολογία. Ελέγχει αν το event έχει συνδεθεί με ένα expert review και αν το review αυτό οδηγεί σε review decision τύπου `True Alarm` [4]. Με αυτόν τον τρόπο επιβεβαιώνεται ότι το event δεν χαρακτηρίζεται απευθείας ως true alarm, αλλά εντάσσεται στη σημασιολογική αλυσίδα:

Alert Event → Expert Review → Review Decision

Ο αντίστοιχος κώδικας DL Query που εκτελέστηκε στο Protégé είναι:

```
'Ventricular Tachycardia Alarm' and ('has review' some ('has decision' some 'True Alarm'))
```

The screenshot shows the DL Query interface in Protégé. At the top, a yellow bar indicates the query type. Below it, the query is displayed in a text box. There are two buttons: 'Execute' and 'Add to ontology'. Below the buttons, the 'Query results' section shows a list of 6 instances.

DL query:

Query (class expression)

`'Ventricular Tachycardia Alarm' and ('has review' some ('has decision' some 'True Alarm'))`

Execute **Add to ontology**

Query results

Instances (6 of 6)

- vtac_event_0060de_0560
- vtac_event_015f23_0512
- vtac_event_0179fd_1194
- vtac_event_0179fd_1265
- vtac_event_0179fd_2773
- vtac_event_01a952_1067

Σχήμα 5.4 DL Query 2 — VT alarm events που έχουν αξιολογηθεί ως true alarms.

Όπως φαίνεται στο Σχήμα 5.4, το query επιστρέφει τα VT alarm events που έχουν τελική απόφαση `True Alarm`. Συγκεκριμένα, επιστρέφονται 6 instances, τα οποία αντιστοιχούν στα true alarm cases της οντολογίας.

Το αποτέλεσμα αυτό επιβεβαιώνει ότι τα true alarm events έχουν συνδεθεί σωστά με τη διαδικασία expert review και με την αντίστοιχη review decision. Επομένως, η οντολογία μπορεί να ανακτήσει τα true alarm cases όχι απλώς επειδή υπάρχουν ως ονόματα individuals, αλλά επειδή συμμετέχουν στη σωστή σημασιολογική διαδρομή.

5.5.3 DL Query 3 – Ποια VT alarm events έχουν αξιολογηθεί ως false alarms;

Το τρίτο DL query χρησιμοποιήθηκε για να ελεγχθεί ποια από τα VTaC alarm events της οντολογίας έχουν αξιολογηθεί ως false alarms. Το query αυτό είναι αντίστοιχο με το προηγούμενο, αλλά αντί για τα true alarm cases αναζητά τα events που οδηγούν σε review decision τύπου False Alarm.

Το ερώτημα που εξετάζεται είναι το εξής:

Ποια από τα events της class Ventricular Tachycardia Alarm συνδέονται, μέσω της διαδικασίας review, με τελική απόφαση False Alarm;

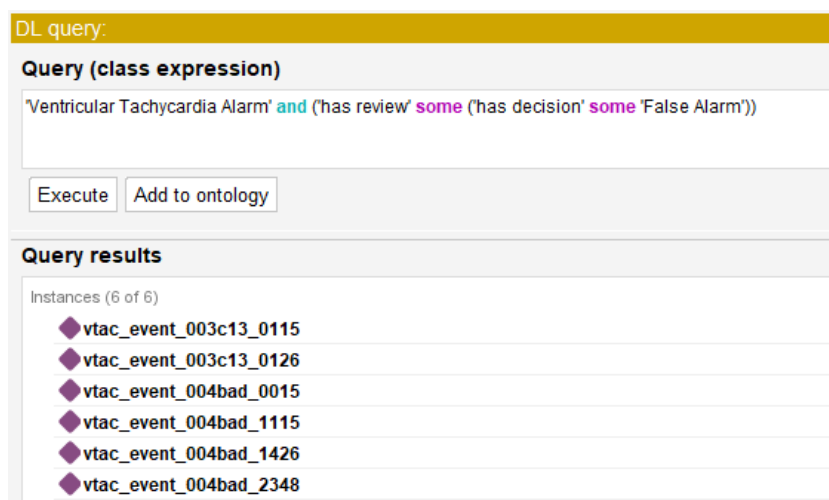
Το query ελέγχει τη σημασιολογική διαδρομή:

Alert Event → Expert Review → Review Decision

Στην περίπτωση αυτή, η review decision πρέπει να ανήκει στην class False Alarm. Με αυτόν τον τρόπο, το false alarm δεν αντιμετωπίζεται απλώς ως λεκτική τιμή, αλλά ως οντολογικό αποτέλεσμα της διαδικασίας αξιολόγησης.

Ο αντίστοιχος κώδικας DL Query που εκτελέστηκε στο Protégé είναι:

```
'Ventricular Tachycardia Alarm' and ('has review' some ('has decision' some 'False Alarm'))
```



Σχήμα 5.5 DL Query 3 — VT alarm events που έχουν αξιολογηθεί ως false alarms.

Όπως φαίνεται στο Σχήμα 5.5, το query επιστρέφει τα VT alarm events που έχουν τελική απόφαση False Alarm. Συγκεκριμένα, επιστρέφονται 6 instances, τα οποία αντιστοιχούν στα false alarm cases της οντολογίας.

Το αποτέλεσμα επιβεβαιώνει ότι τα false alarm events έχουν συνδεθεί σωστά με expert review και με αντίστοιχη false alarm decision. Η ανάκτηση αυτών των events είναι σημαντική, επειδή τα false alarm cases αποτελούν τη βάση για το lightweight risk layer που έχει υλοποιηθεί στην οντολογία.

5.5.4 DL Query 4 – Ποια VT alarm events διαθέτουν evidence για logging;

Το τέταρτο DL query χρησιμοποιήθηκε για να εξεταστεί ποια από τα VTaC alarm events της οντολογίας συνδέονται με Evidence Record το οποίο έχει χαρτογραφηθεί προς το AI Act-related σημείο αναφοράς req_logging.

Το ερώτημα που εξετάζεται είναι:

Ποια events της class Ventricular Tachycardia Alarm συνδέονται με Evidence Record που υποστηρίζει το requirement individual req_logging;

Στην οντολογία, το event δεν συνδέεται απευθείας με το requirement individual. Η διαδρομή που ελέγχεται είναι:

Alert Event --isDocumentedBy→ Evidence Record --supportsRequirement→ req_logging

Ο αντίστοιχος DL Query που εκτελέστηκε στο Protégé είναι:

```
'Ventricular Tachycardia Alarm' and ('is documented by' some ('supports requirement' value req_logging))
```

The screenshot shows the DL Query interface in Protégé. At the top, there is a yellow bar labeled 'DL query:'. Below it, a text box contains the query: `'Ventricular Tachycardia Alarm' and ('is documented by' some ('supports requirement' value req_logging))`. Below the text box are two buttons: 'Execute' and 'Add to ontology'. Below the buttons is a section titled 'Query results'. Under this section, it says 'Instances (12 of 12)'. Below this, there is a list of 12 instances, each preceded by a purple diamond icon. The instances are: vtac_event_003c13_0115, vtac_event_003c13_0126, vtac_event_004bad_0015, vtac_event_004bad_1115, vtac_event_004bad_1426, vtac_event_004bad_2348, vtac_event_0060de_0560, vtac_event_015f23_0512, vtac_event_0179fd_1194, vtac_event_0179fd_1265, vtac_event_0179fd_2773, and vtac_event_01a952_1067.

Σχήμα 5.6 DL Query 4 — VT alarm events με evidence για logging.

Το query επιστρέφει 12 instances, δηλαδή όλα τα επιλεγμένα VT alarm events της proof-of-concept υλοποίησης.

Το αποτέλεσμα δείχνει ότι για καθένα από τα 12 events υπάρχει Evidence Record το οποίο έχει συνδεθεί μέσω supportsRequirement με το req_logging. Με αυτόν τον τρόπο επιβεβαιώνεται λειτουργικά ότι το συγκεκριμένο requirement-to-evidence mapping έχει υλοποιηθεί για όλα τα επιλεγμένα events.

Η συγκεκριμένη ανάκτηση δεν αποδεικνύει ότι ένα πραγματικό AI system διαθέτει πλήρη automatic logging capability σύμφωνα με το άρθρο 12 του AI Act. Επιβεβαιώνει μόνο ότι, μέσα στο proof-of-concept μοντέλο, τα 12 events διαθέτουν evidence records που έχουν χαρτογραφηθεί προς το req_logging.

5.5.5 DL Query 5 – Ποια VT alarm events διαθέτουν evidence για human review;

Το πέμπτο DL query χρησιμοποιήθηκε για να εξεταστεί ποια από τα VTaC alarm events της οντολογίας συνδέονται με Evidence Record που έχει χαρτογραφηθεί προς το requirement individual req_human_review.

Το ερώτημα που εξετάζεται είναι:

Ποια events της class Ventricular Tachycardia Alarm συνδέονται με Evidence Record που υποστηρίζει το req_human_review;

Η διαδρομή που ελέγχεται είναι:

Alert Event --isDocumentedBy→ Evidence Record --supportsRequirement→ req_human_review

Ο αντίστοιχος DL Query είναι:

'Ventricular Tachycardia Alarm' and ('is documented by' some ('supports requirement' value req_human_review))

The screenshot shows a web-based interface for a DL query. At the top, there is a yellow header labeled 'DL query:'. Below it, a section titled 'Query (class expression)' contains the query: 'Ventricular Tachycardia Alarm' and ('is documented by' some ('supports requirement' value req_human_review)). Below the query text are two buttons: 'Execute' and 'Add to ontology'. Underneath, a section titled 'Query results' shows 'Instances (12 of 12)'. A list of 12 instances is displayed, each preceded by a purple diamond icon. The instances are: vtac_event_003c13_0115, vtac_event_003c13_0126, vtac_event_004bad_0015, vtac_event_004bad_1115, vtac_event_004bad_1426, vtac_event_004bad_2348, vtac_event_0060de_0560, vtac_event_015f23_0512, vtac_event_0179fd_1194, vtac_event_0179fd_1265, vtac_event_0179fd_2773, and vtac_event_01a952_1067.

Σχήμα 5.7 DL Query 5 — VT alarm events με evidence για human review.

Το query επιστρέφει 12 instances, δηλαδή όλα τα επιλεγμένα VT alarm events.

Το αποτέλεσμα δείχνει ότι για όλα τα events υπάρχει Evidence Record που έχει χαρτογραφηθεί προς το req_human_review. Η πληροφορία αυτή λειτουργεί συμπληρωματικά προς τη διαφορετική διαδρομή:

Alert Event --hasReview→ Expert Review --hasDecision→ Review Decision

μέσω της οποίας αναπαρίσταται η πραγματική expert αξιολόγηση των VTaC events.

Η επιτυχής ανάκτηση δεν σημαίνει ότι έχει αποδειχθεί πλήρης human oversight σύμφωνα με το άρθρο 14 του AI Act. Το req_human_review αποτελεί εφαρμοσμένο AI Act-related σημείο αναφοράς του proof-of-concept, ενώ η Expert Review αναπαριστά την ανθρώπινη αξιολόγηση που είναι διαθέσιμη για τα VTaC events [4].

5.5.6 DL Query 6 – Ποια VT alarm events διαθέτουν evidence για traceability;

Το έκτο DL query χρησιμοποιήθηκε για να εξεταστεί ποια από τα VTaC alarm events της οντολογίας συνδέονται με Evidence Record που έχει χαρτογραφηθεί προς το requirement individual req_traceability.

Το ερώτημα που εξετάζεται είναι:

Ποια events της class Ventricular Tachycardia Alarm συνδέονται με Evidence Record που υποστηρίζει το req_traceability;

Η διαδρομή που ελέγχεται είναι:

Alert Event --isDocumentedBy→ Evidence Record --supportsRequirement→ req_traceability

Ο αντίστοιχος DL Query είναι:

```
'Ventricular Tachycardia Alarm' and ('is documented by' some ('supports requirement' value req_traceability))
```

The screenshot shows a web interface for a DL query. At the top, a yellow bar contains the text "DL query:". Below this, a section titled "Query (class expression)" contains the query: "'Ventricular Tachycardia Alarm' and ('is documented by' some ('supports requirement' value req_traceability))". There are two buttons: "Execute" and "Add to ontology". Below the query section, a section titled "Query results" shows "Instances (12 of 12)". A list of 12 instances is displayed, each with a purple diamond icon and a text label: vtac_event_003c13_0115, vtac_event_003c13_0126, vtac_event_004bad_0015, vtac_event_004bad_1115, vtac_event_004bad_1426, vtac_event_004bad_2348, vtac_event_0060de_0560, vtac_event_015f23_0512, vtac_event_0179fd_1194, vtac_event_0179fd_1265, vtac_event_0179fd_2773, and vtac_event_01a952_1067.

Σχήμα 5.8 DL Query 6 — VT alarm events με evidence για traceability.

Το query επιστρέφει 12 instances, δηλαδή όλα τα επιλεγμένα VT alarm events της οντολογίας.

Το αποτέλεσμα δείχνει ότι κάθε ένα από τα 12 events συνδέεται με Evidence Record που έχει χαρτογραφηθεί προς το req_traceability. Επιβεβαιώνεται επομένως ότι το συγκεκριμένο requirement-to-evidence mapping έχει υλοποιηθεί για όλα τα επιλεγμένα events.

Το συγκεκριμένο query ελέγχει μόνο τη διαδρομή Event → Evidence Record → req_traceability. Δεν ελέγχει από μόνο του ολόκληρη τη διαδρομή ιχνηλασιμότητας που περιλαμβάνει record, review και decision. Ο πληρέστερος αυτός έλεγχος πραγματοποιείται στο DL Query 11.

Κατά συνέπεια, το αποτέλεσμα δεν πρέπει να ερμηνεύεται ως απόδειξη πλήρους traceability ενός πραγματικού AI system ή πλήρους συμμόρφωσης με το άρθρο 12 του AI Act. Αποδεικνύει λειτουργικά ότι το mapping προς το req_traceability μπορεί να ανακτηθεί για όλα τα 12 proof-of-concept events.

5.5.7 DL Query 7 – Ποια Evidence Records υποστηρίζουν ταυτόχρονα τα τρία βασικά requirement mappings;

Το έβδομο DL query χρησιμοποιήθηκε για να εξεταστεί ποια Evidence Record individuals έχουν χαρτογραφηθεί ταυτόχρονα προς τα τρία βασικά AI Act-related σημεία αναφοράς req_logging, req_human_review και req_traceability.

Σε αντίθεση με τα προηγούμενα queries, τα οποία αναζητούν VTaC alarm events, το συγκεκριμένο query εστιάζει στα individuals της class Evidence Record.

Το ερώτημα που εξετάζεται είναι:

Ποια Evidence Record individuals υποστηρίζουν ταυτόχρονα τα req_logging, req_human_review και req_traceability;

Η σημασιολογική λογική είναι:

Evidence Record --supportsRequirement→ Requirement individual

με ταυτόχρονη παρουσία των τριών συγκεκριμένων requirement mappings.

Ο DL Query που εκτελέστηκε στο Protégé είναι:

```
'Evidence Record'
and ('supports requirement' value req_logging)
and ('supports requirement' value req_human_review)
and ('supports requirement' value req_traceability)
```

The screenshot shows the DL Query interface in Protégé. At the top, there is a yellow bar labeled 'DL query:'. Below it, the query is displayed in a text area: 'Evidence Record' followed by three 'and' conditions: ('supports requirement' value req_logging), ('supports requirement' value req_human_review), and ('supports requirement' value req_traceability). Below the query area are two buttons: 'Execute' and 'Add to ontology'. Below the buttons is a section titled 'Query results' which shows a list of 12 instances of the 'Evidence Record' class. Each instance is preceded by a purple diamond icon. The instances are: vtac_evidence_003c13_0115, vtac_evidence_003c13_0126, vtac_evidence_004bad_0015, vtac_evidence_004bad_1115, vtac_evidence_004bad_1426, vtac_evidence_004bad_2348, vtac_evidence_0060de_0560, vtac_evidence_015f23_0512, vtac_evidence_0179fd_1194, vtac_evidence_0179fd_1265, vtac_evidence_0179fd_2773, and vtac_evidence_01a952_1067.

Σχήμα 5.9 DL Query 7 — Evidence Records που υποστηρίζουν τα τρία βασικά requirement mappings.

Το αποτέλεσμα δείχνει ότι και τα 12 evidence records της proof-of-concept υλοποίησης έχουν συνδεθεί ταυτόχρονα με τα req_logging, req_human_review και req_traceability. Με αυτόν τον τρόπο επιβεβαιώνεται λειτουργικά ότι το βασικό requirement-to-evidence mapping έχει εφαρμοστεί με συνεπή τρόπο στο σύνολο των επιλεγμένων events.

Η ανάκτηση αυτή αφορά τη δομή που έχει μοντελοποιηθεί στην οντολογία. Δεν σημαίνει ότι τα τρία αντίστοιχα σημεία του AI Act έχουν νομικά ικανοποιηθεί σε ένα πραγματικό AI system.

5.5.8 DL Query 8 – Ποια false alarm events έχουν ενταχθεί στο risk layer;

Το όγδοο DL query χρησιμοποιήθηκε για να εξεταστεί ποια από τα VTaC alarm events έχουν ταυτόχρονα αναπαρασταθεί ως false alarm cases και έχουν ενταχθεί στο lightweight risk-related μέρος της οντολογίας.

Το query δεν ελέγχει μόνο την ύπαρξη τελικής Review Decision τύπου False Alarm. Εξετάζει επιπλέον αν το ίδιο Alert Event έχει συνδεθεί με Risk Indicator και Risk-Relevant Event.

Οι σημασιολογικές διαδρομές που εξετάζονται είναι:

Alert Event --hasReview→ Expert Review --hasDecision→ False Alarm

και παράλληλα:

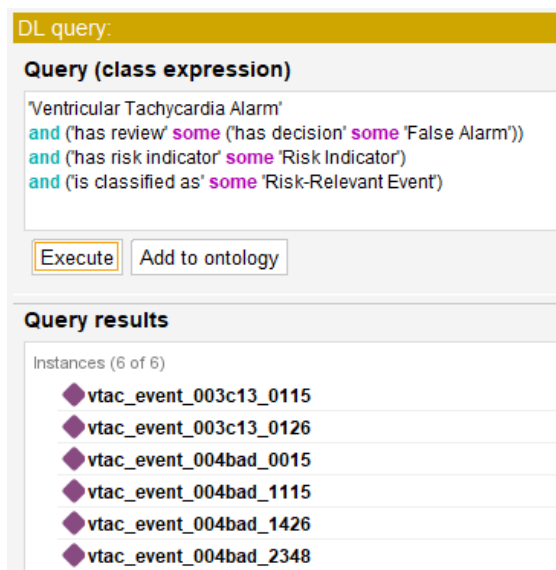
Alert Event --hasRiskIndicator→ Risk Indicator

Alert Event --isClassifiedAs→ Risk-Relevant Event

Με αυτόν τον τρόπο εξετάζεται αν τα false alarm cases δεν καταγράφονται μόνο ως αποτελέσματα αξιολόγησης, αλλά συμμετέχουν και στο πρόσθετο risk-related επίπεδο της οντολογίας.

Ο αντίστοιχος κώδικας DL Query που εκτελέστηκε στο Protégé είναι:

```
'Ventricular Tachycardia Alarm'
and ('has review' some ('has decision' some 'False Alarm'))
and ('has risk indicator' some 'Risk Indicator')
and ('is classified as' some 'Risk-Relevant Event')
```



Σχήμα 5.10 DL Query 8 — False alarm events στο risk layer.

Όπως φαίνεται στο Σχήμα 5.10, το query επιστρέφει 6 instances, δηλαδή τα false alarm events που έχουν ενταχθεί στο lightweight risk layer.

Το αποτέλεσμα δείχνει ότι όλα τα false alarm cases της οντολογίας έχουν συνδεθεί με risk-related στοιχεία. Επομένως, το lightweight risk layer έχει εφαρμοστεί στα events που έχουν αξιολογηθεί ως false alarms, χωρίς όμως αυτό να σημαίνει ότι τεκμηριώνεται πραγματική βλάβη ή πλήρης διαδικασία risk management.

5.5.9 DL Query 9 – Ποια risk-relevant events συνδέονται με potential harm;

Το ένατο DL query χρησιμοποιήθηκε για να εξεταστεί αν τα Risk-Relevant Event individuals της οντολογίας συνδέονται με Potential Harm.

Το ερώτημα που εξετάζεται είναι:

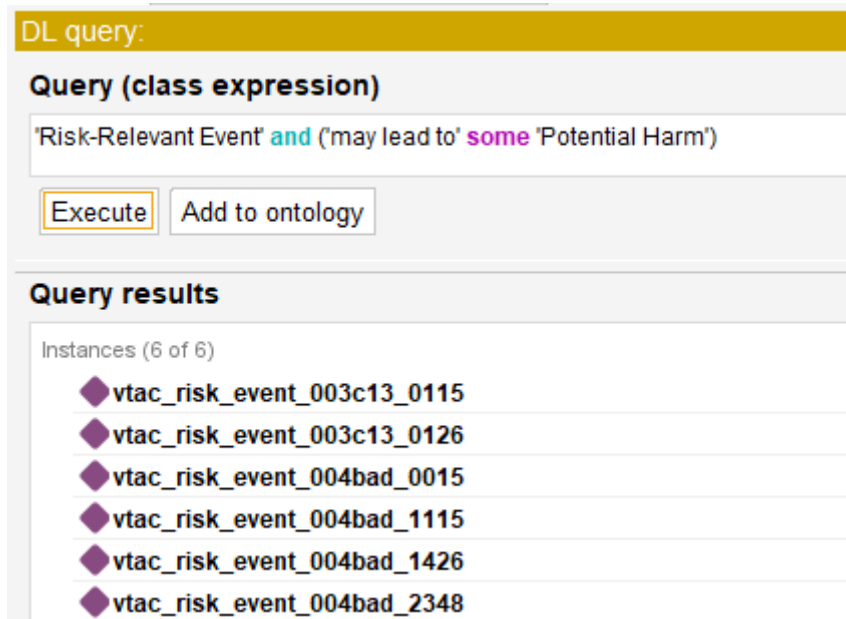
Ποια individuals της class Risk-Relevant Event συνδέονται με κάποιο Potential Harm;

Η σημασιολογική διαδρομή είναι:

Risk-Relevant Event --mayLeadTo→ Potential Harm

Ο DL Query που εκτελέστηκε στο Protégé είναι:

```
'Risk-Relevant Event' and ('may lead to' some 'Potential Harm')
```



Σχήμα 5.11 DL Query 9 — Risk-relevant events με potential harm.

Το query επιστρέφει 6 Risk-Relevant Event individuals. Τα συγκεκριμένα individuals αντιστοιχούν στα έξι false alarm cases που έχουν ενταχθεί στο lightweight risk-related επίπεδο.

Στην παρούσα υλοποίηση, τα Risk-Relevant Event individuals συνδέονται με το Potential Harm individual vtac_harm_alert_fatigue.

Το αποτέλεσμα επιβεβαιώνει ότι η σχέση `mayLeadTo` μπορεί να ανακτηθεί λειτουργικά για όλα τα `risk-relevant cases` που έχουν δημιουργηθεί. Η σχέση αυτή εκφράζει πιθανή σύνδεση με το `alert fatigue` και όχι ότι το συγκεκριμένο `harm` έχει πράγματι συμβεί. Η γενικότερη σύνδεση των συχνών ή μη χρήσιμων `alarms` με φαινόμενα `alarm fatigue` υποστηρίζεται από τη σχετική βιβλιογραφία [6], [13].

5.5.10 DL Query 10 – Ποια Evidence Records έχουν mapping προς το req_risk_management;

Το δέκατο DL query χρησιμοποιήθηκε για να εξεταστεί ποια Evidence Record individuals έχουν συνδεθεί με το requirement individual `req_risk_management`.

Το ερώτημα που εξετάζεται είναι:

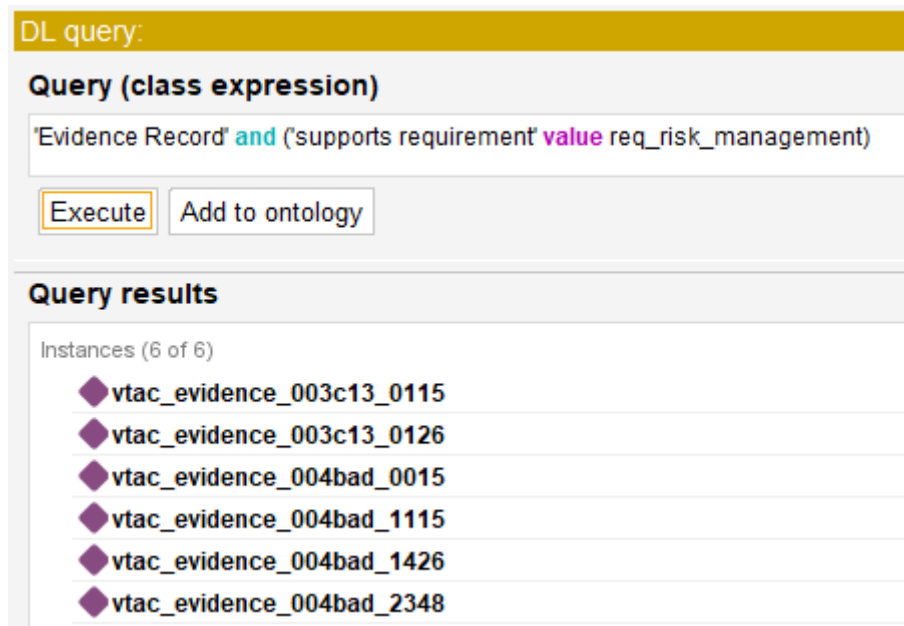
Ποια Evidence Record individuals υποστηρίζουν το req_risk_management;

Η σημασιολογική διαδρομή είναι:

Evidence Record --supportsRequirement→ req_risk_management

Ο DL Query που εκτελέστηκε στο Protégé είναι:

'Evidence Record' and ('supports requirement' value req_risk_management)



Σχήμα 5.12 DL Query 10 — Evidence Records με mapping προς `req_risk_management`.

Το query επιστρέφει 6 Evidence Record individuals.

Τα έξι αποτελέσματα αντιστοιχούν στα evidence records των έξι false alarm cases, τα οποία στο συγκεκριμένο proof-of-concept έχουν χαρτογραφηθεί επιπλέον προς το `req_risk_management` λόγω της συμμετοχής των αντίστοιχων events στο lightweight risk-related επίπεδο.

Το αποτέλεσμα επιβεβαιώνει ότι το συγκεκριμένο mapping έχει υλοποιηθεί και μπορεί να ανακτηθεί λειτουργικά. Δεν σημαίνει ότι τα evidence records αποδεικνύουν την ύπαρξη πλήρους risk management system σύμφωνα με το άρθρο 9 του AI Act [3]. Το req_risk_management αποτελεί περιορισμένο AI Act-related σημείο αναφοράς της παρούσας υλοποίησης.

5.5.11 DL Query 11 – Ποια events διαθέτουν πλήρη σχεδιασμένη διαδρομή ιχνηλασιμότητας;

Το ενδέκατο DL query αποτελεί έναν από τους πιο σύνθετους λειτουργικούς ελέγχους της οντολογίας. Σκοπός του είναι να εξεταστεί αν κάθε επιλεγμένο VT alarm event συμμετέχει ταυτόχρονα στις βασικές σημασιολογικές σχέσεις που συγκροτούν τη διαδρομή ιχνηλασιμότητας του proof-of-concept.

Το ερώτημα που εξετάζεται είναι:

Ποια VT alarm events συνδέονται με waveform record, Expert Review και Review Decision και ταυτόχρονα διαθέτουν Evidence Record που έχει χαρτογραφηθεί προς το req_traceability;

Ο DL Query που εκτελέστηκε είναι:

```
'Ventricular Tachycardia Alarm'
and ('belongs to record' some 'Waveform Record')
and ('has review' some ('has decision' some 'Review Decision'))
and ('is documented by' some ('supports requirement' value req_traceability))
```

DL query:

Query (class expression)

```
'Ventricular Tachycardia Alarm'
and ('belongs to record' some 'Waveform Record')
and ('has review' some ('has decision' some 'Review Decision'))
and ('is documented by' some ('supports requirement' value req_traceability))
```

Execute

Add to ontology

Query results
 Instances (12 of 12)

◆ vtac_event_003c13_0115
◆ vtac_event_003c13_0126
◆ vtac_event_004bad_0015
◆ vtac_event_004bad_1115
◆ vtac_event_004bad_1426
◆ vtac_event_004bad_2348
◆ vtac_event_0060de_0560
◆ vtac_event_015f23_0512
◆ vtac_event_0179fd_1194
◆ vtac_event_0179fd_1265
◆ vtac_event_0179fd_2773
◆ vtac_event_01a952_1067

Σχήμα 5.13 DL Query 11 — Events με πλήρη σχεδιασμένη διαδρομή ιχνηλασιμότητας.

Το query επιστρέφει 12 instances, δηλαδή όλα τα επιλεγμένα VT alarm events.

Το αποτέλεσμα δείχνει ότι για καθένα από τα 12 events ισχύουν ταυτόχρονα οι ακόλουθες συνθήκες:

Alert Event --belongsToRecord→ Waveform Record

Alert Event --hasReview→ Expert Review --hasDecision→ Review Decision

και

Alert Event --isDocumentedBy→ Evidence Record --supportsRequirement→ req_traceability

Επομένως, το query δεν ελέγχει μία μοναδική γραμμική object-property αλυσίδα, αλλά τη συνύπαρξη διαφορετικών σχέσεων γύρω από το ίδιο VT alarm event: σύνδεση με Waveform Record, Expert Review και Review Decision, καθώς και Evidence Record με mapping προς req_traceability.

Η επιτυχής ανάκτηση και των 12 events παρέχει λειτουργική επιβεβαίωση ότι η proof-of-concept οντολογία μπορεί να υποστηρίξει την προβλεπόμενη διαδρομή ιχνηλασιμότητας για το σύνολο των επιλεγμένων cases.

5.5.12 DL Query 12 – Ποια waveform records συνδέονται με VT alarm events;

Το δωδέκατο DL query χρησιμοποιήθηκε για να ελεγχθεί ποια waveform records της οντολογίας συνδέονται με τουλάχιστον ένα VTaC alarm event. Το query αυτό αφορά το αρχικό σημείο της operational αλυσίδας της οντολογίας, δηλαδή τη σύνδεση ανάμεσα στα waveform records και στα αντίστοιχα alert events.

Το ερώτημα που εξετάζεται είναι το εξής:

Ποια waveform records της οντολογίας συνδέονται με τουλάχιστον ένα event της class Ventricular Tachycardia Alarm;

Το query ελέγχει τη σημασιολογική διαδρομή:

Waveform Record → Alert Event

Η σύνδεση αυτή υλοποιείται μέσω του object property has alert event. Με αυτόν τον τρόπο εξετάζεται αν τα VTaC alarm events δεν εμφανίζονται απομονωμένα μέσα στην οντολογία, αλλά έχουν συνδεθεί με τα waveform records από τα οποία προέρχονται.

Ο αντίστοιχος κώδικας DL Query που εκτελέστηκε στο Protégé είναι:

```
'Waveform Record' and ('has alert event' some 'Ventricular Tachycardia Alarm')
```

DL query:

Query (class expression)

'Waveform Record' and ('has alert event' some 'Ventricular Tachycardia Alarm')

Execute Add to ontology

Query results

Instances (6 of 6)

vtac_record_003c13
vtac_record_004bad
vtac_record_0060de
vtac_record_015f23
vtac_record_0179fd
vtac_record_01a952

Σχήμα 5.14 DL Query 12 — Waveform records με VT alarm events.

Όπως φαίνεται στο Σχήμα 5.14, το query επιστρέφει 6 instances, δηλαδή τα waveform records που συνδέονται με VT alarm events.

Το αποτέλεσμα επιβεβαιώνει ότι και τα έξι waveform records της οντολογίας συνδέονται με τουλάχιστον ένα VT alarm event. Επομένως, η οντολογία υποστηρίζει σωστά τη σύνδεση των πραγματικών VTaC records με τα αντίστοιχα alert events, ενισχύοντας την αρχική βάση της ιχνηλασιμότητας.

5.5.13 Συνολική αποτίμηση των DL competency questions

Η εκτέλεση των 12 DL competency questions έδειξε ότι η οντολογία μπορεί να ανακτήσει τα αναμενόμενα individuals και να αξιοποιήσει τις βασικές σημασιολογικές σχέσεις που έχουν υλοποιηθεί στο proof-of-concept.

Τα πρώτα τρία queries ελέγχουν το operational μέρος του μοντέλου, ανακτώντας τα 12 VTaC alert events και διακρίνοντας τα έξι true και έξι false alarm cases με βάση τις αντίστοιχες διαδρομές Expert Review και Review Decision.

Τα Queries 4–7 εξετάζουν το requirement-to-evidence mapping. Επιβεβαιώνουν ότι τα evidence records έχουν χαρτογραφηθεί προς τα req_logging, req_human_review και req_traceability και ότι και τα 12 Evidence Record individuals υποστηρίζουν ταυτόχρονα τα τρία αυτά βασικά mappings.

Τα Queries 8–10 εξετάζουν το lightweight risk-related μέρος της οντολογίας. Ανακτούν τα έξι false alarm events που έχουν ενταχθεί σε αυτό, τα έξι Risk-Relevant Event individuals που συνδέονται με Potential Harm και τα έξι Evidence Record individuals που έχουν χαρτογραφηθεί επιπλέον προς το req_risk_management.

Το Query 11 ελέγχει τη συνύπαρξη των βασικών σχέσεων ιχνηλασιμότητας γύρω από κάθε VT alarm event, ενώ το Query 12 επιβεβαιώνει ότι και τα έξι Waveform Record individuals συνδέονται με τα αντίστοιχα επιλεγμένα events.

Συνολικά, τα αποτελέσματα δείχνουν ότι οι classes, τα object properties και τα individuals που απαιτούνται για τις συγκεκριμένες competency questions έχουν υλοποιηθεί με τρόπο που επιτρέπει την αναμενόμενη σημασιολογική ανάκτηση. Το μοντέλο επομένως δεν λειτουργεί μόνο ως στατική ταξινομική αναπαράσταση, αλλά υποστηρίζει ερωτήματα πάνω στις σχέσεις προέλευσης, αξιολόγησης, τεκμηρίωσης, requirement mapping και risk-related πληροφορίας.

Η επιτυχής εκτέλεση των 12 DL queries αποτελεί λειτουργική επικύρωση μέσα στα όρια του συγκεκριμένου proof-of-concept. Δεν αποτελεί απόδειξη πληρότητας της οντολογίας για ολόκληρο το healthcare domain, κλινική επικύρωση, νομική αξιολόγηση των mappings ή απόδειξη πλήρους συμμόρφωσης με τον AI Act.

Τα DL competency questions παρέχουν επομένως το πρώτο επίπεδο συστηματικού λειτουργικού ελέγχου των σημασιολογικών σχέσεων. Στην επόμενη ενότητα ο έλεγχος συμπληρώνεται με SPARQL queries, τα οποία επιτρέπουν την ταυτόχρονη ανάκτηση πολλών μεταβλητών και την παρουσίαση περισσότερο σύνθετων διαδρομών σε μορφή συγκεντρωτικών αποτελεσμάτων.

5.6 Έλεγχος και αναφορά αποτελεσμάτων μέσω SPARQL queries

Μετά την ολοκλήρωση των DL Queries, η λειτουργική επικύρωση της οντολογίας συνεχίστηκε με τη χρήση SPARQL queries. Τα SPARQL queries αξιοποιήθηκαν ως πρόσθετος μηχανισμός ελέγχου και αναφοράς, καθώς επιτρέπουν την ανάκτηση πληροφορίας από την οντολογία σε πιο δομημένη και συγκεντρωτική μορφή.

Σε αντίθεση με τα DL Queries, τα οποία χρησιμοποιήθηκαν κυρίως για την ανάκτηση συγκεκριμένων ομάδων individuals, τα SPARQL queries μπορούν να εμφανίσουν ταυτόχρονα περισσότερα στοιχεία σε μορφή πίνακα. Για παράδειγμα, μπορούν να παρουσιαστούν μαζί το event, το waveform record προέλευσης, η απόφαση αξιολόγησης, το evidence record, οι απαιτήσεις που υποστηρίζονται και τα risk-related στοιχεία.

Στην παρούσα εργασία χρησιμοποιήθηκαν τέσσερα SPARQL queries. Τα queries αυτά δεν αλλάζουν τη δομή της οντολογίας και δεν δημιουργούν νέα γνώση. Αντίθετα, αξιοποιούν τις σχέσεις που έχουν ήδη μοντελοποιηθεί, ώστε να εξαχθούν πιο αναγνώσιμα και συγκεντρωτικά αποτελέσματα.

Ο Πίνακας 5.3 παρουσιάζει συνοπτικά τα SPARQL queries που χρησιμοποιήθηκαν.

Πίνακας 5-3 SPARQL queries της οντολογίας.

SPARQL Query	Σκοπός
SPARQL Query 1	Συνολική διαδρομή event–record–decision–evidence
SPARQL Query 2	Χαρτογράφηση event–evidence–requirement
SPARQL Query 3	Risk-related mapping για false alarm cases
SPARQL Query 4	Πλήθος VT alarm events ανά waveform record

Η χρήση των SPARQL queries είναι χρήσιμη, επειδή δείχνει ότι η οντολογία μπορεί να αξιοποιηθεί όχι μόνο για λογικό έλεγχο μέσω DL Queries, αλλά και για παραγωγή συγκεντρωτικών αναφορών. Με αυτόν τον τρόπο, το μοντέλο μπορεί να υποστηρίξει πιο πρακτική παρουσίαση της τεκμηρίωσης, της ιχνηλασιμότητας και του requirement-to-evidence mapping.

Στις επόμενες υποενότητες παρουσιάζονται τα τέσσερα SPARQL queries με τη σειρά που χρησιμοποιήθηκαν. Για κάθε query δίνεται το ερώτημα, ο κώδικας SPARQL, το αντίστοιχο screenshot από το Protégé και η ερμηνεία του αποτελέσματος.

5.6.1 SPARQL Query 1 – Συγκεντρωτική ανάκτηση event–record–decision–evidence

Το πρώτο SPARQL query χρησιμοποιήθηκε για τη συγκεντρωτική ανάκτηση διαφορετικών πληροφοριών που συνδέονται με κάθε VT alarm event μέσα στην οντολογία. Σε αντίθεση με τα DL Queries, τα οποία επιστρέφουν κυρίως ομάδες από individuals, το SPARQL query εμφανίζει περισσότερα στοιχεία ταυτόχρονα σε μορφή πίνακα.

Το ερώτημα που εξετάζεται είναι το εξής:

Πώς συνδέεται κάθε VT alarm event με το waveform record προέλευσής του, την απόφαση αξιολόγησης, το evidence record και την ένδειξη risk relevance;

Το query εμφανίζει για κάθε event το αντίστοιχο waveform record, τη review decision, την τιμή του decisionLabel, το evidence record και την τιμή του isRiskRelevant. Με αυτόν τον τρόπο, το αποτέλεσμα δεν δείχνει απλώς ότι τα events υπάρχουν στην οντολογία, αλλά παρουσιάζει συγκεντρωτικά πώς συνδέονται με βασικά στοιχεία τεκμηρίωσης και αξιολόγησης.

Ο αντίστοιχος κώδικας SPARQL Query που εκτελέστηκε στο Protégé είναι:

PREFIX : <http://www.semanticweb.org/aiact-healthcare-vtac#>

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

```
SELECT ?event ?record ?decision ?decisionLabel ?evidence ?isRiskRelevant
```

```
WHERE {
```

```
  ?event rdf:type :VTAlarm .
```

```
  ?event :belongsToRecord ?record .
```

```
  ?event :hasReview ?review .
```

```
  ?review :hasDecision ?decision .
```

```
  ?decision :decisionLabel ?decisionLabel .
```

```
  ?event :isDocumentedBy ?evidence .
```

```
  ?event :isRiskRelevant ?isRiskRelevant .
```

```
}
```

```
ORDER BY ?event
```

SPARQL query:

```

PREFIX : <http://www.semanticweb.org/aiact-healthcare-vtac#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

SELECT ?event ?record ?decision ?decisionLabel ?evidence ?isRiskRelevant
WHERE {
  ?event rdf:type :VTAlarm .
  ?event :belongsToRecord ?record .
  ?event :hasReview ?review .
  ?review :hasDecision ?decision .
  ?decision :decisionLabel ?decisionLabel .
  ?event :isDocumentedBy ?evidence .
  ?event :isRiskRelevant ?isRiskRelevant .
}
ORDER BY ?event

```

event	record	decision	decisionLabel	evidence	isRiskRelevant
vtac_event_003c13_0115	vtac_record_003c13	vtac_decision_003c13_0115_false	"false_alarm"	vtac_evidence_003c13_0115	"true"<http://www.w3.org/2001/XMLSchema#boolean>
vtac_event_003c13_0126	vtac_record_003c13	vtac_decision_003c13_0126_false	"false_alarm"	vtac_evidence_003c13_0126	"true"<http://www.w3.org/2001/XMLSchema#boolean>
vtac_event_004bad_0015	vtac_record_004bad	vtac_decision_004bad_0015_false	"false_alarm"	vtac_evidence_004bad_0015	"true"<http://www.w3.org/2001/XMLSchema#boolean>
vtac_event_004bad_1115	vtac_record_004bad	vtac_decision_004bad_1115_false	"false_alarm"	vtac_evidence_004bad_1115	"true"<http://www.w3.org/2001/XMLSchema#boolean>
vtac_event_004bad_1426	vtac_record_004bad	vtac_decision_004bad_1426_false	"false_alarm"	vtac_evidence_004bad_1426	"true"<http://www.w3.org/2001/XMLSchema#boolean>
vtac_event_004bad_2348	vtac_record_004bad	vtac_decision_004bad_2348_false	"false_alarm"	vtac_evidence_004bad_2348	"true"<http://www.w3.org/2001/XMLSchema#boolean>
vtac_event_0060de_0560	vtac_record_0060de	vtac_decision_0060de_0560_true	"true_alarm"	vtac_evidence_0060de_0560	"false"<http://www.w3.org/2001/XMLSchema#boolean>
vtac_event_015f23_0512	vtac_record_015f23	vtac_decision_015f23_0512_true	"true_alarm"	vtac_evidence_015f23_0512	"false"<http://www.w3.org/2001/XMLSchema#boolean>
vtac_event_0179fd_1194	vtac_record_0179fd	vtac_decision_0179fd_1194_true	"true_alarm"	vtac_evidence_0179fd_1194	"false"<http://www.w3.org/2001/XMLSchema#boolean>
vtac_event_0179fd_1265	vtac_record_0179fd	vtac_decision_0179fd_1265_true	"true_alarm"	vtac_evidence_0179fd_1265	"false"<http://www.w3.org/2001/XMLSchema#boolean>
vtac_event_0179fd_2773	vtac_record_0179fd	vtac_decision_0179fd_2773_true	"true_alarm"	vtac_evidence_0179fd_2773	"false"<http://www.w3.org/2001/XMLSchema#boolean>
vtac_event_01a952_1067	vtac_record_01a952	vtac_decision_01a952_1067_true	"true_alarm"	vtac_evidence_01a952_1067	"false"<http://www.w3.org/2001/XMLSchema#boolean>

Execute

Reasoner active ☒ Show Inferences

Σχήμα 5.15 SPARQL Query 1 — Συγκεντρωτική ανάκτηση event–record–decision–evidence.

Όπως φαίνεται στο Σχήμα 5.15, το query επιστρέφει 12 γραμμές, μία για κάθε VT alarm event της οντολογίας.

Τα πρώτα έξι events έχουν decisionLabel με τιμή false_alarm και isRiskRelevant με τιμή true. Αυτό δείχνει ότι τα false alarm cases έχουν συνδεθεί με το risk-related επίπεδο της οντολογίας.

Τα υπόλοιπα έξι events έχουν decisionLabel με τιμή true_alarm και isRiskRelevant με τιμή false. Αυτό δείχνει ότι τα true alarm cases ακολουθούν τη βασική τεκμηριωτική πορεία, χωρίς πρόσθετη risk-related αναπαράσταση.

Το αποτέλεσμα επιβεβαιώνει ότι κάθε VT alarm event συνδέεται με record προέλευσης, decision και evidence record. Επομένως, η οντολογία μπορεί να παράγει συγκεντρωτική αναφορά για τα βασικά στοιχεία τεκμηρίωσης κάθε event.

5.6.2 SPARQL Query 2 – Χαρτογράφηση event–evidence–requirement

Το δεύτερο SPARQL query χρησιμοποιήθηκε για να παρουσιαστεί η χαρτογράφηση ανάμεσα στα VT alarm events, στα evidence records και στις απαιτήσεις που υποστηρίζονται από αυτά. Ενώ το πρώτο SPARQL query παρουσίασε τη συνολική πορεία κάθε event, το παρόν query εστιάζει ειδικά στο requirement-to-evidence mapping.

Το ερώτημα που εξετάζεται είναι το εξής:

Ποιες απαιτήσεις υποστηρίζονται από το evidence record κάθε VT alarm event;

Το query εμφανίζει για κάθε event το αντίστοιχο evidence record και το requirement που υποστηρίζεται από αυτό. Με αυτόν τον τρόπο γίνεται ορατό ότι η οντολογία δεν συνδέει απευθείας ένα event με μια απαίτηση, αλλά χρησιμοποιεί το evidence record ως ενδιάμεσο τεκμηριωτικό στοιχείο.

Ο αντίστοιχος κώδικας SPARQL Query που εκτελέστηκε στο Protégé είναι:

PREFIX : <http://www.semanticweb.org/aiact-healthcare-vtac#>

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

```

SELECT ?event ?evidence ?requirement
WHERE {
  ?event rdf:type :VTAlarm .
  ?event :isDocumentedBy ?evidence .
  ?evidence :supportsRequirement ?requirement .
}
ORDER BY ?event ?requirement

```

SPARQL query:

```

PREFIX : <http://www.semanticweb.org/aiact-healthcare-vtac#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

SELECT ?event ?evidence ?requirement
WHERE {
  ?event rdf:type :VTAlarm .
  ?event :isDocumentedBy ?evidence .
  ?evidence :supportsRequirement ?requirement .
}
ORDER BY ?event ?requirement

```

event	evidence	requirement
vtac_event_003c13_0115	vtac_evidence_003c13_0115	req_human_review
vtac_event_003c13_0115	vtac_evidence_003c13_0115	req_logging
vtac_event_003c13_0115	vtac_evidence_003c13_0115	req_risk_management
vtac_event_003c13_0115	vtac_evidence_003c13_0115	req_traceability
vtac_event_003c13_0126	vtac_evidence_003c13_0126	req_human_review
vtac_event_003c13_0126	vtac_evidence_003c13_0126	req_logging
vtac_event_003c13_0126	vtac_evidence_003c13_0126	req_risk_management
vtac_event_003c13_0126	vtac_evidence_003c13_0126	req_traceability
vtac_event_004bad_0015	vtac_evidence_004bad_0015	req_human_review
vtac_event_004bad_0015	vtac_evidence_004bad_0015	req_logging
vtac_event_004bad_0015	vtac_evidence_004bad_0015	req_risk_management
vtac_event_004bad_0015	vtac_evidence_004bad_0015	req_traceability
vtac_event_004bad_1115	vtac_evidence_004bad_1115	req_human_review
vtac_event_004bad_1115	vtac_evidence_004bad_1115	req_logging
vtac_event_004bad_1115	vtac_evidence_004bad_1115	req_risk_management
vtac_event_004bad_1115	vtac_evidence_004bad_1115	req_traceability
vtac_event_004bad_1426	vtac_evidence_004bad_1426	req_human_review
vtac_event_004bad_1426	vtac_evidence_004bad_1426	req_logging

Execute

Reasoner active ☒ Show Inferences

Σχήμα 5.16 SPARQL Query 2 — Χαρτογράφηση event–evidence–requirement.

Όπως φαίνεται στο Σχήμα 5.16, το query επιστρέφει τις σχέσεις ανάμεσα σε κάθε VT alarm event, το αντίστοιχο evidence record και τις απαιτήσεις που αυτό υποστηρίζει.

Το αποτέλεσμα δείχνει ότι τα evidence records συνδέονται με τις επιλεγμένες απαιτήσεις req_logging, req_human_review και req_traceability. Επιπλέον, για τα false alarm cases εμφανίζεται και η σύνδεση με την απαίτηση req_risk_management.

Με αυτόν τον τρόπο επιβεβαιώνεται ότι το requirement-to-evidence mapping έχει υλοποιηθεί σωστά στην οντολογία. Τα evidence records λειτουργούν ως ενδιάμεσο επίπεδο τεκμηρίωσης και επιτρέπουν τη σύνδεση των πραγματικών VTaC events με τις selected AI Act-related requirements.

5.6.3 SPARQL Query 3 – Risk-related mapping για false alarm cases

Το τρίτο SPARQL query χρησιμοποιήθηκε για να παρουσιαστεί συγκεντρωτικά η σύνδεση των false alarm events με τα risk-related στοιχεία της οντολογίας. Το query αυτό εστιάζει στο lightweight risk layer και εμφανίζει για κάθε σχετικό event τον αντίστοιχο risk indicator, το risk-relevant event και το potential harm.

Το ερώτημα που εξετάζεται είναι το εξής:

Ποια VT alarm events έχουν ενταχθεί στο lightweight risk layer και με ποια risk-related στοιχεία συνδέονται;

Το query ανακτά μόνο τα events που έχουν τις απαραίτητες risk-related συνδέσεις. Με αυτόν τον τρόπο παρουσιάζεται καθαρά ποια false alarm cases έχουν αποκτήσει πρόσθετη σημασιολογική αναπαράσταση σε σχέση με τον κίνδυνο, χωρίς να επαναλαμβάνεται όλη η βασική τεκμηριωτική διαδρομή της οντολογίας.

Ο αντίστοιχος κώδικας SPARQL Query που εκτελέστηκε στο Protégé είναι:

PREFIX : <http://www.semanticweb.org/aiact-healthcare-vtac#>

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

```
SELECT ?event ?riskIndicator ?riskEvent ?potentialHarm
```

```
WHERE {
```

```
  ?event rdf:type :VTAlarm .
```

```
  ?event :hasRiskIndicator ?riskIndicator .
```

```
  ?event :isClassifiedAs ?riskEvent .
```

```
  ?riskEvent :mayLeadTo ?potentialHarm .
```

```
}
```

```
ORDER BY ?event
```

SPARQL query:

```

PREFIX : <http://www.semanticweb.org/aiact-healthcare-vtac#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

SELECT ?event ?riskIndicator ?riskEvent ?potentialHarm
WHERE {
  ?event rdf:type :VTAlarm .
  ?event :hasRiskIndicator ?riskIndicator .
  ?event :isClassifiedAs ?riskEvent .
  ?riskEvent :mayLeadTo ?potentialHarm .
}
ORDER BY ?event

```

event	riskIndicator	riskEvent	potentialHarm
vtac_event_003c13_0115	vtac_indicator_003c13_0115_false_alarm	vtac_risk_event_003c13_0115	vtac_harm_alert_fatigue
vtac_event_003c13_0126	vtac_indicator_003c13_0126_false_alarm	vtac_risk_event_003c13_0126	vtac_harm_alert_fatigue
vtac_event_004bad_0015	vtac_indicator_004bad_0015_false_alarm	vtac_risk_event_004bad_0015	vtac_harm_alert_fatigue
vtac_event_004bad_1115	vtac_indicator_004bad_1115_false_alarm	vtac_risk_event_004bad_1115	vtac_harm_alert_fatigue
vtac_event_004bad_1426	vtac_indicator_004bad_1426_false_alarm	vtac_risk_event_004bad_1426	vtac_harm_alert_fatigue
vtac_event_004bad_2348	vtac_indicator_004bad_2348_false_alarm	vtac_risk_event_004bad_2348	vtac_harm_alert_fatigue

Execute

Reasoner active ☒ Show Inferences

Σχήμα 5.17 SPARQL Query 3 — Risk-related mapping για false alarm cases.

Όπως φαίνεται στο Σχήμα 5.17, το query επιστρέφει 6 γραμμές, μία για κάθε false alarm event που έχει ενταχθεί στο lightweight risk layer.

Κάθε event συνδέεται με το αντίστοιχο risk indicator, με ένα risk-relevant event και με το potential harm vtac_harm_alert_fatigue. Το αποτέλεσμα επιβεβαιώνει ότι τα false alarm cases έχουν οργανωθεί οντολογικά ως περιπτώσεις με πιθανή risk-related σημασία, χωρίς αυτό να σημαίνει ότι τεκμηριώνεται πραγματική βλάβη ή πλήρης διαδικασία risk management.

5.6.4 SPARQL Query 4 – Πλήθος VT alarm events ανά waveform record

Το τέταρτο SPARQL query χρησιμοποιήθηκε για να παρουσιαστεί συγκεντρωτικά πόσα VT alarm events συνδέονται με κάθε waveform record της οντολογίας. Το query αυτό δεν εστιάζει σε ένα μεμονωμένο event, αλλά παρέχει συνοπτική εικόνα της κατανομής των events στα αντίστοιχα records.

Το ερώτημα που εξετάζεται είναι το εξής:

Πόσα VT alarm events αντιστοιχούν σε κάθε waveform record της οντολογίας;

Το query αξιοποιεί τη σχέση hasAlertEvent, μέσω της οποίας κάθε waveform record συνδέεται με ένα ή περισσότερα VT alarm events. Με τη χρήση της συνάρτησης COUNT, το αποτέλεσμα εμφανίζεται συγκεντρωτικά ανά record.

Ο αντίστοιχος κώδικας SPARQL Query που εκτελέστηκε στο Protégé είναι:

PREFIX : <http://www.semanticweb.org/aiact-healthcare-vtac#>

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

SELECT ?record (COUNT(?event) AS ?numberOfEvents)

WHERE {

 ?record rdf:type :WaveformRecord .

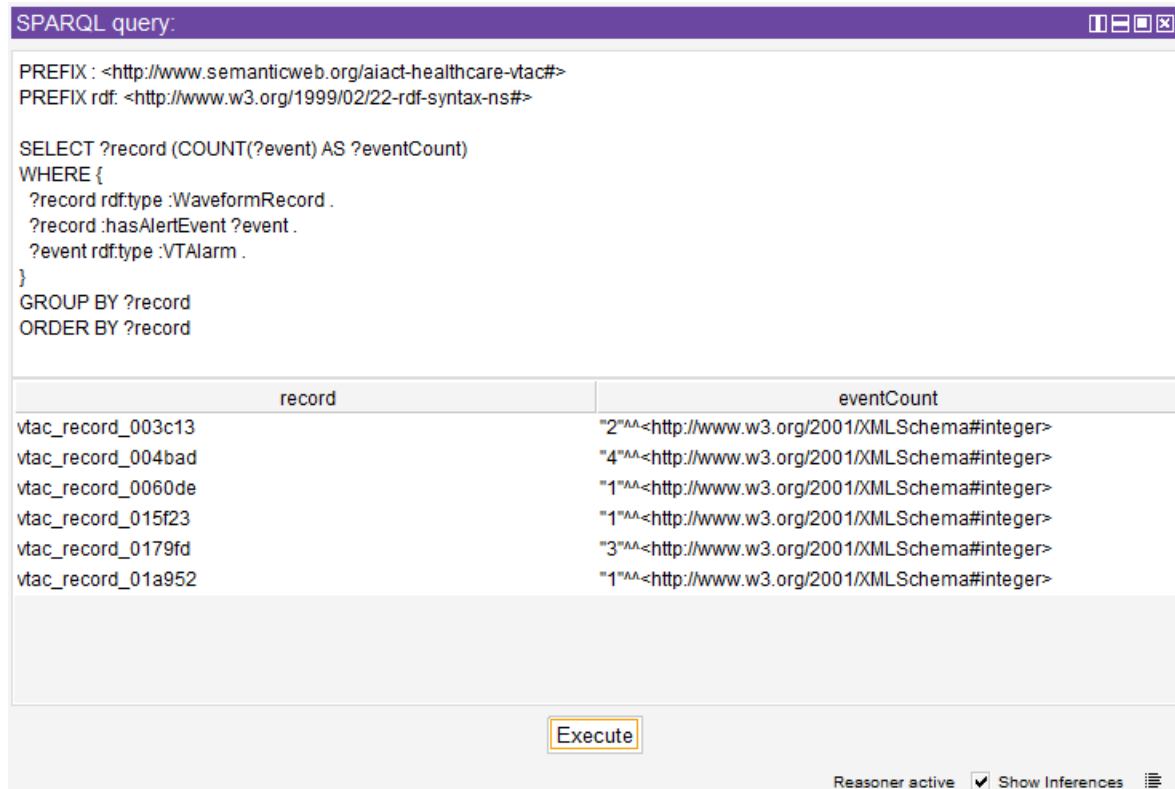
 ?record :hasAlertEvent ?event .

```
?event rdf:type :VTAlarm .

}

GROUP BY ?record

ORDER BY ?record
```



Σχήμα 5.18 SPARQL Query 4 — Πλήθος VT alarm events ανά waveform record.

Όπως φαίνεται στο Σχήμα 5.18, το query επιστρέφει 6 γραμμές, μία για κάθε waveform record που έχει εισαχθεί στην οντολογία.

Το αποτέλεσμα δείχνει την ακόλουθη κατανομή:

Πίνακας 5-4 Πλήθος VT alarm events ανά waveform record

Waveform record	Πλήθος VT alarm events
vtac_record_003c13	2
vtac_record_004bad	4
vtac_record_0060de	1
vtac_record_015f23	1
vtac_record_0179fd	3
vtac_record_01a952	1

Το αποτέλεσμα επιβεβαιώνει ότι τα 12 VT alarm events της οντολογίας έχουν συνδεθεί με τα 6 waveform records από τα οποία προέρχονται. Με αυτόν τον τρόπο, το query παρέχει μια συνοπτική

εικόνα της σχέσης record–event και ενισχύει την τεκμηρίωση της προέλευσης των VTaC alarm events μέσα στο μοντέλο.

5.6.5 Συνολική αποτίμηση των SPARQL queries

Η εκτέλεση των τεσσάρων SPARQL queries έδειξε ότι οι πληροφορίες που έχουν αναπαρασταθεί στην οντολογία μπορούν να ανακτηθούν σε δομημένη και συγκεντρωτική μορφή.

Το SPARQL Query 1 συνδύασε για κάθε VT alarm event πληροφορίες σχετικά με το waveform record προέλευσης, τη Review Decision, το Evidence Record και την ένδειξη risk relevance. Το Query 2 παρουσίασε το requirement-to-evidence mapping και κατέστησε εμφανή τον ρόλο του Evidence Record ως ενδιάμεσου τεκμηριωτικού κόμβου μεταξύ των πραγματικών events και των requirement individuals.

Το Query 3 ανέκτησε συγκεντρωτικά την πρόσθετη risk-related αναπαράσταση των έξι false alarm cases, ενώ το Query 4 παρουσίασε την κατανομή των 12 επιλεγμένων events στα έξι waveform records της υλοποίησης.

Σε σύγκριση με τα DL competency questions, τα SPARQL queries δεν αντικαθιστούν τον λογικό έλεγχο αλλά τον συμπληρώνουν. Επιτρέπουν την ταυτόχρονη ανάκτηση πολλών σχετικών μεταβλητών και την παραγωγή περισσότερο αναγνώσιμων αναφορών πάνω στις ήδη υλοποιημένες σημασιολογικές σχέσεις [25].

Τα αποτελέσματα επιβεβαιώνουν λειτουργικά ότι μπορούν να ανακτηθούν οι βασικές πληροφορίες που απαιτούνται από τον σχεδιασμό του proof-of-concept: η προέλευση των events, οι expert decisions, τα evidence records, τα requirement mappings και οι risk-related συνδέσεις των false alarm cases.

Η επιτυχής εκτέλεση των SPARQL queries δεν αποτελεί απόδειξη πληρότητας της οντολογίας, κλινικής εγκυρότητας ή νομικής συμμόρφωσης με τον AI Act. Δείχνει ότι, για τα συγκεκριμένα instances και τις σχέσεις που έχουν υλοποιηθεί, το μοντέλο μπορεί να υποστηρίξει οργανωμένα, αναζητήσιμη και επαναλήψιμη ανάκτηση της σημασιολογικά δομημένης πληροφορίας.

Σε συνδυασμό με τον έλεγχο μέσω HermiT reasoner και τα 12 DL competency questions, τα τέσσερα SPARQL queries ολοκληρώνουν τη λειτουργική επικύρωση της παρούσας proof-of-concept υλοποίησης.

5.7 Συνολική αποτίμηση και επίλογος κεφαλαίου

Στο παρόν κεφάλαιο παρουσιάστηκε ο έλεγχος και η λειτουργική επικύρωση της οντολογίας που υλοποιήθηκε στο Protégé. Στόχος ήταν να εξεταστεί αν το proof-of-concept μοντέλο παρουσιάζει λογική συνέπεια και αν οι σημασιολογικές σχέσεις που σχεδιάστηκαν και υλοποιήθηκαν μπορούν να ανακτηθούν με ελεγχόμενο και επαναλήψιμο τρόπο.

Αρχικά πραγματοποιήθηκε έλεγχος με τον HermiT reasoner. Η εκτέλεσή του ολοκληρώθηκε χωρίς να εντοπιστεί inconsistency στην εξεταζόμενη έκδοση της οντολογίας και χωρίς να εμφανιστούν

unsatisfiable classes. Παράλληλα, η εξέταση της inferred class hierarchy δεν ανέδειξε ταξινομήσεις που να έρχονται σε αντίθεση με τη βασική σχεδιασμένη οργάνωση του μοντέλου. Τα αποτελέσματα αυτά παρέχουν ένδειξη λογικής συνέπειας ως προς τα δηλωμένα axioms, χωρίς να συνεπάγονται από μόνα τους πληρότητα ή συνολική ορθότητα του domain model.

Στη συνέχεια εκτελέστηκαν 12 DL competency questions. Τα πρώτα ερωτήματα ανέκτησαν τα 12 πραγματικά VTaC event individuals και επέτρεψαν τη διάκριση των έξι true και έξι false alarm cases μέσω των σχέσεων Expert Review και Review Decision. Τα επόμενα queries εξέτασαν το requirement-to-evidence mapping και έδειξαν ότι τα Evidence Record individuals μπορούν να ανακτηθούν με βάση τις συνδέσεις τους προς τα req_logging, req_human_review και req_traceability. Για τα false alarm cases ελέγχθηκε επιπλέον η πρόσθετη risk-related αναπαράσταση και το mapping προς το req_risk_management.

Ιδιαίτερη σημασία είχε το DL Query 11, το οποίο εξέτασε τη συνύπαρξη των βασικών σχέσεων ιχνηλασιμότητας γύρω από κάθε VT alarm event. Και τα 12 events συνδέθηκαν με waveform record, Expert Review και Review Decision και, παράλληλα, με Evidence Record που είχε χαρτογραφηθεί προς το req_traceability. Το αποτέλεσμα αυτό δείχνει ότι η προβλεπόμενη διαδρομή ιχνηλασιμότητας του proof-of-concept μπορεί να ανακτηθεί λειτουργικά για όλα τα επιλεγμένα περιστατικά.

Η λειτουργική επικύρωση συμπληρώθηκε με τέσσερα SPARQL queries. Τα αποτελέσματά τους επέτρεψαν τη συγκεντρωτική παρουσίαση της σχέσης μεταξύ event, waveform record, review decision και evidence, την ανάκτηση του requirement-to-evidence mapping, την παρουσίαση της risk-related δομής των έξι false alarm cases και την καταμέτρηση των 12 επιλεγμένων events στα έξι waveform records. Με αυτόν τον τρόπο, η οντολογία έδειξε ότι μπορεί να υποστηρίξει όχι μόνο λογικά ερωτήματα πάνω στα individuals, αλλά και δομημένη ανάκτηση συνδυασμένων πληροφοριών.

Τα αποτελέσματα του κεφαλαίου επιβεβαιώνουν επομένως τη λειτουργικότητα της οντολογίας μέσα στο συγκεκριμένο και περιορισμένο πλαίσιο στο οποίο σχεδιάστηκε. Το μοντέλο μπορεί να οργανώσει πραγματικά VTaC healthcare monitoring events, να διατηρήσει τη σύνδεσή τους με την προέλευση και την expert αξιολόγησή τους, να τα τεκμηριώσει μέσω Evidence Records και να υποστηρίξει σημασιολογικά mappings προς επιλεγμένα AI Act-related σημεία αναφοράς. Για τα false alarm cases μπορεί επιπλέον να αναπαραστήσει την πρόσθετη lightweight risk-related πληροφορία που προβλέφθηκε στον σχεδιασμό.

Παράλληλα, είναι σημαντικό να διατηρούνται σαφή τα όρια αυτής της επικύρωσης. Η επιτυχής εκτέλεση του reasoner, των DL Queries και των SPARQL queries δεν αποδεικνύει ότι η οντολογία αποτελεί πλήρη αναπαράσταση του healthcare domain, ούτε συνιστά κλινική ή νομική επικύρωση. Επίσης, τα requirement-to-evidence mappings δεν αποτελούν απόδειξη πλήρους συμμόρφωσης με τον AI Act. Η αξιολόγηση που πραγματοποιήθηκε είναι λογική και λειτουργική και αφορά αποκλειστικά τη συμπεριφορά της συγκεκριμένης proof-of-concept υλοποίησης.

Συνολικά, το Κεφάλαιο 5 ολοκληρώνει το πρακτικό μέρος της ανάπτυξης και του ελέγχου της οντολογίας. Τα αποτελέσματα παρέχουν τη βάση για το επόμενο κεφάλαιο, στο οποίο δεν εξετάζεται πλέον αν οι σημασιολογικές σχέσεις μπορούν να ανακτηθούν, αλλά τι σημαίνουν τα αποτελέσματα αυτά σε σχέση με τον αρχικό ερευνητικό σκοπό. Στο Κεφάλαιο 6 ακολουθεί η συνολική ερμηνεία των αποτελεσμάτων, η απάντηση στα ερευνητικά ερωτήματα, η αποτίμηση της συμβολής και των περιορισμών της εργασίας και η παρουσίαση των πιθανών μελλοντικών επεκτάσεων.

Κεφάλαιο 6ο: Συζήτηση, συμπεράσματα και μελλοντικές επεκτάσεις

6.1 Εισαγωγή

Το παρόν κεφάλαιο ολοκληρώνει τη διπλωματική εργασία μέσα από τη συνολική συζήτηση και ερμηνεία των αποτελεσμάτων που προέκυψαν από τον σχεδιασμό, την υλοποίηση και τη λειτουργική επικύρωση της προτεινόμενης οντολογίας. Ενώ τα προηγούμενα κεφάλαια επικεντρώθηκαν στο θεωρητικό υπόβαθρο, στη μεθοδολογία, στην ανάπτυξη του μοντέλου στο Protégé και στον λειτουργικό του έλεγχο, στο παρόν κεφάλαιο το ενδιαφέρον μετατοπίζεται στη σημασία των αποτελεσμάτων και στον βαθμό στον οποίο αυτά απαντούν στον αρχικό ερευνητικό σκοπό της εργασίας.

Κεντρικό ζητούμενο της διπλωματικής ήταν να εξεταστεί αν μια οντολογική προσέγγιση μπορεί να υποστηρίξει τη σημασιολογική αναπαράσταση και τεκμηρίωση πραγματικών healthcare monitoring alarm events και να δημιουργήσει μια ελέγξιμη σύνδεση μεταξύ των πραγματικών περιστατικών, της ανθρώπινης αξιολόγησής τους, των διαθέσιμων evidence records και επιλεγμένων AI Act-related σημείων αναφοράς. Η εργασία δεν είχε ως στόχο την ανάπτυξη ενός συστήματος μηχανικής μάθησης για την πρόβλεψη true ή false alarms ούτε την απόδειξη νομικής συμμόρφωσης ενός πραγματικού συστήματος τεχνητής νοημοσύνης. Ο στόχος ήταν η ανάπτυξη και ο λειτουργικός έλεγχος ενός εφαρμοσμένου οντολογικού proof-of-concept.

Για την εξέταση του παραπάνω στόχου χρησιμοποιήθηκαν πραγματικά ventricular tachycardia alarm events από το VTaC dataset. Δώδεκα επιλεγμένα events, έξι με τελική απόφαση true alarm και έξι με τελική απόφαση false alarm, χρησιμοποιήθηκαν ως πραγματικές περιπτώσεις πάνω στις οποίες εφαρμόστηκε η προτεινόμενη σημασιολογική δομή. Τα πραγματικά στοιχεία του dataset συνδυάστηκαν με οντολογικές έννοιες που δημιουργήθηκαν για την αναπαράσταση της προέλευσης, της expert αξιολόγησης, της τελικής απόφασης, της τεκμηρίωσης και του requirement-to-evidence mapping.

Η λειτουργική επικύρωση που παρουσιάστηκε στο προηγούμενο κεφάλαιο έδειξε ότι η υλοποιημένη οντολογία μπορεί να ανακτήσει τις βασικές σχέσεις για τις οποίες σχεδιάστηκε. Ο HermiT reasoner δεν εντόπισε λογική ασυνέπεια στην εξεταζόμενη έκδοση του μοντέλου, ενώ τα 12 DL competency questions και τα 4 SPARQL queries επέτρεψαν την ανάκτηση των πραγματικών VTaC events, των true και false alarm cases, των waveform records, των Expert Reviews και Review Decisions, των Evidence Records, των requirement mappings και των risk-related σχέσεων που είχαν ενσωματωθεί στο proof-of-concept.

Ιδιαίτερη σημασία έχει ότι τα αποτελέσματα δεν προκύπτουν από μία μοναδική γραμμική σημασιολογική αλυσίδα. Η ιχνηλασιμότητα οργανώνεται μέσα από διαφορετικές αλλά συνδεδεμένες διαδρομές γύρω από το ίδιο Alert Event: τη σύνδεση με το waveform record προέλευσης, τη διαδρομή Expert Review–Review Decision και τη διαδρομή Event–Evidence Record–Requirement Mapping. Για τα false alarm cases προστίθεται επιπλέον το lightweight risk-related επίπεδο, μέσω του οποίου τα επιλεγμένα events συνδέονται με Risk Indicator, Risk-Relevant Event και Potential Harm.

Τα αποτελέσματα αυτά πρέπει να ερμηνεύονται μέσα στα όρια της συγκεκριμένης εφαρμογής. Η επιτυχής λειτουργία των queries δεν αποδεικνύει πλήρη κάλυψη του healthcare domain, κλινική εγκυρότητα, πλήρη εφαρμογή ενός risk management system ή νομική συμμόρφωση με τον AI Act.

Δείχνει όμως ότι, για τα συγκεκριμένα πραγματικά instances και τις σχέσεις που έχουν σχεδιαστεί, μια οντολογική προσέγγιση μπορεί να χρησιμοποιηθεί για τη δομημένη οργάνωση, την τεκμηρίωση, την ιχνηλασιμότητα και την αναζήτηση πληροφοριών που διαφορετικά θα παρέμεναν αποσπασματικές.

Με βάση τα παραπάνω, το κεφάλαιο οργανώνεται σε έξι επιμέρους ενότητες. Αρχικά πραγματοποιείται συνολική ερμηνεία των αποτελεσμάτων και στη συνέχεια δίνονται απαντήσεις στα ερευνητικά ερωτήματα που τέθηκαν στην αρχή της εργασίας. Ακολουθεί η αποτίμηση της επιστημονικής, μεθοδολογικής και πρακτικής συμβολής της προσέγγισης, η παρουσίαση των βασικών περιορισμών και η εξέταση πιθανών μελλοντικών επεκτάσεων. Το κεφάλαιο ολοκληρώνεται με το τελικό συμπέρασμα της διπλωματικής εργασίας.

6.2 Συνολική ερμηνεία των αποτελεσμάτων

Τα αποτελέσματα της εργασίας δείχνουν ότι η αξία της προτεινόμενης οντολογικής προσέγγισης δεν βρίσκεται σε μία μεμονωμένη class, property ή query, αλλά στη δυνατότητα να οργανώνονται διαφορετικές κατηγορίες πληροφορίας γύρω από το ίδιο πραγματικό healthcare monitoring event. Το alarm event αποκτά έτσι ένα ευρύτερο σημασιολογικό πλαίσιο, στο οποίο μπορούν να συνυπάρχουν πληροφορίες σχετικά με την προέλευσή του, την ανθρώπινη αξιολόγηση, την τελική απόφαση, την τεκμηρίωση και τα επιλεγμένα AI Act-related mappings.

Το στοιχείο αυτό είναι ιδιαίτερα σημαντικό επειδή τα δεδομένα που χρησιμοποιήθηκαν δεν αποτελούν υποθετικά παραδείγματα που δημιουργήθηκαν αποκλειστικά για τις ανάγκες της οντολογίας. Τα waveform records, τα alarm event identifiers και οι τελικές expert αποφάσεις έχουν πραγματική βάση στο VTaC dataset. Η οντολογική μοντελοποίηση προσθέτει πάνω σε αυτά ένα σημασιολογικό επίπεδο οργάνωσης, χωρίς να μεταβάλλει ή να επαναπροσδιορίζει την αρχική κλινική ταξινόμηση των περιστατικών. Με αυτή την έννοια, η οντολογία δεν αντικαθιστά τα δεδομένα αλλά λειτουργεί ως δομή που επιτρέπει στα διαφορετικά στοιχεία που τα συνοδεύουν να συνδεθούν με ρητό και αναζητήσιμο τρόπο.

Ένα από τα βασικότερα αποτελέσματα αφορά τον ρόλο του Evidence Record. Η υλοποίηση έδειξε ότι η ύπαρξη ενός διακριτού τεκμηριωτικού κόμβου επιτρέπει να αποφεύγεται μια υπερβολικά άμεση σύνδεση μεταξύ ενός πραγματικού event και ενός κανονιστικού σημείου αναφοράς. Το event περιγράφει το πραγματικό περιστατικό, ενώ το Evidence Record αναπαριστά το επίπεδο της τεκμηρίωσης που έχει δημιουργηθεί στο proof-of-concept. Το requirement mapping πραγματοποιείται στη συνέχεια πάνω στο evidence και όχι απευθείας πάνω στο ίδιο το περιστατικό. Η διάκριση αυτή είναι ουσιαστική, επειδή διαχωρίζει σημασιολογικά το γεγονός, το τεκμήριο και το κανονιστικό σημείο προς το οποίο το τεκμήριο έχει χαρτογραφηθεί.

Η λογική αυτή παρουσιάζει σαφή συνάφεια με προηγούμενες ontology-based προσεγγίσεις για AI governance. Η TAIR, για παράδειγμα, οργανώνει concepts και normative requirements από κανονιστικά και τεχνικά κείμενα και τα καθιστά traceable, queryable και navigable [5]. Αντίστοιχα, η FRIA ontology υποστηρίζει τη μετατροπή πληροφοριών σχετικών με διαδικασίες του AI Act σε δομημένη και machine-readable μορφή, με στόχο την τεκμηρίωση και την ανάπτυξη υποστηρικτικών εργαλείων [17]. Η παρούσα εργασία κινείται στην ίδια γενικότερη κατεύθυνση της σημασιολογικής οργάνωσης, αλλά

εφαρμόζει τη λογική αυτή σε διαφορετικό επίπεδο: στη σύνδεση πραγματικών healthcare monitoring events με event-level evidence και επιλεγμένα AI Act-related σημεία αναφοράς.

Ένα δεύτερο σημαντικό αποτέλεσμα αφορά την ιχνηλασιμότητα. Ο λειτουργικός έλεγχος έδειξε ότι η ιχνηλασιμότητα δεν χρειάζεται να αναπαρίσταται ως μία μοναδική γραμμική αλυσίδα. Στην προτεινόμενη οντολογία προκύπτει από τη σύνθεση διαφορετικών σχέσεων γύρω από το ίδιο Alert Event. Η σύνδεση με το Waveform Record παρέχει την προέλευση του περιστατικού, η διαδρομή Expert Review–Review Decision οργανώνει την πληροφορία της ανθρώπινης αξιολόγησης και η σύνδεση Event–Evidence Record οργανώνει την τεκμηρίωση. Το mapping του Evidence Record προς το `req_traceability` προσθέτει στη συνέχεια το αντίστοιχο AI Act-related σημείο αναφοράς.

Η συγκεκριμένη μορφή ιχνηλασιμότητας έχει πρακτική σημασία επειδή επιτρέπει διαφορετικές κατευθύνσεις αναζήτησης. Μπορεί να ξεκινήσει κανείς από ένα συγκεκριμένο event και να αναζητήσει το record, την αξιολόγηση, την απόφαση και το evidence που σχετίζονται με αυτό. Μπορεί επίσης να ξεκινήσει από ένα waveform record και να εντοπίσει τα events που συνδέονται με αυτό ή από ένα requirement individual και να εξετάσει ποια evidence records έχουν χαρτογραφηθεί προς αυτό. Το γεγονός ότι οι διαδρομές αυτές ανακτήθηκαν τόσο μέσω DL competency questions όσο και μέσω SPARQL δείχνει ότι η πληροφορία δεν έχει απλώς αποθηκευτεί, αλλά έχει οργανωθεί με τρόπο που επιτρέπει διαφορετικές σημασιολογικές αναζητήσεις.

Η συνδυασμένη χρήση DL Queries και SPARQL ανέδειξε επίσης δύο συμπληρωματικές μορφές αξιοποίησης της οντολογίας. Τα DL competency questions χρησιμοποιήθηκαν κυρίως για τον έλεγχο συγκεκριμένων λογικών συνθηκών, όπως η ανάκτηση true ή false alarm cases και η ύπαρξη συγκεκριμένων requirement mappings. Τα SPARQL queries επέτρεψαν περισσότερο συγκεντρωτική παρουσίαση των ίδιων πληροφοριών, συνδυάζοντας πολλαπλές μεταβλητές σε ένα αποτέλεσμα. Η δυνατότητα αυτή ενισχύει τον χαρακτήρα της οντολογίας ως αναζητήσιμου μοντέλου γνώσης και συμφωνεί με τη γενικότερη κατεύθυνση προηγούμενων semantic approaches, όπου η δυνατότητα querying αποτελεί βασικό πλεονέκτημα της οντολογικής αναπαράστασης [5].

Το lightweight risk-related επίπεδο έδωσε ένα επιπλέον αποτέλεσμα διαφορετικής φύσης. Η βασική οντολογία μπορούσε να αναπαραστήσει την προέλευση, την αξιολόγηση και την τεκμηρίωση τόσο των true όσο και των false alarm cases. Για τα false alarms, όμως, δοκιμάστηκε αν το ίδιο μοντέλο μπορεί να δεχθεί επιπλέον risk-related σημασιολογική πληροφορία χωρίς να αλλάξει η βασική δομή του event. Η δημιουργία Risk Indicator και Risk-Relevant Event individuals και η σύνδεσή τους με Potential Harm έδειξε ότι ένα πραγματικό event μπορεί να αποκτήσει πρόσθετη σημασιολογική διάσταση όταν αυτό απαιτείται από το use case.

Η συγκεκριμένη επέκταση είναι σημαντική όχι επειδή αποδεικνύει ότι ένα συγκεκριμένο false alarm προκαλεί alert fatigue, αλλά επειδή δείχνει ότι το μοντέλο μπορεί να διαχωρίσει το πραγματικό γεγονός από την πιθανή risk-related ερμηνεία που προστίθεται πάνω σε αυτό. Το Alert Event παραμένει το πραγματικό VTaC event, ενώ το Risk-Relevant Event και το Potential Harm αποτελούν πρόσθετες οντολογικές έννοιες. Με αυτή τη διάκριση αποφεύγεται η μετατροπή μιας πιθανής σχέσης σε δήλωση βέβαιης αιτιότητας.

Σε ευρύτερο επίπεδο, τα αποτελέσματα ενισχύουν την άποψη ότι οι οντολογίες μπορούν να λειτουργήσουν ως ενδιάμεσο σημασιολογικό επίπεδο ανάμεσα σε γενικότερες αρχές ή απαιτήσεις και σε συγκεκριμένα στοιχεία ενός συστήματος. Η ontology-based requirements engineering βιβλιογραφία

έχει ήδη δείξει ότι σύνθετες ethicality έννοιες μπορούν να αποσαφηνιστούν και να χρησιμοποιηθούν για πιο συγκεκριμένη ανάλυση απαιτήσεων [27], ενώ άλλες Ethical AI ontologies οργανώνουν αρχές και συναφείς έννοιες σε δομημένη μορφή [28]. Στην παρούσα εφαρμογή, η ίδια γενικότερη λογική μεταφέρεται σε επίπεδο πραγματικών healthcare monitoring instances και διαθέσιμων τεκμηρίων.

Παράλληλα, τα αποτελέσματα δείχνουν και κάτι που αφορά τη σχέση μεταξύ οντολογικής αναπαράστασης και κανονιστικής τεκμηρίωσης. Η μετατροπή ενός κανονιστικού σημείου σε requirement individual και η δημιουργία ενός mapping προς evidence δεν αρκούν για να μετατρέψουν την οντολογία σε μηχανισμό νομικής συμμόρφωσης. Αυτό που επιτυγχάνεται είναι διαφορετικό: δημιουργείται μια ρητή σημασιολογική σχέση που μπορεί να ελεγχθεί, να αναζητηθεί και να επεκταθεί. Η αξία βρίσκεται επομένως στην οργανωμένη τεκμηρίωση και στην ικανότητα να απαντάται το ερώτημα «ποιο evidence έχει χαρτογραφηθεί προς ποιο σημείο αναφοράς και με ποιο πραγματικό event συνδέεται;».

Η ερμηνεία αυτή είναι σημαντική και για τον ρόλο του AI Act στην εργασία. Ο Κανονισμός δεν χρησιμοποιείται για να χαρακτηριστούν τα VTaC δεδομένα ως συμμορφούμενα ή μη συμμορφούμενα. Χρησιμοποιείται ως κανονιστικό πλαίσιο αναφοράς, από το οποίο επιλέχθηκαν συγκεκριμένες διαστάσεις που μπορούσαν να συνδεθούν με το διαθέσιμο use case και την οντολογική τεκμηρίωση. Με αυτόν τον τρόπο, το proof-of-concept εξετάζει τη δυνατότητα σημασιολογικής γεφύρωσης μεταξύ πραγματικών operational δεδομένων και κανονιστικά σχετικών εννοιών χωρίς να υπερβαίνει το επίπεδο πληροφορίας που είναι πραγματικά διαθέσιμο.

Συνολικά, τα αποτελέσματα δείχνουν ότι η προτεινόμενη οντολογία λειτουργεί κυρίως ως μηχανισμός οργάνωσης και σύνδεσης γνώσης. Η βασική της αξία δεν είναι η παραγωγή νέων κλινικών αποφάσεων, αλλά η δυνατότητα να μετατρέψει ένα σύνολο διαφορετικών και κατά τα άλλα αποσπασματικών πληροφοριών σε ένα συνεκτικό σημασιολογικό μοντέλο. Μέσα σε αυτό το μοντέλο, η προέλευση, η αξιολόγηση, η απόφαση, η τεκμηρίωση, τα requirement mappings και η πρόσθετη risk-related πληροφορία μπορούν να εξεταστούν ως διακριτές αλλά αλληλοσυνδεδεμένες διαστάσεις του ίδιου πραγματικού περιστατικού.

Η συνολική αυτή ερμηνεία παρέχει τη βάση για την επόμενη ενότητα, όπου τα αποτελέσματα αντιστοιχίζονται πλέον άμεσα στα ερευνητικά ερωτήματα που τέθηκαν στην αρχή της διπλωματικής εργασίας.

6.3 Απάντηση στα ερευνητικά ερωτήματα

Με βάση τα αποτελέσματα του σχεδιασμού, της υλοποίησης και της λειτουργικής επικύρωσης της οντολογίας, είναι πλέον δυνατό να δοθούν συγκεκριμένες απαντήσεις στο κεντρικό ερευνητικό ερώτημα και στα πέντε επιμέρους ερευνητικά ερωτήματα της εργασίας. Οι απαντήσεις που ακολουθούν περιορίζονται στα αποτελέσματα του συγκεκριμένου proof-of-concept και δεν επαναλαμβάνουν αναλυτικά τη διαδικασία υλοποίησης που παρουσιάστηκε στα προηγούμενα κεφάλαια.

Κεντρικό Ερευνητικό Ερώτημα (ΚΕΕ): Με ποιον τρόπο μπορεί μια οντολογία να υποστηρίξει τη σημασιολογική αναπαράσταση, την τεκμηρίωση και την ιχνηλασιμότητα πραγματικών

healthcare monitoring alarm events, καθώς και τη σύνδεσή τους με επιλεγμένα AI Act-related σημεία αναφοράς;

Η εργασία έδειξε ότι μια οντολογία μπορεί να υποστηρίξει τον παραπάνω στόχο οργανώνοντας διαφορετικές κατηγορίες πληροφορίας γύρω από το ίδιο πραγματικό alarm event και συνδέοντάς τις μέσω ρητών σημασιολογικών σχέσεων. Το event δεν αντιμετωπίζεται ως απομονωμένη εγγραφή, αλλά ως κεντρικό αντικείμενο γνώσης που μπορεί να συνδεθεί με το waveform record προέλευσής του, την ανθρώπινη expert αξιολόγηση, την τελική απόφαση και το Evidence Record που χρησιμοποιείται για την τεκμηρίωσή του.

Η σύνδεση με τα επιλεγμένα AI Act-related σημεία αναφοράς πραγματοποιείται μέσω του Evidence Record και όχι απευθείας από το event. Με αυτόν τον τρόπο δημιουργείται ένα requirement-to-evidence mapping στο οποίο διατηρείται σαφής διάκριση μεταξύ πραγματικού περιστατικού, διαθέσιμης τεκμηρίωσης και κανονιστικού σημείου αναφοράς.

Η λειτουργική επικύρωση έδειξε ότι οι συγκεκριμένες σχέσεις μπορούν να ανακτηθούν μέσω DL competency questions και SPARQL queries. Επομένως, η απάντηση στο κεντρικό ερευνητικό ερώτημα είναι θετική μέσα στα όρια της συγκεκριμένης εφαρμογής: η οντολογική προσέγγιση μπορεί να λειτουργήσει ως σημασιολογικό επίπεδο οργάνωσης και ιχνηλασιμότητας πραγματικών healthcare monitoring events και να υποστηρίξει ελέγξιμα mappings μεταξύ event-level evidence και επιλεγμένων AI Act-related σημείων αναφοράς.

EE1: Πώς μπορεί να αναπαρασταθεί οντολογικά ένα πραγματικό healthcare monitoring alarm event, ώστε να συνδέεται με το waveform record προέλευσής του, την expert αξιολόγηση, την τελική απόφαση και την τεκμηρίωσή του;

Η υλοποίηση έδειξε ότι ένα πραγματικό alarm event μπορεί να αναπαρασταθεί ως individual της class Ventricular Tachycardia Alarm και να λειτουργήσει ως κεντρικό σημείο σύνδεσης διαφορετικών πληροφοριών.

Η προέλευσή του αναπαρίσταται μέσω της σχέσης με το αντίστοιχο Waveform Record. Η ανθρώπινη αξιολόγηση οργανώνεται μέσω της σύνδεσης του Alert Event με Expert Review και στη συνέχεια με Review Decision. Παράλληλα, το ίδιο event συνδέεται με Evidence Record, το οποίο αναπαριστά το επίπεδο της τεκμηρίωσης.

Σημαντικό αποτέλεσμα είναι ότι οι πληροφορίες αυτές δεν χρειάζεται να ενσωματώνονται όλες ως απλές τιμές πάνω στο event. Η αναπαράστασή τους ως διακριτά individuals επιτρέπει να διατηρούνται χωριστά η προέλευση, η αξιολόγηση, η απόφαση και η τεκμηρίωση, ενώ οι μεταξύ τους σχέσεις παραμένουν ρητές και αναζητήσιμες.

Τα αποτελέσματα των DL Queries για τα true και false alarms και το SPARQL Query 1 επιβεβαίωσαν ότι οι αντίστοιχες πληροφορίες μπορούν να ανακτηθούν για τα 12 επιλεγμένα VTaC events. Επομένως, το EE1 απαντάται μέσω μιας event-centered οντολογικής αναπαράστασης, στην οποία το πραγματικό alarm event λειτουργεί ως κοινό σημείο αναφοράς για τις διαφορετικές σημασιολογικές διαδρομές.

EE2: Πώς μπορούν επιλεγμένα AI Act-related σημεία αναφοράς, όπως logging, human review, traceability και risk management, να συνδεθούν με πραγματικά VTaC events μέσω evidence records;

Η σύνδεση επιτυγχάνεται μέσω της εισαγωγής του Evidence Record ως ενδιάμεσου τεκμηριωτικού κόμβου.

Η βασική λογική είναι:

Alert Event → Evidence Record → Requirement individual

Το event συνδέεται με το Evidence Record μέσω του isDocumentedBy, ενώ το Evidence Record συνδέεται με τα επιλεγμένα requirement individuals μέσω του supportsRequirement.

Με αυτόν τον σχεδιασμό, τα req_logging, req_human_review, req_traceability και req_risk_management δεν αποδίδονται απευθείας στο πραγματικό event. Το mapping πραγματοποιείται πάνω στο evidence που έχει δημιουργηθεί στο πλαίσιο της οντολογικής αναπαράστασης.

Τα λειτουργικά αποτελέσματα έδειξαν ότι και τα 12 Evidence Record individuals έχουν χαρτογραφηθεί προς τα req_logging, req_human_review και req_traceability, ενώ τα evidence records των έξι false alarm cases έχουν συνδεθεί επιπλέον με το req_risk_management.

Η απάντηση στο EE2 είναι επομένως ότι τα AI Act-related σημεία αναφοράς μπορούν να συνδεθούν με πραγματικά events μέσω ενός ρητού requirement-to-evidence mapping. Το mapping αυτό παρέχει οργανωμένη και αναζητήσιμη τεκμηρίωση, χωρίς όμως να αποτελεί απόδειξη πλήρους νομικής συμμόρφωσης με τον AI Act.

EE3: Πώς μπορεί η οντολογία να υποστηρίξει την ιχνηλασιμότητα μεταξύ του waveform record, του alarm event, της expert αξιολόγησης, του Evidence Record και του requirement-to-evidence mapping;

Τα αποτελέσματα έδειξαν ότι η ιχνηλασιμότητα υποστηρίζεται μέσω της σύνθεσης πολλαπλών σημασιολογικών σχέσεων γύρω από το ίδιο Alert Event.

Η προέλευση του event διατηρείται μέσω της σύνδεσης με το Waveform Record. Η ανθρώπινη αξιολόγηση και η τελική απόφαση αναπαρίστανται μέσω της διαδρομής Expert Review–Review Decision. Παράλληλα, το Event συνδέεται με Evidence Record, το οποίο στη συνέχεια χαρτογραφείται προς το αντίστοιχο requirement individual.

Η ιχνηλασιμότητα επομένως δεν εξαρτάται από μία μοναδική γραμμική σχέση, αλλά από τη δυνατότητα να συνδυάζονται και να ανακτώνται οι διαφορετικές αυτές διαδρομές.

Το DL Query 11 αποτέλεσε τον βασικό λειτουργικό έλεγχο αυτού του ερευνητικού ερωτήματος. Και τα 12 VT alarm events ικανοποίησαν ταυτόχρονα τις συνθήκες σύνδεσης με Waveform Record, Expert Review και Review Decision και με Evidence Record που έχει χαρτογραφηθεί προς το

`req_traceability`. Παράλληλα, τα SPARQL queries έδειξαν ότι οι πληροφορίες μπορούν να παρουσιαστούν συγκεντρωτικά.

Επομένως, το EE3 απαντάται θετικά για το συγκεκριμένο proof-of-concept: η ontology μπορεί να υποστηρίξει ιχνηλασιμότητα όχι μόνο μέσω identifiers, αλλά μέσω ρητών και ανακτήσιμων σημασιολογικών σχέσεων μεταξύ της προέλευσης, της αξιολόγησης, της τεκμηρίωσης και του requirement mapping.

EE4: Πώς μπορούν τα false alarm cases να αποκτήσουν πρόσθετη risk-related σημασιολογική αναπαράσταση μέσα στην οντολογία;

Η εργασία έδειξε ότι τα false alarm cases μπορούν να αποκτήσουν πρόσθετη risk-related αναπαράσταση χωρίς να μεταβάλλεται η βασική τους ταυτότητα ως πραγματικά VTaC alarm events.

Για κάθε ένα από τα έξι επιλεγμένα false alarm cases δημιουργήθηκαν Risk Indicator και Risk-Relevant Event individuals. Το Alert Event συνδέεται χωριστά με τα δύο αυτά στοιχεία, ενώ το Risk-Relevant Event μπορεί στη συνέχεια να συνδέεται μέσω `mayLeadTo` με Potential Harm.

Στην υλοποίηση χρησιμοποιήθηκε το `vtac_harm_alert_fatigue` ως Potential Harm. Η συγκεκριμένη σχέση έχει σκόπιμα διατυπωθεί ως πιθανή και όχι ως βέβαιη αιτιώδης σύνδεση. Κατά συνέπεια, η οντολογία δεν δηλώνει ότι ένα συγκεκριμένο false alarm προκάλεσε alert fatigue. Δηλώνει ότι το επιλεγμένο case έχει ενταχθεί σε μία πρόσθετη σημασιολογική δομή μέσα στην οποία μπορεί να αναπαρασταθεί πιθανή risk-related σημασία.

Τα DL Queries 8 και 9 και το SPARQL Query 3 ανέκτησαν τα έξι αντίστοιχα cases και τις risk-related σχέσεις τους. Το αποτέλεσμα δείχνει ότι η βασική event representation μπορεί να επεκταθεί με πρόσθετες έννοιες κινδύνου χωρίς να απαιτείται αλλαγή της αρχικής πληροφορίας του dataset.

Επομένως, το EE4 απαντάται μέσω ενός lightweight risk-related επιπέδου το οποίο εμπλουτίζει σημασιολογικά τα false alarm cases, χωρίς να ισχυρίζεται πραγματική βλάβη ή πλήρη υλοποίηση risk management system.

EE5: Με ποιον τρόπο μπορεί να ελεγχθεί λειτουργικά η προτεινόμενη οντολογία μέσω reasoner, DL competency questions και SPARQL queries;

Η λειτουργική επικύρωση πραγματοποιήθηκε σε τρία συμπληρωματικά επίπεδα.

Πρώτον, ο HermiT reasoner χρησιμοποιήθηκε για τον έλεγχο της λογικής συνέπειας και της inferred class hierarchy. Στην εξεταζόμενη έκδοση δεν εντοπίστηκε inconsistency ή unsatisfiable class.

Δεύτερον, τα 12 DL competency questions χρησιμοποιήθηκαν για την εξέταση συγκεκριμένων λογικών συνθηκών του μοντέλου. Μέσω αυτών ανακτήθηκαν τα VTaC alarm events, οι κατηγορίες true και false, τα requirement mappings, οι risk-related σχέσεις και η προβλεπόμενη διαδρομή ιχνηλασιμότητας.

Τρίτον, τα τέσσερα SPARQL queries χρησιμοποιήθηκαν για τη συγκεντρωτική ανάκτηση πολλών μεταβλητών και σχέσεων. Με αυτά παρουσιάστηκαν η σύνδεση event–record–decision–evidence, το requirement-to-evidence mapping, η risk-related πληροφορία των false alarm cases και η κατανομή των events στα waveform records.

Οι τρεις μορφές ελέγχου δεν εξετάζουν ακριβώς το ίδιο στοιχείο και γι' αυτό λειτουργούν συμπληρωματικά. Ο reasoner εξετάζει τη λογική συνέπεια, τα DL competency questions ελέγχουν συγκεκριμένες σημασιολογικές συνθήκες και τα SPARQL queries υποστηρίζουν περισσότερο σύνθετη και συγκεντρωτική ανάκτηση δεδομένων.

Επομένως, το EE5 απαντάται με μια συνδυασμένη διαδικασία λογικού και λειτουργικού ελέγχου. Τα αποτελέσματα έδειξαν ότι η ontology συμπεριφέρεται σύμφωνα με τον σχεδιασμό της για τα συγκεκριμένα proof-of-concept instances. Η επικύρωση αυτή αφορά τη λογική και λειτουργική συμπεριφορά του μοντέλου και δεν πρέπει να συγχέεται με κλινική, νομική ή πλήρη εξωτερική επικύρωση.

Συνολικά, οι απαντήσεις στα πέντε επιμέρους ερευνητικά ερωτήματα συγκλίνουν στην απάντηση του κεντρικού ερευνητικού ερωτήματος. Η οντολογική προσέγγιση που αναπτύχθηκε επιτρέπει την οργανωμένη αναπαράσταση πραγματικών VTaC alarm events, τη σύνδεσή τους με την προέλευση και την ανθρώπινη αξιολόγησή τους, την τεκμηρίωσή τους μέσω Evidence Records, τη δημιουργία ελέγξιμων requirement mappings και, όπου αυτό έχει προβλεφθεί, την προσθήκη risk-related πληροφορίας.

Η λειτουργική επικύρωση έδειξε ότι οι συγκεκριμένες σχέσεις μπορούν να ανακτηθούν με συνέπεια στο πλαίσιο της proof-of-concept υλοποίησης. Με βάση αυτό το αποτέλεσμα, η εργασία παρέχει θετική απάντηση στο βασικό ερευνητικό ερώτημα, με την απαραίτητη προϋπόθεση ότι η απάντηση αφορά το συγκεκριμένο περιορισμένο use case και όχι γενική απόδειξη εφαρμοσιμότητας ή κανονιστικής συμμόρφωσης για οποιοδήποτε πραγματικό σύστημα τεχνητής νοημοσύνης.

6.4 Συμβολή της εργασίας

Η συμβολή της παρούσας διπλωματικής πρέπει να αξιολογηθεί μέσα στο συγκεκριμένο πεδίο και στα όρια του proof-of-concept που αναπτύχθηκε. Η εργασία δεν προτείνει μια γενική οντολογία για ολόκληρο τον AI Act, ούτε ένα πλήρες σύστημα διαχείρισης κανονιστικής συμμόρφωσης. Η κύρια συμβολή της βρίσκεται στον εφαρμοσμένο συνδυασμό πραγματικών healthcare monitoring alarm events, οντολογικής τεκμηρίωσης και επιλεγμένων AI Act-related σημείων αναφοράς σε ένα ενιαίο, λειτουργικά ελέγξιμο σημασιολογικό μοντέλο.

Μια πρώτη διάσταση της συμβολής αφορά τη χρήση πραγματικών δεδομένων. Αντί να βασιστεί αποκλειστικά σε αφηρημένα ή υποθετικά παραδείγματα, η εργασία αξιοποίησε πραγματικά VTaC waveform records, ventricular tachycardia alarm events και διαθέσιμες expert αποφάσεις. Η επιλογή αυτή επιτρέπει η προτεινόμενη οντολογική δομή να εφαρμόζεται πάνω σε πραγματικές περιπτώσεις και να εξετάζεται αν οι σχεδιασμένες σχέσεις μπορούν να υποστηριχθούν από συγκεκριμένα instances. Η

συμβολή δεν βρίσκεται επομένως στην παραγωγή νέων κλινικών δεδομένων, αλλά στη σημασιολογική αξιοποίηση και οργάνωση ήδη διαθέσιμων πραγματικών πληροφοριών.

Δεύτερη σημαντική διάσταση αποτελεί η εισαγωγή του Evidence Record ως διακριτού τεκμηριωτικού κόμβου. Η σχεδιαστική αυτή επιλογή επιτρέπει να διαχωρίζεται το πραγματικό alarm event από το evidence που χρησιμοποιείται για την τεκμηρίωσή του και από το κανονιστικό σημείο αναφοράς προς το οποίο το evidence έχει χαρτογραφηθεί. Με τον τρόπο αυτό, το requirement-to-evidence mapping αποκτά σαφή και ελέγξιμη σημασιολογική δομή και αποφεύγεται η απευθείας σύνδεση ενός πραγματικού περιστατικού με έναν ισχυρισμό κανονιστικής συμμόρφωσης.

Η διάσταση αυτή διαφοροποιεί το συγκεκριμένο proof-of-concept από οντολογικές προσεγγίσεις που επικεντρώνονται κυρίως στην αναπαράσταση κανονιστικών concepts και requirements. Η TAIR, για παράδειγμα, παρέχει μηχανισμό σημασιολογικής οργάνωσης και χαρτογράφησης concepts και normative requirements από τον AI Act και standards [5]. Αντίστοιχα, άλλες προσεγγίσεις οργανώνουν ethicality requirements ή διαδικασίες σχετικές με τον AI Act σε δομημένη μορφή [17], [27], [28]. Η παρούσα εργασία δεν αντικαθιστά αυτές τις προσεγγίσεις, αλλά εξετάζει ένα διαφορετικό εφαρμοσμένο επίπεδο: τη σύνδεση πραγματικού event-level operational evidence με επιλεγμένα κανονιστικά σημεία αναφοράς.

Μια τρίτη διάσταση της συμβολής είναι η event-centered αναπαράσταση της ιχνηλασιμότητας. Αντί η traceability να αντιμετωπίζεται μόνο ως identifier ή ως μία απλή γραμμική ακολουθία, οργανώνεται μέσω διαφορετικών αλλά συνδεδεμένων σημασιολογικών διαδρομών γύρω από το ίδιο Alert Event. Η προέλευση συνδέεται με το Waveform Record, η ανθρώπινη αξιολόγηση με Expert Review και Review Decision και η τεκμηρίωση με Evidence Record και requirement mappings. Η δομή αυτή επιτρέπει διαφορετικές κατευθύνσεις αναζήτησης και διατηρεί σαφή διάκριση μεταξύ των διαφορετικών τύπων πληροφορίας.

Τέταρτη διάσταση αποτελεί η προσθήκη του lightweight risk-related επιπέδου για τα false alarm cases. Η εργασία δείχνει πώς η βασική αναπαράσταση ενός πραγματικού event μπορεί να επεκταθεί με Risk Indicator, Risk-Relevant Event και Potential Harm χωρίς να τροποποιείται η αρχική expert απόφαση του dataset και χωρίς να δηλώνεται μη τεκμηριωμένη αιτιώδης σχέση. Η επιλογή αυτή παρέχει ένα παράδειγμα του τρόπου με τον οποίο ένα βασικό event model μπορεί να εμπλουτίζεται με πρόσθετες σημασιολογικές διαστάσεις ανάλογα με τις ανάγκες του use case.

Πέμπτη διάσταση της συμβολής αφορά τη λειτουργική δυνατότητα ελέγχου του μοντέλου. Η ontology δεν περιορίστηκε στη δημιουργία μιας ιεραρχίας classes και στην οπτική παρουσίασή της στο Protégé. Σχεδιάστηκε ώστε συγκεκριμένα ερωτήματα να μπορούν να εξεταστούν μέσω reasoner, DL competency questions και SPARQL queries. Με αυτόν τον τρόπο, οι βασικές σχεδιαστικές επιλογές μετατράπηκαν σε ελέγξιμες σημασιολογικές συνθήκες και σε αναζητήσιμες διαδρομές πληροφορίας.

Ιδιαίτερη μεθοδολογική αξία έχει επίσης η σαφής διάκριση ανάμεσα σε πραγματικά δεδομένα και σε οντολογική μοντελοποίηση. Τα VTaC records, events και expert decisions έχουν πραγματική βάση στο dataset, ενώ τα Evidence Records, τα requirement individuals και τα risk-related individuals αποτελούν σημασιολογικές κατασκευές της παρούσας εργασίας. Η διάκριση αυτή επιτρέπει να είναι σαφές ποια πληροφορία προέρχεται από το πραγματικό use case και ποια προστίθεται από το προτεινόμενο οντολογικό μοντέλο.

Με βάση τη βιβλιογραφία που εξετάστηκε στην παρούσα εργασία, δεν εντοπίστηκε προσέγγιση που να συνδυάζει με τον ίδιο ακριβώς τρόπο πραγματικά VTaC alarm events, expert decisions, event-level Evidence Records, requirement-to-evidence mapping προς επιλεγμένα AI Act-related σημεία αναφοράς και lightweight risk-related αναπαράσταση. Η διαπίστωση αυτή δεν συνεπάγεται ότι δεν υπάρχουν ή δεν μπορούν να αναπτυχθούν άλλες συναφείς προσεγγίσεις. Οριοθετεί τον ειδικό συνδυασμό στοιχείων στον οποίο τοποθετείται η συμβολή της συγκεκριμένης διπλωματικής.

Συνολικά, η συμβολή της εργασίας μπορεί να συνοψιστεί στη μετάβαση από ένα πραγματικό αλλά σχετικά αποσπασματικό σύνολο operational πληροφοριών σε ένα ενιαίο και αναζητήσιμο σημασιολογικό πλαίσιο. Το πλαίσιο αυτό συνδέει πραγματικά healthcare monitoring events με την προέλευση, την ανθρώπινη αξιολόγηση, την τεκμηρίωση, επιλεγμένα AI Act-related mappings και, όπου εφαρμόζεται, πρόσθετη risk-related πληροφορία. Η αξία της προσέγγισης βρίσκεται επομένως κυρίως στη δομημένη σημασιολογική σύνδεση αυτών των επιπέδων και στη δυνατότητα λειτουργικού ελέγχου των μεταξύ τους σχέσεων.

6.5 Περιορισμοί της εργασίας

Τα αποτελέσματα της παρούσας εργασίας πρέπει να ερμηνεύονται μέσα στο συγκεκριμένο πεδίο εφαρμογής και στα όρια του proof-of-concept που αναπτύχθηκε. Η λειτουργική επικύρωση έδειξε ότι η προτεινόμενη οντολογική δομή μπορεί να αναπαραστήσει και να ανακτήσει τις σημασιολογικές σχέσεις για τις οποίες σχεδιάστηκε. Ωστόσο, το αποτέλεσμα αυτό δεν επιτρέπει γενίκευση πέρα από το συγκεκριμένο use case χωρίς πρόσθετη μελέτη και επέκταση του μοντέλου.

Ένας πρώτος περιορισμός αφορά το εύρος των δεδομένων που χρησιμοποιήθηκαν. Η υλοποίηση εφαρμόστηκε σε 12 επιλεγμένα VTaC alarm events και όχι στο σύνολο του dataset. Η επιλογή αυτή ήταν επαρκής για τον λειτουργικό έλεγχο διαφορετικών περιπτώσεων μέσα στο proof-of-concept, αλλά δεν επιτρέπει στατιστικά συμπεράσματα σχετικά με τη συνολική κατανομή ή συμπεριφορά των alarms. Τα αποτελέσματα αφορούν επομένως τη δυνατότητα σημασιολογικής αναπαράστασης των συγκεκριμένων cases και όχι την αντιπροσωπευτικότητα ενός μεγαλύτερου πληθυσμού δεδομένων.

Αντίστοιχα, η εργασία περιορίζεται σε ventricular tachycardia alarm events και δεν εξετάζει άλλα είδη clinical alarms ή διαφορετικά healthcare monitoring scenarios. Παρότι η αρχιτεκτονική έχει σχεδιαστεί με τρόπο που επιτρέπει μελλοντική επέκταση, δεν μπορεί να θεωρηθεί δεδομένο ότι το ίδιο μοντέλο θα εφαρμόζοταν χωρίς τροποποιήσεις σε διαφορετικούς τύπους alarms, άλλες κλινικές ειδικότητες ή διαφορετικές πηγές δεδομένων.

Ένας δεύτερος σημαντικός περιορισμός αφορά το επίπεδο της διαθέσιμης κλινικής πληροφορίας. Η οντολογία αξιοποιεί τα waveform record identifiers, τα alarm event identifiers και τις διαθέσιμες expert αποφάσεις του VTaC, αλλά δεν πραγματοποιεί ανάλυση των ίδιων των waveform signals ούτε διαθέτει πλήρες κλινικό πλαίσιο για κάθε περιστατικό. Για τον λόγο αυτό, τα αποτελέσματα δεν μπορούν να χρησιμοποιηθούν για κλινική ερμηνεία, αξιολόγηση της πραγματικής επίπτωσης ενός alarm ή εκτίμηση της κατάστασης συγκεκριμένου ασθενούς.

Ο περιορισμός αυτός είναι ιδιαίτερα σημαντικός και για το lightweight risk-related επίπεδο. Η σύνδεση των false alarm cases με Risk Indicator, Risk-Relevant Event και Potential Harm αποτελεί

σημασιολογική αναπαράσταση πιθανής risk-related σημασίας. Δεν τεκμηριώνει ότι ένα συγκεκριμένο false alarm προκάλεσε alert fatigue ή άλλη πραγματική βλάβη. Επομένως, το risk-related μέρος πρέπει να θεωρείται παράδειγμα επέκτασης του μοντέλου και όχι κλινική ή ποσοτική αξιολόγηση κινδύνου.

Τρίτος περιορισμός αφορά το κανονιστικό εύρος της εργασίας. Η οντολογία χρησιμοποιεί τέσσερα εφαρμοσμένα AI Act-related σημεία αναφοράς — `req_logging`, `req_human_review`, `req_traceability` και `req_risk_management` — και δεν επιχειρεί να αναπαραστήσει το σύνολο του AI Act. Κατά συνέπεια, τα αποτελέσματα δεν μπορούν να χρησιμοποιηθούν για την αξιολόγηση συνολικής συμμόρφωσης ενός πραγματικού AI system με τον Κανονισμό. Το requirement-to-evidence mapping δείχνει ότι μπορούν να δημιουργηθούν ρητές και αναζητήσιμες σημασιολογικές συνδέσεις, όχι ότι οι αντίστοιχες νομικές υποχρεώσεις έχουν πλήρως ικανοποιηθεί.

Στο ίδιο πλαίσιο, η εργασία δεν αξιολογεί το αρχικό VTaC monitoring system ως AI system ούτε ως high-risk AI system. Τα VTaC events χρησιμοποιούνται ως πραγματικά operational δεδομένα πάνω στα οποία δοκιμάζεται η οντολογική προσέγγιση. Επομένως, η σύνδεση με τον AI Act αφορά το επίπεδο της σημασιολογικής τεκμηρίωσης και του requirement-to-evidence mapping και όχι τον νομικό χαρακτηρισμό του ίδιου του dataset ή του αρχικού monitoring environment.

Ένας ακόμη περιορισμός αφορά τη φύση της επικύρωσης. Η αξιολόγηση που πραγματοποιήθηκε ήταν λογική και λειτουργική. Ο HerMiT reasoner χρησιμοποιήθηκε για τον έλεγχο συνέπειας, ενώ τα DL competency questions και τα SPARQL queries εξέτασαν αν οι βασικές σχέσεις και τα αναμενόμενα instances μπορούν να ανακτηθούν σύμφωνα με τον σχεδιασμό. Η επιτυχής εκτέλεση αυτών των ελέγχων δείχνει ότι το μοντέλο λειτουργεί όπως προβλέφθηκε στο συγκεκριμένο proof-of-concept, αλλά δεν αποτελεί πλήρη ontology quality evaluation.

Δεν πραγματοποιήθηκε ανεξάρτητη αξιολόγηση του μοντέλου από επαγγελματίες υγείας, νομικούς ειδικούς στον AI Act ή ανεξάρτητους ontology engineers. Επομένως, η κλινική καταλληλότητα των εννοιών, η νομική επάρκεια των mappings και η ευρύτερη ποιότητα της οντολογίας δεν έχουν επικυρωθεί εξωτερικά. Η διάσταση αυτή αποτελεί σημαντικό πεδίο για μελλοντική αξιολόγηση.

Τέλος, η υλοποίηση πραγματοποιήθηκε σε περιβάλλον Protégé και σε περιορισμένο αριθμό instances. Δεν εξετάστηκαν ζητήματα κλιμάκωσης σε πολύ μεγαλύτερο όγκο δεδομένων, αυτοματοποιημένης εισαγωγής νέων events, real-world integration με healthcare information systems ή συνεχούς ενημέρωσης της οντολογίας. Συνεπώς, τα αποτελέσματα δεν παρέχουν ακόμη στοιχεία σχετικά με την απόδοση και την πρακτική χρηστικότητα μιας αντίστοιχης προσέγγισης σε παραγωγικό περιβάλλον.

Οι παραπάνω περιορισμοί δεν αναιρούν τα αποτελέσματα του proof-of-concept, αλλά καθορίζουν το εύρος μέσα στο οποίο αυτά μπορούν να υποστηριχθούν. Η εργασία δείχνει ότι η συγκεκριμένη οντολογική προσέγγιση είναι λειτουργικά εφαρμόσιμη στα επιλεγμένα VTaC cases και ότι μπορεί να οργανώσει με συνεκτικό τρόπο event-level information, evidence και AI Act-related mappings. Η περαιτέρω αξιολόγηση της γενικευσιμότητας, της κλινικής καταλληλότητας, της κανονιστικής πληρότητας και της εφαρμογής σε πραγματικές υποδομές απαιτεί πρόσθετη έρευνα.

6.6 Μελλοντικές επεκτάσεις

Η παρούσα εργασία σχεδιάστηκε ως proof-of-concept και επομένως δημιουργεί τη βάση για αρκετές κατευθύνσεις περαιτέρω ανάπτυξης. Οι μελλοντικές επεκτάσεις δεν αφορούν μόνο την αύξηση του αριθμού των individuals, αλλά και τη διεύρυνση του κανονιστικού, κλινικού και τεχνικού πεδίου της οντολογίας.

Μια πρώτη και άμεσα εφαρμόσιμη επέκταση είναι η αξιοποίηση μεγαλύτερου μέρους του VTaC dataset. Η παρούσα υλοποίηση βασίστηκε σε 12 επιλεγμένα events, ώστε να είναι δυνατός ο ελεγχόμενος σχεδιασμός και η λειτουργική επικύρωση του proof-of-concept. Σε επόμενη φάση, η ίδια βασική δομή θα μπορούσε να εφαρμοστεί σε σημαντικά μεγαλύτερο αριθμό από τα 5.037 διαθέσιμα annotated events του dataset [4]. Μια τέτοια επέκταση θα επέτρεπε να εξεταστούν ζητήματα κλιμάκωσης της οντολογίας, αποτελεσματικότητας των queries και διαχείρισης μεγαλύτερου όγκου individuals και σχέσεων.

Η επέκταση αυτή θα μπορούσε να συνδυαστεί με αυτοματοποιημένη εισαγωγή δεδομένων. Στην παρούσα εργασία τα individuals και οι βασικές σχέσεις δημιουργήθηκαν στο πλαίσιο μιας ελεγχόμενης proof-of-concept υλοποίησης. Μελλοντικά, θα μπορούσε να αναπτυχθεί μηχανισμός μετατροπής των διαθέσιμων VTaC metadata σε ontology instances, ώστε νέα waveform records, alarm events, expert decisions και οι βασικές μεταξύ τους σχέσεις να δημιουργούνται συστηματικά. Με αυτόν τον τρόπο θα μπορούσε να εξεταστεί η μετάβαση από μια χειροκίνητα οργανωμένη εφαρμογή σε μια περισσότερο αυτοματοποιημένη διαδικασία σημασιολογικής ενσωμάτωσης δεδομένων.

Μια δεύτερη σημαντική κατεύθυνση αφορά τη διεύρυνση του κανονιστικού επιπέδου. Η σημερινή ontology χρησιμοποιεί τέσσερα επιλεγμένα AI Act-related σημεία αναφοράς που σχετίζονται με logging, human review, traceability και risk management. Μελλοντικά, η αναπαράσταση θα μπορούσε να επεκταθεί σε πρόσθετες διαστάσεις του AI Act, όπως τεχνική τεκμηρίωση, data governance, transparency, accuracy, robustness και cybersecurity, όπου αυτές είναι σχετικές με το αντίστοιχο use case [3]. Η επέκταση αυτή θα απαιτούσε λεπτομερέστερη σημασιολογική αποτύπωση των σχετικών διατάξεων και σαφέστερη διάκριση μεταξύ νομικών υποχρεώσεων, τεχνικών απαιτήσεων και διαθέσιμων evidence types.

Παράλληλα, μια πιο ώριμη έκδοση της προσέγγισης θα μπορούσε να συνδεθεί με υφιστάμενες AI governance ontologies και σημασιολογικά λεξιλόγια. Αντί το κανονιστικό επίπεδο να παραμένει αποκλειστικά application-specific, θα μπορούσαν να εξεταστούν mappings προς έννοιες από προσεγγίσεις όπως TAIR, AIRO ή FRIA ontology [5], [16], [17]. Μια τέτοια εξέλιξη θα ενίσχυε τη σημασιολογική επαναχρησιμοποίηση και θα επέτρεπε να εξεταστεί κατά πόσο το event-level evidence model της παρούσας εργασίας μπορεί να συνεργαστεί με γενικότερα μοντέλα AI governance.

Τρίτη κατεύθυνση είναι η εφαρμογή της προσέγγισης σε διαφορετικά healthcare monitoring datasets και σε άλλα είδη alarm events. Η σημερινή υλοποίηση περιορίζεται σε ventricular tachycardia alarms. Η σύνδεση με δεδομένα διαφορετικών τύπων αρρυθμιών, άλλων φυσιολογικών παραμέτρων ή διαφορετικών monitoring environments θα επέτρεπε να εξεταστεί ποια μέρη της ontology μπορούν να επαναχρησιμοποιηθούν και ποια απαιτούν εξειδίκευση. Με τον τρόπο αυτό θα μπορούσε να αξιολογηθεί στην πράξη η γενικευσιμότητα της βασικής event-centered αρχιτεκτονικής.

Ιδιαίτερο ενδιαφέρον παρουσιάζει και η περαιτέρω ανάπτυξη του risk-related επιπέδου. Στην παρούσα μορφή του είναι σκόπιμα lightweight και χρησιμοποιείται κυρίως για να δείξει ότι ένα false alarm case μπορεί να αποκτήσει πρόσθετη risk-related σημασιολογική αναπαράσταση. Μια μελλοντική έκδοση θα μπορούσε να περιλαμβάνει περισσότερους τύπους Risk Indicator και Potential Harm, πληροφορίες για τη σοβαρότητα ή το πλαίσιο ενός κινδύνου και σαφέστερες σχέσεις μεταξύ event, πιθανής συνέπειας και αντίστοιχης τεκμηρίωσης. Η ανάπτυξη αυτή θα πρέπει να πραγματοποιηθεί με ιδιαίτερη προσοχή, ώστε η οντολογία να μην μετατρέψει πιθανές risk-related σχέσεις σε μη τεκμηριωμένους ισχυρισμούς πραγματικής αιτιότητας.

Μια ακόμη σημαντική εξέλιξη αφορά την εξωτερική αξιολόγηση της οντολογίας. Η παρούσα εργασία επικύρωσε τη λογική και λειτουργική συμπεριφορά του μοντέλου, αλλά μια επόμενη φάση θα μπορούσε να περιλαμβάνει αξιολόγηση από διαφορετικές ομάδες ειδικών. Επαγγελματίες υγείας θα μπορούσαν να εξετάσουν την κλινική καταλληλότητα της event representation, ειδικοί σε θέματα AI Act τη σημασιολογική αντιστοίχιση των regulatory concepts και ontology engineers τη συνοχή, την επαναχρησιμοποίηση και την ποιότητα του μοντέλου. Με αυτόν τον τρόπο η αξιολόγηση θα μπορούσε να επεκταθεί πέρα από τον εσωτερικό λειτουργικό έλεγχο.

Σε πιο τεχνικό επίπεδο, θα μπορούσε να εξεταστεί η σύνδεση της ontology με πραγματικές ροές δεδομένων από healthcare monitoring systems. Σε ένα τέτοιο σενάριο, νέα monitoring events θα μπορούσαν να εισάγονται στο σημασιολογικό μοντέλο μαζί με metadata σχετικά με την προέλευση, την αξιολόγηση και την τεκμηρίωσή τους. Τα queries θα μπορούσαν στη συνέχεια να χρησιμοποιούνται για την αναζήτηση evidence, την υποστήριξη audit διαδικασιών ή την ανίχνευση περιπτώσεων στις οποίες απουσιάζει κάποια αναμενόμενη πληροφορία. Η κατεύθυνση αυτή θα απαιτούσε σημαντικά μεγαλύτερη τεχνική υποδομή και δεν αποτελεί μέρος της παρούσας υλοποίησης.

Τέλος, ως πιο μακροπρόθεσμη κατεύθυνση μπορεί να εξεταστεί η εφαρμογή αντίστοιχων σημασιολογικών μηχανισμών σε connected healthcare environments που αξιοποιούν IoT, edge/cloud υποδομές και μελλοντικά δίκτυα επικοινωνιών 5G/6G. Τα AI-native 6G healthcare environments αναμένεται να υποστηρίζουν εφαρμογές όπως continuous patient monitoring, όπου δεδομένα μπορούν να διακινούνται μεταξύ αισθητήρων, edge devices, δικτυακών υποδομών και cloud platforms [31]. Η αυξημένη κατανομή και διασύνδεση αυτών των περιβαλλόντων ενισχύει παράλληλα την ανάγκη για διαλειτουργικά και machine-readable πλαίσια που μπορούν να συνδέουν τεχνικές πληροφορίες με απαιτήσεις κανονιστικής και δεοντολογικής διακυβέρνησης [31]. Η επέκταση προς 6G δεν αποτελεί μέρος του σημερινού proof-of-concept, αλλά μπορεί να θεωρηθεί μια μακροπρόθεσμη ερευνητική κατεύθυνση στην οποία η βασική λογική event–evidence–normative mapping της παρούσας εργασίας θα μπορούσε να διερευνηθεί σε περισσότερο σύνθετα και κατανεμημένα healthcare environments.

Οι παραπάνω κατευθύνσεις δείχνουν ότι η παρούσα ontology μπορεί να αντιμετωπιστεί όχι ως ολοκληρωμένο τελικό σύστημα, αλλά ως αρχικό εφαρμοσμένο μοντέλο πάνω στο οποίο μπορούν να προστεθούν περισσότερα δεδομένα, πλουσιότερη κανονιστική γνώση, μεγαλύτερη αυτοματοποίηση, εξωτερική αξιολόγηση και σύνδεση με πραγματικές τεχνικές υποδομές. Η αξία των μελλοντικών επεκτάσεων θα εξαρτηθεί από τη διατήρηση της βασικής αρχής της εργασίας: τη σαφή διάκριση μεταξύ πραγματικού event, διαθέσιμου evidence, σημασιολογικής μοντελοποίησης και κανονιστικής ερμηνείας.

6.7 Τελικό συμπέρασμα

Η παρούσα διπλωματική εργασία διερεύνησε τη δυνατότητα αξιοποίησης μιας οντολογικής προσέγγισης για τη σημασιολογική αναπαράσταση και τεκμηρίωση πραγματικών healthcare monitoring alarm events και τη σύνδεσή τους με επιλεγμένα AI Act-related σημεία αναφοράς. Η προσέγγιση εφαρμόστηκε σε πραγματικά δεδομένα από το VTaC dataset και υλοποιήθηκε ως proof-of-concept ontology στο Protégé.

Το βασικό συμπέρασμα της εργασίας είναι ότι μια οντολογία μπορεί να λειτουργήσει ως ενδιάμεσο σημασιολογικό επίπεδο μεταξύ πραγματικών operational events, της ανθρώπινης αξιολόγησής τους, της διαθέσιμης τεκμηρίωσης και κανονιστικά σχετικών εννοιών. Μέσα από τη συγκεκριμένη υλοποίηση κατέστη δυνατό τα alarm events να μην αντιμετωπίζονται ως απομονωμένες εγγραφές, αλλά ως στοιχεία ενός ευρύτερου πλαισίου γνώσης στο οποίο η προέλευση, η expert αξιολόγηση, η τελική απόφαση, το evidence και τα requirement mappings μπορούν να διατηρούνται ως διακριτές αλλά συνδεδεμένες πληροφορίες.

Κεντρική σχεδιαστική επιλογή αποτέλεσε η χρήση του Evidence Record ως ενδιάμεσου τεκμηριωτικού κόμβου. Με αυτόν τον τρόπο διαχωρίστηκε το πραγματικό event από το evidence που δημιουργήθηκε στο πλαίσιο της οντολογικής μοντελοποίησης και από το AI Act-related σημείο αναφοράς προς το οποίο το evidence χαρτογραφήθηκε. Η προσέγγιση αυτή επιτρέπει την ανάπτυξη ελέγξιμων requirement-to-evidence mappings χωρίς να μετατρέπει τις σημασιολογικές συνδέσεις σε ισχυρισμούς νομικής συμμόρφωσης.

Παράλληλα, η εργασία ανέδειξε ότι η ιχνηλασιμότητα μπορεί να οργανωθεί μέσω πολλαπλών συνδεδεμένων σημασιολογικών διαδρομών γύρω από το ίδιο alarm event. Η προέλευση από το waveform record, η ανθρώπινη αξιολόγηση, η τελική απόφαση και η τεκμηρίωση μπορούν να ανακτηθούν μέσα από το ίδιο μοντέλο, παρέχοντας μια πιο οργανωμένη εικόνα του κύκλου πληροφορίας που συνοδεύει ένα πραγματικό monitoring event.

Η πρόσθετη risk-related αναπαράσταση των false alarm cases έδειξε επίσης ότι η βασική ontology μπορεί να επεκτείνεται με νέες σημασιολογικές διαστάσεις χωρίς να αλλοιώνεται η αρχική πληροφορία του dataset. Η επέκταση αυτή παραμένει σκόπιμα περιορισμένη και δεν συνιστά πλήρες risk management framework, αλλά λειτουργεί ως ένδειξη της δυνατότητας εμπλουτισμού του μοντέλου με πρόσθετα επίπεδα γνώσης.

Η λειτουργική επικύρωση μέσω HermiT reasoner, DL competency questions και SPARQL queries επιβεβαίωσε ότι οι βασικές σχέσεις που σχεδιάστηκαν μπορούν να ανακτηθούν στο συγκεκριμένο proof-of-concept. Η σημασία του αποτελέσματος αυτού βρίσκεται κυρίως στο ότι η ontology δεν αποτελεί μόνο μια στατική ταξινόμηση εννοιών, αλλά ένα αναζητήσιμο μοντέλο στο οποίο μπορούν να εξεταστούν συγκεκριμένες σημασιολογικές συνθήκες και διαδρομές πληροφορίας.

Η εργασία δεν ανέπτυξε μοντέλο τεχνητής νοημοσύνης και δεν αξιολόγησε το VTaC dataset ή το αρχικό monitoring system ως AI system. Η σύνδεση με την τεχνητή νοημοσύνη και τον AI Act αφορά το επίπεδο της σημασιολογικής τεκμηρίωσης και του governance-oriented mapping που θα μπορούσε να υποστηρίξει αντίστοιχα events σε ένα AI-enabled healthcare monitoring περιβάλλον. Η διάκριση αυτή είναι ουσιαστική, καθώς επιτρέπει στην εργασία να αξιοποιεί πραγματικά δεδομένα χωρίς να αποδίδει σε αυτά χαρακτηρισμούς που δεν υποστηρίζονται από την πηγή τους.

Συνολικά, η διπλωματική δείχνει ότι η οντολογική μοντελοποίηση μπορεί να αποτελέσει χρήσιμο εργαλείο για τη μετάβαση από αποσπασματικές operational πληροφορίες σε ένα περισσότερο συνεκτικό, αναζητήσιμο και τεκμηριωμένο σημασιολογικό πλαίσιο. Η αξία της προσέγγισης βρίσκεται όχι στην αυτόματη λήψη κλινικών ή νομικών αποφάσεων, αλλά στην οργανωμένη σύνδεση δεδομένων, evidence και κανονιστικά σχετικών εννοιών με τρόπο που μπορεί να ελεγχθεί, να επεκταθεί και να αποτελέσει βάση για πιο ολοκληρωμένες μελλοντικές εφαρμογές.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] High-Level Expert Group on Artificial Intelligence, *Ethics Guidelines for Trustworthy AI*. European Commission, 2019, doi: 10.2759/346720.
- [2] The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*, 1st ed. IEEE, 2019.
- [3] European Parliament and Council of the European Union, “Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations and Directives (Artificial Intelligence Act),” *Official Journal of the European Union*, Jul. 12, 2024.
- [4] L.-w. Lehman, B. Moody, L. McCullum, H. Saeed, H. Deep, D. Perry, T. Struja, Q. Li, G. Clifford, and R. Mark, “VTaC: A Benchmark Dataset of Ventricular Tachycardia Alarms from ICU Monitors,” *PhysioNet*, ver. 1.0, 2024, doi: 10.13026/8td2-g363.
- [5] J. Hernandez, D. Golpayegani, and D. Lewis, “Ontology-Based Approach for Mapping Concepts and Requirements from Regulations and Standards: The Case of the EU AI Act and International Standards,” in *Legal Knowledge and Information Systems: JURIX 2024*, J. Savelka, J. Harasta, T. Novotná, and J. Mišek, Eds., *Frontiers in Artificial Intelligence and Applications*, vol. 395. IOS Press, 2024, pp. 295–300, doi: 10.3233/FAIA241258.
- [6] J. Chromik, S. A. I. Klopfenstein, B. Pfitzner, Z.-C. Sinno, B. Arnrich, F. Balzer, and A.-S. Poncette, “Computational approaches to alleviate alarm fatigue in intensive care medicine: A systematic literature review,” *Frontiers in Digital Health*, vol. 4, Art. no. 843747, 2022, doi: 10.3389/fdgh.2022.843747.
- [7] E. J. Topol, “High-performance medicine: The convergence of human and artificial intelligence,” *Nature Medicine*, vol. 25, no. 1, pp. 44–56, 2019, doi: 10.1038/s41591-018-0300-7.
- [8] P. Rajpurkar, E. Chen, O. Banerjee, and E. J. Topol, “AI in health and medicine,” *Nature Medicine*, vol. 28, no. 1, pp. 31–38, 2022, doi: 10.1038/s41591-021-01614-0.
- [9] World Health Organization, *Ethics and Governance of Artificial Intelligence for Health: WHO Guidance*. Geneva, Switzerland: World Health Organization, 2021.
- [10] H. Ruppel, M. Funk, and R. Whittemore, “Measurement of physiological monitor alarm accuracy and clinical relevance in intensive care units,” *American Journal of Critical Care*, vol. 27, no. 1, pp. 11–21, 2018, doi: 10.4037/ajcc2018385.
- [11] K. R. Johnson, J. I. Hagadorn, and D. W. Sink, “Alarm safety and alarm fatigue,” *Clinics in Perinatology*, vol. 44, no. 3, pp. 713–728, 2017, doi: 10.1016/j.clp.2017.05.005.
- [12] B. Li, L. Yue, H. Nie, Z. Cao, X. Chai, B. Peng, T. Zhang, and W. Huang, “The effect of intelligent management interventions in intensive care units to reduce false alarms: An integrative review,” *International Journal of Nursing Sciences*, vol. 11, no. 1, pp. 133–142, 2024, doi: 10.1016/j.ijnss.2023.12.008.
- [13] E. A. M. Michels, S. Gilbert, I. Koval, and M. K. Wekenborg, “Alarm fatigue in healthcare: A scoping review of definitions, influencing factors, and mitigation strategies,” *BMC Nursing*, vol. 24, Art. no. 664, 2025, doi: 10.1186/s12912-025-03369-2.

- [14] N. Díaz-Rodríguez, J. Del Ser, M. Coeckelbergh, M. López de Prado, E. Herrera-Viedma, and F. Herrera, “Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation,” *Information Fusion*, vol. 99, Art. no. 101896, 2023, doi: 10.1016/j.inffus.2023.101896.
- [15] C. Huang, Z. Zhang, B. Mao, and X. Yao, “An overview of artificial intelligence ethics,” *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 4, pp. 799–819, Aug. 2023, doi: 10.1109/TAI.2022.3194503.
- [16] D. Golpayegani, H. J. Pandit, and D. Lewis, “AIRO: An ontology for representing AI risks based on the proposed EU AI Act and ISO risk management standards,” in *Towards a Knowledge-Aware AI*, Studies on the Semantic Web, vol. 55. IOS Press, 2022, pp. 51–65, doi: 10.3233/SSW220008.
- [17] T. Rintamäki and H. J. Pandit, “Developing an ontology for AI Act Fundamental Rights Impact Assessments,” OSF Preprints, 2024, doi: 10.31219/osf.io/tm74p.
- [18] European Commission, “AI Act,” *Shaping Europe’s Digital Future*, 2026. [Online]. Accessed: Sep. 3, 2026.
- [19] T. R. Gruber, “A translation approach to portable ontology specifications,” *Knowledge Acquisition*, vol. 5, no. 2, pp. 199–220, 1993, doi: 10.1006/knac.1993.1008.
- [20] R. Studer, V. R. Benjamins, and D. Fensel, “Knowledge engineering: Principles and methods,” *Data & Knowledge Engineering*, vol. 25, nos. 1–2, pp. 161–197, 1998, doi: 10.1016/S0169-023X(97)00056-6.
- [21] N. Guarino, “Formal ontology and information systems,” in *Formal Ontology in Information Systems: Proceedings of FOIS ’98*, N. Guarino, Ed. Amsterdam, The Netherlands: IOS Press, 1998, pp. 3–15.
- [22] N. F. Noy and D. L. McGuinness, *Ontology Development 101: A Guide to Creating Your First Ontology*, Stanford Knowledge Systems Laboratory, Tech. Rep. KSL-01-05, Stanford, CA, USA, Mar. 2001.
- [23] R. Cyganiak, D. Wood, and M. Lanthaler, “RDF 1.1 Concepts and Abstract Syntax,” W3C Recommendation, Feb. 25, 2014.
- [24] W3C OWL Working Group, “OWL 2 Web Ontology Language Document Overview, Second Edition,” W3C Recommendation, Dec. 11, 2012.
- [25] S. Harris and A. Seaborne, “SPARQL 1.1 Query Language,” W3C Recommendation, Mar. 21, 2013.
- [26] M. A. Musen and the Protégé Team, “The Protégé project: A look back and a look forward,” *AI Matters*, vol. 1, no. 4, pp. 4–12, 2015, doi: 10.1145/2757001.2757003.
- [27] R. Guizzardi, G. Amaral, G. Guizzardi, and J. Mylopoulos, “An ontology-based approach to engineering ethicality requirements,” *Software and Systems Modeling*, vol. 22, no. 6, pp. 1897–1923, 2023, doi: 10.1007/s10270-023-01115-3.
- [28] A. Harrison, D. Spagnuolo, and I. Tiddi, “An ontology for Ethical AI Principles,” unpublished manuscript, 2021.
- [29] T. M. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill, 1997.

- [30] B. Glimm, I. Horrocks, B. Motik, G. Stoilos, and Z. Wang, “HermiT: An OWL 2 reasoner,” *Journal of Automated Reasoning*, vol. 53, no. 3, pp. 245–269, 2014, doi: 10.1007/s10817-014-9305-1.
- [31] E. Kanaki, E. Andrikopoulou, P. Chatzimisios, and A. Kanta, “Aligning AI-native 6G healthcare systems with EU ethical and legal frameworks,” *Studies in Health Technology and Informatics*, vol. 336, pp. 755–759, 2026, doi: 10.3233/SHTI260272.

ΠΑΡΑΡΤΗΜΑ Α: Συγκεντρωτική τεχνική προδιαγραφή της οντολογίας

Το παρόν παράρτημα παρουσιάζει συγκεντρωτικά τη βασική τεχνική προδιαγραφή της οντολογίας που αναπτύχθηκε στο Protégé. Στόχος είναι να παρέχεται σε ένα ενιαίο σημείο η συνολική εικόνα των βασικών δομικών στοιχείων του μοντέλου, χωρίς να επαναλαμβάνεται η αναλυτική παρουσίαση του Κεφαλαίου 4.

Η τελική proof-of-concept έκδοση της οντολογίας περιλαμβάνει:

- 23 classes
- 12 object properties
- 9 data properties
- 76 individuals

Η class owl:Thing αποτελεί τη ρίζα της OWL ιεραρχίας και δεν προσμετράται στις 23 θεματικές classes.

A.1 Classes της οντολογίας

A/A	Class	Άμεση υπερκλάση
1	Evidence and Assessment Concept	owl:Thing
2	Compliance Assessment	Evidence and Assessment Concept
3	Control Mechanism	Evidence and Assessment Concept
4	Evidence Record	Evidence and Assessment Concept
5	Normative Concept	owl:Thing
6	Compliance Requirement	Normative Concept
7	Human Review Requirement	Compliance Requirement
8	Logging Requirement	Compliance Requirement
9	Risk Management Requirement	Compliance Requirement
10	Traceability Requirement	Compliance Requirement
11	Operational Concept	owl:Thing
12	Alert Event	Operational Concept
13	Ventricular Tachycardia Alarm	Alert Event
14	Expert Review	Operational Concept
15	Monitoring System	Operational Concept
16	Review Decision	Operational Concept

17	False Alarm	Review Decision
18	True Alarm	Review Decision
19	Waveform Record	Operational Concept
20	Risk Concept	owl:Thing
21	Potential Harm	Risk Concept
22	Risk Indicator	Risk Concept
23	Risk-Relevant Event	Risk Concept

Οι classes οργανώνονται σε τέσσερις βασικές θεματικές περιοχές: operational, evidence and assessment, normative και risk-related επίπεδο. Η Compliance Assessment διατηρείται ως μέρος του schema και ως σημείο μελλοντικής επέκτασης, χωρίς direct individuals στην παρούσα proof-of-concept έκδοση.

A.2 Object properties

Object property	Domain	Range	Βασικός ρόλος
assessesEvent	Compliance Assessment	Alert Event	Συνδέει μια πιθανή αξιολόγηση με το event που εξετάζεται.
assessesRequirement	Compliance Assessment	Compliance Requirement	Συνδέει μια πιθανή αξιολόγηση με requirement concept.
belongsToRecord	Alert Event	Waveform Record	Συνδέει ένα alarm event με το waveform record προέλευσής του.
hasAlertEvent	Waveform Record	Alert Event	Συνδέει ένα waveform record με τα alarm events που περιλαμβάνει.
hasDecision	Expert Review	Review Decision	Συνδέει την expert αξιολόγηση με την τελική απόφαση.
hasReview	Alert Event	Expert Review	Συνδέει ένα alarm event με την expert αξιολόγησή του.
hasRiskIndicator	Alert Event	Risk Indicator	Συνδέει επιλεγμένα events με risk indicator.
isClassifiedAs	Alert Event	Risk-Relevant Event	Συνδέει επιλεγμένο event με την πρόσθετη risk-related αναπαράστασή του.

isDocumentedBy	Alert Event	Evidence Record	Συνδέει ένα event με το Evidence Record που το τεκμηριώνει.
mayLeadTo	Risk-Relevant Event	Potential Harm	Αναπαριστά πιθανή σύνδεση risk-relevant event με potential harm.
supportsRequirement	Evidence Record	Compliance Requirement	Συνδέει Evidence Record με το αντίστοιχο requirement mapping.
usesEvidence	Compliance Assessment	Evidence Record	Συνδέει μια πιθανή αξιολόγηση με Evidence Record.

Οι object properties οργανώνουν τις βασικές συνδεδεμένες σημασιολογικές διαδρομές του μοντέλου. Ιδιαίτερα σημαντικές είναι οι διαδρομές:

Alert Event → Waveform Record

Alert Event → Expert Review → Review Decision

Alert Event → Evidence Record → Compliance Requirement

και, για το lightweight risk-related επίπεδο:

Alert Event → Risk Indicator

Alert Event → Risk-Relevant Event → Potential Harm

A.3 Data properties

Data property	Βασική χρήση
assessmentOutcome	Έχει προβλεφθεί για την απόδοση περιγραφικού αποτελέσματος σε μελλοντικά Compliance Assessment individuals. Δεν χρησιμοποιείται σε instance-level assertion στην παρούσα έκδοση.
decisionLabel	Αποδίδει την περιγραφική τιμή της expert απόφασης, όπως true alarm ή false alarm.
eventId	Διατηρεί το identifier του πραγματικού VTaC alarm event.
hasHumanReview	Boolean ένδειξη ότι στο proof-of-concept έχει καταγραφεί human expert review για το event.
isLogged	Boolean χαρακτηριστικό της proof-of-concept αναπαράστασης σχετικά με την καταγραφή του event.
isRiskRelevant	Boolean ένδειξη ότι το event έχει ενταχθεί στο lightweight risk-related επίπεδο.
isTraceable	Boolean χαρακτηριστικό της proof-of-concept αναπαράστασης σχετικά με την ιχνηλασιμότητα του event.
potentialHarmLabel	Παρέχει περιγραφικό label για το Potential Harm individual.
recordId	Διατηρεί το identifier του αντίστοιχου VTaC waveform record.

Τα boolean data properties hasHumanReview, isLogged, isRiskRelevant και isTraceable αποτελούν χαρακτηριστικά της συγκεκριμένης proof-of-concept αναπαράστασης. Δεν αντικαθιστούν τις object-property σχέσεις και δεν αποτελούν από μόνα τους απόδειξη εκπλήρωσης νομικών ή τεχνικών απαιτήσεων του AI Act.

A.4 Individuals της οντολογίας

Η τελική έκδοση περιλαμβάνει συνολικά **76 individuals**, κατανεμημένα στις ακόλουθες κατηγορίες:

Κατηγορία	Πλήθος
Requirement individuals	4
Control Mechanism individuals	4
Monitoring System individual	1
Waveform Record individuals	6
VTaC Alert Event individuals	12
Expert Review individuals	12
Review Decision individuals	12
Evidence Record individuals	12
Risk Indicator individuals	6
Risk-Relevant Event individuals	6
Potential Harm individual	1
Σύνολο	76

Requirement individuals

req_logging
req_human_review
req_traceability
req_risk_management

Τα παραπάνω χρησιμοποιούνται ως εφαρμοσμένα AI Act-related σημεία αναφοράς στο proof-of-concept.

Control Mechanism individuals

logging_control
human_review_control
traceability_control
risk_management_control

Monitoring System individual

vtac_icu_monitoring_system

Waveform Record individuals

vtac_record_003c13
vtac_record_004bad
vtac_record_0060de
vtac_record_015f23
vtac_record_0179fd
vtac_record_01a952

VTaC Alert Event individuals

vtac_event_003c13_0115
vtac_event_003c13_0126
vtac_event_004bad_0015
vtac_event_004bad_1115
vtac_event_004bad_1426
vtac_event_004bad_2348
vtac_event_0060de_0560
vtac_event_015f23_0512
vtac_event_0179fd_1194
vtac_event_0179fd_1265
vtac_event_0179fd_2773
vtac_event_01a952_1067

Expert Review individuals

vtac_review_003c13_0115
vtac_review_003c13_0126
vtac_review_004bad_0015
vtac_review_004bad_1115
vtac_review_004bad_1426
vtac_review_004bad_2348
vtac_review_0060de_0560
vtac_review_015f23_0512
vtac_review_0179fd_1194
vtac_review_0179fd_1265
vtac_review_0179fd_2773
vtac_review_01a952_1067

Review Decision individuals – False Alarm

vtac_decision_003c13_0115_false
vtac_decision_003c13_0126_false
vtac_decision_004bad_0015_false
vtac_decision_004bad_1115_false
vtac_decision_004bad_1426_false
vtac_decision_004bad_2348_false

Review Decision individuals – True Alarm

vtac_decision_0060de_0560_true
vtac_decision_015f23_0512_true
vtac_decision_0179fd_1194_true
vtac_decision_0179fd_1265_true
vtac_decision_0179fd_2773_true
vtac_decision_01a952_1067_true

Evidence Record individuals

Για κάθε ένα από τα 12 επιλεγμένα VTaC events έχει δημιουργηθεί ένα αντίστοιχο Evidence Record individual. Τα Evidence Records λειτουργούν ως ενδιάμεσοι τεκμηριωτικοί κόμβοι μεταξύ των πραγματικών alarm events και των επιλεγμένων AI Act-related mappings.

Risk-related individuals

Για τα έξι false alarm cases έχουν δημιουργηθεί:

- 6 Risk Indicator individuals,
- 6 Risk-Relevant Event individuals,
- 1 Potential Harm individual.

To Potential Harm individual είναι:

vtac_harm_alert_fatigue

Το risk-related επίπεδο αποτελεί lightweight σημασιολογική επέκταση και δεν δηλώνει ότι κάποιο συγκεκριμένο false alarm προκάλεσε πραγματική βλάβη.

Συνολικά, το Παράρτημα Α παρέχει μια συγκεντρωτική τεχνική εικόνα της τελικής proof-of-concept οντολογίας και λειτουργεί ως σημείο αναφοράς για τη δομή που αναλύθηκε στο Κεφάλαιο 4 και ελέγχθηκε λειτουργικά στο Κεφάλαιο 5.

ΠΑΡΑΡΤΗΜΑ Β: DL competency questions και SPARQL queries

Το παρόν παράρτημα συγκεντρώνει τις DL Query expressions και τα SPARQL queries που χρησιμοποιήθηκαν για τη λειτουργική επικύρωση της οντολογίας. Η αναλυτική παρουσίαση της λογικής, της εκτέλεσης και των αποτελεσμάτων των ερωτημάτων περιλαμβάνεται στο Κεφάλαιο 5.

Οι DL expressions παρουσιάζονται με τις ονομασίες που χρησιμοποιούνται στο περιβάλλον Protégé. Τα SPARQL queries χρησιμοποιούν το namespace της υλοποιημένης οντολογίας.

B.1 DL competency questions

DL Query 1 – Εντοπισμός των VTaC alert events

Ερώτημα:

Ποια individuals έχουν εισαχθεί στην οντολογία ως VTaC alert events τύπου ventricular tachycardia alarm;

DL expression:

'Ventricular Tachycardia Alarm'

Αποτέλεσμα: 12 VTaC alert event individuals.

DL Query 2 – VT alarm events που έχουν αξιολογηθεί ως true alarms

Ερώτημα:

Ποια VT alarm events συνδέονται, μέσω της διαδικασίας review, με τελική απόφαση True Alarm;

DL expression:

'Ventricular Tachycardia Alarm'

and ('has review' some ('has decision' some 'True Alarm'))

Αποτέλεσμα: 6 true alarm events.

DL Query 3 – VT alarm events που έχουν αξιολογηθεί ως false alarms

Ερώτημα:

Ποια VT alarm events συνδέονται, μέσω της διαδικασίας review, με τελική απόφαση False Alarm;

DL expression:

'Ventricular Tachycardia Alarm'

and ('has review' some ('has decision' some 'False Alarm'))

Αποτέλεσμα: 6 false alarm events.

DL Query 4 – VT alarm events με evidence για logging

Ερώτημα:

Ποια VT alarm events συνδέονται με Evidence Record που έχει χαρτογραφηθεί προς το req_logging;

DL expression:

'Ventricular Tachycardia Alarm'

and ('is documented by' some ('supports requirement' value req_logging))

Αποτέλεσμα: 12 VT alarm events

DL Query 5 – VT alarm events με evidence για human review

Ερώτημα:

Ποια VT alarm events συνδέονται με Evidence Record που έχει χαρτογραφηθεί προς το req_human_review;

DL expression:

'Ventricular Tachycardia Alarm'

and ('is documented by' some ('supports requirement' value req_human_review))

Αποτέλεσμα: 12 VT alarm events.

DL Query 6 – VT alarm events με evidence για traceability

Ερώτημα:

Ποια VT alarm events συνδέονται με Evidence Record που έχει χαρτογραφηθεί προς το req_traceability;

DL expression:

'Ventricular Tachycardia Alarm'

and ('is documented by' some ('supports requirement' value req_traceability))

Αποτέλεσμα: 12 VT alarm events.

DL Query 7 – Evidence Records με τα τρία βασικά requirement mappings

Ερώτημα:

Ποια Evidence Record individuals υποστηρίζουν ταυτόχρονα τα req_logging, req_human_review και req_traceability;

DL expression:

'Evidence Record'

and ('supports requirement' value req_logging)

and ('supports requirement' value req_human_review)

and ('supports requirement' value req_traceability)

Αποτέλεσμα: 12 Evidence Record individuals.

DL Query 8 – False alarm events στο lightweight risk-related επίπεδο**Ερώτημα:**

Ποια false alarm events έχουν συνδεθεί ταυτόχρονα με Risk Indicator και Risk-Relevant Event;

DL expression:

'Ventricular Tachycardia Alarm'

and ('has review' some ('has decision' some 'False Alarm'))

and ('has risk indicator' some 'Risk Indicator')

and ('is classified as' some 'Risk-Relevant Event')

Αποτέλεσμα: 6 false alarm events.

DL Query 9 – Risk-relevant events με potential harm**Ερώτημα:**

Ποια Risk-Relevant Event individuals συνδέονται με κάποιο Potential Harm;

DL expression:

'Risk-Relevant Event'

and ('may lead to' some 'Potential Harm')

Αποτέλεσμα: 6 Risk-Relevant Event individuals.

DL Query 10 – Evidence Records με mapping προς req_risk_management**Ερώτημα:**

Ποια Evidence Record individuals έχουν χαρτογραφηθεί προς το req_risk_management;

DL expression:

'Evidence Record'

and ('supports requirement' value req_risk_management)

Αποτέλεσμα: 6 Evidence Record individuals.

DL Query 11 – Events με πλήρη σχεδιασμένη διαδρομή ιχνηλασιμότητας

Ερώτημα:

Ποια VT alarm events συνδέονται με waveform record, Expert Review και Review Decision και ταυτόχρονα διαθέτουν Evidence Record που έχει χαρτογραφηθεί προς το req_traceability;

DL expression:

'Ventricular Tachycardia Alarm'

and ('belongs to record' some 'Waveform Record')

and ('has review' some ('has decision' some 'Review Decision'))

and ('is documented by' some ('supports requirement' value req_traceability))

Αποτέλεσμα: 12 VT alarm events.

Το συγκεκριμένο query ελέγχει τη συνύπαρξη πολλαπλών σημασιολογικών σχέσεων γύρω από το ίδιο event και όχι μία μοναδική γραμμική object-property αλυσίδα.

DL Query 12 – Waveform records με VT alarm events

Ερώτημα:

Ποια waveform records συνδέονται με τουλάχιστον ένα event της class Ventricular Tachycardia Alarm;

DL expression:

'Waveform Record'

and ('has alert event' some 'Ventricular Tachycardia Alarm')

Αποτέλεσμα: 6 Waveform Record individuals.

B.2 SPARQL queries

Για τα SPARQL queries χρησιμοποιούνται τα ακόλουθα prefixes:

PREFIX : <http://www.semanticweb.org/aiact-healthcare-vtac#>

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

SPARQL Query 1 – Συγκεντρωτική ανάκτηση event–record–decision–evidence

Ερώτημα:

Πώς συνδέεται κάθε VT alarm event με το waveform record προέλευσής του, την απόφαση αξιολόγησης, το Evidence Record και την ένδειξη risk relevance;

PREFIX : <http://www.semanticweb.org/aiact-healthcare-vtac#>

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

```
SELECT ?event ?record ?decision ?decisionLabel ?evidence ?isRiskRelevant
```

```
WHERE {
```

```
    ?event rdf:type :VTAlarm .
```

```
    ?event :belongsToRecord ?record .
```

```
    ?event :hasReview ?review .
```

```
    ?review :hasDecision ?decision .
```

```
    ?decision :decisionLabel ?decisionLabel .
```

```
    ?event :isDocumentedBy ?evidence .
```

```
    ?event :isRiskRelevant ?isRiskRelevant .
```

```
}
```

```
ORDER BY ?event
```

Αποτέλεσμα: 12 γραμμές, μία για κάθε επιλεγμένο VT alarm event.

SPARQL Query 2 – Χαρτογράφηση event–evidence–requirement**Ερώτημα:**

Ποια requirement mappings συνδέονται με το Evidence Record κάθε VT alarm event;

PREFIX : <http://www.semanticweb.org/aiact-healthcare-vtac#>

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

```
SELECT ?event ?evidence ?requirement
```

```
WHERE {
```

```
    ?event rdf:type :VTAlarm .
```

```
    ?event :isDocumentedBy ?evidence .
```

```
    ?evidence :supportsRequirement ?requirement .
```

```
}
```

```
ORDER BY ?event ?requirement
```

Το query παρουσιάζει συγκεντρωτικά τη διαδρομή:

Alert Event → Evidence Record → Requirement individual

και καθιστά ορατό το requirement-to-evidence mapping της οντολογίας.

SPARQL Query 3 – Risk-related mapping για false alarm cases

Ερώτημα:

Ποια VT alarm events έχουν ενταχθεί στο lightweight risk-related επίπεδο και με ποια risk-related στοιχεία συνδέονται;

PREFIX : <http://www.semanticweb.org/aiact-healthcare-vtac#>

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

```
SELECT ?event ?riskIndicator ?riskEvent ?potentialHarm
```

```
WHERE {
```

```
  ?event rdf:type :VTAlarm .
```

```
  ?event :hasRiskIndicator ?riskIndicator .
```

```
  ?event :isClassifiedAs ?riskEvent .
```

```
  ?riskEvent :mayLeadTo ?potentialHarm .
```

```
}
```

```
ORDER BY ?event
```

Αποτέλεσμα: 6 γραμμές, που αντιστοιχούν στα έξι false alarm cases τα οποία έχουν ενταχθεί στο lightweight risk-related επίπεδο.

Η σύνδεση με το Potential Harm εκφράζει πιθανή risk-related σχέση και όχι τεκμηρίωση πραγματικής πρόκλησης βλάβης.

SPARQL Query 4 – Πλήθος VT alarm events ανά waveform record

Ερώτημα:

Πόσα VT alarm events αντιστοιχούν σε κάθε waveform record της οντολογίας;

PREFIX : <http://www.semanticweb.org/aiact-healthcare-vtac#>

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

```
SELECT ?record (COUNT(?event) AS ?numberOfEvents)
```

```
WHERE {
```

```
  ?record rdf:type :WaveformRecord .
```

```

?record :hasAlertEvent ?event .
?event rdf:type :VTAlarm .
}

```

GROUP BY ?record

ORDER BY ?record

Αποτέλεσμα: 6 waveform records με την ακόλουθη κατανομή:

Waveform record	Πλήθος VT alarm events
vtac_record_003c13	2
vtac_record_004bad	4
vtac_record_0060de	1
vtac_record_015f23	1
vtac_record_0179fd	3
vtac_record_01a952	1

Σύνολο: 12 VT alarm events.

Τα 12 DL competency questions και τα 4 SPARQL queries αποτελούν το βασικό σύνολο λειτουργικών ελέγχων της proof-of-concept οντολογίας. Τα DL queries ελέγχουν συγκεκριμένες λογικές συνθήκες και ανακτούν τα αντίστοιχα individuals, ενώ τα SPARQL queries επιτρέπουν τη συγκεντρωτική ανάκτηση και παρουσίαση συνδυασμένων σχέσεων από τον RDF γράφο της οντολογίας. Τα αποτελέσματα των ερωτημάτων χρησιμοποιούνται για λειτουργική επικύρωση του σχεδιασμένου μοντέλου και δεν αποτελούν κλινική αξιολόγηση ή απόδειξη νομικής συμμόρφωσης με τον AI Act.