



ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ στα
ΠΟΛΥΠΛΟΚΑ ΣΥΣΤΗΜΑΤΑ και ΔΙΚΤΥΑ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ
ΤΜΗΜΑ ΒΙΟΛΟΓΙΑΣ
ΤΜΗΜΑ ΓΕΩΛΟΓΙΑΣ
ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ



ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Σημασιολογική Αναπαράσταση Δεδομένων του ΕΣΠΑ. Στοχεύοντας σε ένα Εννοιολογικό Δίκτυο Οικονομικών Δεδομένων.

Ευάγγελος Α. Χονδροκώστας

ΕΠΙΒΛΕΠΩΝ: Χαράλαμπος Μπράτσας, ΕΔΙΠ Α.Π.Θ.

Υπό την εποπτεία: Ιωάννη Αντωνίου, Καθηγητή Α.Π.Θ.

Θεσσαλονίκη, Δεκέμβριος 2016



ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ στα
ΠΟΛΥΠΛΟΚΑ ΣΥΣΤΗΜΑΤΑ και ΔΙΚΤΥΑ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ
ΤΜΗΜΑ ΒΙΟΛΟΓΙΑΣ
ΤΜΗΜΑ ΓΕΩΛΟΓΙΑΣ
ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ



ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Σημασιολογική Αναπαράσταση Δεδομένων του ΕΣΠΑ. Στοχεύοντας σε ένα Εννοιολογικό Δίκτυο Οικονομικών Δεδομένων.

Ευάγγελος Α. Χουδροκώστας

ΕΠΙΒΛΕΠΩΝ: Χαράλαμπος Μπράτσας, ΕΔΙΠ Α.Π.Θ.

Υπό την εποπτεία: Ιωάννη Αντωνίου, Καθηγητή Α.Π.Θ.

Εγκρίθηκε από την Τριμελή Εξεταστική Επιτροπή την 20η Δεκεμβρίου 2016.

.....
Χαράλαμπος Μπράτσας
ΕΔΙΠ Α.Π.Θ.

.....
Ιωάννης Αντωνίου
Καθηγητής Α.Π.Θ.

.....
Νικόλαος Φαρμάκης
Αν. Καθηγητής Α.Π.Θ.

Θεσσαλονίκη, Δεκέμβριος 2016

.....
Ευάγγελος Χονδροκώστας
Πτυχιούχος Μαθηματικών
Α.Π.Θ.

Copyright © Ευάγγελος Α. Χονδροκώστας, 2016
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευτεί ότι εκφράζουν τις επίσημες θέσεις του Α.Π.Θ.

Περίληψη

Τα δεδομένα του Εταιρικού Συμφώνου για το Πλαίσιο Ανάπτυξης (ΕΣΠΑ) της Ελλάδας, πρώην Εθνικό Στρατηγικό Πλαίσιο Αναφοράς για την προγραμματική περίοδο 2007-2013, δίνονται υπό τη μορφή ανοιχτών δεδομένων μέσω της ιστοσελίδας ANAPTYXI.gov.gr. Στόχος της εργασίας είναι η σημασιολογική αναπαράσταση αυτών των δεδομένων. Για την επίτευξη αυτού του σκοπού γίνεται επαναχρησιμοποίηση του σημασιολογικού μοντέλου OpenBudgets.eu, το οποίο έχει αναπτυχθεί στα πλαίσια του Ευρωπαϊκού προγράμματος OpenBudgets.eu. Επιπλέον, για την πληρέστερη περιγραφή των δεδομένων των έργων, δημιουργείται νέα οντολογία που μπορεί να επαναχρησιμοποιηθεί στην δημιουργία ενός εννοιολογικού δικτύου οικονομικών δεδομένων. Τέλος, έχοντας ολοκληρώσει το σκέλος της σημασιολογικής περιγραφής των δεδομένων και τη δημοσίευσή τους, δίνουμε ένα παράδειγμα για το πώς μπορούν να αναλυθούν τα δεδομένα αυτά και να χρησιμοποιηθούν στον εντοπισμό πιθανής διαφθοράς και απάτης, με την δημιουργία χαρακτηριστικών δεικτών για τον εντοπισμό παράτυπων σημείων και την εύρεση *κόκκινων σημαιών (red flags)*.

Λέξεις-κλειδιά: Σημασιολογικός Ιστός, Οντολογίες, RDF, Δημοσιονομικά δεδομένα, Κόκκινες σημαίες

Abstract

The data of the Greek Partnership Agreement (PA), former National Strategic Reference Framework for the programming period 2007-2013, are provided as Open Data by the official website of the Greek Ministry for Development and Competitiveness, ANAPTYXI.gov.gr. The purpose of this master thesis is the semantic representation of said data. To achieve this goal the semantic data model of OpenBudgets.eu is being reused. OpenBudgets.eu data model has been developed within the OpenBudgets.eu european project. Moreover, to fully describe the project data, a new ontology is being created, so it can be reused to create a conceptual economic data network. Finally, having completed the semantic representation of our data and its publishing, we give an example of how the data can be analyzed and be used to detect probable corruption and fraud by creating feature indicators for identifying *red flags*.

Keywords: Semantic Web, Ontologies, RDF, Fiscal data, Red flags

Ευχαριστίες

Θα ήθελα να ευχαριστήσω όλους τους διδάσκοντες του Διατμηματικού Προγράμματος Μεταπτυχιακών Σπουδών *Πολύπλοκα Συστήματα και Δίκτυα* για την γνώση που μου προσέφεραν μέσα από τα μαθήματα και τις διαλέξεις τους. Ιδιαίτερα, θα ήθελα να ευχαριστήσω τον Δρ. Χαράλαμπο Μπράτσα, όχι μόνο για την πολύτιμη γνώση που μου προσέφερε, αλλά κυρίως για την στήριξη που μου παρείχε και την εμπιστοσύνη που μου έδειξε ώστε να καταφέρω να ολοκληρώσω τη διπλωματική μου εργασία. Καθοριστική ήταν επίσης η συμβολή του Καθηγητή Ιωάννη Αντωνίου, ο οποίος με την εμπειρία και τον μοναδικό τρόπο σκέψης του, κατάφερε να διευρύνει τους πνευματικούς μου ορίζοντες.

Δε θα γινόταν να ξεχάσω όμως όλους τους συμφοιτητές μου, με τους οποίους πέρασα ένα κουραστικό αλλά συνάμα υπέροχο χρόνο.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένεια μου για την υπομονή που έδειξε και την υποστήριξη που μου παρείχε καθ' όλη τη διάρκεια των σπουδών μου.

Περιεχόμενα

Περίληψη	5
Abstract	7
Ευχαριστίες	9
Περιεχόμενα	13
1 Πρόλογος	15
2 Εθνικό Στρατηγικό Πλαίσιο Αναφοράς	17
2.1 Η Πολιτική Συνοχής της ΕΕ	17
2.2 Ευρωπαϊκά Διαρθρωτικά και Επενδυτικά Ταμεία	18
2.2.1 Ευρωπαϊκό Ταμείο Περιφερειακής Ανάπτυξης	18
2.2.2 Ευρωπαϊκό Κοινωνικό Ταμείο	19
2.2.3 Ταμείο Συνοχής	19
2.2.4 Ευρωπαϊκό Γεωργικό Ταμείο Αγροτικής Ανάπτυξης	20
2.2.5 Ευρωπαϊκό Ταμείο Θάλασσας και Αλιείας	20
2.3 Προγραμματισμός και Υλοποίηση	20
2.4 Το ΕΣΠΑ στην Ελλάδα	21
3 Σημαιολογικός Ιστός, Διασυνδεδεμένα Δεδομένα και Οντολογίες	25
3.1 Σημαιολογικός ιστός	25
3.1.1 Τυποποίηση και Αρχιτεκτονική Σημαιολογικού Ιστού	26
3.2 Διασυνδεδεμένα Δεδομένα	28
3.3 Οντολογίες	30
4 Διαδικασία Απόκτησης Δεδομένων	33

4.1	ΑΝΑΠΤΥΞΗ.gov.gr	33
4.1.1	Τι είναι	33
4.1.2	API Ανοιχτών Δεδομένων	35
4.1.3	Κατηγορίες Πράξεων	35
4.2	Τοπική Βάση Δεδομένων	36
4.3	Εξαγωγή και Αποθήκευση Δεδομένων	39
5	Σημασιολογική Αναπαράσταση Δεδομένων	43
5.1	Εργαλεία που χρησιμοποιήθηκαν	43
5.1.1	Protégé	43
5.1.2	UnifiedViews	44
5.2	OpenBudgets.eu	46
5.2.1	Μοντέλο Δεδομένων OpenBudgets.eu	47
5.2.1.1	To RDF Data Cube Vocabulary	47
5.2.1.2	Ο Κύβος στο Μοντέλο Δεδομένων OpenBudgets.eu	48
5.2.1.3	Δημιουργία της δομής του Κύβου για τα Δεδομένα του ΕΣΠΑ	50
5.3	Vocabulary for Describing Fiscal Projects	51
5.3.1	Γιατί νέα οντολογία	51
5.3.2	Ορισμός κλάσεων και ιδιοτήτων	52
5.3.3	Επέκταση της Οντολογίας για τα Δεδομένα ΕΣΠΑ	55
5.4	Μετατροπή των Δεδομένων σε RDF	59
6	Κόκκινες σημαίες	63
6.1	Ορισμοί	65
6.2	Διαδικασία	65
6.2.1	Προσδιορισμός συνόλων δεδομένων	65
6.2.2	Ορισμός δεικτών	65
6.2.2.1	Έργα	65
6.2.2.2	Απαλλοτριώσεις	66
6.2.3	Ανάλυση - Αποτελέσματα - Σχολιασμός	67
6.2.3.1	Έργα	67
6.2.3.2	Απαλλοτριώσεις	73

6.2.4 Σύνοψη Αποτελεσμάτων	82
7 Επίλογος	83
7.1 Συμπεράσματα	83
7.2 Μελλοντική Εργασία	84
Βιβλιογραφία	85

Κεφάλαιο 1

Πρόλογος

Τα δημοσιονομικά σύνολα δεδομένων από την φύση τους χαρακτηρίζονται από ποικιλομορφία. Σπάνια θα βρεθούν δημοσιονομικά σύνολα δεδομένων που να έχουν ίδια δομή, αν και πολλά εξ' αυτών μπορεί να έχουν κοινά στοιχεία. Συνεπώς, παρά το γεγονός ότι σημειώνεται αύξηση στην ποσότητα των δημοσιονομικών δεδομένων που γίνονται διαθέσιμα, η ετερογένεια τους όπως είναι λογικό προκαλεί προβλήματα στην ενοποίηση αυτών των δεδομένων (data integration) και δυσκολεύει την ανάλυση τους.

Λύση σε αυτό το πρόβλημα έρχεται να δώσει ο Σημασιολογικός Ιστός. Μέσα από τις τεχνολογίες του Σημασιολογικού Ιστού μπορούν να συνδυαστούν σύνολα δεδομένων από διαφορετικές πηγές, να περιγραφούν εννοιολογικά και να σχηματιστούν συνδέσεις μεταξύ τους.

Με την ανάλυση των εννοιολογικών δεδομένων μπορούν να οριστούν δείκτες παρακολούθησης των δημοσιονομικών δεδομένων, οι οποίοι μπορούν να χρησιμοποιηθούν για την αναγνώριση πιθανής παρατυπίας.

Στο πλαίσιο αυτό, δηλαδή της σημασιολογικής περιγραφής και ανάλυσης των δημοσιονομικών δεδομένων του ελληνικού ΕΣΠΑ, κινείται η παρούσα διπλωματική εργασία.

Η εργασία δομείται ως εξής:

Στο δεύτερο κεφάλαιο γίνεται μια εισαγωγή της έννοιας της πολιτικής συνοχής της Ευρωπαϊκής Ένωσης και των σκοπών που θέλει να πετύχει με τις επενδύσεις της. Στη συνέχεια, περιγράφονται τα Ευρωπαϊκά Διαρθρωτικά και Επενδυτικά Ταμεία και πώς υλοποιείται η πολιτική της περιφερειακής ανάπτυξης της ΕΕ μέσα από τη δημιουργία των ΕΣΠΑ. Τέλος, γίνεται μια αναφορά στο ελληνικό ΕΣΠΑ.

Στο τρίτο κεφάλαιο περιγράφεται ο Σημασιολογικός Ιστός, τα Διασυνδεδεμένα

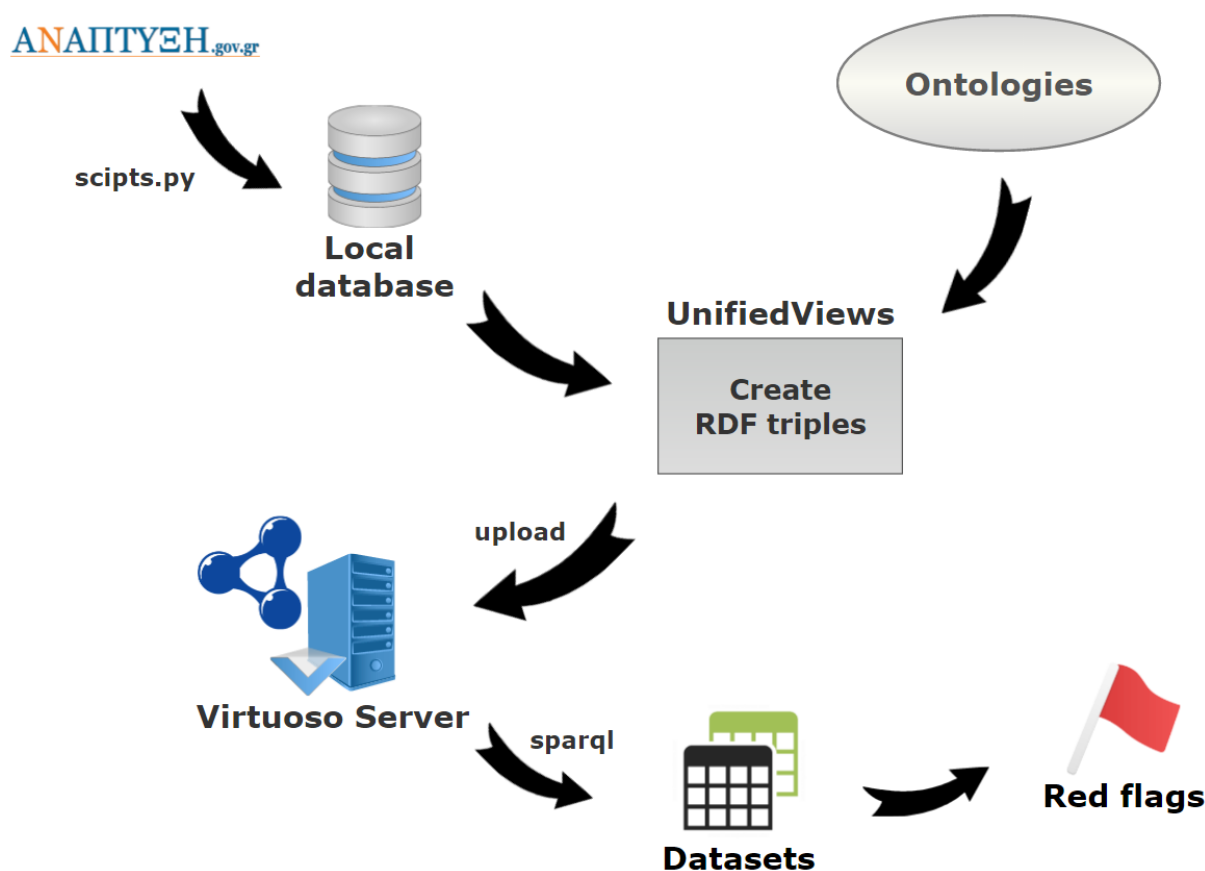
δεδομένα καθώς και η έννοια της Οντολογίας.

Στο τέταρτο κεφάλαιο περιγράφεται η διαδικασία απόκτησης των δεδομένων του ΕΣΠΑ από την ιστοσελίδα ANAITYEH.gov.gr.

Στο πέμπτο κεφάλαιο περιγράφεται η διαδικασία της σημασιολογικής αναπαράστασης των δεδομένων του ΕΣΠΑ. Αναλύονται οι οντολογίες που χρησιμοποιήθηκαν καθώς και τα εργαλεία που βοήθησαν τόσο στην κατασκευή της νέας οντολογίας όσο και στην δημιουργία των RDF τριπλέτων.

Στο έκτο κεφάλαιο ορίζεται η έννοια της απάτης, παρατυπίας και της κόκκινης σημαίας. Δημιουργούνται χαρακτηριστικοί δείκτες για τα δεδομένα έργων και απαλλοτριώσεων, ενώ στη συνέχεια γίνεται ανάλυση τους για την εύρεση των κόκκινων σημαιών. Ακολουθεί παρουσίαση και σχολιασμός των αποτελεσμάτων.

Τέλος στο έβδομο κεφάλαιο παρουσιάζονται τα συμπεράσματα και αναφέρονται ιδέες και προτάσεις για μελλοντική εργασία.



Σχήμα 1.1: Η αρχιτεκτονική του συστήματος

Κεφάλαιο 2

Εθνικό Στρατηγικό Πλαίσιο Αναφοράς

Η Ελλάδα τα τελευταία χρόνια βρίσκεται σε μία δύσκολη περίοδο της ιστορίας της. Μια περίοδο βαθιάς κρίσης που χαρακτηρίζεται από οικονομική ύφεση, περικοπές σε μισθούς και συντάξεις, ανεργία, φτώχεια και στενότητα πόρων. Μέσα σε αυτή την κατάσταση και την προσπάθεια που κάνει η χώρα για να ανακάμψει, τα κοινοτικά κονδύλια του Εθνικού Στρατηγικού Πλαισίου Αναφοράς (ΕΣΠΑ) αποτελούν το βασικότερο χρηματοδοτικό μοχλό για την ανάπτυξη της οικονομίας και την υποστήριξη της κοινωνικής συνοχής στην χώρα μας.

Τί ακριβώς είναι όμως το ΕΣΠΑ και πώς προέκυψε; Για να μάθει κανείς θα πρέπει να ανατρέξει στην πολιτική της συνοχής της Ευρωπαϊκής Ένωσης (ΕΕ).

2.1 Η Πολιτική Συνοχής της ΕΕ

Η πολιτική συνοχής της ΕΕ είναι η πολιτική πίσω από τα εκατοντάδες χιλιάδες έργα ανά την Ευρώπη τα οποία χρηματοδοτούνται από το Ευρωπαϊκό Ταμείο Περιφερειακής Ανάπτυξης (ΕΤΠΑ), το Ευρωπαϊκό Κοινωνικό Ταμείο (ΕΚΤ) και το Ταμείο Συνοχής (ΤΣ) [17].

Η οικονομική και κοινωνική συνοχή - όπως ορίζεται στην ενιαία ευρωπαϊκή πράξη του 1986 - αφορά τη "μείωση των ανισοτήτων μεταξύ των διαφόρων περιφερειών και την καθυστέρηση των λιγότερο ευνοημένων" [18]. Στη Συνθήκη της Λισαβόνας, προστίθεται μία ακόμα παράμετρος στη συνοχή καθώς γίνεται λόγος και για "εδαφική συνοχή" δίνοντας ιδιαίτερη έμφαση «στις αγροτικές περιοχές, τις ζώνες που συντελείται βιομηχανική μετάβαση και τις περιοχές που πλήττονται από σοβαρά και μόνιμα φυσικά ή δημογραφικά προβλήματα, όπως οι

υπερβόρειες περιοχές που είναι ιδιαίτερα αραιοκατοικημένες και οι νησιωτικές, διασυνωριακές και ορεινές περιοχές» [17].

Συνεπώς, σκοπός της πολιτικής συνοχής είναι να στηρίζει τη δημιουργία θέσεων εργασίας, την ανταγωνιστικότητα των επιχειρήσεων, την οικονομική ανάπτυξη, την αειφόρο ανάπτυξη και τη βελτίωση της ποιότητας ζωής των πολιτών που ζουν σε χώρες της ΕΕ. Επίσης, στοχεύει στην αλληλεγγύη μεταξύ των Κρατών-Μελών της ΕΕ, στη σύγκλιση των Περιφερειών μεταξύ τους και στη μείωση των αναπτυξιακών ανισοτήτων και γενικότερα των οικονομικών, κοινωνικών και εδαφικών διαφορών τους [19].

Κατά την περίοδο 2014-2020, ο συντονισμός και η συνεκτικότητα μεταξύ της πολιτικής συνοχής και των άλλων πολιτικών της ΕΕ που συμβάλλουν στην περιφερειακή ανάπτυξη, έχουν ενισχυθεί με τον καθορισμό κοινών διατάξεων για το ΕΤΠΑ, το ΕΚΤ, το ΤΣ, το Ευρωπαϊκό Γεωργικό Ταμείο Αγροτικής Ανάπτυξης (ΕΓΤΑΑ) και το Ευρωπαϊκό Ταμείο Θάλασσας και Αλιείας (ΕΤΘΑ). Τα πέντε αυτά ταμεία είναι γνωστά ως τα Ευρωπαϊκά Διαρθρωτικά και Επενδυτικά Ταμεία (ΕΔΕΤ) [17].

2.2 Ευρωπαϊκά Διαρθρωτικά και Επενδυτικά Ταμεία

Τα Ευρωπαϊκά Διαρθρωτικά και Επενδυτικά Ταμεία αποτελούν το βασικό επενδυτικό εργαλείο της Ευρωπαϊκής Ένωσης και συνεργάζονται για τη στήριξη της οικονομικής ανάπτυξης όλων των χωρών της [1, 2].

Το ΕΤΠΑ και το ΕΚΤ είναι τα Διαρθρωτικά Ταμεία. Αυτά μαζί με το ταμείο συνοχής (ΤΣ) αποτελούν τα βασικά ταμεία μέσω των οποίων υλοποιείται η περιφερειακή πολιτική της ΕΕ [3].

2.2.1 Ευρωπαϊκό Ταμείο Περιφερειακής Ανάπτυξης

Ιδρύθηκε το 1975 και χρηματοδοτεί την ανάπτυξη των περιφερειακών οικονομιών, τις οικονομικές αλλαγές, την προώθηση της ανταγωνιστικότητας καθώς και την εδαφική συνεργασία σε όλη την ΕΕ [5].

Το ΕΤΠΑ¹ με τη χρηματοδότηση του επιδιώκει την ενίσχυση της οικονομικής και κοινωνικής συνοχής στην ΕΕ διορθώνοντας ανισορροπίες μεταξύ περιφερειών.

¹http://ec.europa.eu/regional_policy/el/funding/erdf/

Οι επενδύσεις του επικεντρώνονται σε αρκετούς βασικούς τομείς προτεραιότητας όπως είναι η Έρευνα και καινοτομία, το ψηφιακό θεματολόγιο, η στήριξη των μικρομεσαίων επιχειρήσεων (ΜΜΕ) και η οικονομία με μειωμένη χρήση άνθρακα [4].

Ανάλογα με την κατηγορία της περιφέρειας κατανέμονται και οι πόροι σε αυτές τις προτεραιότητες. Στις πιο ανεπτυγμένες περιφέρειες, τουλάχιστον το 80% της χρηματοδότησης πρέπει να εστιάζει τουλάχιστον σε δύο από αυτές τις προτεραιότητες. Στις περιφέρειες μετάβασης, αυτό ισχύει για 60% της χρηματοδότησης ενώ το ποσοστό για τις λιγότερο ανεπτυγμένες είναι 50% [4].

2.2.2 Ευρωπαϊκό Κοινωνικό Ταμείο

Το ΕΚΤ¹ ιδρύθηκε το 1958. Αποτελεί ένα από τα κυριότερα χρηματοδοτικά όργανα της ΕΕ για τη στήριξη των εθνικών πολιτικών που αποσκοπούν στην αύξηση της απασχόλησης και των ευκαιριών σε αυτή, στη βελτίωση της ποιότητας και παραγωγικότητας της εργασίας, καθώς και στη μείωση του κοινωνικού αποκλεισμού και των περιφερειακών ανισοτήτων σε θέματα απασχόλησης [7].

Επενδύει σε ανθρώπους, και εκτός από την απασχόληση προσπαθεί να αυξήσει τις ευκαιρίες και στην εκπαίδευση σε όλη την Ευρωπαϊκή Ένωση. Επίσης έχει σκοπό να βελτιώσει την κατάσταση των ανθρώπων που διατρέχουν τον κίνδυνο φτώχειας. Οι παρεμβάσεις του ΕΚΤ καλύπτουν όλες τις περιφέρειες της ΕΕ [6].

2.2.3 Ταμείο Συνοχής

Το Ταμείο Συνοχής² λειτουργεί από το 1994 με στόχο την μείωση των οικονομικών και κοινωνικών ανισοτήτων και την προώθηση της αειφόρου ανάπτυξης στα Κράτη Μέλη με ακαθάριστο εθνικό εισόδημα (ΑΕΕ) ανά κάτοικο κάτω του 90% του ευρωπαϊκού μέσου όρου [8, 9].

Η χρηματοδότηση από το ΤΣ πηγαίνει σε δραστηριότητες που στηρίζουν τη δημιουργία διευρωπαϊκών δικτύων μεταφορών ή σε έργα που σχετίζονται με την ενέργεια ή τις μεταφορές εφόσον επωφελούν το περιβάλλον [8].

Το ταμείο αυτό ενισχύει τα 13 Κράτη Μέλη που προσχώρησαν στην ΕΕ από το 2004 και μετά, καθώς και την Ελλάδα και την Πορτογαλία [9].

¹<http://ec.europa.eu/esf/home.jsp?langId=el>

²http://ec.europa.eu/regional_policy/index.cfm/el/funding/cohesion-fund/

2.2.4 Ευρωπαϊκό Γεωργικό Ταμείο Αγροτικής Ανάπτυξης

Το Ευρωπαϊκό γεωργικό ταμείο αγροτικής ανάπτυξης (ΕΓΤΑΑ)¹ στηρίζει την πολιτική αγροτικής ανάπτυξης της ΕΕ. Για την επίτευξη αυτού του σκοπού, χρηματοδοτεί προγράμματα αγροτικής ανάπτυξης σε όλα τις χώρες της ΕΕ. Τα προγράμματα εκπονούνται με συνεργασία της Ευρωπαϊκής Επιτροπής και των Κρατών Μελών και λαμβάνουν υπόψη τις στρατηγικές κατευθύνσεις της πολιτικής αγροτικής ανάπτυξης καθώς και τις προτεραιότητες που ορίζονται στα εθνικά σχέδια στρατηγικής [10].

2.2.5 Ευρωπαϊκό Ταμείο Θάλασσας και Αλιείας

Το ΕΤΘΑ² είναι το Ταμείο Θαλάσσιας και Αλιευτικής Πολιτικής της ΕΕ για την περίοδο 2014-2020. Βοηθάει τους αλιείς με τη μετάβαση στη βιώσιμη αλιεία, στηρίζει τις παράκτιες κοινότητες στη διαφοροποίηση των οικονομιών τους, χρηματοδοτεί έργα που δημιουργούν νέες θέσεις εργασίας και βελτιώνουν την ποιότητα ζωής κατά μήκος των Ευρωπαϊκών ακτών και διευκολύνει την πρόσβαση στη χρηματοδότηση.

2.3 Προγραμματισμός και Υλοποίηση

Η περιφερειακή πολιτική της ΕΕ αγγίζει όλους τους τομείς και τα επίπεδα της - τόσο σε ευρωπαϊκό και εθνικό επίπεδο μέχρι και το επίπεδο των τοπικών κοινοτήτων. Η στρατηγική προώθησης και στήριξης της αρμονικής ανάπτυξης των Κρατών Μελών και των περιφερειών της ΕΕ αποτελεί μέρος της πολιτικής της συνοχής. Η πολιτική αυτή πραγματοποιείται μέσα από τη στενή συνεργασία της Ευρωπαϊκής Επιτροπής με τους εθνικούς και περιφερειακούς φορείς [12].

Το πλαίσιο της πολιτικής συνοχής έχει καθοριστεί για μια περίοδο 7 ετών. Η τρέχουσα περίοδος καλύπτει τα έτη 2014-2020.

Στην πρώτη φάση υλοποίησης, το Ευρωπαϊκό Συμβούλιο και το Ευρωπαϊκό Κοινοβούλιο αποφασίζουν από κοινού για τον προϋπολογισμό της πολιτικής και τους κανόνες για τη χρήση της, βάσει μιας πρότασης της Ευρωπαϊκής Επιτροπής. Εκτός των κοινών κανόνων για τα ΕΔΕΤ υπάρχουν επίσης συγκεκριμένοι κανόνες για κάθε ταμείο ξεχωριστά [12].

¹https://ec.europa.eu/agriculture/rural-development-2014-2020_el

²https://ec.europa.eu/fisheries/cfp/emff_el

Στην αμέσως επόμενη φάση, ξεκινάει μια διαδικασία διαβούλευσης μεταξύ της Ευρωπαϊκής Επιτροπής και των χωρών της ΕΕ, όπου αναλύονται σχολαστικά οι αρχές και οι προτεραιότητες της πολιτικής συνοχής. Ακολούθως, κάθε Κράτος Μέλος υποβάλλει ένα προσχέδιο Σύμβασης Εταιρικής Σχέσης, το οποίο εξετάζει συνοπτικά την στρατηγική της χώρας και προτείνει μια λίστα από προγράμματα. Επιπλέον, τα Κράτη Μέλη παρουσιάζουν προσχέδια Επιχειρησιακών Προγραμμάτων (ΕΠ). Αυτά τα προγράμματα μπορούν να καλύπτουν ολόκληρη τη χώρα ή/και περιφέρειες. Σε αυτά δύναται να συμπεριλαμβάνονται και διακρατικά προγράμματα [12].

Στη συνέχεια, η Επιτροπή διαπραγματεύεται με τους φορείς της κάθε χώρας για το τελικό περιεχόμενο της Σύμβασης, καθώς και για κάθε πρόγραμμα. Η σύμβαση αυτή αποτελεί το Εταιρικό Σύμφωνο για το Πλαίσιο Ανάπτυξης (ΕΣΠΑ) και κάθε Κράτος Μέλος έχει το δικό του ΕΣΠΑ. Στον προγραμματισμό και τη διαχείριση των ΕΠ μπορούν να συμμετέχουν εργαζόμενοι και εργοδότες καθώς και φορείς της κοινωνίας των πολιτών [12].

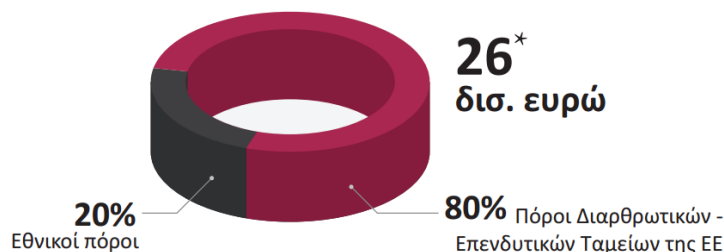
Η υλοποίηση των προγραμμάτων γίνεται από τα ίδια τα Κράτη Μέλη και τις περιφέρειες τους. Την οργάνωση της επιλογής, παρακολούθησης και αξιολόγησης των χιλιάδων έργων αναλαμβάνουν οι *Διαχειριστικές Αρχές* που ορίζονται σε κάθε χώρα ή περιφέρεια. Στη συνέχεια, αφού η Επιτροπή δεσμεύσει τα απαραίτητα κεφάλαια για να επιτρέψει στις χώρες να αρχίσουν να τα δαπανούν στα προγράμματά τους, πληρώνει την πιστοποιημένη δαπάνη σε κάθε χώρα [12].

Τέλος, κάθε πρόγραμμα παρακολουθείται από την Επιτροπή μαζί με την ενδιαφερόμενη χώρα και καθ' όλη τη διάρκεια της προγραμματικής περιόδου υποβάλλονται αναφορές για την εξέλιξη της πορείας τους [12].

2.4 Το ΕΣΠΑ στην Ελλάδα

Η Ελλάδα όπως και κάθε Κράτος Μέλος έχει το δικό της ΕΣΠΑ. Μέσω της εφαρμογής του ΕΣΠΑ "επιδιώκεται η αντιμετώπιση των διαρθρωτικών αδυναμιών της χώρας που συνετέλεσαν στην εμφάνιση της οικονομικής κρίσης αλλά και των προβλημάτων, οικονομικών και κοινωνικών, που αυτή δημιούργησε" [13]. Το Εταιρικό Σύμφωνο για το Πλαίσιο Ανάπτυξης (ΕΣΠΑ) 2014-2020 είναι το στρατηγικό σχέδιο για την ανάπτυξη της χώρας που επιδιώκει την επίτευξη των στόχων της πολιτικής συνοχής της ΕΕ και της Στρατηγικής "Ευρώπη 2020". Η Ελλάδα ενεργοποίησε τα προγράμματα της νέας περιόδου πριν από κάθε άλλη χώρα της ΕΕ. Για την προγραμματική περίοδο 2014-2020 αναμένεται να διατεθούν μέσω

του ΕΣΠΑ περί τα 26 δις ευρώ. Από αυτά όπως φαίνεται και στο σχήμα 2.1 το 80% προέρχεται από πόρους της ΕΕ και το υπόλοιπο 20% καλύπτεται από εθνικά κεφάλαια [14].



Σχήμα 2.1: Ποσοστό χρηματοδότησης του ΕΣΠΑ 2014-2020 από ΕΕ και Ελλάδα¹

Τα χρήματα αυτά θα διατεθούν για την υλοποίηση 20 Επιχειρησιακών Προγραμμάτων (ΕΠ), από τα οποία τα 7 είναι Τομεακά και τα 13 Περιφερειακά. Τα Τομεακά Προγράμματα αφορούν έναν ή περισσότερους τομείς και έχουν ως γεωγραφικό πεδίο εφαρμογής όλη τη χώρα ενώ τα 13 Περιφερειακά Επιχειρησιακά Προγράμματα (ΠΕΠ), αφορούν το καθένα μία από τις ελληνικές Περιφέρειες και περιλαμβάνουν δράσεις περιφερειακής εμβέλειας [15]. Η κατανομή των πόρων του ΕΣΠΑ ανά ΕΠ φαίνεται στο σχήμα 2.2.

ΕΠ Ανταγωνιστικότητα, Επιχειρηματικότητα και Καινοτομία (4,7 δισ. ευρώ)	ΕΠ Υποδομές Μεταφορών, Περιβάλλον και Αειφόρος Ανάπτυξη (5,2 δισ. ευρώ)
ΕΠ Ανάπτυξη Ανθρώπινου Δυναμικού, Εκπαίδευση και Διά Βίου Μάθηση (2,7 δισ. ευρώ)	ΕΠ Μεταρρύθμιση Δημόσιου Τομέα (0,5 δισ. ευρώ)
Πρόγραμμα Αγροτικής Ανάπτυξης (6 δισ. ευρώ)	ΕΠ Τεχνική Βοήθεια (0,4 δισ. ευρώ)
ΕΠ Αλιείας και Θάλασσας (0,5 δισ. ευρώ)	13 ΠΕΠ, ένα για κάθε μία από τις 13 διοικητικές Περιφέρειες της χώρας (5,8 δισ. ευρώ)

Σχήμα 2.2: Κατανομή κονδυλίων του ΕΣΠΑ 2014-2020 ανά Επιχειρησιακό Πρόγραμμα³

³<https://www.espa.gr/el/pages/Factsheet.aspx>

Κατά την προηγούμενη προγραμματική περίοδο, 2007-2013, τα χρήματα που απορροφήθηκαν από την Ελλάδα ήταν 20,1 δις ευρώ. Αυτά συνέβαλαν μεταξύ άλλων στην δημιουργία 21.000 θέσεων εργασίας σε μικρομεσαίες επιχειρήσεις, τη σύσταση 2.611 νέων επιχειρήσεων και τη κατασκευή 144 χιλιομέτρων νέου οδικού δικτύου [16].

Κεφάλαιο 3

Σημασιολογικός Ιστός, Διασυνδεδεμένα Δεδομένα και Οντολογίες

3.1 Σημασιολογικός ιστός

Ο Παγκόσμιος Ιστός έχει αλλάξει τόσο τον τρόπο με τον οποίο οι άνθρωποι επικοινωνούν μεταξύ τους, όσο και τον τρόπο με τον οποίο ασκούν τις επαγγελματικές τους δραστηριότητες. Βρίσκεται στην καρδιά μιας επανάστασης που μεταμορφώνει τον ανεπτυγμένο κόσμο προς μία οικονομία γνώσης και πιο γενικά σε μία κοινωνία γνώσης [20].

Η εξέλιξη αυτή έχει αλλάξει τον τρόπο με τον οποίο αντιλαμβανόμαστε τους υπολογιστές, οι οποίοι αρχικά χρησιμοποιούνταν για αριθμητικούς υπολογισμούς. Αυτή τη στιγμή η κυρίαρχη χρήση τους έγκειται στην επεξεργασία πληροφοριών, σε τυπικές εφαρμογές όπως είναι τα συστήματα βάσεων δεδομένων, την επεξεργασία κειμένου και το gaming. Σήμερα παρατηρείται μια μετάβαση που ως επίκεντρο έχει την άποψη ότι οι υπολογιστές αποτελούν σημεία εισόδου στις λεωφόρους της πληροφορίας [20].

Από το 1998 [21] (ημέρα εφεύρεσης του Παγκόσμιου Ιστού από τον Βρετανό επιστήμονα Tim Berners-Lee) μέχρι σήμερα, ο Ιστός εξελίχθηκε σε ένα παγκόσμιο χώρο πληροφοριών που μετρά πάνω από ένα δισεκατομμύριο ιστοσελίδες¹, οι οποίες καλύπτουν σχεδόν κάθε θέμα. Το πρόβλημα που εγείρεται όμως είναι η δυσκολία εύρεσης της επιθυμητής πληροφορίας. Επειδή το περιεχόμενο των ιστοσελίδων παρουσιάζεται σε γλώσσα κατανοητή μόνο από τον άνθρωπο, οι μη-

¹<http://www.internetlivestats.com/total-number-of-websites/>

χανές αναζήτησης δεν έχουν πρόσβαση στην σημασιολογία του περιεχομένου, με αποτέλεσμα οι δυνατότητες τους να περιορίζονται κυρίως στην αναζήτηση λέξεων-κλειδιών [20]. Ο χρήστης παίρνει μια λίστα από ιστοσελίδες που περιέχουν τις λέξεις-κλειδιά και καλείται να περιηγηθεί σε αυτές για να βρει την πληροφορία που ψάχνει. Σε περίπτωση που δεν βρει την πληροφορία που ζητά αναγκάζεται να κάνει πολλές αναζητήσεις και πολλές φορές να συνδυάσει τις πληροφορίες που θα βρει.

Μία λύση στο πρόβλημα αυτό δίνει η αναπαράσταση του περιεχομένου του Ιστού σε μορφή κατανοητή από τις μηχανές [20]. Το εγχείρημα του Σημασιολογικού Ιστού που οραματίστηκε ο ίδιος ο Tim Berners-Lee στοχεύει στο να κάνει τα δεδομένα που υπάρχουν στον Ιστό να έχουν νόημα για τους υπολογιστές, με άλλα λόγια να καταστεί δυνατό οι υπολογιστές να ερμηνεύουν και να επεξεργάζονται τις πληροφορίες [22]. Ο Tim Berners-Lee ορίζει το Σημασιολογικό Ιστό ως εξής:

“The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation.” [22]

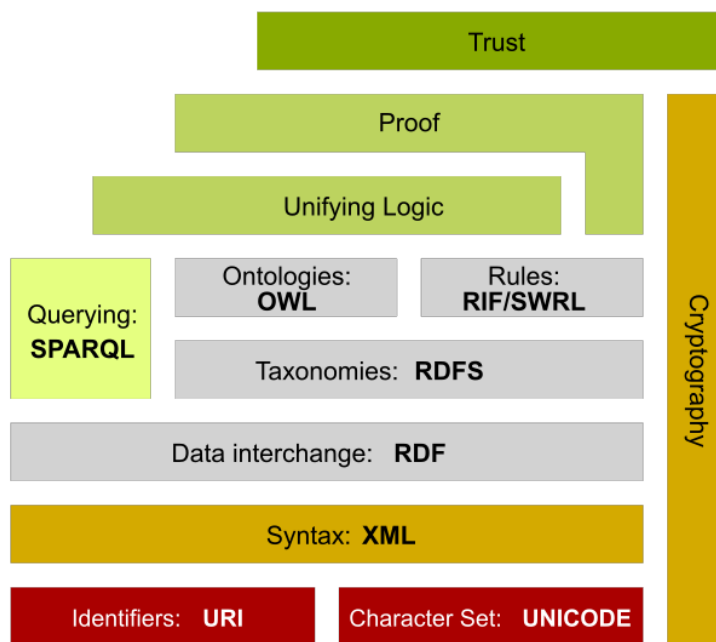
3.1.1 Τυποποίηση και Αρχιτεκτονική Σημασιολογικού Ιστού

Η Τυποποίηση του Σημασιολογικού Ιστού βρίσκεται υπό την εποπτεία του World Wide Web Consortium (W3C) [23].

Οι τεχνολογίες που έχουν αναπτυχθεί και αποτελούν πρότυπα του W3C περιλαμβάνουν [24]:

- Resource Description Framework (RDF)
- RDF Schema (RDFS)
- Simple Knowledge Organization System (SKOS)
- SPARQL, an RDF query language
- Notation3 (N3)
- N-Triples
- Turtle
- Web Ontology Language (OWL)
- Rule Interchange Format (RIF)

Η αρχιτεκτονική του Σημασιολογικού Ιστού βασίζεται σε επίπεδα (layers) όπως φαίνεται στο σχήμα 3.1. Κάθε επίπεδο εκμεταλλεύεται και χρησιμοποιεί τις δυνατότητες των από κάτω επιπέδων [25].



Σχήμα 3.1: Η αρχιτεκτονική του Σημασιολογικού Ιστού ως μια στοίβα επιπέδων

Το περιεχόμενο των επιπέδων αυτών αναλύεται παρακάτω:

- **Universal Resource Identifier (URI):** Το Ενιαίο Αναγνωστικό Πόρου είναι μια σειρά από χαρακτήρες που χρησιμοποιούνται για τον προσδιορισμό λογικών ή φυσικών πόρων. Το Internationalized Resource Identifier (IRI), γενίκευση του URI, παρέχει τα μέσα για την μοναδική αναγνώριση πόρων του Σημασιολογικού Ιστού.
- **UNICODE:** Το UNICODE χρησιμεύει στην αναπαράσταση και χειραγώγηση κειμένων από διάφορες γλώσσες.
- **XML:** Το XML είναι μια γλώσσα σήμανσης που επιτρέπει τη δημιουργία εγγράφων τα οποία αποτελούνται από δομημένα δεδομένα. Ο Σημασιολογικός Ιστός δίνει νόημα (σημασιολογία) σε δομημένα δεδομένα.
- **RDF:** Το RDF (Resource Description Framework) είναι ένα πλαίσιο για την δημιουργία τριπλέτων. Η πληροφορία που σχετίζεται με έναν πόρο αναπαριστάται ως ένας γράφος. Για τον λόγο αυτό ο Σημασιολογικός Ιστός καλείται πολλές φορές ως γιγαντιαίος παγκόσμιος γράφος.

- **RDFS:** Το RDF Schema παρέχει το βασικό λεξιλόγιο για το RDF. Δίνει τη δυνατότητα να δημιουργηθούν κλάσεις, ιδιότητες, ιεραρχίες κλάσεων και ιδιοτήτων.
- **OWL:** Η Web Ontology Language επεκτείνει το RDFS προσθέτοντας πιο προηγμένες δομές για την περιγραφή της σημασιολογίας των προτάσεων RDF. Βασίζεται στην περιγραφική λογική και προσθέτει τη δυνατότητα συλλογισμών στον σημασιολογικό ιστό.
- **SPARQL:** Είναι η γλώσσα ερωτημάτων για RDF. Μπορεί να χρησιμοποιηθεί για την ανάκτηση πληροφοριών από RDF δεδομένα.
- **RIF:** Το Rule Interchange Format είναι μια μορφή ανταλλαγής κανόνων. Χρησιμοποιείται κυρίως για να περιγραφούν σχέσεις οι οποίες είναι δύσκολο να περιγραφούν άμεσα μέσω της OWL.

3.2 Διασυνδεδεμένα Δεδομένα

Ο Σημασιολογικός Ιστός δεν αφορά μόνο την δημοσίευση δεδομένων στον ιστό αλλά και την δημιουργία συνδέσεων μεταξύ τους, έτσι ώστε οποιοσδήποτε άνθρωπος ή μηχανή να μπορεί να εξερευνεί τον ιστό των δεδομένων. Με τα διασυνδεδεμένα δεδομένα μπορεί να βρεθούν άλλα σχετικά δεδομένα [26].

Όπως ο ιστός υπερκειμένου, έτσι και ο ιστός δεδομένων αποτελείται από έγγραφα. Η διαφορά τους είναι ότι στον δεύτερο οι σχέσεις μεταξύ των εγγράφων περιγράφονται εννοιολογικά.

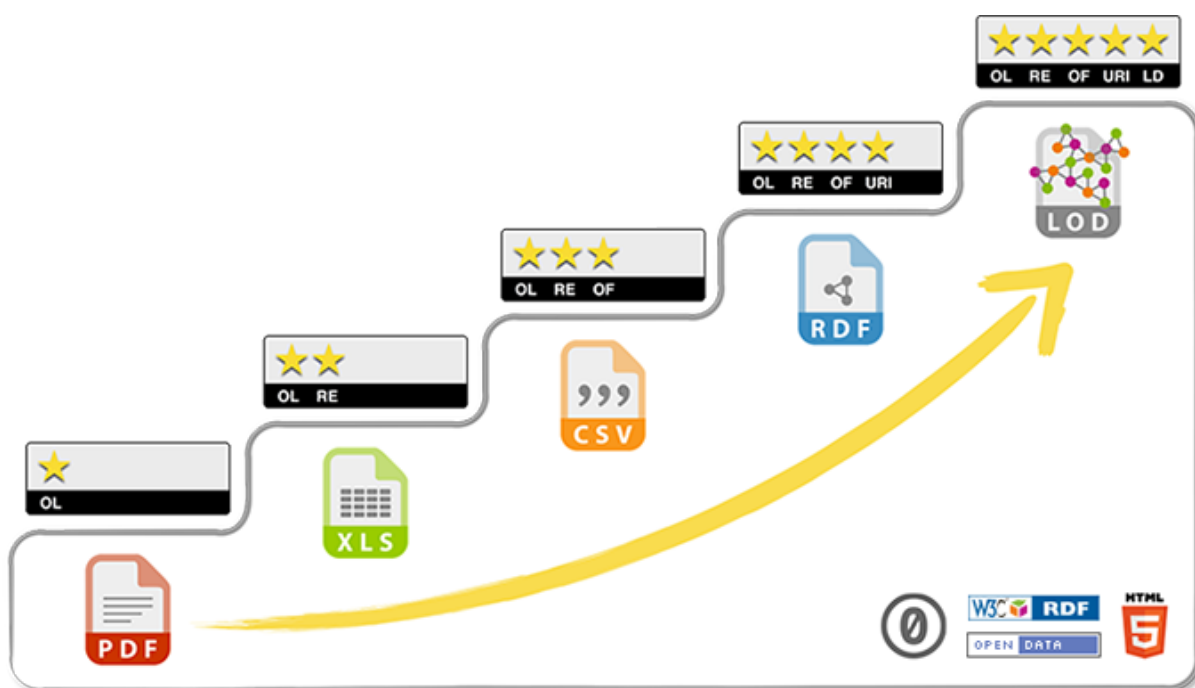
Ο Tim Berners-Lee περιέγραψε μια σειρά από κανόνες για τη δημοσίευση δεδομένων στο διαδίκτυο με έναν τρόπο που όλα τα δημοσιευμένα δεδομένα να γίνουν μέρος ενός ενιαίου παγκόσμιου χώρου δεδομένων. Οι τέσσερις αυτοί κανόνες που είναι γνωστοί ως Αρχές των Διασυνδεδεμένων Δεδομένων είναι οι εξής [26]:

1. Χρήση URIs για την ονομασία των πόρων.
2. Χρήση HTTP URIs ώστε οι πόροι να μπορούν να αναζητηθούν από τους χρήστες. Η διαδικασία της επίσκεψης ενός URI στον ιστό με σκοπό την ανάκτηση της πληροφορίας σχετικά με τον αναφερόμενο πόρο λέγεται dereferencing HTTP URIs.
3. Παροχή χρήσιμης πληροφορίας με τη χρήση των προτύπων διασυνδεδεμένων δεδομένων (RDF/SPARQL).

4. Χρήση συνδέσμων προς άλλα URIs ώστε ο χρήστης να μπορεί να ανακαλύψει περισσότερα δεδομένα.

Το 2010 [26], ο Tim Berners-Lee πρότεινε ένα σύστημα για την αξιολόγηση των Ανοιχτών Διασυνδεδεμένων Δεδομένων. Στόχος του ήταν να ενθαρρύνει τους ιδιοκτήτες των δεδομένων αυτών να δημοσιεύουν τα δεδομένα τους με βάση τις Αρχές των Διασυνδεδεμένων Δεδομένων. Σύμφωνα με αυτό το σύστημα θα πρέπει:

- Τα δεδομένα να είναι διαθέσιμα στο διαδίκτυο (σε οποιαδήποτε μορφή), αλλά με ανοιχτή άδεια χρήσης ώστε να είναι - πραγματικά - *Ανοιχτά Δεδομένα*.
- Τα δεδομένα να είναι διαθέσιμα σε μηχαναγνώσιμη μορφή (π.χ Excel αντί για μια σαρωμένη εικόνα ενός πίνακα σε pdf).
- Τα δεδομένα να είναι από τη μια διαθέσιμα σε μηχαναγνώσιμη και επιπλέον σε μη-ιδιόκτητη μορφή (π.χ. CSV αντί Excel).
- Επιπλέον των προηγούμενων, να γίνεται χρήση ανοιχτών προτύπων από το W3C (RDF και SPARQL) για τον εντοπισμό των πόρων, ώστε οι χρήστες να στραφούν στα δεδομένα αυτά.
- Όλα τα προηγούμενα και επιπρόσθετα σύνδεση τους με άλλα δεδομένα (Linked Open Data).



Σχήμα 3.2: Το σύστημα 5 αστερών των Ανοιχτών Διασυνδεδεμένων Δεδομένων

3.3 Οντολογίες

Οντολογία στο πλαίσιο της μεταφυσικής, κλάδου της φιλοσοφίας, είναι η φιλοσοφική μελέτη της φύσης των όντων, της ύπαρξης ή της πραγματικότητας καθώς και των βασικών κατηγοριών των όντων και των μεταξύ τους σχέσεων [27]. Στην επιστήμη των υπολογιστών, η οντολογία είναι ένας τυπικός και σαφής ορισμός μιας κοινής και συμφωνημένης εννοιολογικής μορφοποίησης που αφορά σε ένα πεδίο ενδιαφέροντος. Αυτή η τυπική αναπαράσταση γνώσης ως ένα σύνολο εννοιών, σχέσεων και ιδιοτήτων μπορεί να χρησιμοποιηθεί για συλλογιστική (εξαγωγή συμπερασμάτων/νέας γνώσης) και για την δομημένη περιγραφή γνώσης ενός πεδίου ενδιαφέροντος [28].

Όπως αναφέρθηκε, οι σχέσεις αυτές αποτυπώνονται στον σημασιολογικό ιστό μέσω των RDF τριπλέτων. Για την κατανόηση αυτών των σχέσεων από τις μηχανές δημιουργούνται τα λεξιλόγια (vocabularies) - οντολογίες (ontologies). Οι σύγχρονες οντολογίες μοιράζονται πολλές δομικές ομοιότητες, ανεξαρτήτως της γλώσσας στην οποία εκφράζονται. Τα πιο κοινά στοιχεία που συναντά κανείς σε μία οντολογία είναι [29]:

- **Άτομα (Individuals):** Στιγμιότυπα ή αντικείμενα.
- **Κλάσεις (Classes):** Σύνολα, συλλογές, έννοιες, κλάσεις στον προγραμματισμό, τύπους αντικειμένων ή είδη πραγμάτων.
- **Ιδιότητες (Attributes):** Πτυχές, ιδιότητες, χαρακτηριστικά και παραμέτρους που μπορεί να έχουν τα αντικείμενα και οι κλάσεις.
- **Σχέσεις (Relations):** Τρόποι με τους οποίους οι κλάσεις και τα άτομα μπορούν να συσχετίζονται μεταξύ τους.
- **Λειτουργικοί όροι (Function terms):** Σύνθετες δομές που σχηματίζονται από ορισμένες συσχετίσεις και μπορούν να χρησιμοποιηθούν στη θέση ενός ατομικού όρου σε μια πρόταση.
- **Περιορισμοί (Restrictions):** Αυστηρά δηλωμένες περιγραφές για το τι πρέπει να ισχύει προκειμένου ένας ισχυρισμός να γίνει αποδεκτός ως είσοδος.
- **Κανόνες (Rules):** Προτάσεις της μορφής εάν-τότε (προγενέστερη - συνακόλουθη) που περιγράφουν τα λογικά συμπεράσματα που μπορούν να εξαχθούν από έναν ισχυρισμό.

- **Αξιώματα (Axioms):** Ισχυρισμοί (συμπεριλαμβανομένων των κανόνων) σε μια λογική μορφή, που μαζί αποτελούν τη συνολική θεωρία που περιγράφει η οντολογία στον τομέα εφαρμογής της.
- **Γεγονότα (Events):** Αλλαγές των χαρακτηριστικών ή σχέσεων

Υπάρχουν τρία είδη οντολογιών, η Domain ontology, η Upper ontology και η Hybrid ontology. Η Domain ontology αναπαριστά έννοιες που ανήκουν στον πραγματικό κόσμο. Η Upper ontology είναι μια οντολογία που αποτελείται από γενικούς όρους (όπως αντικείμενο, ιδιότητα, σχέση) που είναι κοινοί σε όλους τους τομείς. Μια σημαντική λειτουργία μιας Upper ontology είναι να υποστηρίξει την ευρεία σημασιολογική διαλειτουργικότητα μεταξύ ενός μεγάλου αριθμού Domain ontologies, παρέχοντας μια κοινή αφετηρία για τη διατύπωση των ορισμών. Η Hybrid ontology ουσιαστικά είναι συνδυασμός των δύο προηγούμενων οντολογιών [29].

Για την κωδικοποίηση των οντολογιών χρησιμοποιούνται τυποποιημένες γλώσσες, οι γλώσσες οντολογιών (ontology languages). Υπάρχουν πολλές τέτοιες γλώσσες για οντολογίες, άλλες είναι ιδιόκτητες (proprietary) και άλλες βασίζονται σε πρότυπα (standards-based). Η πιο γνωστή αυτών είναι η OWL [30].

Κεφάλαιο 4

Διαδικασία Απόκτησης Δεδομένων

Τα δεδομένα τα οποία χρησιμοποιήθηκαν στην εργασία αυτή προήλθαν από την ιστοσελίδα ANAPTYXH.gov.gr¹, η οποία αποτελεί τον επίσημο διαδικτυακό τόπο του Υπουργείου Ανάπτυξης και Ανταγωνιστικότητας (πλέον Υπουργείο Οικονομίας και Ανάπτυξης)² για την παροχή αναλυτικής πληροφόρησης της πορείας υλοποίησης των έργων του ΕΣΠΑ [31].

Με τη χρήση κώδικα που συντάχθηκε στη γλώσσα προγραμματισμού Python τα δεδομένα εξήχθησαν από το Open Data API³ της ιστοσελίδας και αποθηκεύτηκαν σε τοπική βάση δεδομένων MySQL που δημιουργήθηκε.

4.1 ANAPTYXH.gov.gr

4.1.1 Τι είναι

Ο ιστοχώρος, ANAPTYXH.gov.gr, αναπτύχθηκε από την Ειδική Υπηρεσία Ολοκληρωμένου Πληροφοριακού Συστήματος (ΟΠΣ)⁴, σε συνεργασία με την Ειδική Υπηρεσία Στρατηγικής Σχεδιασμού και Αξιολόγησης Αναπτυξιακών Προγραμμάτων (ΕΥΣΣΑΑΠ)⁵ και τη Μονάδα Οργάνωσης της Διαχείρισης Αναπτυξιακών Προγραμμάτων (ΜΟΔ Α.Ε.)⁶. Πρωταρχικός στόχος του είναι τόσο η διασφάλιση της διαφάνειας, όσο και η διάχυση πληροφόρησης σχετικά με την υλοποίηση των Επιχειρησιακών Προγραμμάτων του ΕΣΠΑ [31].

¹<http://anaptyxi.gov.gr/>

²<http://www.mindev.gov.gr/>

³<http://anaptyxi.gov.gr/ODHelp.aspx>

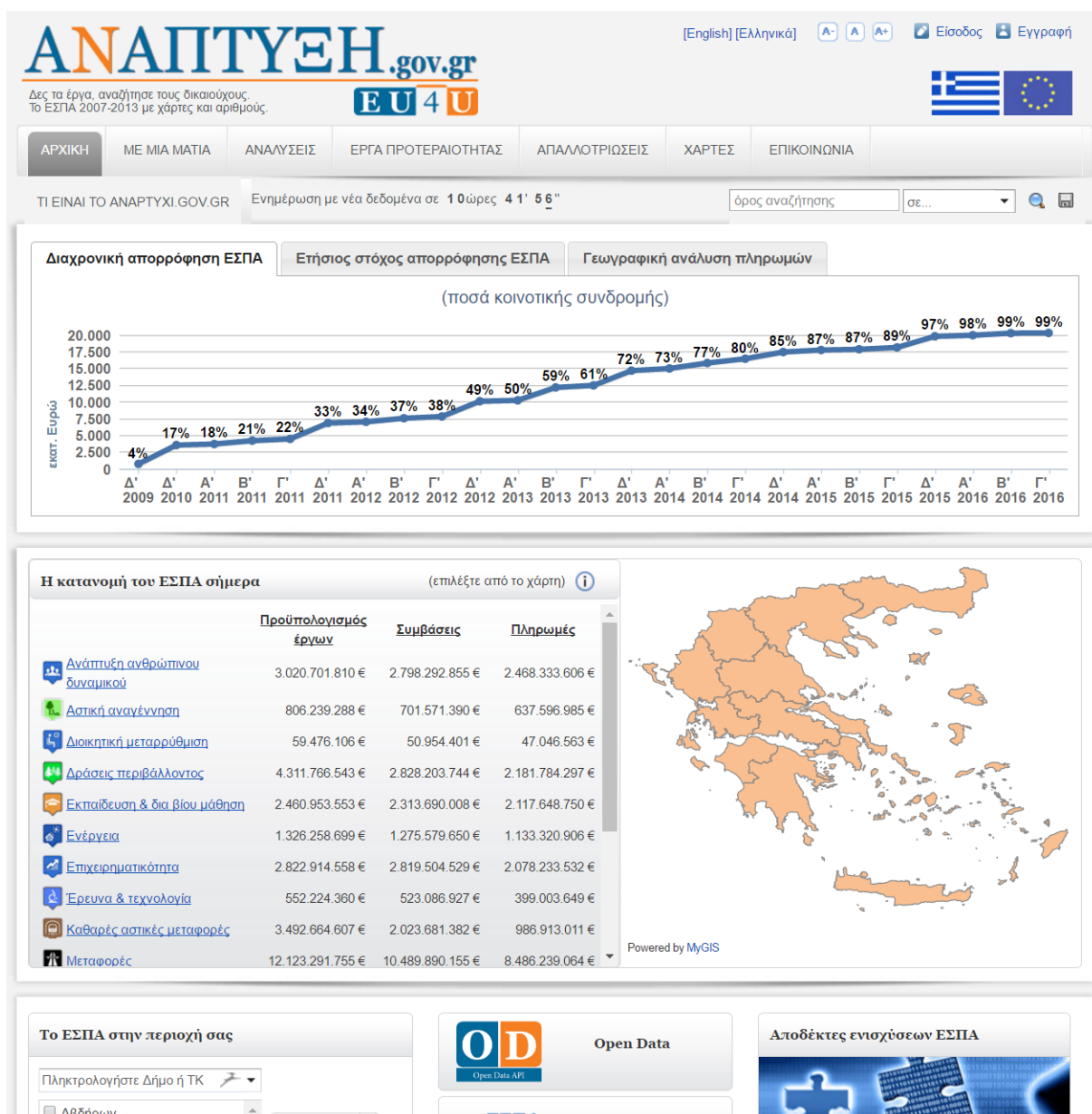
⁴<http://www.ops.gr>

⁵http://www.ggea.gr/geniki_grammateia/ypiresies10.htm

⁶<https://www.mou.gr/el/Pages/Default.aspx>

Το ΟΠΣ αποτελεί το κεντρικό σύστημα πληροφόρησης, διαχείρισης και συνεχούς παρακολούθησης της πορείας υλοποίησης των συγχρηματοδοτούμενων έργων. Μέσω αυτού, οι Διαχειριστικές Αρχές, οι Ενδιάμεσες Διαχειριστικές Αρχές, οι Ενδιάμεσοι Φορείς Διαχείρισης και οι Δικαιούχοι των Επιχειρησιακών Προγραμμάτων τροφοδοτούν το ΑΝΑΠΤΥΞΗ.gov.gr με τις σχετικές πληροφορίες [31].

Όπως φαίνεται και από το σχήμα 4.1, ο καθένας μπορεί να δει καθημερινά την εικόνα του ΕΣΠΑ μέσα από διαδραστικούς χάρτες, καταλόγους έργων και στατιστικές αναλύσεις [31].



Σχήμα 4.1: Αρχική σελίδα του ΑΝΑΠΤΥΞΗ.gov.gr

4.1.2 API Ανοιχτών Δεδομένων

Η Γενική Γραμματεία Δημοσίων Επενδύσεων – ΕΣΠΑ¹ μέσω του διαδικτυακού χώρου ANAΠTYΞH.gov.gr και του API ανοιχτών δεδομένων του, παρέχει σε όλους τους ενδιαφερόμενους, μηχανογραφική πρόσβαση σε δεδομένα του ΕΣΠΑ. Η δημόσια διάθεση αυτών των δεδομένων γίνεται κατ' εφαρμογή των διατάξεων του Κεφαλαίου Α' του ν. 4305/2014 (ΦΕΚ 237/Α)² σχετικά με την ανοικτή διάθεση και περαιτέρω χρήση εγγράφων, πληροφοριών και δεδομένων του δημόσιου τομέα και στοχεύει στην ενίσχυση της διαφάνειας σε αυτόν [32].

Με την υποβολή κλήσεων στο API μπορούν να ανακτηθούν σε μορφή JSON και κατά περίπτωση σε CSV αλλά και HTML, σύνολα δεδομένων που αφορούν πράξεις, υποέργα πράξεων καθώς και λίστα πράξεων [32].

4.1.3 Κατηγορίες Πράξεων

Οι πράξεις αυτές χωρίζονται σε δύο βασικές κατηγορίες:

1. **Έργα:** "Μια ομάδα δραστηριοτήτων που αποσκοπεί στην υλοποίηση ενός ολοκληρωμένου και λειτουργικά αυτοτελούς αντικειμένου/αποτελέσματος. Τα έργα με τη σειρά τους μπορεί να αποτελούνται από υποέργα." [33]
2. **Ενισχύσεις:** "Η επιλεκτική ενίσχυση από το κράτος προς φορείς που ασκούν οικονομική δραστηριότητα (ιδιωτικούς ή δημόσιους), η οποία επηρεάζει τον ανταγωνισμό σε Ευρωπαϊκό επίπεδο. Η ενίσχυση μπορεί να έχει τη μορφή επιχορήγησης, φοροαπαλλαγής, εγγύησης, παραχώρησης χρήσης δημόσιας περιουσίας κ.α." [34]

Στην ιστοσελίδα υπάρχει επίσης μία ενότητα με 181 **έργα προτεραιότητας**. "Τα έργα αυτά επιλέχθηκαν από τις ελληνικές αρχές σε συνεργασία με τις αρμόδιες υπηρεσίες της Ευρωπαϊκής Επιτροπής και με βάση κριτήρια που σχετίζονται με την ωριμότητα, το μέγεθος και τη σημασία των κοινωνικών και οικονομικών επιπτώσεών τους" [35]. Τα έργα προτεραιότητας περιλαμβάνουν άλλα έργα ή ενισχύσεις.

¹<http://www.ggea.gr/default.htm>

²<https://mef.diaugeia.gov.gr/attachments/3b400d10-458b-4640-8eb4-7985f229d76a.pdf>

ΑΝΑΠΤΥΞΗ.gov.gr

Δες τα έργα, αναζήτησε τους δικαιούχους
Το ΕΣΠΑ 2007-2013 με χάρτες και αριθμούς.

EU 4 U

[English] [Ελληνικά] A- A A+

ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ

ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ

ΑΝΑΠΤΥΞΗ

ΤΙ ΕΙΝΑΙ ΤΟ ΑΝΑΠΤΥΧΙ.GOV.GR Τελευταία ενημέρωση: < Οκτώβριος 2016 >

Έργα προτεραιότητας

ΟΛΟΚΛΗΡΩΜΕΝΟ ΟΜΑΛΗ ΠΟΡΕΙΑ ΥΛΟΠΟΙΗΣΗΣ ΑΠΑΙΤΕΙΤΑΙ ΕΠΙΤΑΧΥΝΣΗ ΕΡΓΑ ΣΕ ΚΙΝΔΥΝΟ ΑΚΥΡΩΘΕΝ

ΚΑΤ.	ΚΩΔ.	ΤΙΤΛΟΣ	ΤΟΜΕΑΣ	ΠΡΟΓΡΑΜΜΑ	ΠΕΡΙΟΧΗ	Εγκεκριμένος Π/Υ (€)	Π/Υ Συγχρ/νος από ΕΕ (€)	ΣΥΜΒΑΣΕΙΣ(€)	ΠΛΗΡΩΜΕΣ(€)
	1	Σταθμός συμπίεσης φυσικού αερίου στη Νέα Μεσημβρία (Θεσσαλονίκη)	ΕΝΕΡΓΕΙΑ	Ανταγωνιστικότητα και Επιχειρηματικότητα	Κεντρική Μακεδονία	21.393.383	23.098.741	21.393.383	18.840.531
	2	Αγωγός Φυσικού Αερίου Υψηλής Πίεσης προς Αλιβέρι	ΕΝΕΡΓΕΙΑ	Ανταγωνιστικότητα και Επιχειρηματικότητα	Στερεά Ελλάδα, Αττική	23.449.170	24.681.926	23.449.170	11.898.976
	3	Ενεργόπολις - Κοζάνη	ΕΝΕΡΓΕΙΑ	Μακεδονία - Θράκη	Δυτική Μακεδονία	15.260.950	20.000.000	8.395.507	6.656.604
	4	Β' Αναβάθμιση Του Τερματικού Σταθμού Υδροποιημένου Φυσικού Αερίου (ΥΦΑ) της Νήσου Ρεβυθούσας (Κατασκευή 3ης δεξαμενής)	ΕΝΕΡΓΕΙΑ	Ανταγωνιστικότητα και Επιχειρηματικότητα	Αττική	14.028.615	54.962.757	14.028.615	14.028.614
	5	Πιλοτική αξιοποίηση του γεωθερμικού πεδίου Ερατεινού στο Δήμο Νέστου	ΕΝΕΡΓΕΙΑ	Μακεδονία - Θράκη	Ανατολική Μακεδονία - Θράκη	8.374.619	5.000.000	8.374.619	8.023.137
	6	Αγωγός Φυσικού Αερίου Υψηλής Πίεσης από τους Αγ. Θεοδώρους έως τη Μεγαλόπολη	ΕΝΕΡΓΕΙΑ	Ανταγωνιστικότητα και Επιχειρηματικότητα	Αττική, Πελοπόννησος	26.418.087	44.670.162	26.418.087	26.418.086
		Διασύνδεση των Κυκλάδων με το ηπειρωτικό		Ανταγωνιστικότητα					

Σχήμα 4.2: Λίστα των έργων προτεραιότητας

4.2 Τοπική Βάση Δεδομένων

Το πρώτο βήμα που χρειάστηκε να γίνει για την συλλογή των δεδομένων από το API ανοιχτών δεδομένων ήταν η δημιουργία μιας τοπικής βάσης δεδομένων (NSRFdb) όπου αρχικά θα αποθηκεύονταν και μετέπειτα θα επεξεργάζονταν τα εν λόγω δεδομένα.

Αφού μελετήθηκαν προσεκτικά τα δεδομένα που ήταν διαθέσιμα, αποφασίστηκε να δημιουργηθούν πίνακες για τα έργα (Projects), τα έργα προτεραιότητας (Priorityprojects), τα υποέργα (Subprojects), τις ενισχύσεις (Support-grants), τους φορείς (Bodies), τις απαλλοτριώσεις (Expropriations), τα επιχειρησιακά προγράμματα (Operational Programmes), τις θεματικές προτεραιότητες (Thematic

Priorities) καθώς και ότι άλλο κρίθηκε απαραίτητο. Το διάγραμμα της βάσης δεδομένων, με τους πίνακες και τις σχέσεις που τους συνδέουν, φαίνεται αναλυτικά στο σχήμα 4.3.

Όπως γίνεται αντιληπτό από το σχήμα, ο κεντρικός πίνακας είναι ο "MIScodes" όπου αποθηκεύονται όλοι οι κωδικοί αριθμοί MIS (Monitoring Information System). Οι κωδικοί είναι μοναδικοί για κάθε έργο, ενίσχυση και έργο προτεραιότητας.

Ιδιαίτερη προσοχή δόθηκε επίσης, στον τύπο δεδομένων που ορίστηκε σε κάθε πεδίο των πινάκων, όπου θα αποθηκεύονταν τα δεδομένα. Στους πίνακες 4.1 και 4.2, αναφέρονται ενδεικτικά οι τύποι που ορίστηκαν στα πεδία των πινάκων της βάσης, καθώς και η περιγραφή των τύπων αυτών.

Πίνακας 4.1: Τύπος των δεδομένων για τον πίνακα των έργων

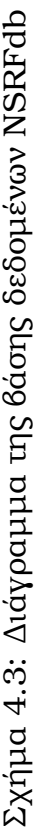
Πεδίο	Τύπος	Περιγραφή
projMIS	int(11)	Ακέραιος μέγιστου μήκους 11 ψηφίων
projTitle	varchar(500)	Αλφαριθμητικό μεγέθους 500 χαρακτήρων
projBody	varchar(300)	Αλφαριθμητικό μεγέθους 300 χαρακτήρων
currentstatusCode	tinyint(4)	Ακέραιος μέγιστου μήκους 4 ψηφίων
regionCode	smallint(6)	Ακέραιος μέγιστου μήκους 6 ψηφίων
projBudget	int(11)	Ακέραιος μέγιστου μήκους 11 ψηφίων
projContracts	int(11)	Ακέραιος μέγιστου μήκους 11 ψηφίων
projPayments	int(11)	Ακέραιος μέγιστου μήκους 11 ψηφίων
projHasImages	tinyint(4)	Ακέραιος μέγιστου μήκους 4 ψηφίων
projCompletion	float	Κινητής υποδιαστολής
projStartDate	date	Ημερομηνία με μορφή ΕΕΕΕ-ΜΜ-ΗΗ
projEndDate	date	Ημερομηνία με μορφή ΕΕΕΕ-ΜΜ-ΗΗ
projDescription	text	Πεδίο κειμένου με μέγιστο μήκος 65535
projStatusReportDate	date	Ημερομηνία με μορφή ΕΕΕΕ-ΜΜ-ΗΗ
projStatusReport	text	Πεδίο κειμένου με μέγιστο μήκος 65535
projCountSubproject	smallint(6)	Ακέραιος μέγιστου μήκους 6 ψηφίων
operationCode	tinyint(4)	Ακέραιος μέγιστου μήκους 4 ψηφίων
thematicCode	varchar(2)	Αλφαριθμητικό μεγέθους 2 χαρακτήρων
projURL	varchar(60)	Αλφαριθμητικό μεγέθους 60 χαρακτήρων
projPartofMajor	tinyint(4)	Ακέραιος μέγιστου μήκους 4 ψηφίων
projCCI	int(11)	Ακέραιος μέγιστου μήκους 11 ψηφίων

Μετά τον προσδιορισμό των πινάκων και τύπων των πεδίων, ακολούθησε η διαδικασία δημιουργίας των πινάκων στην βάση. Οι εντολές sql για τη δημιουργία των πινάκων εκτελέστηκαν από scripts που συντάχθηκαν στη γλώσσα προγραμμα-

Πίνακας 4.2: Τύπος των δεδομένων για τον πίνακα των φορέων

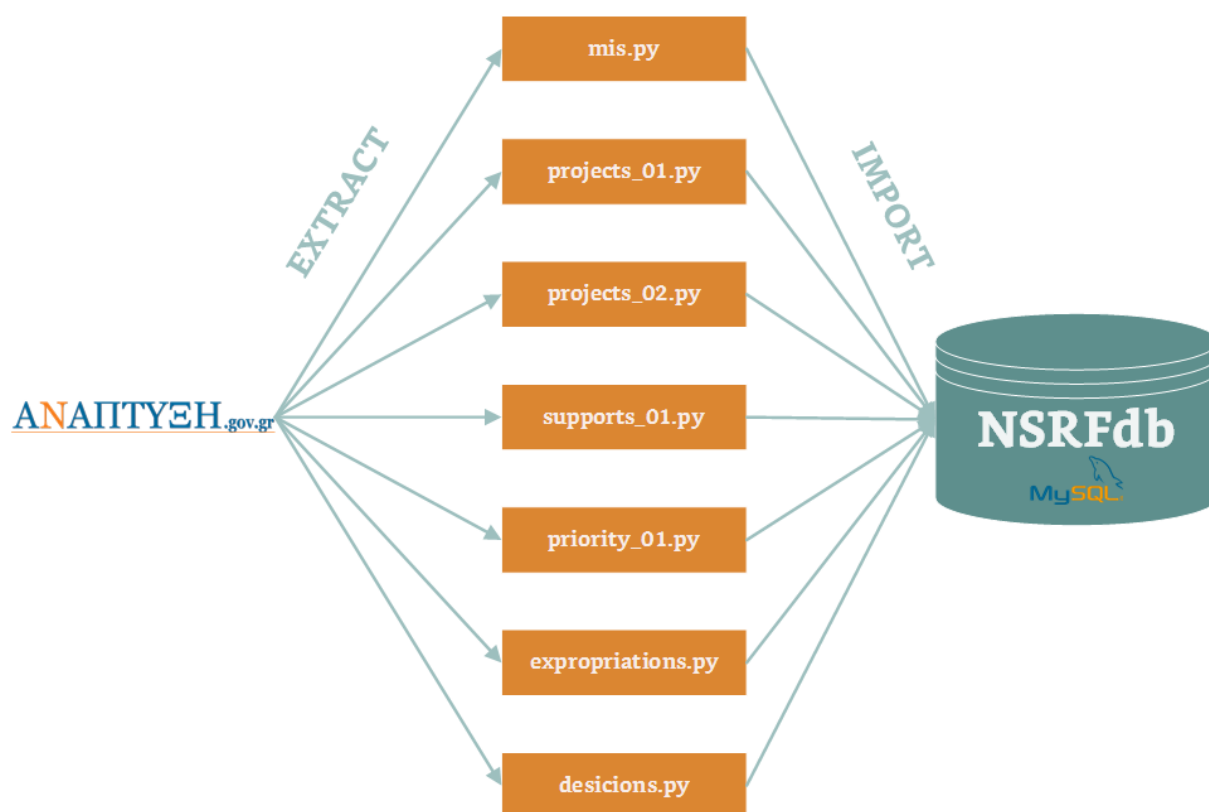
Πεδίο	Τύπος	Περιγραφή
MIScode	int(11)	Ακέραιος μέγιστου μήκους 11 ψηφίων
bodyID	int(11)	Ακέραιος μέγιστου μήκους 11 ψηφίων
bodyCode	tinyint(4)	Ακέραιος μέγιστου μήκους 4 ψηφίων
bodyName	varchar(1000)	Αλφαριθμητικό μεγέθους 1000 χαρακτήρων
bodyRepresentative	text	Πεδίο κειμένου με μέγιστο μήκος 65535
bodyAddress	varchar(1000)	Αλφαριθμητικό μεγέθους 1000 χαρακτήρων
bodyTelephone	varchar(1000)	Αλφαριθμητικό μεγέθους 1000 χαρακτήρων
bodyEmail	varchar(1000)	Αλφαριθμητικό μεγέθους 1000 χαρακτήρων
bodyFax	varchar(1000)	Αλφαριθμητικό μεγέθους 1000 χαρακτήρων

ματισμού Python.



4.3 Εξαγωγή και Αποθήκευση Δεδομένων

Όπως και για τη δημιουργία των πινάκων στη βάση έτσι και για την ανάκτηση των δεδομένων από την ΑΝΑΠΤΥΞΗ και εισαγωγή τους στη βάση δεδομένων αναπτύχθηκε κώδικας σε Python και εκτελέστηκε μια σειρά από scripts. Η διαδικασία φαίνεται στο σχήμα 4.4. Ακολουθεί μια σύντομη περιγραφή των 20 scripts που χρησιμοποιήθηκαν.



Σχήμα 4.4: Εξαγωγή δεδομένων από την ΑΝΑΠΤΥΞΗ και αποθήκευση τους στην βάση δεδομένων

1. **mis.py:** Μαζεύει όλους τους κωδικούς MIS των έργων, ενισχύσεων και έργων προτεραιότητας και τους εισάγει στον πίνακα "MIScodes".
2. **projects_01.py:** Ανακτά από το API δεδομένα που αφορούν το έργο και τα εισάγει στον πίνακα της βάσης "Projects". Τα δεδομένα περιλαμβάνουν τον κωδικό MIS, τον τίτλο του έργου, τον δικαιούχο του έργου, την τρέχουσα κατάσταση του έργου, την περιφέρεια στην οποία έχει χωροθετηθεί, τον προϋπολογισμό, τις συμβάσεις, τις πληρωμές και αν διαθέτει ή όχι κάποιο έγγραφο.

3. **projects_02.py:** Εμπλουτίζει τον πίνακα των έργων με πληροφορίες σχετικά με τον αριθμό των υποέργων που έχει το έργο, το επιχειρησιακό πρόγραμμα στο οποίο ανήκει το έργο, την ημερομηνία έναρξης και ολοκλήρωσης του έργου.
4. **projects_03.py:** Προσθέτει την περιγραφή του έργου, το ποσοστό ολοκλήρωσης, την αναφορά της εξέλιξης του έργου και την πιο πρόσφατη ημερομηνία ανανέωσης της αναφοράς. Επιπλέον σε περίπτωση που το έργο έχει υποέργα, όλες οι πληροφορίες αυτών (τίτλο, ανάδοχος κ.α.) προστίθονται στον πίνακα "Subprojects". Το ίδιο γίνεται και για τους φορείς και τα έγγραφα όπου τα δεδομένα εισάγονται στους πίνακες "Bodies" και "Files" αντίστοιχα.
5. **projects_04.py:** Προσθέτει την θεματική προτεραιότητα και το URL του έργου.
6. **supports_01.py:** Παρόμοια με το script *projects.py*, ανακτά και εισάγει στη βάση, τα δεδομένα για τις ενισχύσεις ("Supportgrants"). Αυτά αφορούν τον κωδικό MIS της ενίσχυσης, τον τίτλο, της ενισχυόμενης επιχείρησης, την τρέχουσα κατάσταση, την περιφέρεια, τον προϋπολογισμό, το ποσό των συμβάσεων, τις πληρωμές και το αν διαθέτουν έγγραφα.
7. **supports_02.py:** Προσθέτει στον πίνακα "Supportgrants" δεδομένα για το επιχειρησιακό πρόγραμμα και το URL της ενίσχυσης.
8. **supports_03.py:** Εισάγει την περιγραφή της ενίσχυσης. Ακόμη προσθέτει στους πίνακες "Bodies" και "Files" τα δεδομένα για τους φορείς και τα έγγραφα αντίστοιχα.
9. **supports_04.py:** Προσθέτει τις ημερομηνίες έναρξης και λήξης, καθώς και τη θεματική προτεραιότητα της ενίσχυσης.
10. **priority_01.py:** Αποθηκεύει στον πίνακα "Priorityprojects" τα δεδομένα για τα έργα προτεραιότητας. Αυτά περιέχουν τον κωδικό, τον τίτλο, τον αναπτυξιακό τομέα, το επιχειρησιακό πρόγραμμα, την περιφέρεια, τον προϋπολογισμό, το ποσό συγχρηματοδότησης από την Ε.Ε., τις συμβάσεις, τις πληρωμές και την τρέχουσα κατάσταση του έργου προτεραιότητας.
11. **priority_02.py:** Προσθέτει επιπλέον τη θεματική προτεραιότητα, το σύνολο των έργων που περιλαμβάνονται, το ποσοστό ολοκλήρωσης, την περιγραφή, την ημερομηνία έναρξης και λήξης, την ημερομηνία αξιολόγησης της εξέλιξης, την αναφορά της αξιολόγησης και το URL του έργου προτεραιότητας.

12. **ppincludep.py**: Γεμίζει τον πίνακα "ppIncludedProjects" με τους κωδικούς των έργων/ενισχύσεων που ανήκουν σε κάθε έργο προτεραιότητας.
13. **expropriations.py**: Εισάγει τα δεδομένα των απαλλοτριώσεων. Αυτά περιλαμβάνουν τον κωδικό και τίτλο του έργου το οποίο αφορά η απαλλοτρίωση, την εκτιμώμενη, κηρυχθείσα και αποδοθείσα επιφάνεια για απαλλοτρίωση, τον αρχικό προϋπολογισμό, τον προϋπολογισμό, τις πληρωμές το ποσοστό ολοκλήρωσης, τον αριθμό των αποφάσεων, την ημερομηνία της τελευταίας ενημέρωσης των δεδομένων, την ημερομηνία κήρυξης της πρώτης απόφασης για την απαλλοτρίωση των απαραίτητων εκτάσεων για την υλοποίηση του έργου, την ημερομηνία στην οποία ολοκληρώθηκε η απόδοση των απαλλοτριωμένων εκτάσεων στον Κύριο του έργου, καθώς και το URL της απαλλοτρίωσης.
14. **desicions.py**: Εξάγει από το API και εισάγει στον πίνακα "DecisionData" τα δεδομένα για τις αποφάσεις των απαλλοτριώσεων. Σε αυτά περιέχονται ο κωδικός του έργου και ο αριθμός της απόφασης, το είδος διαδικασίας απόκτησης της γης, ο αριθμός και η ημερομηνία του ΦΕΚ κήρυξης, ο αριθμός πρωτοκόλλου της απόφασης κήρυξης, το θεσμικό πλαίσιο, η κηρυχθείσα επιφάνεια, ο αρχικός προϋπολογισμός, ο αριθμός, κωδικός, εξέλιξη και ημερομηνία του τρέχοντος σταδίου, η αποδοθείσα επιφάνεια, ο προϋπολογισμός, οι πληρωμές και τα όλα τα προηγούμενα στάδια με ΑΑ, ημερομηνία και περιγραφή.
15. **projperdevelop.py**: Εισάγει ποια έργα περιέχονται σε κάθε αναπτυξιακό πρόγραμμα
16. **projperpref.py**: Εισάγει ποια έργα περιέχονται ανά Νομό
17. **projpermun.py**: Εισάγει ποια έργα περιέχονται ανά Δήμο
18. **grantsperdevelop.py**: Εισάγει ποιες ενισχύσεις περιέχονται σε κάθε αναπτυξιακό πρόγραμμα
19. **grantsperpref.py**: Εισάγει ποιες ενισχύσεις περιέχονται ανά Νομό
20. **grantspermun.py**: Εισάγει ποιες ενισχύσεις περιέχονται ανά Δήμο

Τονίζεται ότι το σύνολο των δεδομένων συλλέχθηκε τον Απρίλιο του 2016, ενώ μόνο τα δεδομένα των απαλλοτριώσεων και των σχετικών με αυτά αποφάσεων, ανανεώθηκαν τον Νοέμβριο του 2016.

Κεφάλαιο 5

Σημασιολογική Αναπαράσταση Δεδομένων

Το επόμενο στάδιο ήταν η σημασιολογική αναπαράσταση των δεδομένων, δηλαδή η μετατροπή τους σε τριπλέτες RDF με τη χρήση κατάλληλων λεξιλογίων. Πιο συγκεκριμένα, για την περιγραφή των συνόλων δεδομένων προϋπολογισμών και δαπανών, επαναχρησιμοποιήθηκε το σημασιολογικό μοντέλο OpenBudgets.eu, ενώ για την πληρέστερη περιγραφή της εξέλιξης της πορείας των έργων, τις μεταξύ τους σχέσεις καθώς και τη σχέση τους με άλλους φορείς χρησιμοποιήθηκε μια νέα οντολογία, που δημιουργήθηκε για την περιγραφή τέτοιου είδους δεδομένων.

Για τη δημιουργία της οντολογίας χρησιμοποιήθηκε το Protégé¹ ενώ για τον μετασχηματισμό των δεδομένων σε RDF το UnifiedViews².

Η παραπάνω διαδικασία απεικονίζεται στο σχήμα 5.1.

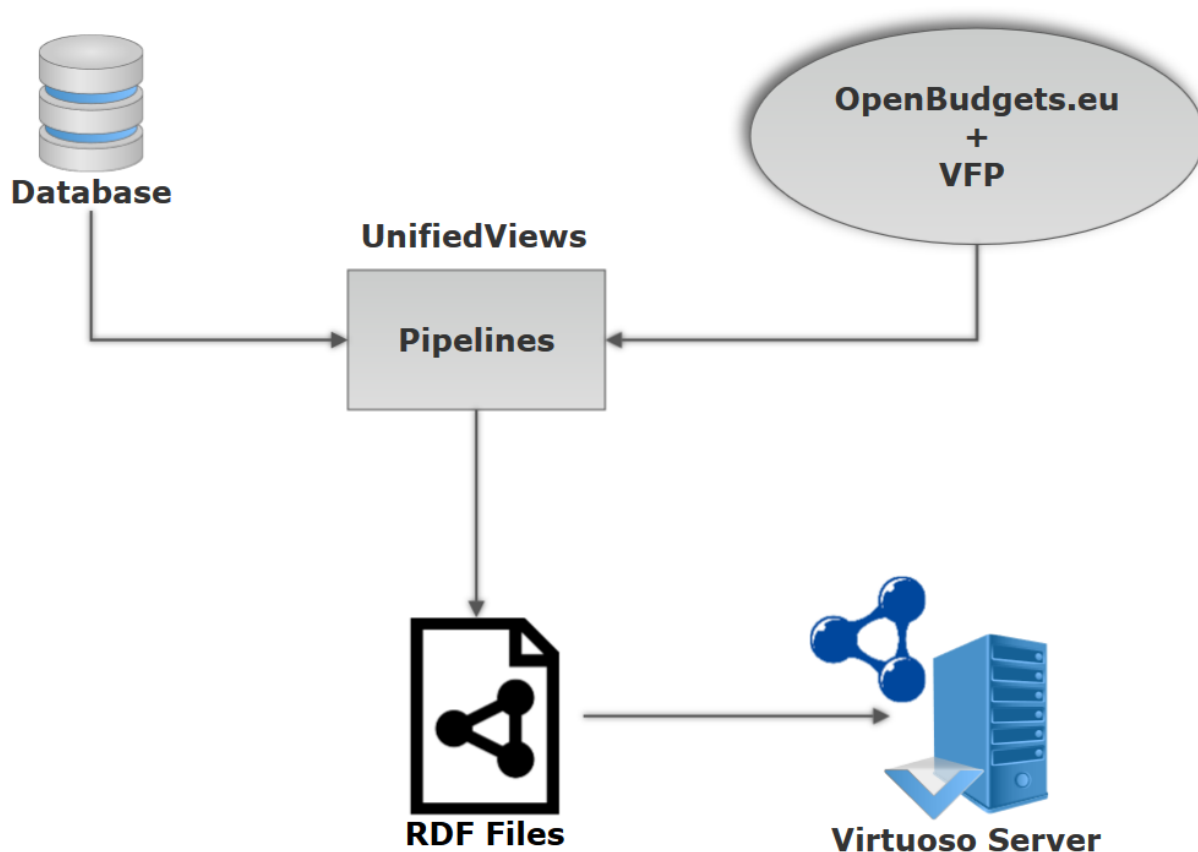
5.1 Εργαλεία που χρησιμοποιήθηκαν

5.1.1 Protégé

Το Protégé είναι ένα δωρεάν εργαλείο ανοιχτού κώδικα για την δημιουργία και επεξεργασία οντολογιών. Υποστηρίζει πλήρως την γλώσσα OWL-2 καθώς και όλες τις RDF προδιαγραφές του W3C [40]. Είναι εύκολο στη χρήση ώστε να μπορεί χρησιμοποιηθεί από τον οποιοδήποτε, ενώ διαθέτει παράλληλα και επιλογές για προχωρημένους χρήστες. Μπορεί να αποθηκεύει τις οντολογίες σε διάφορες μορφές, όπως RDF, OWL, Turtle κ.α. Το πρόγραμμα περιέχει και επέκταση για

¹<http://protege.stanford.edu/>

²<https://unifiedviews.eu/>



Σχήμα 5.1: Μεθοδολογία σημασιολογικής αναπαράστασης των δεδομένων

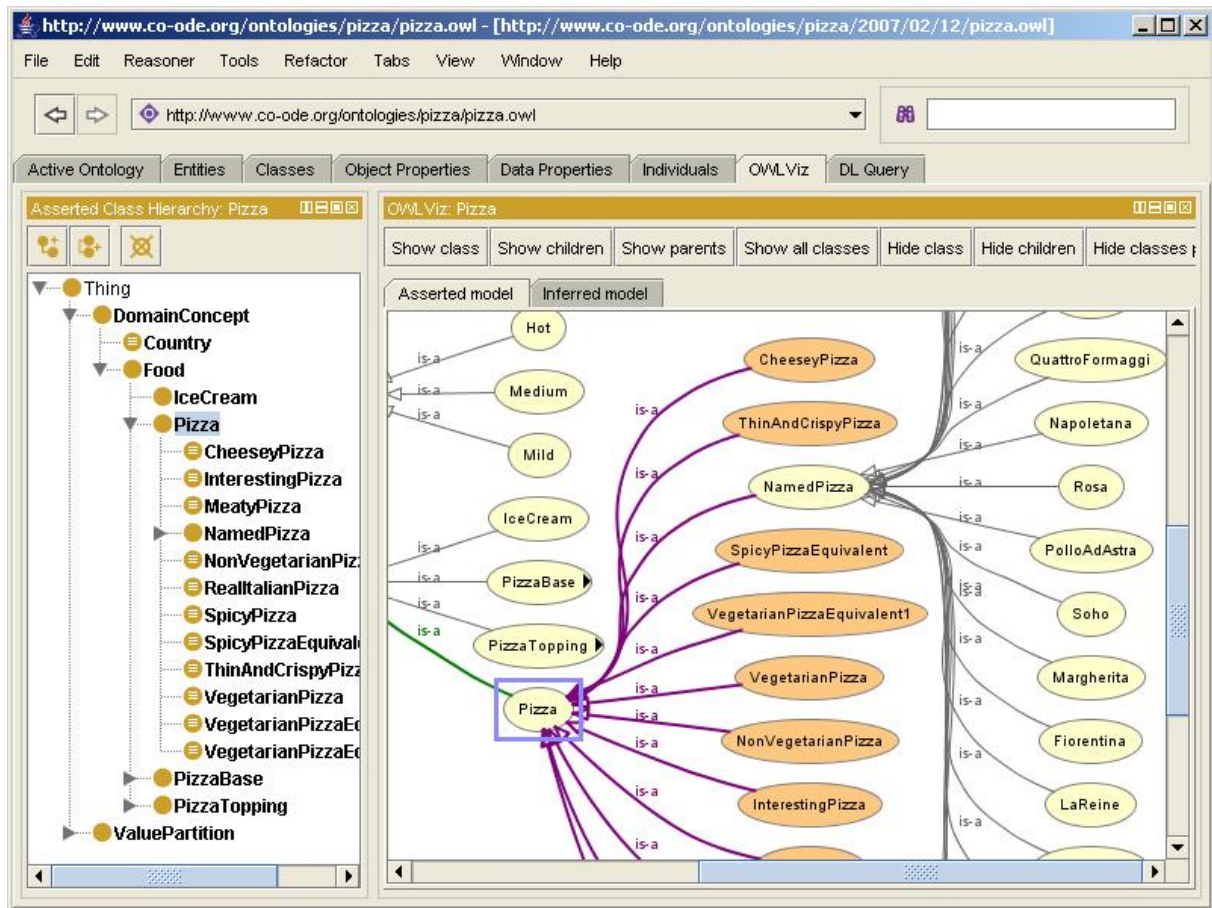
την αναπαράσταση οντολογιών με τη μορφή γράφων.

5.1.2 UnifiedViews

Το UnifiedViews είναι ένα ανοιχτού κώδικα, ETL (Extract-Transform-Load) εργαλείο για τη διαχείριση δεδομένων σε μορφή RDF. Ο χρήστης έχει τη δυνατότητα να ορίζει, εκτελεί, ελέγχει, αποσφαλματώνει, προγραμματίζει και μοιράζεται εργασίες για την επεξεργασία RDF δεδομένων (RDF data processing tasks). Οι εργασίες αυτές αποκαλούνται *pipelines* [41].

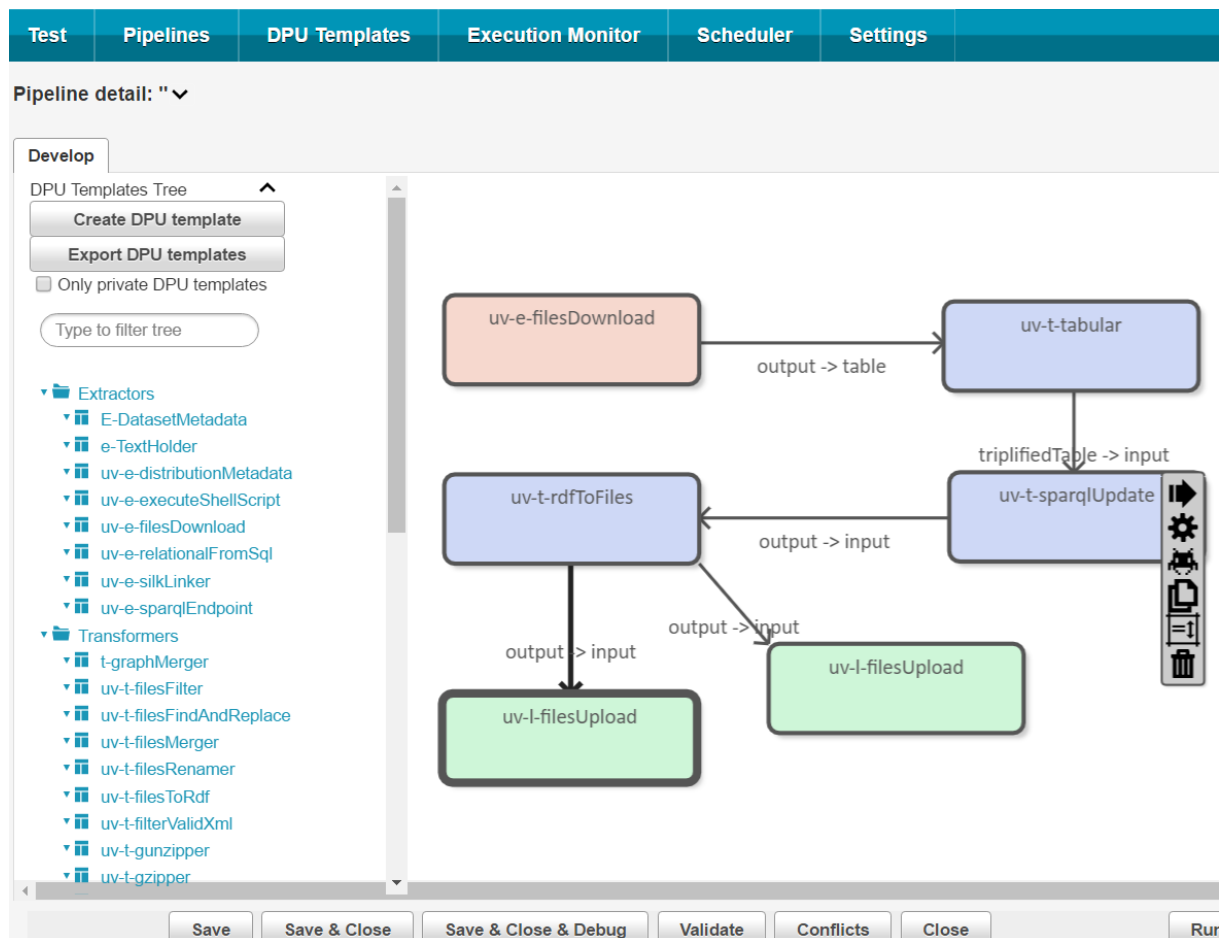
Κάθε *pipeline* αποτελείται από ένα ή περισσότερα *data processing units (DPUs)*, μεταξύ των οποίων γίνεται η διακίνηση των δεδομένων. Κάθε *DPU* μπορεί να δηλώνει ορισμένες υποχρεωτικές ή προαιρετικές εισόδους, να ενσωματώνει κάποια λογική επεξεργασίας δεδομένων και να παράγει εξόδους.

Με βάση τον αριθμό των εισόδων και εξόδων που δηλώνουν καθώς και με βάση τον επιδιωκόμενο σκοπό τους, διακρίνονται 4 τύποι DPU:



Σχήμα 5.2: Protégé Desktop

1. **Extractor:** Δεν δηλώνει είσοδο. Δέχεται τα δεδομένα από κάποια εξωτερική πηγή. Έχει κόκκινο πλαίσιο.
2. **Transformer:** Δηλώνει είσοδο και έξοδο. Μετασχηματίζει τα δεδομένα. Έχει μπλε πλαίσιο.
3. **Loader:** Δηλώνει μόνο είσοδο. Τα δεδομένα που παράγονται δεν διατηρούνται στο UnifiedViews (π.χ. εξάγονται στον σκληρό μας δίσκο). Έχει πράσινο πλαίσιο.
4. **Quality Assessor:** Αξιολογεί την ποιότητα των δεδομένων εισόδου και εξάγει μια αναφορά. Έχει κίτρινο πλαίσιο.



Σχήμα 5.3: Παράδειγμα μιας *pipeline* και των *DPUs* που την αποτελούν στο UnifiedViews

5.2 OpenBudgets.eu

Το OpenBudgets¹ είναι ένα έργο χρηματοδοτούμενο από το Horizon 2020, το μεγαλύτερο πρόγραμμα έρευνας και καινοτομίας της Ευρωπαϊκής Ένωσης. Σκοπό του έχει τη στήριξη των δημοσιογράφων, των οργανώσεων της κοινωνίας των πολιτών, τις ΜΚΟ, τους πολίτες και τις δημόσιες διοικήσεις, παρέχοντας μια επισκόπηση των δημοσίων δαπανών, καθώς επίσης εργαλεία και κατάλληλα δεδομένα και ιστορίες, υποστηρίζοντας τον αγώνα για διαφάνεια στον δημόσιο τομέα [36].

Στο πλαίσιο υποστήριξης της οικονομικής διαφάνειας που θα ενισχύει τη λογοδοσία στο δημόσιο τομέα και θα οδηγεί σε πρόληψη της διαφθοράς, ένας από τους βασικούς στόχους του OpenBudgets.eu ήταν η δημιουργία ενός σημασιολογικού μοντέλου δεδομένων [37].

¹<http://openbudgets.eu/>

5.2.1 Μοντέλο Δεδομένων OpenBudgets.eu

Το μοντέλο δεδομένων OpenBudgets.eu σχεδιάστηκε για την περιγραφή δημοσιονομικών δεδομένων που δημοσιεύονται από τον δημόσιο τομέα και καλύπτει τόσο τους προγραμματισμένους προϋπολογισμούς όσο και τις εκτελεσθέντες συναλλαγές [38].

Βασίζεται στο Data Cube Vocabulary (DCV), ένα λεξιλόγιο - πρότυπο του W3C - για την περιγραφή και δημοσίευση πολυδιάστατων στατιστικών δεδομένων [39].

5.2.1.1 Το RDF Data Cube Vocabulary

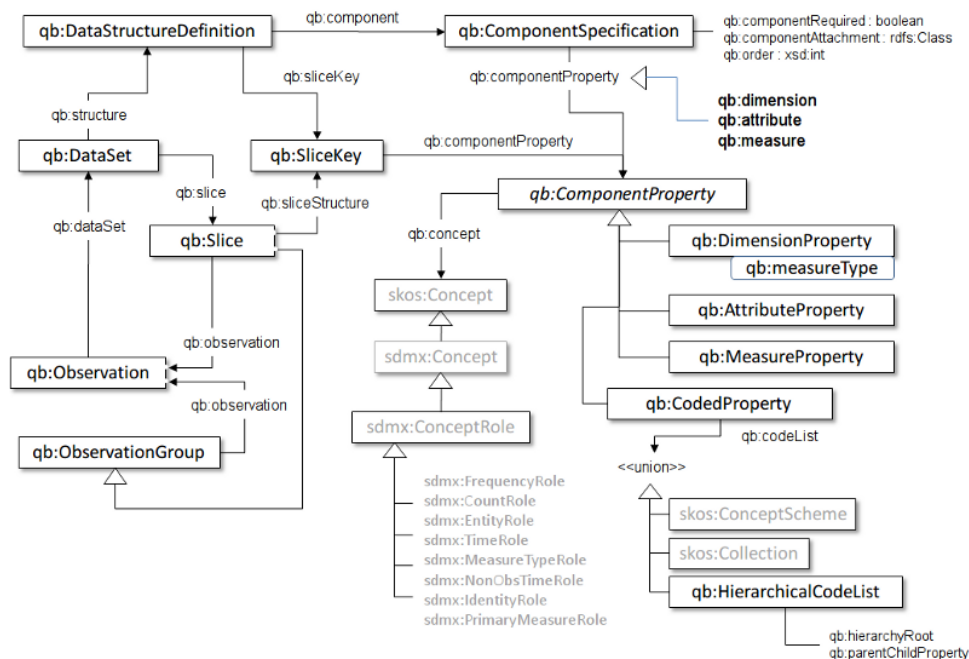
Είναι ένα λεξιλόγιο που χρησιμοποιείται ευρέως για την αναπαράσταση πολυδιάστατων στατιστικών δεδομένων συμβατά με το SDMX¹ (Statistical Data and Metadata eXchange), ένα πρότυπο ISO για την ανταλλαγή και κοινή χρήση στατιστικών δεδομένων και μεταδεδομένων μεταξύ οργανισμών [39].

Το DCV αναπαριστά τα σύνολα δεδομένων ως κύβους δεδομένων, δηλαδή συλλογές δεδομένων που αποτελούνται από παρατηρήσεις, συναφείς διαστάσεις και μεταδεδομένα. Το λεξιλόγιο παρέχει ένα σύνολο από κλάσεις και ιδιότητες για την αναπαράσταση των κύβων δεδομένων σε μορφή RDF και την δημοσίευση τους σύμφωνα με τις αρχές των διασυνδεδεμένων δεδομένων. Οι κλάσεις, οι ιδιότητες και οι σχέσεις τους απεικονίζονται στο Σχήμα 5.4.

Για κάθε σύνολο δεδομένων (qb:DataSet) πρέπει να οριστεί και κατασκευαστεί η δομή του (qb:DataStructureDefinition). Η δομή αυτή αποτελείται από τον προσδιορισμό των συστατικών της ιδιοτήτων (qb:ComponentProperty). Υπάρχουν τρεις τύποι συστατικών ιδιοτήτων:

- **Διαστάσεις - Dimensions** (qb:DimensionProperty): Καθορίζουν τις διαστάσεις που χρησιμοποιούνται στο σύνολο δεδομένων για την οργάνωση των παρατηρούμενων τιμών σε ένα πολυδιάστατο χώρο. Ένα σύνολο τιμών όλων των διαστάσεων είναι ικανό για να εντοπίσει μια μοναδική παρατήρηση. Παραδείγματα αποτελούν ο χρόνος που αφορά την παρατήρηση ή η γεωγραφική περιοχή την οποία καλύπτει αυτή.
- **Μέτρα - Measures** (qb:MeasureProperty): Καθορίζουν τους τύπους των παρατηρούμενων τιμών στο σύνολο δεδομένων.
- **Χαρακτηριστικά - Attributes** (qb:AttributeProperty): Καθορίζουν πρόσθετες ιδιότητες των παρατηρούμενων τιμών, όπως για παράδειγμα τη μο-

¹<http://sdmx.org/>



Σχήμα 5.4: Οι βασικοί όροι του DCV και οι μεταξύ τους σχέσεις

νάδα μέτρησης τους.

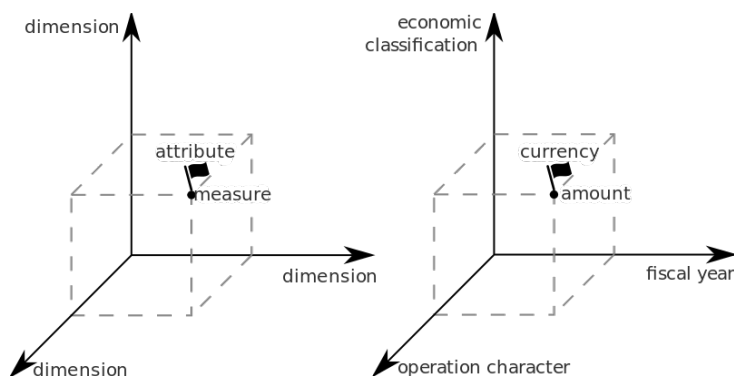
Τα σύνολα δεδομένων που περιγράφονται με το DCV αποτελούνται από παρατηρήσεις (qb:Observation). Μια παρατήρηση θα μπορούσε να θεωρηθεί ως μια καταγραφή των μέτρων (μίας ή περισσότερων παρατηρούμενων τιμών) και των αντίστοιχων τιμών καθορισμένων διαστάσεων και χαρακτηριστικών της. Επιλέγοντας συγκεκριμένες τιμές από μία ή περισσότερες διαστάσεις, μπορεί να οριστεί μία *φέτα* (qb:Slice).

5.2.1.2 Ο Κύβος στο Μοντέλο Δεδομένων OpenBudgets.eu

Τα δημοσιονομικά δεδομένα αποτελούνται συνήθως από χρηματικά ποσά, τα οποία καταχωρίζονται με βάση τις τιμές των διάφορων διαστάσεων, όπως το φορολογικό έτος ή το χρηματοδοτούμενο έργο. Τα ποσά αποτελούν τα μέτρα, που συντονίζονται από τις διαστάσεις σε κύβους δημοσιονομικών δεδομένων.

Το μοντέλο δεδομένων OpenBudgets.eu, χρησιμοποιεί την τεχνολογία RDF για την σημασιολογική θεμελίωση του οικονομικού τομέα και την δημιουργία όρων που τον περιγράφουν, με τελικό σκοπό την ενοποίηση ετερογενών δεδομένων. Η ενοποίηση επιτυγχάνεται με την επαναχρησιμοποίηση συστατικών ιδιοτήτων, η οποία επιτρέπει με τη σειρά της, τον συνδυασμό δημοσιονομικών δεδομένων και την διενέργεια συγκριτικών αναλύσεων [38, 46].

Οι κύριες συστατικές ιδιότητες του μοντέλου OpenBudgets.eu καταγράφο-



Σχήμα 5.5: Ο γενικός κύβος δεδομένων και ένα παράδειγμα κύβου στο μοντέλο OpenBudgets.eu

νται στον πίνακα 5.1. Από αυτές, ανάλογα με το αν τα δεδομένα προς περιγραφή αφορούν προϋπολογισμούς ή δαπάνες έχουμε και διαφορετικά υποχρεωτικές συστατικές ιδιότητες Υποχρεωτικές ιδιότητες για προϋπολογισμούς είναι τα `budgetaryUnit`, `fiscalPeriod`, `amount` και `currency` ενώ, για τις δαπάνες είναι το `organization`, `operationCharacter`, `date`, `amount` και `currency` (attribute) [43, 44, 45, 47].

Πίνακας 5.1: Συστατικές ιδιότητες του μοντέλου OpenBudgets.eu

Dimensions	Attributes	Optional properties	Measures
<code>accountingRecord</code>	<code>currency</code>	<code>contract</code>	<code>amount</code>
<code>administrativeClassification</code>	<code>taxesIncluded</code>	<code>location</code>	
<code>budgetLine</code>			
<code>budgetPhase</code>			
<code>classification</code>			
<code>currency</code>			
<code>date</code>			
<code>economicClassification</code>			
<code>fiscalPeriod</code>			
<code>fiscalYear</code>			
<code>functionalClassification</code>			
<code>operationCharacter</code>			
<code>organization</code>			
<code>partner</code>			
<code>paymentPhase</code>			
<code>programmeClassification</code>			
<code>project</code>			
<code>taxesIncluded</code>			

Το γεγονός ότι το μοντέλο δεν ακολουθεί ένα σταθερό σχήμα για τα δεδομένα, αλλά υιοθετεί την επαναχρησιμοποίηση συστατικών ιδιοτήτων, οι οποίες μπορούν είτε να επαναχρησιμοποιηθούν άμεσα είτε να επεκταθούν μέσω υποϊδιοτήτων, καθιστά το μοντέλο ως το πλέον κατάλληλο για την περιγραφή ετερογενών συνόλων δημοσιονομικών δεδομένων, όπως αυτά του ΕΣΠΑ [38].

5.2.1.3 Δημιουργία της δομής του Κύβου για τα Δεδομένα του ΕΣΠΑ

Για την εφαρμογή του μοντέλου OpenBudgets στα δεδομένα του ΕΣΠΑ χρειάστηκε να δημιουργηθεί η δομή του κύβου, δηλαδή να οριστεί το data structure definition (DSD). Ύστερα από μελέτη της δομής των δεδομένων, το DSD που χρησιμοποιήθηκε έχει την ακόλουθη μορφή:

Κώδικας 5.1: Data structure definition για τα δεδομένα του ΕΣΠΑ

```
<http://data.openbudgets.eu/ontology/dsd/nsrf-GR> a
qb:DataStructureDefinition ;
  rdfs:label "Data structure definition for the Greek
National Strategic Reference Framework data"@en ;
  qb:component
[ qb:dimension obeu-dimension:organization ],
  [ qb:dimension obeu-dimension:fiscalPeriod ;
    qb:componentAttachment qb:DataSet ],
  [ qb:dimension obeu-dimension:operationCharacter ;
    qb:componentAttachment qb:DataSet ],
  [ qb:componentProperty
nsrf-gr-optional:functionalClassification ],
  [ qb:dimension obeu-dimension:project ],
  [ qb:componentProperty
nsrf-gr-optional:programmeClassification ],
  [ qb:dimension obeu-dimension:paymentPhase ],
  [ qb:attribute obeu-attribute:currency ;
    qb:componentAttachment qb:DataSet ],
  [ qb:measure obeu-measure:amount ] .
```

Εδώ, το DSD αποτελείται από πέντε διαστάσεις που επαναχρησιμοποιούνται από το μοντέλο OpenBudgets (organization, fiscalPeriod, operationCharacter, paymentPhase, project), δύο νέες συστατικές ιδιότητες (functionalClassification και programmeClassification), ένα χαρακτηριστικό που επαναχρησιμοποιείται

(currency) και ένα μέτρο (amount) που επίσης επαναχρησιμοποιείται από το μοντέλο.

Για τις ιδιότητες `fiscalPeriod`, `operationCharacter` και `currency` η ιδιότητα `qb:componentAttachment` τέθηκε στο `qb:Dataset`, αφού η κάθε μια από τις τρεις ιδιότητες έχει μοναδική τιμή σε όλες τις παρατηρήσεις και δεν υπάρχει λόγος να επαναλαμβάνεται σε κάθε μία από αυτές. Οι δύο νέες συστατικές ιδιότητες ορίστηκαν ως `obeu:OptionalProperty` και δε χρησιμοποιήθηκαν οι αντίστοιχες ιδιότητες του μοντέλου, καθώς τα αντίστοιχα δεδομένα που θα περιγράψουν έχουν κενές τιμές¹.

Αφού οριστεί το DSD, μπορεί να προχωρήσει η μετατροπή των δεδομένων σε μορφή RDF με τη χρήση του `UnifiedViews`. Η διαδικασία αυτή περιγράφεται στο κεφάλαιο 5.4

5.3 Vocabulary for Describing Fiscal Projects

Για την πληρέστερη περιγραφή των δεδομένων ΕΣΠΑ κατασκευάστηκε ένα λεξιλόγιο το οποίο θα περιγράφει γενικά οικονομικά/δημοσιονομικά έργα. Το όνομα της οντολογίας στα αγγλικά είναι `Vocabulary for Describing Fiscal Projects` ή αλλιώς `VFP` για συντομία.

5.3.1 Γιατί νέα οντολογία

Ύστερα από εκτεταμένη έρευνα που πραγματοποιήθηκε στο διαδίκτυο και ειδικά στην ιστοσελίδα `Linked Open Vocabularies (LOV)`², βρέθηκαν λεξιλόγια που περιγράφουν έργα λογισμικού (`DOAP - Description of a Project`)³ ή διοικητικές πληροφορίες για ερευνητικά έργα (`FRAPO`)⁴, αλλά όχι κάποιο που να μπορεί να χρησιμοποιηθεί για την λεπτομερή περιγραφή οικονομικών έργων και της πορείας υλοποίησής τους.

Αυτός ήταν και ο κύριος λόγος δημιουργίας μιας οντολογίας που θα περιγράφει οικονομικά έργα γενικότερα και θα μπορούσε να χρησιμοποιηθεί σαν βάση για την περιγραφή των δεδομένων του ΕΣΠΑ.

¹Σύμφωνα με τα `Integrity constraints` που προτείνει το `W3C` για καλά "σχηματισμένους κύβους" τόσο τα μέτρα όσο και οι διαστάσεις θα πρέπει να έχουν τιμή για κάθε παρατήρηση. Μόνο τα χαρακτηριστικά είναι προαιρετικά [42].

²<https://lov.okfn.org/dataset/lov/>

³<http://usefulinc.com/ns/doap>

⁴<http://www.sparantologies.net/ontologies/frapo/source.html>

5.3.2 Ορισμός κλάσεων και ιδιοτήτων

Το πρώτο βήμα για την δημιουργία της οντολογίας ήταν να οριστούν οι κλάσεις της και οι ιδιότητές τους.

Για να γίνει σωστά αυτό ερευνήθηκαν σε βάθος ιστοσελίδες άλλων Ευρωπαϊκών χωρών που δίνουν - αντίστοιχες με το ANAΠΤΥΞΗ.gov.gr - πληροφορίες για τα έργα και την εξέλιξή τους.

Συγκεκριμένα, εκτός από την Ελλάδα και το ANAΠΤΥΞΗ.gov.gr , έγινε περιήγηση σε ιστοσελίδες με έργα στην Γαλλία¹, την Φινλανδία², τη Δανία³, το Λουξεμβούργο⁴, την Τσεχία⁵ και την Ιταλία⁶.

Έχοντας μελετήσει όλες τις διαθέσιμες πληροφορίες καταλήξαμε στις έννοιες που θα αποτελέσουν τις κλάσεις της οντολογίας, τις μεταξύ τους σχέσεις καθώς και τις ιδιότητες τους.

Κλάσεις:

- Έργο (vfp:Project): Ένα οικονομικό ή δημοσιονομικό έργο ή ενίσχυση. Αποτελεί την κύρια κλάση της οντολογίας, υποκλάση (rdfs:subClassOf) της foaf:Project.
- Οργανισμός (vfp:Organization): Η κλάση αυτή αναπαριστά όλους τους οργανισμούς/φορείς που αφορούν το έργο (αναδοχοι, δικαιούχοι κ.α.). Αποτελεί υποκλάση της org:Organization.
- Περιοχή (vfp:Place): Υποκλάση της dbo:Place, αναπαριστά την γεωγραφική θέση που σχετίζεται με το έργο. Η περιοχή μπορεί να αφορά περιφέρεια, νομό, δήμο ή οποιαδήποτε άλλη γεωγραφική κατηγοριοποίηση.
- Έγγραφο (vfp:Document): Η κλάση όλων των εγγράφων του έργου. Στα έγγραφα περιέχονται και οι εικόνες. Αποτελεί υποκλάση της foaf:Document.

Ιδιότητες:

- Τίτλος (vfp:title): Ο τίτλος του έργου. DatatypeProperty με εύρος τιμών του τύπου xsd:string.

¹<http://en.europe-en-france.gouv.fr/Just-realize-!/Focus-on-best-projects>

²<https://www.eura2007.fi/rrtiepa/projektilista.php?rahasto=ALL>

³<http://eustrukturfonde.dk/>

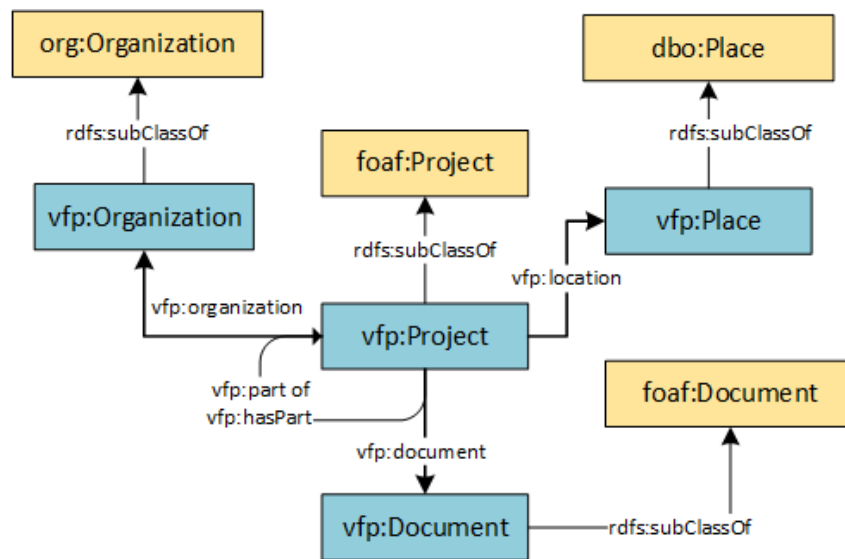
⁴<http://www.fonds-europeens.public.lu/fr/projets-cofinances/index.php>

⁵<http://www.esfcr.cz/>

⁶<http://www.opencoesione.gov.it/>

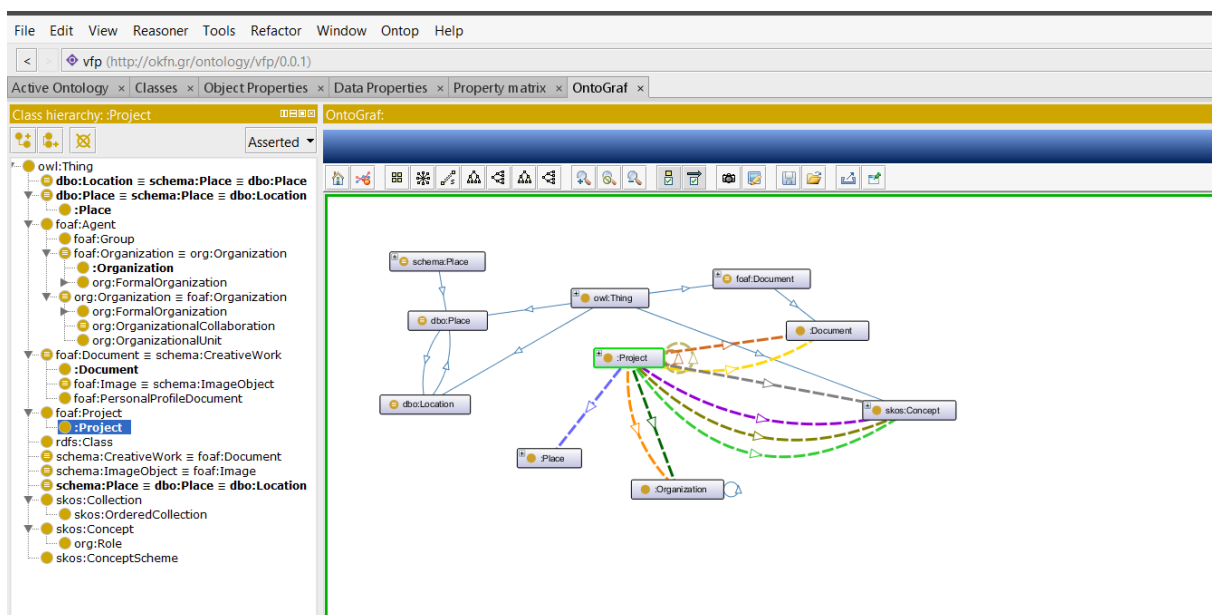
- Περιγραφή (vfp:description): Περιγραφή του έργου. DatatypeProperty με εύρος τιμών του τύπου xsd:string.
- Κατάσταση (vfp:currentStatus): Κατάσταση της πορείας υλοποίησης του έργου. ObjectProperty με εύρος τιμών του τύπου skos:Concept.
- Προϋπολογισμός (vfp:budget): Προϋπολογισμός του έργου. Ορίζεται ως owl:equivalentProperty της ιδιότητας obeu-measure:amount καθώς λαμβάνει ίδιες τιμές. Έχει εύρος xsd:decimal.
- Συμβάσεις (vfp:contracts): Συμβάσεις του έργου. Ορίζεται ως owl:equivalentProperty της ιδιότητας obeu-measure:amount καθώς λαμβάνει ίδιες τιμές. Έχει εύρος xsd:decimal.
- Πληρωμές (vfp:payments): Πληρωμές του έργου. Ορίζεται ως owl:equivalentProperty της ιδιότητας obeu-measure:amount καθώς λαμβάνει ίδιες τιμές. Έχει εύρος xsd:decimal.
- Ολοκλήρωση (vfp:completion): Ποσοστό ολοκλήρωσης των πληρωμών όσον αφορά τον προϋπολογισμό. DatatypeProperty με εύρος τιμών του τύπου xsd:float.
- Ημερομηνία έναρξης (vfp:startDate): Ημερομηνία έναρξης του έργου. Έχει εύρος τιμών xsd:date.
- Ημερομηνία ολοκλήρωσης (vfp:endDate): Ημερομηνία ολοκλήρωσης του έργου. Έχει εύρος τιμών xsd:date.
- Αναφορά (vfp:statusReport): Αναφορά εξέλιξης της πορείας του έργου. Έχει εύρος τιμών xsd:string.
- Ημερομηνία αναφοράς (vfp:statusReportDate): Ημερομηνία τελευταίας ανανέωσης της αναφοράς εξέλιξης του έργου. Έχει εύρος τιμών xsd:date.
- URL (vfp:url): URL του έργου. Είναι rdfs:subPropertyOf της foaf:page και έχει εύρος τιμών vfp:Document.

Στο σχήμα 5.6 περιγράφονται οι σχέσεις μεταξύ των κλάσεων. Ένα έργο αφορά έναν οργανισμό/φορέα και αυτό περιγράφεται με την ιδιότητα vfp:organization. Με την ιδιότητα vfp:location δηλώνεται η περιοχή στην οποία έχει χωροθετηθεί το έργο. Επίσης τα έγγραφα που αφορούν το έργο συνδέονται με την ιδιότητα vfp:document. Τέλος, ορίζονται και οι ιδιότητες vfp:partOf και vfp:hasPart για την περίπτωση που ένα έργο αποτελεί μέρος ενός άλλου.



Σχήμα 5.6: Οι κλάσεις του VFP και οι μεταξύ τους σχέσεις

Έχοντας ολοκληρωθεί η διαδικασία εντοπισμού των κλάσεων, των μεταξύ τους σχέσεων και των ιδιοτήτων τους, ακολουθεί η κατασκευή της οντολογίας. Το εργαλείο που χρησιμοποιήθηκε για το σκοπό αυτό, όπως αναφέρθηκε, ήταν το Protégé. Ως χώρος ονομάτων χρησιμοποιήθηκε το <http://okfn.gr/ontology/vfp#> με προτεινόμενο πρόθεμα (prefix) το `vfp:`. Όλα τα IRIs των όρων της οντολογίας θα ξεκινάνε με αυτό ως βάση. Αφού εισήχθησαν οι απαραίτητες οντολογίες (ORG, FOAF, DBO), ορίστηκαν οι κλάσεις και ιδιότητες τους (Datatype και Object) και στο τέλος αποθηκεύτηκε σε ένα αρχείο κατάληξης `.owl`.



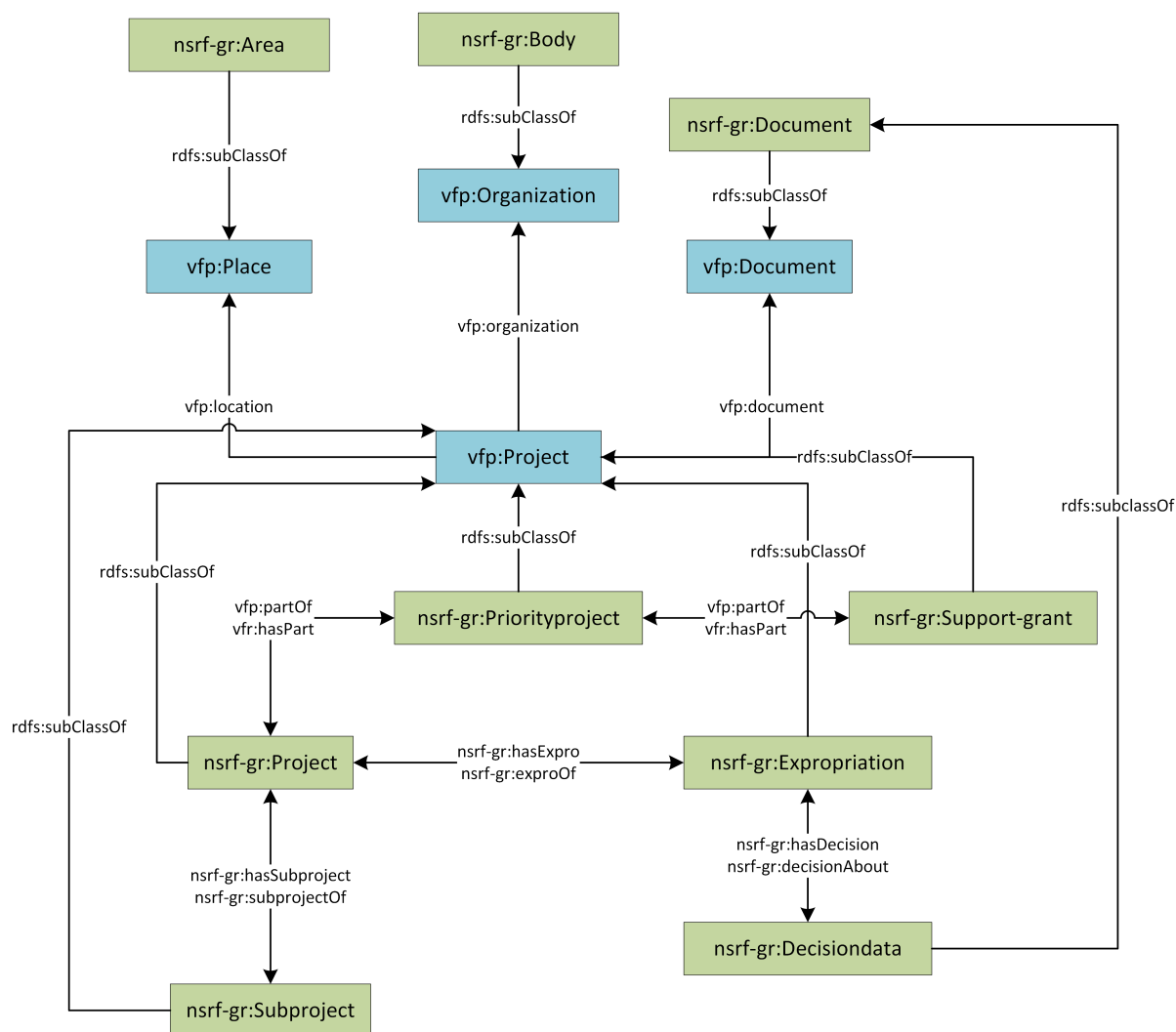
Σχήμα 5.7: Ορισμός των κλάσεων στο Protégé και αναπαράσταση τους με το plug-in OntoGraf

5.3.3 Επέκταση της Οντολογίας για τα Δεδομένα ΕΣΠΑ

Η οντολογία VFP που έχει δημιουργηθεί και προτείνεται για την περιγραφή οικονομικών έργων, θα αποτελέσει την βάση για τη σημασιολογική περιγραφή των διαθέσιμων δεδομένων του ΕΣΠΑ.

Ουσιαστικά, ορίζεται ένα νέο λεξιλόγιο όρων (κλάσεων-ιδιοτήτων) με βάση την οντολογία VFP την οποία επεκτείνει, ακολουθώντας παρόμοια διαδικασία με την προηγούμενη.

Τα δεδομένα του ΕΣΠΑ περιλαμβάνουν έννοιες όπως έργο, έργο προτεραιότητας, υποέργο, ενισχύσεις, απαλλοτριώσεις, έγγραφα, φορείς κ.α. Για κάθε μια από αυτές, ορίζονται κλάσεις που αποτελούν υποκλάσεις αυτών που έχουν οριστεί στην οντολογία VFP. Στο σχήμα 5.8 φαίνονται οι βασικές σχέσεις και ιδιότητες των κλάσεων. Περιγραφή των ιδιοτήτων που δημιουργήθηκαν υπάρχει στους πίνακες που ακολουθούν.



Σχήμα 5.8: Κλάσεις των δεδομένων του ΕΣΠΑ και οι μεταξύ τους σχέσεις

Πίνακας 5.2: Οι ιδιότητες που συνδέουν τις κλάσεις μεταξύ τους

Property name	Label
nsrf-gr:body	Φορέας του έργου ΕΣΠΑ
nsrf-gr:hasExpro	Έχει απαλλοτρίωση
nsrf-gr:subprojectOf	Υποέργο του
nsrf-gr:hasSubproject	Έχει υποέργο
nsrf-gr:hasDecision	Η απαλλοτρίωση έχει απόφαση
nsrf-gr:decisionAbout	Απόφαση σχετικά με την απαλλοτρίωση
nsrf-gr:location	Τοποθεσία
nsrf-gr:exproOf	Απαλλοτρίωση του

Πίνακας 5.3: Οι ιδιότητες που ορίστηκαν για τους φορείς και τα έγγραφα

Property name	Label
nsrf-gr:bodyRepresentative	Αντιπρόσωπος φορέα
nsrf-gr:bodyFax	Φαξ φορέα
nsrf-gr:bodyType	Είδος φορέα
nsrf-gr:fileType	Είδος εγγράφου

Πίνακας 5.4: Οι ιδιότητες που ορίστηκαν για τις απαλλοτριώσεις

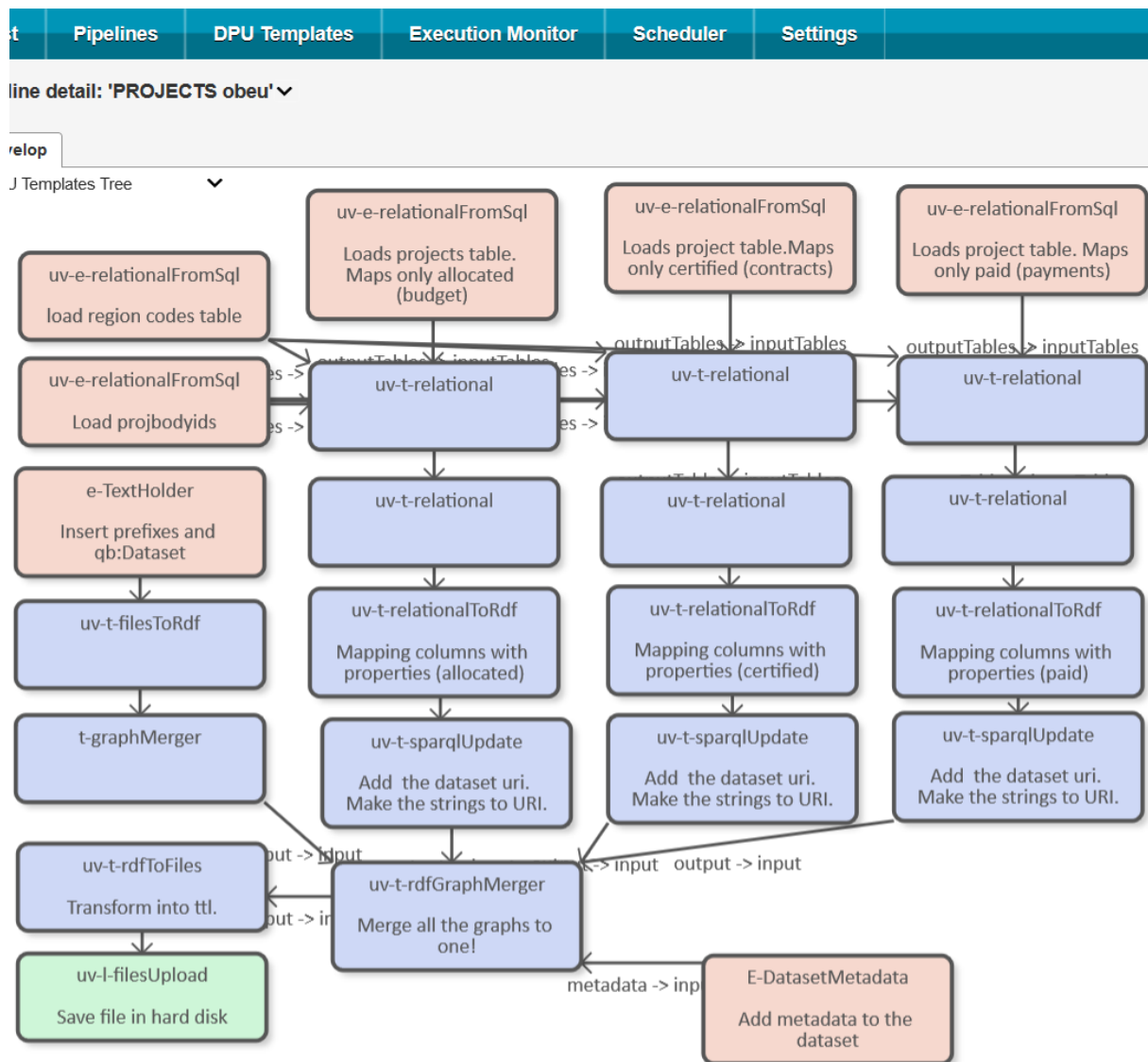
Property name	Label
nsrf-gr:exproAssignedArea	Αποδοθείσα επιφάνεια
nsrf-gr:exproBudget	Προϋπολογισμός
nsrf-gr:exproCompleteDate	Ημερομηνία ολοκλήρωσης απαλλοτρίωσης
nsrf-gr:exproCompletion	Ποσοστό ολοκλήρωσης απαλλοτρίωσης
nsrf-gr:exproCountDecision	Αριθμός αποφάσεων
nsrf-gr:exproDeclaredArea	Κηρυχθείσα επιφάνεια
nsrf-gr:exproEstimatedArea	Εκτιμώμενη επιφάνεια προς απαλλοτρίωση
nsrf-gr:exproFirstDate	1η Απόφαση Κήρυξης
nsrf-gr:exproInitialBudget	Αρχικός προϋπολογισμός
nsrf-gr:exproLastUpdate	Τελευταία ενημέρωση
nsrf-gr:unitMeasure	Μονάδα μέτρησης
nsrf-gr:exproURL	URL απαλλοτρίωσης
nsrf-gr:exproPayments	Πληρωμές
nsrf-gr:exproTitle	Τίτλος Απαλλοτρίωσης

Πίνακας 5.5: Οι ιδιότητες που ορίστηκαν για τις αποφάσεις των απαλλοτριώσεων

Property name	Label
nsrf-gr:exproOrigAA	ΑΑ Απόφασης
nsrf-gr:exproaaStadiou	ΑΑ σταδίου
nsrf-gr:exproapodEpifaneia	Αποδοθείσα επιφάνεια
nsrf-gr:exproarProtok	Αριθμός πρωτοκόλλου
nsrf-gr:exproeidosApof	Είδος απόφασης
nsrf-gr:exproepikProyp	Προϋπολογισμός
nsrf-gr:exprofekDate	Ημερομηνία έκδοσης ΦΕΚ
nsrf-gr:exprofekTitle	Τίτλος ΦΕΚ
nsrf-gr:exprofekcomb	Αριθμός & ημ/νία ΦΕΚ κήρυξης
nsrf-gr:exprohmerStadiou	Ημερομηνία σταδίου
nsrf-gr:exprokodStadiou	Κωδικός σταδίου
nsrf-gr:exprokyrEpif	Κηρυχθείσα επιφάνεια
nsrf-gr:exproperStadiou	Τρέχον στάδιο εξέλιξης
nsrf-gr:exproplaisio	Θεσμικό πλαίσιο
nsrf-gr:exproposo	Πληρωμές
nsrf-gr:exproprevStages	Προηγούμενα στάδια
nsrf-gr:exprosypProyp	Αρχικός προϋπολογισμός

Πίνακας 5.6: Οι ιδιότητες που ορίστηκαν για τα έργα, υποέργα, ενισχύσεις και έργα προτεραιότητας

Property name	Label
nsrf-gr:includedProjects	Αριθμός έργων που περιέχονται στο έργο προτεραιότητας
nsrf-gr:euCofinanced	Προϋπολογισμός συγχρηματοδοτούμενος από ΕΕ
nsrf-gr:subprojID	Κωδικός Υποέργου
nsrf-gr:implementor	Ανάδοχος του υποέργου
nsrf-gr:actualSubmissionDate	Πραγματική ημερομηνία υποβολής της πρότασης
nsrf-gr:actualEntranceDate	Πραγματική ημερομηνία έγκρισης για χρηματοδότηση
nsrf-gr:actualAuctionApproveDate	Πραγματική ημερομηνία προέγκρισης δημοπράτησης
nsrf-gr:actualAuctionDate	Πραγματική ημερομηνία δημοπράτησης
nsrf-gr:actualContractApproveDate	Πραγματική ημερομηνία προέγκρισης υπογραφής της σύμβασης
nsrf-gr:actualContractDate	Πραγματική ημερομηνία υπογραφής της Σύμβασης
nsrf-gr:actualCompletionDate	Πραγματική ημερομηνία ολοκλήρωσης
nsrf-gr:estimSubmissionDate	Εκτιμώμενη ημερομηνία υποβολής της πρότασης
nsrf-gr:estimEntranceDate	Εκτιμώμενη ημερομηνία έγκρισης για χρηματοδότηση
nsrf-gr:estimAuctionApproveDate	Εκτιμώμενη ημερομηνία προέγκρισης δημοπράτησης
nsrf-gr:estimAuctionDate	Εκτιμώμενη ημερομηνία δημοπράτησης
nsrf-gr:estimContractApproveDate	Εκτιμώμενη ημερομηνία προέγκρισης υπογραφής της σύμβασης
nsrf-gr:estimContractDate	Εκτιμώμενη ημερομηνία υπογραφής της Σύμβασης
nsrf-gr:estimCompletionDate	Εκτιμώμενη ημερομηνία ολοκλήρωσης
nsrf-gr:thematic	Θεματική προτεραιότητα
nsrf-gr:operational	Επιχειρησιακό πρόγραμμα
nsrf-gr:countSubproject	Αριθμός υποέργων



Σχήμα 5.10: Δομή του pipeline για την μετατροπή των δεδομένων των έργων με βάση το μοντέλο OpenBudgets.eu

κατασκευάστηκαν συνολικά 24 pipelines. Τα pipelines απαρτίζονται από DPUs όπως φαίνεται και στο σχήμα 5.10. Η λογική που ακολουθείται είναι ίδια σε όλα τα pipeline.

Αρχικά γίνεται χρήση του extractor `uv-e-relationalFromSql`, ο οποίος παίρνει και εισάγει (σαν πίνακα) στο περιβάλλον του UnifiedViews, δεδομένα από έναν πίνακα της βάσης. Ανάλογα με το pipeline μπορεί να υπάρχουν περισσότεροι από ένας τέτοιοι extractors.

Στη συνέχεια, ο transformer `uv-t-relational` χρησιμοποιείται για να εκτελέσει sql εντολές με τους πίνακες που δημιουργούνται από τον extractor. Ακολούθως, ο transformer `uv-t-relationalToRdf` χρησιμοποιείται για την μετατροπή

uv-t-relationalToRdf detail

Name: uv-t-relationalToRdf

Parent: uv-t-relationalToRdf

Description: Mapping columns with properties (allocated)

☒ Use custom description ☐ Use template configuration

DPU configuration | Fault tolerance | About

Resource URI base

http://data.openbudgets.eu/resource/dataset/nsrf-GR/projects/observation/

Key column

obs

Class for a row entity

http://purl.org/linked-data/cube#Observation

☐ Full column mapping ☒ Ignore blank cells ☐ Advanced key column

☐ Generate row column ☐ Generate subject for table ☐ Auto type as string

☐ Generate table/row class ☐ Generate labels

Simple mapping

Column name	Output type	Language	Use Dbf types	Property URI
projmisuri	String		☐	http://data.openbudgets.eu/ontology/dsd/dimension/project
projttitle	String	el	☐	http://www.w3.org/2000/01/rdf-schema#label
projbodyuri	String		☐	http://data.openbudgets.eu/ontology/dsd/dimension/organization

Save Cancel Copy from template Save as DPU tem

Σχήμα 5.11: Λεπτομέρειες για την διαμόρφωση του transformer uv-t-relationalToRdf για τα έργα

των δεδομένων σε RDF μορφή. Στον transformer, όπως φαίνεται και στο σχήμα 5.11, ορίζεται το URI, το οποίο σε συνδυασμό με τις τιμές της κύριας στήλης του πίνακα που δέχεται ως είσοδο, θα αποτελέσουν τη βάση για τον σχηματισμό της ονομασίας του πόρου.

Ακολούθως, ορίζεται η κλάση του πόρου και οι υπόλοιπες στήλες αντιστοιχίζονται στις ιδιότητες που χρησιμοποιούνται. Μετά χρησιμοποιείται ο transformer uv-t-sparqlUpdate ο οποίος εκτελεί sparql εντολές και μετασχηματίζει τα δεδομένα τα οποία δέχεται ως είσοδο. Στην πλειονότητα των περιπτώσεων, χρησιμοποιήθηκε για τον μετασχηματισμό τιμών του τύπου xsd:string σε URI. Επειδή οι πίνακες που εισάγονται είναι πολλοί και προκύπτουν αντίστοιχοι RDF γράφοι, χρησιμοποιείται ο uv-t-rdfGraphMerger transformer για να τους ενοποιήσει σε έναν γράφο που να περιέχει όλες τις RDF τριπλέτες. Σε αυτόν εισάγονται και μεταδεδομένα με τον extractor E-DatasetMetadata καθώς και τα prefixes που χρησιμοποιήθηκαν με τον extractor e-TextHolder.

Στο τέλος, ο `uv-t-rdfToFiles transformer` μετατρέπει τον ενιαίο γράφο σε ένα RDF αρχείο (υποστηρίζει πολλά μορφότυπα RDF όπως Turtle, RDF/XML, N-Triples, N3, RDFa κ.α.) και με τον loader `uv-l-filesUpload` το αρχείο εξάγεται στον τοπικό δίσκο.

Μετά την ολοκλήρωση της σημασιολογικής αναπαράστασης των δεδομένων, οι τριπλέτες μεταφορτώθηκαν σε έναν Openlink Virtuoso server.

Κεφάλαιο 6

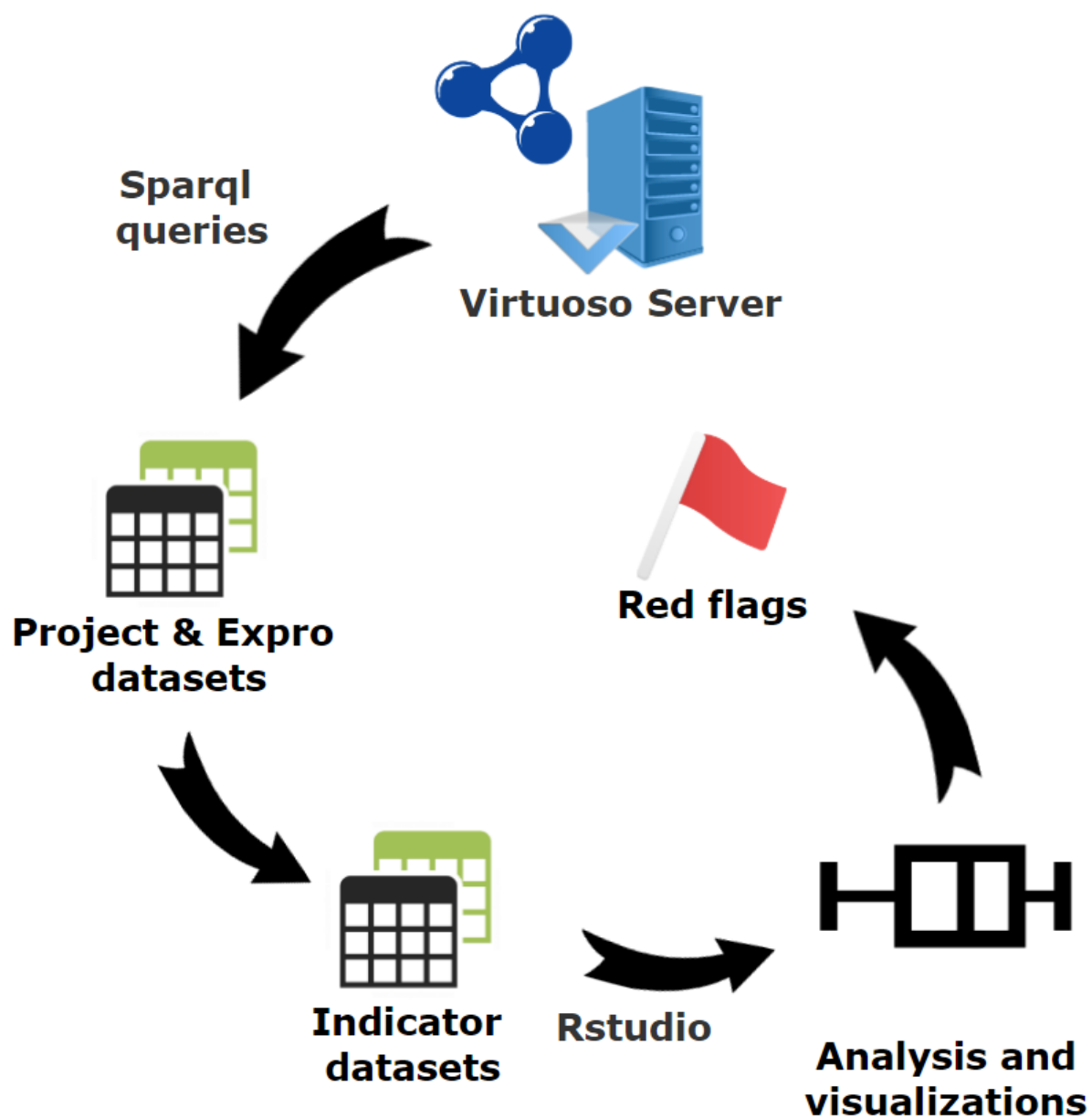
Κόκκινες σημαίες

Η αντιμετώπιση της απάτης, γενικά, αλλά και ειδικά στο πλαίσιο των διαρθρωτικών δράσεων, αποτελεί σημαντικό ζήτημα για την ΕΕ. Ήδη, η Ευρωπαϊκή Υπηρεσία Καταπολέμησης της Απάτης (OLAF)¹, διερευνά περιπτώσεις απάτης και διαφθοράς, καθώς και άλλες παράνομες δραστηριότητες που σκοπό έχουν να πλήξουν τους Ευρωπαίους φορολογούμενους. Επίσης, η ΕΕ χρηματοδοτεί προγράμματα [48] και εκπονεί μελέτες [49] για τον εντοπισμό και μείωση της διαφθοράς στον τομέα των δημοσίων συμβάσεων. Χαρακτηριστικό είναι ότι στην Ευρώπη αυτή τη στιγμή βρίσκεται σε εξέλιξη μία μεγάλη αλλαγή στις ανοιχτές συμβάσεις, με την αναμόρφωση όλων των έντυπων των συμβάσεων που σκοπό έχει, μεταξύ άλλων, την καταπολέμηση της απάτης και της διαφθοράς στα Κράτη Μέλη της ΕΕ [50][51].

Δυστυχώς, η έλλειψη τέτοιων πληροφοριών στα δεδομένα του ΕΣΠΑ που συλλέχθηκαν, δεν επέτρεψε την εφαρμογή των ήδη υπάρχουσών πρακτικών [52] για τον εντοπισμό πιθανής απάτης, οδήγησε όμως στη δημιουργία χαρακτηριστικών δεικτών για ένα υποσύνολο των δεδομένων μας, στα οποία έγινε εντοπισμός των κόκκινων σημαιών.

Η διαδικασία που πραγματοποιήθηκε παριστάνεται γραφικά στο σχήμα 6.1.

¹http://ec.europa.eu/anti-fraud/home_el



Σχήμα 6.1: Μεθοδολογία για την εύρεση των κόκκινων σημαιών

6.1 Ορισμοί

Στο πλαίσιο των διαρθρωτικών δράσεων ορίζονται ως:

- **Παρατυπία:** Μία πράξη ασύμβατη προς τους κοινοτικούς κανόνες, η οποία ενδέχεται να θίγει τα δημοσιονομικά συμφέροντα της ΕΕ. Μπορεί να οφείλεται σε πραγματικά σφάλματα στα οποία υποπίπτουν τόσο δικαιούχοι όσο και αρμόδιοι φορείς [53].
- **Απάτη:** Μια πράξη εσκεμμένου δόλου/παραπλάνησης, με στόχο την απόκομιση προσωπικού οφέλους ή την πρόκληση ζημίας σε άλλον. Μια ηθελημένη παρατυπία συνιστά απάτη [53].
- **Κόκκινες σημαίες (Red flags):** Είναι ένα σύνολο περιπτώσεων που παρουσιάζουν ασυνήθιστη συμπεριφορά. Αποτελούν ενδείξεις για ύπαρξη παράνομης δραστηριότητας, οπότε και απαιτείται ανάληψη περαιτέρω δράσης ώστε να αποκλειστεί η πιθανότητα απάτης. Τονίζεται ότι οι κόκκινες σημαίες δεν υποδεικνύουν ενοχή ή αθωότητα, απλά παρέχουν πιθανά προειδοποιητικά σημάδια απάτης. [54].

6.2 Διαδικασία

6.2.1 Προσδιορισμός συνόλων δεδομένων

Η μεθοδολογία που ακολουθείται για την ανάλυση των δεδομένων ΕΣΠΑ εφαρμόζεται ενδεικτικά για δύο σύνολα δεδομένων, αυτά των Έργων και των Απαλλοτριώσεων. Δημιουργούνται SPARQL ερωτήματα τα οποία στέλνονται στον εξυπηρετητή για να ανακτηθούν από αυτόν. Τα ερωτήματα εκτελούνται μέσα από το RStudio και τα δεδομένα που ανακτώνται, αποθηκεύονται σε μορφή πινάκων. Ακολούθως, αφού μελετηθούν, ορίζονται οι δείκτες.

6.2.2 Ορισμός δεικτών

6.2.2.1 Έργα

Το σύνολο δεδομένων των Έργων αποτελείται από 12509 έργα. Επιλέχθηκαν να μελετηθούν τα συμβασιοποιημένα και ολοκληρωμένα - 11798 στο σύνολο - καθώς τα υπόλοιπα έργα (υπό αξιολόγηση και εγκεκριμένα) είχαν μηδενικές τιμές για στα πεδία του προϋπολογισμού, των συμβάσεων και των πληρωμών.

Παρακάτω ορίζονται οι δείκτες για τα έργα:

1. **Overdue:** Αν το έργο είναι εκπρόθεσμο (παίρνει τιμή 1) ή όχι (παίρνει τιμή 0). Σαν ημερομηνία τίθεται η 31 Δεκεμβρίου 2015¹². Όσα έργα ξεπερνάνε αυτή την ημερομηνία δηλώνονται αυτόματα ως κόκκινες σημαίες.
2. **Completion:** Ποσοστό ολοκλήρωσης χρηματοδότησης του έργου. Ορίζεται ως: $\frac{projPayments}{projBudget}$.
3. **Payment_completion:** Η ολοκλήρωση των πληρωμών ως προς το ποσό των συμβάσεων. Ορίζεται ως: $\frac{projPayments}{projContracts}$.
4. **Contract_completion:** Η μεταβολή του ποσού των συμβάσεων σε σχέση με τον προϋπολογισμό. Ορίζεται ως: $\frac{projContract}{projBudget}$.

6.2.2.2 Απαλλοτριώσεις

Το σύνολο δεδομένων των Απαλλοτριώσεων αποτελείται από 247 απαλλοτριώσεις. Οι δείκτες που ορίστηκαν είναι:

1. **Estimated_area:** Η εκτιμώμενη επιφάνεια για απαλλοτρίωση. Μετριέται σε τετραγωνικά μέτρα.
2. **Completion:** Ποσοστό ολοκλήρωσης της απαλλοτρίωσης. Ορίζεται ως: $\frac{exproAssignedArea}{exproDeclaredArea}$.
3. **Payments:** Η κατάσταση των πληρωμών ως προς τον προϋπολογισμό. Ορίζεται ως: $\frac{exproPayments}{exproBudget}$.
4. **Initial_budget_changes:** Δείχνει την αλλαγή του αρχικού με τον τελικό προϋπολογισμό. Ορίζεται ως: $\frac{exproBudget}{exproInitialBudget}$.
5. **Declared_area_cost:** Το κόστος του 1 m² κηρυχθείσας επιφάνειας. Ορίζεται ως: $\frac{exproBudget}{exproDeclaredArea}$.
6. **Unassigned_declared_area:** Η επιφάνεια που δεν έχει αποδοθεί ακόμα. Ορίζεται ως: $\frac{exproDeclaredArea - exproAssignedArea}{exproDeclaredArea}$.
7. **Decisions:** Ο αριθμός των αποφάσεων.

¹<https://www.epixeirein.com.gr/paratasi-espa/>

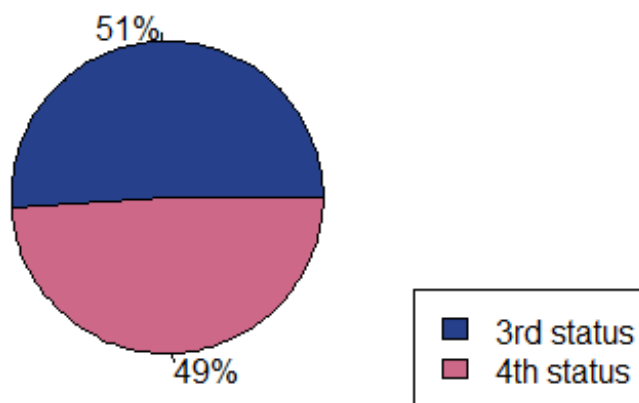
²<http://www.dianeosis.org/2016/02/espa-main/>

6.2.3 Ανάλυση - Αποτελέσματα - Σχολιασμός

Αφού ορίστηκαν οι δείκτες, χρησιμοποιήθηκαν για να δημιουργηθούν τα τελικά σύνολα δεδομένων που θα αναλυθούν. Τα νέα σύνολα δημιουργήθηκαν μέσω εντολών στο Rstudio. Σε όσες περιπτώσεις χρειάστηκε δημιουργήθηκαν θηκογράμματα για την εύρεση των κόκκινων σημαιών.

6.2.3.1 Έργα

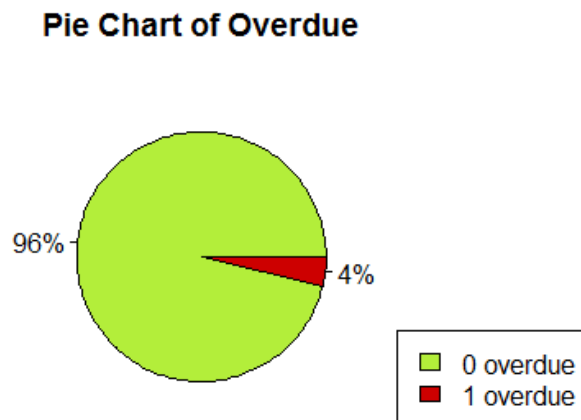
Pie Chart of Current Status



Σχήμα 6.2: Διάγραμμα πίτας με τον διαχωρισμό της κατάστασης των έργων.

Στο σχήμα 6.2 φαίνεται το ποσοστό των συμβασιοποιημένων και των ολοκληρωμένων έργων. Το ποσοστό και των δύο έργων είναι σχεδόν μοιρασμένο. Υπάρχουν 6031 έργα που είναι συμβασιοποιημένα και δεν έχουν ολοκληρωθεί και 5767 ολοκληρωμένα.

Το σχήμα 6.3 δείχνει το ποσοστό των έργων τα οποία δεν έχουν ολοκληρωθεί μέχρι το πέρας του έτους 2015. Αυτά αποτελούν το 4% του συνόλου δεδομένων που μελετήθηκε, το οποίο μεταφράζεται σε 443 έργα. Αυτά χαρακτηρίζονται ως κόκκινη σημαία και αξίζει να διερευνηθούν οι λόγοι που συμβαίνει κάτι τέτοιο. Ένας γρήγορος έλεγχος αυτών έδειξε ότι πολλά από αυτά τα έργα είχανε κατάσταση



Σχήμα 6.3: Διάγραμμα πίτας με το ποσοστό των εκπρόθεσμων έργων (1 overdue).

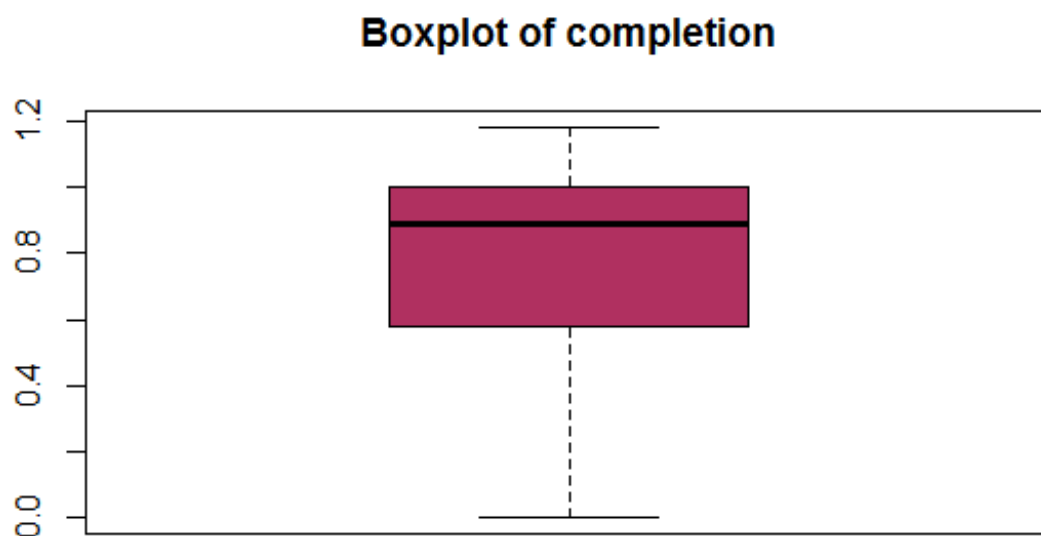
4, δηλαδή ήταν σύμφωνα με τα δεδομένα ολοκληρωμένα! Αυτό λογικά προέκυψε από παρατυπία και λάθος καταγραφή της ημερομηνίας ολοκλήρωσης από τις Διαχειριστικές Αρχές. Πράγματι, μετά από έλεγχο σε μερικά έργα με κατάσταση 4 και ημερομηνία ολοκλήρωσης μεταγενέστερη της 2015-12-31, οι ημερομηνίες τους έχουν διορθωθεί.

Στο σχήμα 6.4 παρουσιάζεται το Θηκόγραμμα για τον δείκτη *Completion*. Δεν παρουσιάζονται ακραίες τιμές, οπότε δεν προκύπτει κάποια κόκκινη σημαία. Παρ' όλα αυτά επειδή τα έργα για τα οποία ο δείκτης είναι μεγαλύτερος από τη μονάδα κρίνονται προβληματικά και τα εκλαμβάνουμε ως κόκκινη σημαία. Αυτά τα 15 έργα καταγράφονται στον πίνακα 6.2.3.1.

Στο σχήμα 6.5 αντίθετα, που αφορά τον δείκτη *Payment_completion*, εντοπίζονται 274 ακραίες τιμές.

Στο σχήμα 6.6 για τον δείκτη *Contract_completion* εντοπίζονται πολλές τιμές εκτός ορίων (1014 στο σύνολο). Ερευνήθηκε η πιο ακραία τιμή και διαπιστώθηκε ότι ήταν ακόμη μια περίπτωση λάθους καταγραφής των στοιχείων του έργου (40000 ευρώ συμβάσεις, έναντι 2362 ευρώ προϋπολογισμού) τα οποία ενημερώθηκαν σε μεταγενέστερη ημερομηνία από αυτήν της εξαγωγής των δεδομένων.

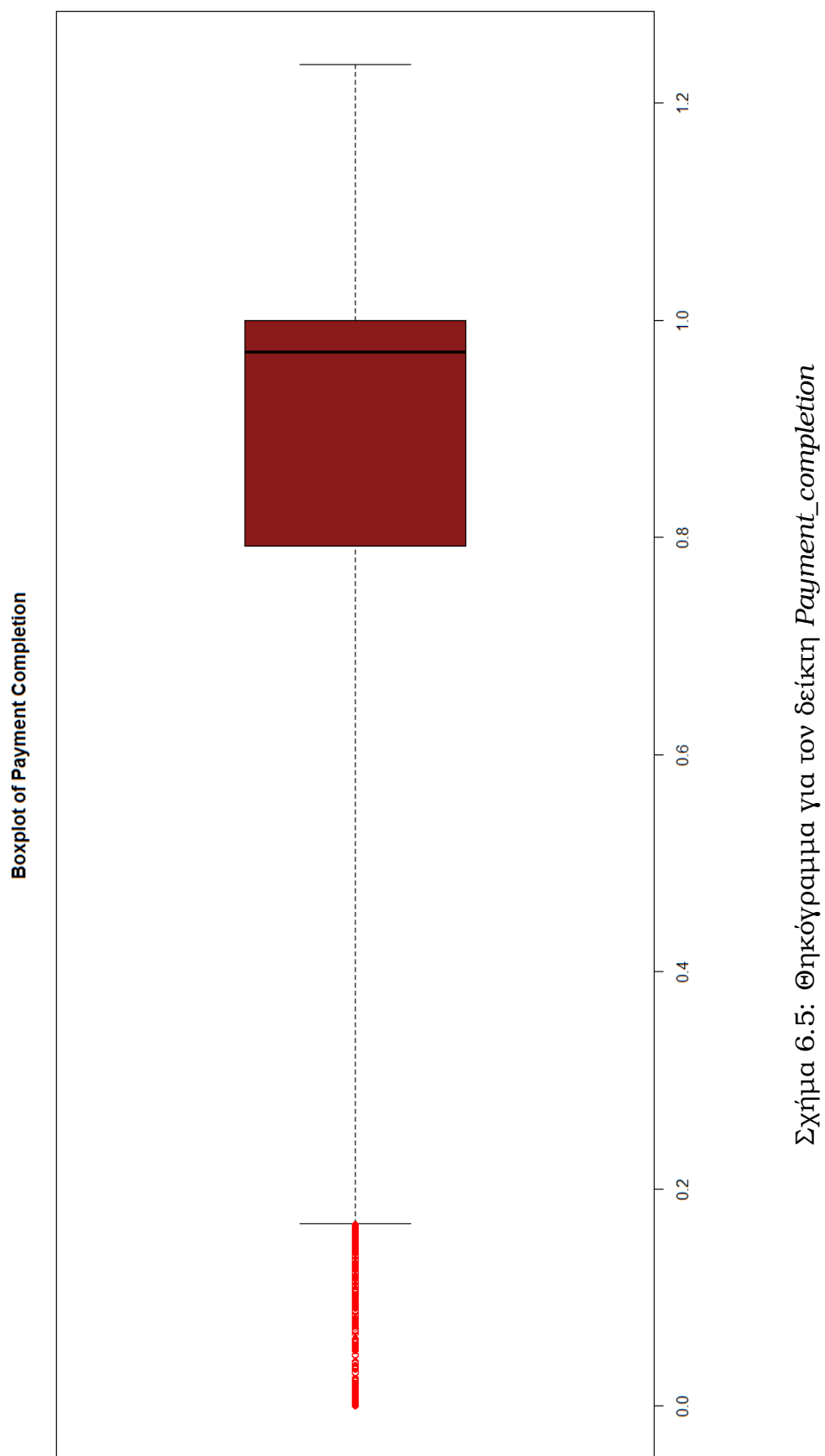
Συνολικά προέκυψαν 1561 καταχωρήσεις έργων με τουλάχιστον μία κόκκινη σημαία, 16 εξ' αυτών είχαν τρεις. Στο σχήμα 6.7 παρουσιάζονται τα 20 έργα με τις περισσότερες κόκκινες σημαίες.

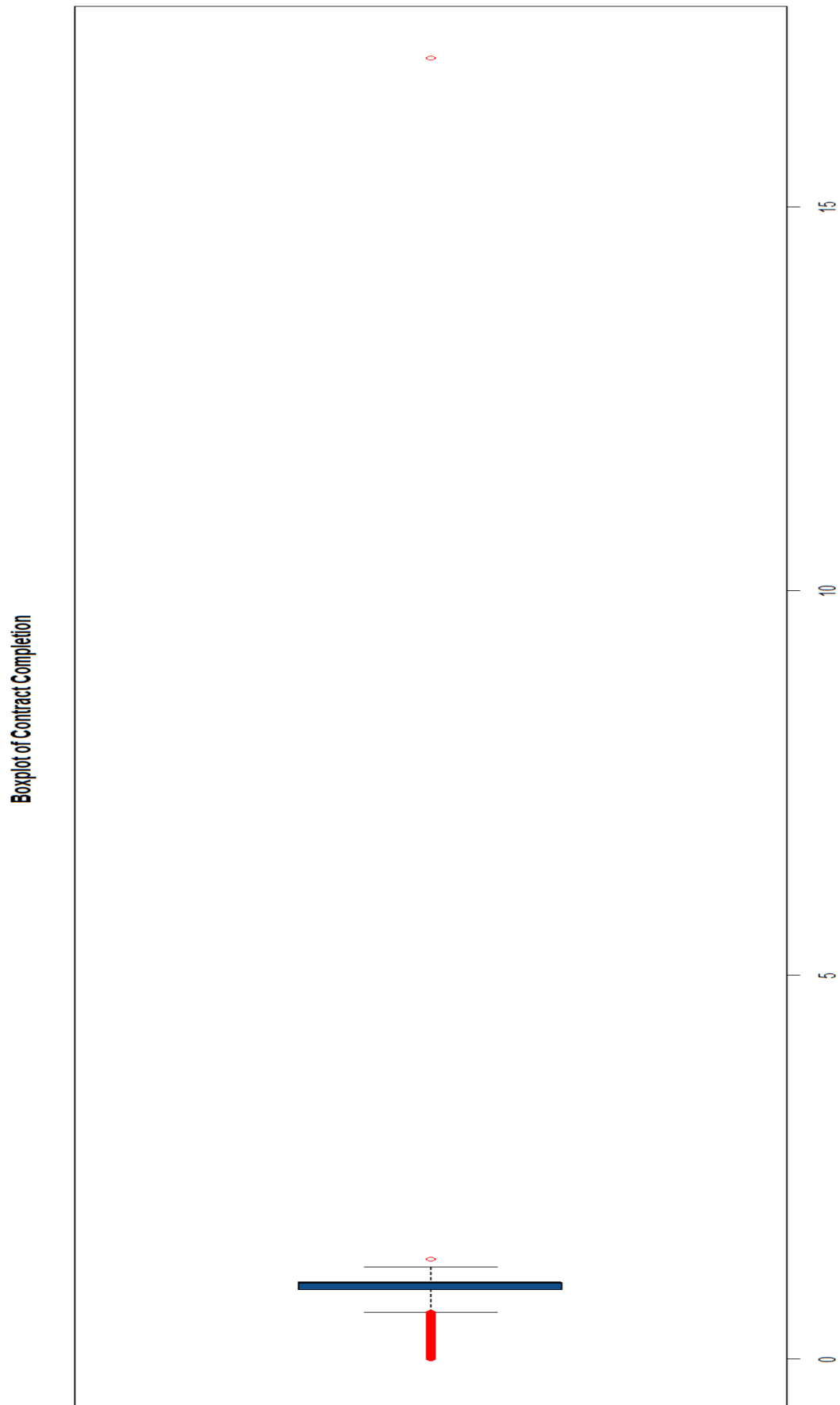


Σχήμα 6.4: Θηκόγραμμα για τον δείκτη *Completion*

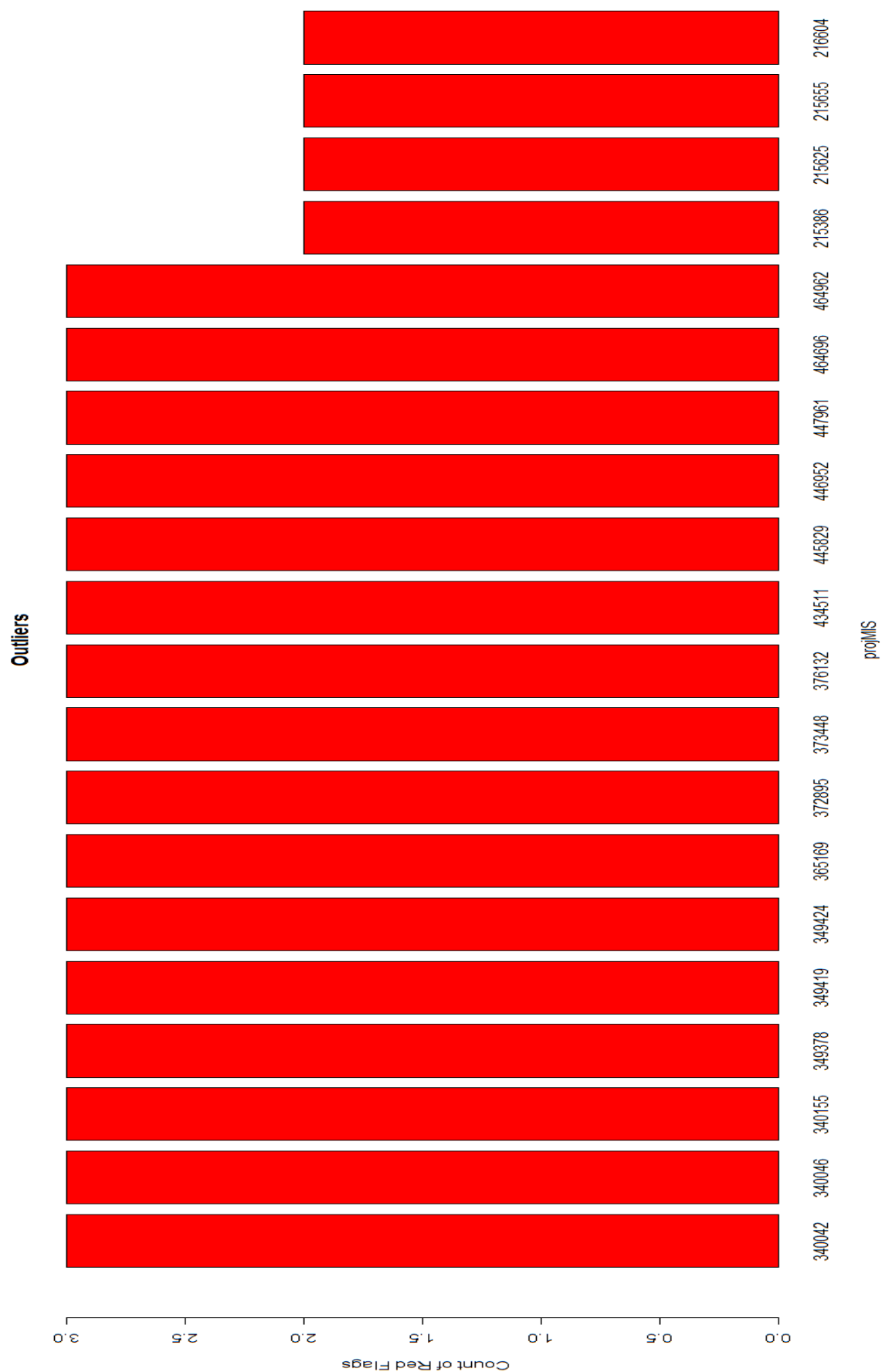
216604	277243	300505
216815	277248	301186
216969	296005	303654
304271	398790	376687
304314	398792	423556

Πίνακας 6.1: Τα έργα που ο δείκτης *Completion* ξεπερνάει το 1





Σχήμα 6.6: Θηκόγραμμα για τον δείκτη *Contract completion*



Σχήμα 6.7: Διάγραμμα συχνότητας των 20 έργων με τις περισσότερες κόκκινες σημαίες

6.2.3.2 Απαλλοτριώσεις

Στο σχήμα 6.8 παρουσιάζεται το Θηκόγραμμα για τον δείκτη *Estimated_area*. Παρατηρούνται 14 τιμές εκτός των ορίων του Θηκογράμματος, οι οποίες αποτελούν κόκκινες σημαίες. Από αυτές, τρεις ακραίες τιμές που αφορούν απαλλοτριώσεις έργων με εκτιμώμενη επιφάνεια για απαλλοτρίωση άνω των 90 εκατομμυρίων τετραγωνικών μέτρων.

Στο σχήμα 6.9, παρουσιάζεται το Θηκόγραμμα για τον δείκτη *Payments*. Δεν παρουσιάζονται ακραίες τιμές. Όμως υπάρχουν 62 έργα που ο δείκτης *Payments* ξεπερνά τη μονάδα και καταγράφονται ως κόκκινες σημαίες.

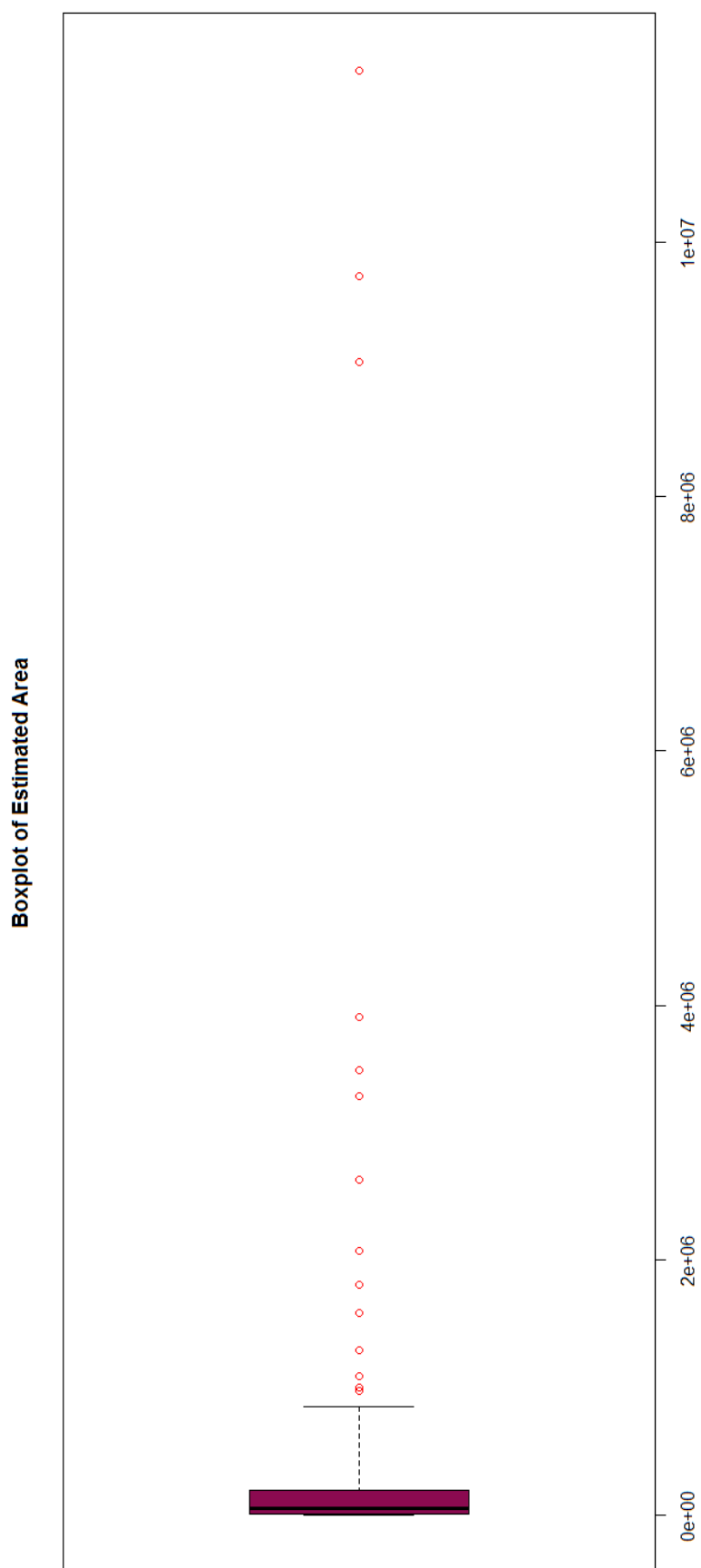
Στα σχήματα 6.10 και 6.11, παρουσιάζονται τα Θηκογράμματα για τους δείκτες *expro_Unassigned_Declared_Area* και *Completion*. Και στις δύο περιπτώσεις εντοπίζονται 7 ακραίες τιμές που αφορούν τις ίδιες απαλλοτριώσεις έργων.

Στο σχήμα 6.12, παρουσιάζεται το Θηκόγραμμα για τον δείκτη *Decisions*. Εντοπίζονται 12 ακραίες τιμές που σημειώνονται ως κόκκινες σημαίες.

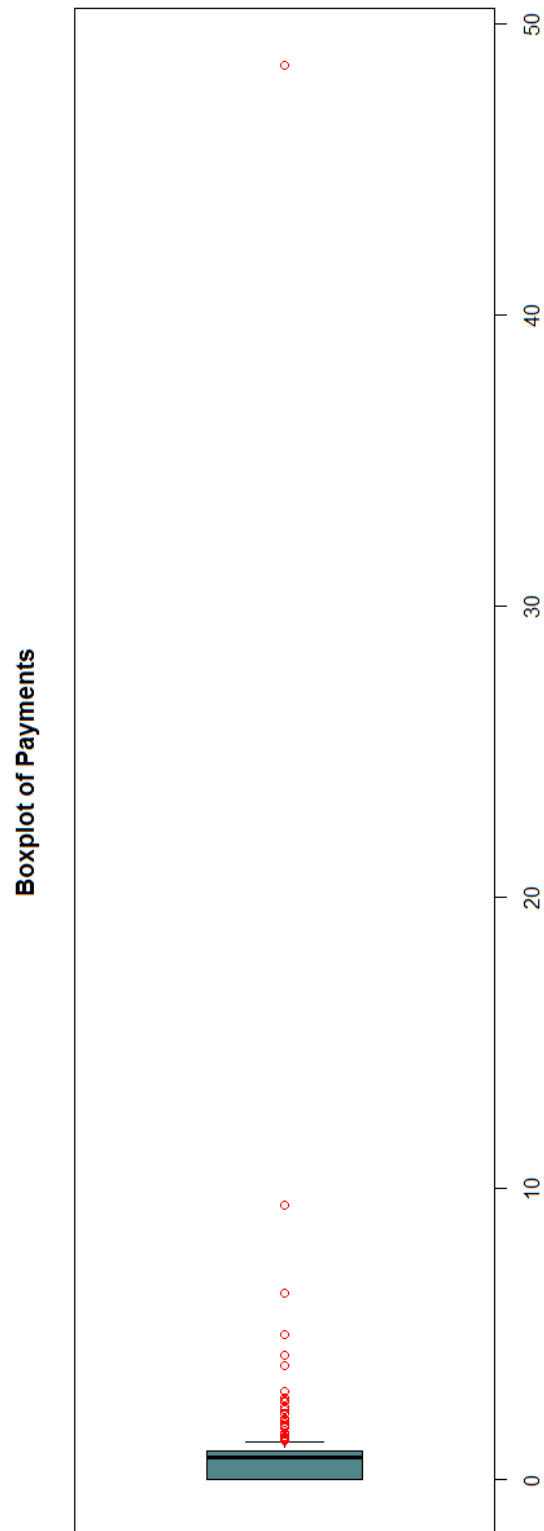
Στο σχήμα 6.13, παρουσιάζεται το Θηκόγραμμα για τον δείκτη *Initial_budget_changes*. Σύνολο παρατηρούνται 76 τιμές εκτός των ορίων.

Τέλος, στο σχήμα 6.14, παρουσιάζεται το Θηκόγραμμα για τον δείκτη *Declared_area_cost*. Εδώ εντοπίζονται 31 ακραίες τιμές που αποτελούν κόκκινες σημαίες.

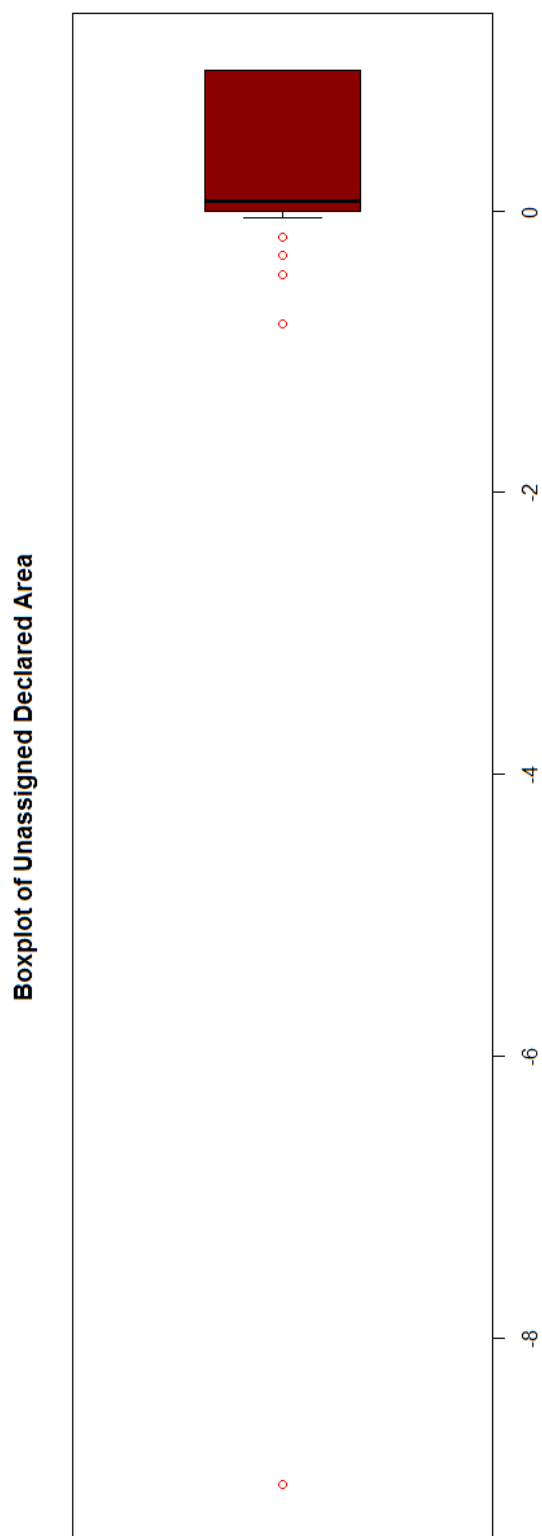
Για το σύνολο δεδομένων των απαλλοτριώσεων προέκυψαν 129 περιπτώσεις με τουλάχιστον 1 κόκκινη σημαία. Τις περισσότερες κόκκινες σημαίες συγκεντρώνει οι απαλλοτριώσεις του έργου 215620, ενώ 6 απαλλοτριώσεις είχαν από 4. Στο σχήμα 6.15 παρουσιάζονται οι 10 απαλλοτριώσεις έργων με τις περισσότερες κόκκινες σημαίες.



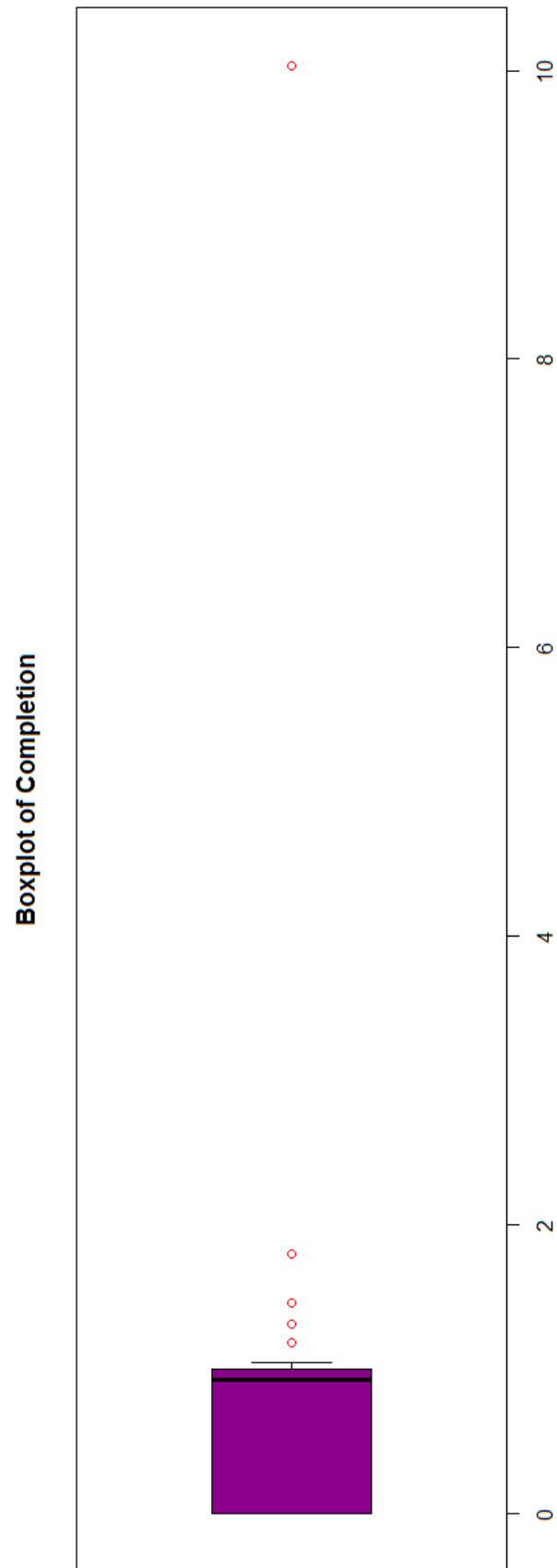
Σχήμα 6.8: Θηκόγραμμα για τον δείκτη *Estimated_area*



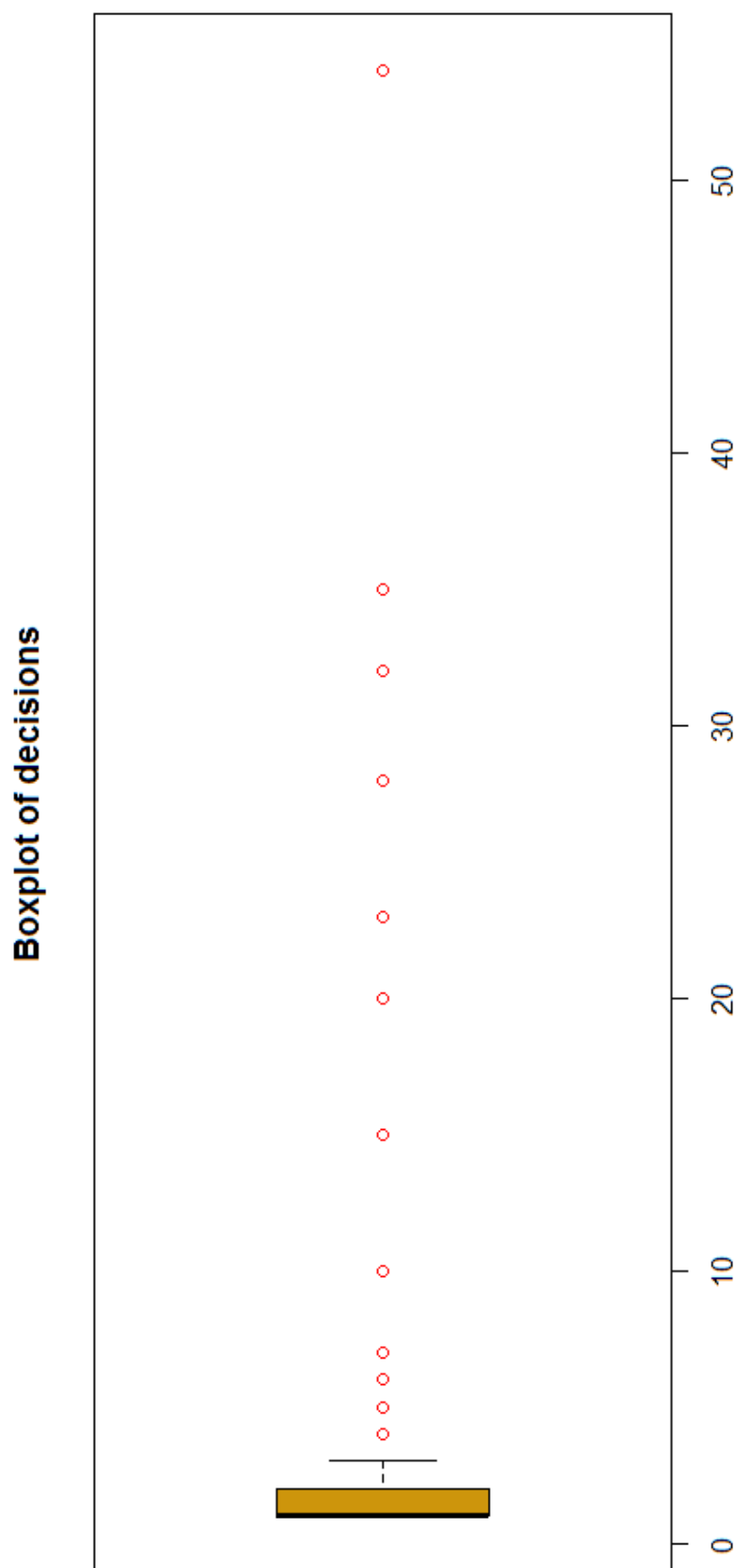
Σχήμα 6.9: Θηκόγραμμα για τον δείκτη *Payments*



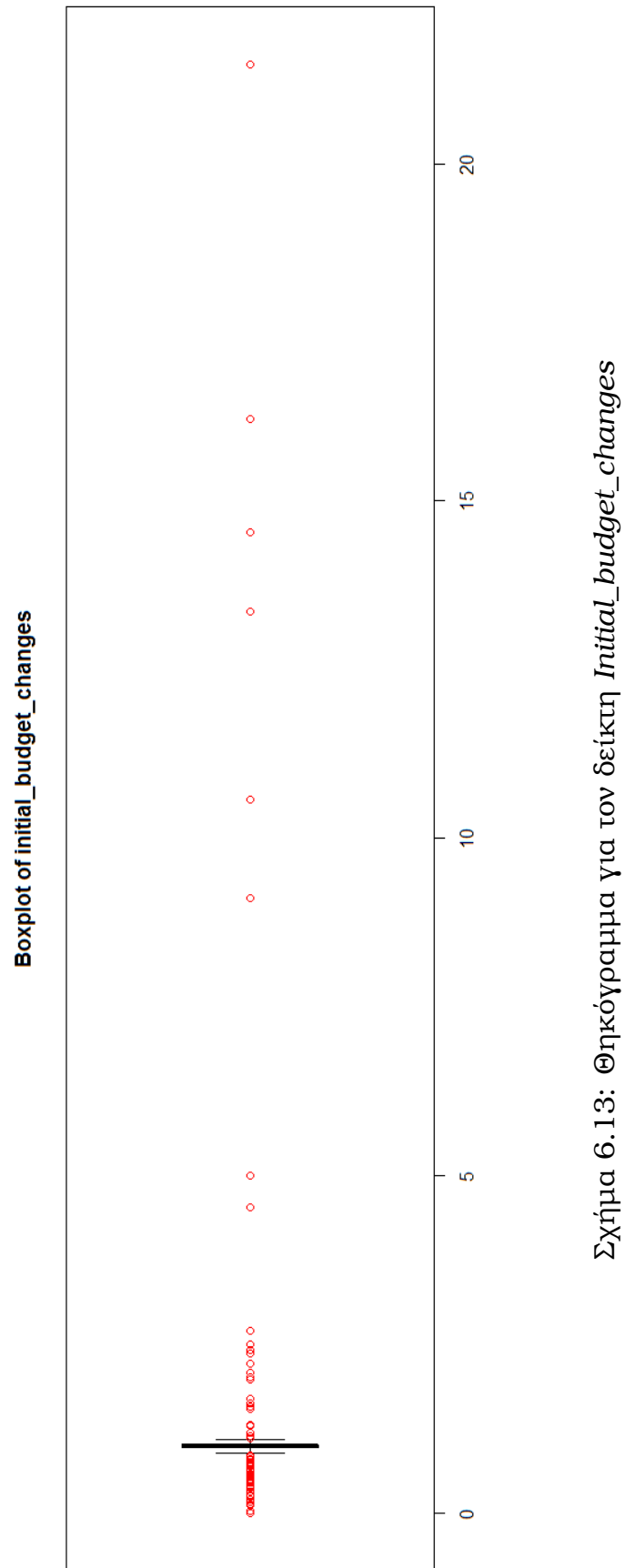
Σχήμα 6.10: Θηκόγραμμα για τον δείκτη *Unassigned_declared_area*

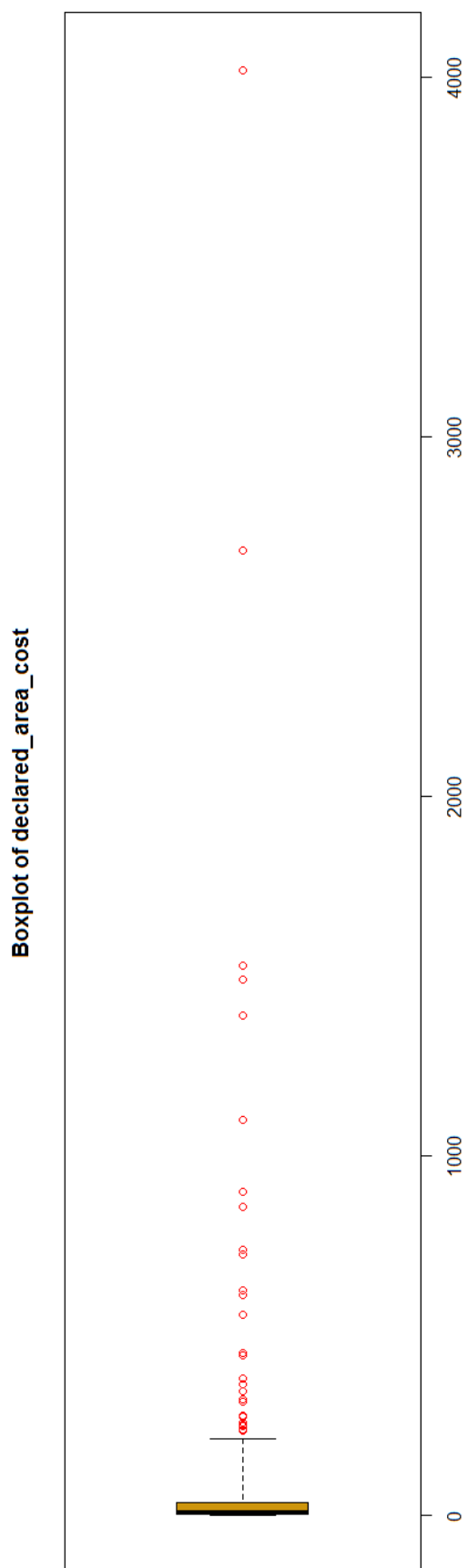


Σχήμα 6.1.1: Θηκόγραμμα για τον δείκτη *Completion*

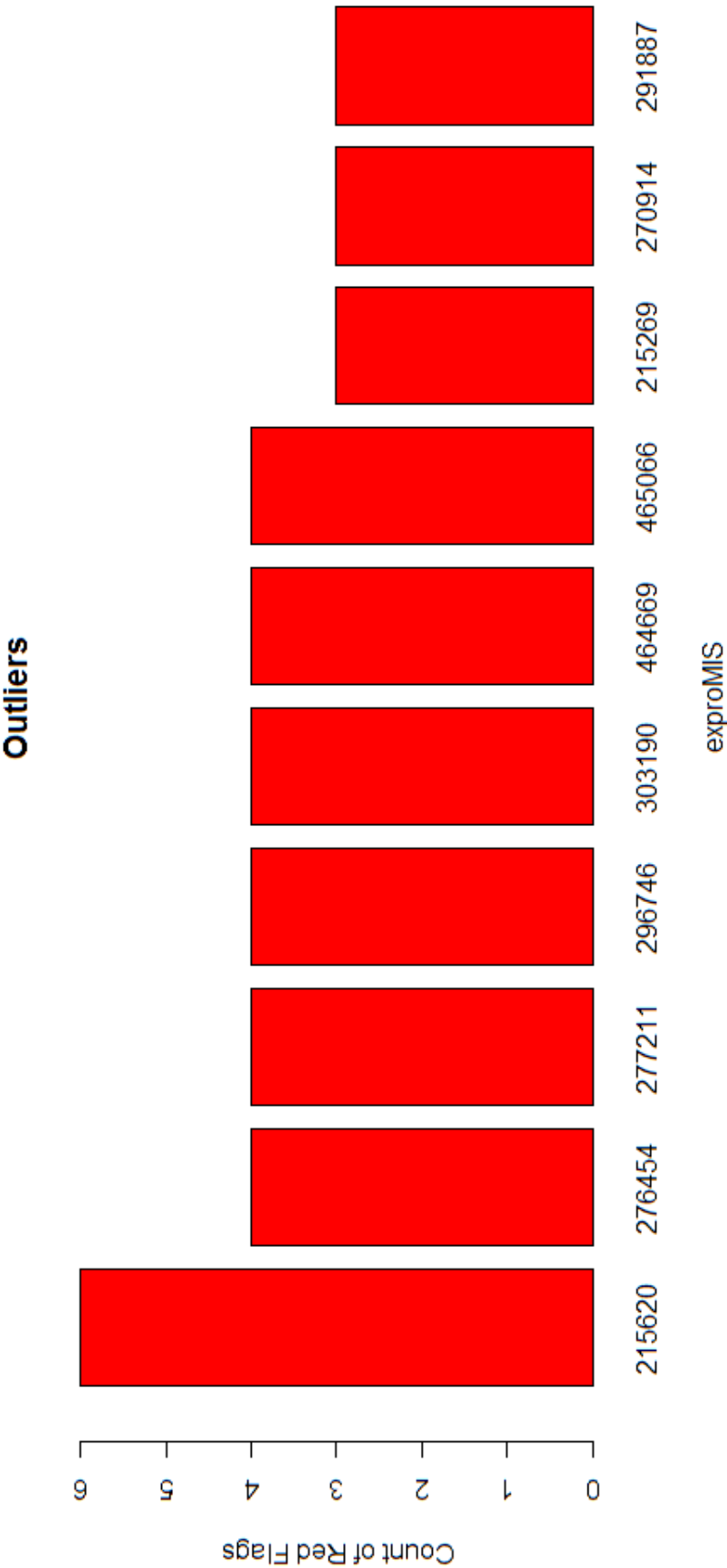


Σχήμα 6.12: Θηκόγραμμα για τον δείκτη *Decisions*





Σχήμα 6.14: Θηκόγραμμα για τον δείκτη *Declared_area_cost*



Σχήμα 6.15: Διάγραμμα συχνότητας των 10 απαλλοτριώσεων έργων με τις περισσότερες κόκκινες σημαίες

6.2.4 Σύνοψη Αποτελεσμάτων

Μετά την διαδικασία εντοπισμού των κόκκινων σημαιών για κάθε δείκτη των δύο συνόλων και αφού υπολογίστηκε ο συνολικός αριθμός κόκκινων σημαιών του κάθε έργου και απαλλοτρίωσης, τα δύο σύνολα δεδομένων συνδυάστηκαν με σκοπό να βρεθεί αν τα έργα που είναι κόκκινες σημαίες στους δείκτες απαλλοτρίωσης εμφανίζονται και στο άλλο σύνολο δεδομένων με τους δείκτες έργων. Τελικά προκύπτουν 39 τέτοιες περιπτώσεις, οι οποίες φαίνονται στον παρακάτω πίνακα.

215067	299666	296564
215268	299882	374349
215269	303190	375518
215386	327464	376118
216581	340111	377755
216706	348541	448235
216769	351885	448323
273513	352504	450589
277649	355373	465403
292398	355494	477937
296089	365912	488662
296294	371534	525503
296410	372944	525880

Πίνακας 6.2: Τα έργα με απαλλοτρίωση τα οποία εμφάνισαν κόκκινες σημαίες και στα δύο σύνολα που αναλύθηκαν.

Κεφάλαιο 7

Επίλογος

7.1 Συμπεράσματα

Η παρούσα εργασία μελέτησε τον τρόπο με τον οποίο μπορεί να περιγραφεί το σύνολο των - ανοιχτών - δεδομένων του ελληνικού ΕΣΠΑ μέσα από τις τεχνολογίες του Σηματολογικού Ιστού, ενώ μέρος των δεδομένων αυτών αναλύθηκε για τον εντοπισμό περιέργων συμπεριφορών, οι οποίες μπορούν να οδηγήσουν στην αντιμετώπιση μιας πιθανής περίπτωσης απάτης ή/και διαφθοράς.

Το θέμα που πραγματεύτηκε η εργασία ήταν άκρως ενδιαφέρον και σημαντικό, καθώς ήταν η πρώτη προσπάθεια για την περιγραφή δεδομένων που αφορά χρηματοδότηση έργων από την Ευρωπαϊκή Ένωση.

Για την υλοποίηση της εργασίας, ήταν αρχικά απαραίτητη η συλλογή και αποθήκευση των δεδομένων αυτών σε τοπική βάση. Αυτό επιτεύχθηκε με την σύνταξη κώδικα σε γλώσσα Python. Όσον αφορά το σηματολογικό σκέλος, για την περιγραφή των δεδομένων χρησιμοποιήθηκε πρώτα το μοντέλο δεδομένων OpenBudgets.eu, το οποίο είναι κατάλληλο για την εννοιολογική αναπαράσταση δημοσιονομικών δεδομένων. Όμως, η αδυναμία του μοντέλου να περιγράψει περισσότερες ιδιότητες των έργων και η απουσία κάποιου άλλου λεξιλογίου που να το πετυχαίνει επαρκώς, κατέστησε αναγκαία την κατασκευή μιας νέας οντολογίας με σκοπό την περιγραφή οικονομικών έργων.

Τέλος, έγινε μια πρώτη ανάλυση ενός υποσυνόλου των δεδομένων του ΕΣΠΑ με σκοπό την ανίχνευση πιθανών αφύσικων συμπεριφορών στα έργα και τις απαλλοτριώσεις του ΕΣΠΑ - κόκκινων σημαιών.

7.2 Μελλοντική Εργασία

Όσον αφορά τη μελλοντική εργασία, μια από τις άμεσες ενέργειες θα είναι η δημοσίευση της οντολογίας που κατασκευάστηκε στην ανοιχτή κοινότητα. Αυτό θα βοηθήσει στην γρήγορη ανάπτυξη της, αφού αφενός θα βρεθούν λάθη και τυχόν παρατυπίες και αφετέρου τα περισσότερα σχόλια και οι προτάσεις (feedback) θα οδηγήσουν στον εμπλουτισμό της με νέες κλάσεις -οι οποίες πιθανόν να αγνοήθηκαν και αποτελούν σημαντικές έννοιες - και άλλες ιδιότητες με αποτέλεσμα την σταθεροποίηση της οντολογίας (stable status). Στο ιδανικό σενάριο, θα μπορεί να χρησιμοποιείται ως πρότυπο για την περιγραφή δημοσιονομικών δεδομένων - έργων. Παράλληλα θα γίνει και πιο εύκολη η σύνδεση της με άλλες οντολογίες.

Ακολουθώντας το παράδειγμα της περιγραφής των δεδομένων του ΕΣΠΑ, μπορεί να γίνει προσπάθεια περιγραφής και άλλων συνόλων δημοσιονομικών δεδομένων και κάνοντας συνδέσεις μεταξύ τους (π.χ. δεδομένα ΕΣΠΑ άλλων χωρών, σύνδεση ΕΣΠΑ - Συμβάσεων) να εμπλουτιστεί ακόμα παραπάνω το παγκόσμιο Νέφος Διασυνδεδεμένων Ανοιχτών Δεδομένων (Linked Open Data Cloud - LOD Cloud).

Ένα στοίχημα θα είναι η δημιουργία pipeline με την προγραμματισμένη ενημέρωση των δεδομένων του ΕΣΠΑ και μετατροπή τους σε τριπλέτες ανά τακτά χρονικά διαστήματα (π.χ ανά τρίμηνο).

Μία ακόμα ενδιαφέρουσα εφαρμογή θα ήταν η κατασκευή ενός ιστοχώρου με σύγχρονο σχεδιασμό, εύκολο και απλό περιβάλλον χρήσης, με σκοπό την οπτικοποίηση των κόκκινων σημαιών.

Τέλος, ο μεγάλος στόχος είναι η δημιουργία/εφαρμογή πιο πολύπλοκων αλγορίθμων για την εύρεση των κόκκινων σημαιών, οι οποίοι θα μπορούν να χρησιμοποιούν μη επιβλεπόμενη μάθηση και θα εφαρμόζονται στο σύνολο των δεδομένων του ΕΣΠΑ.

Βιβλιογραφία

- [1] Ec.europa.eu. (2016). *Ευρωπαϊκά Διαρθρωτικά και Επενδυτικά Ταμεία - Ευρωπαϊκή Επιτροπή*. Διαθέσιμο σε: http://ec.europa.eu/contracts_grants/funds_el.htm (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [2] Ec.europa.eu. (2016). *Ευρωπαϊκά διαρθρωτικά και επενδυτικά ταμεία*. Διαθέσιμο σε: http://ec.europa.eu/regional_policy/el/funding/ (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [3] ΕΣΠΑ 2014-2020. (2009). *Διαρθρωτικά Ταμεία*. [online] Διαθέσιμο σε: <https://www.espa.gr/el/pages/DictionaryFS.aspx?item=99> (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [4] Ec.europa.eu. (2014). *Ευρωπαϊκό Ταμείο Περιφερειακής Ανάπτυξης*. Διαθέσιμο σε: http://ec.europa.eu/regional_policy/el/funding/erdf/ (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [5] Ec.europa.eu. (2016). *Γλωσσάριο*. Διαθέσιμο σε: http://ec.europa.eu/regional_policy/el/policy/what/glossary/e/european-regional-development-fund (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [6] Ec.europa.eu. (2014). *Ευρωπαϊκό Κοινωνικό Ταμείο*. Διαθέσιμο σε: http://ec.europa.eu/regional_policy/el/funding/social-fund/ (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [7] Ec.europa.eu. (2016). *Γλωσσάριο*. Διαθέσιμο σε: http://ec.europa.eu/regional_policy/el/policy/what/glossary/e/european-social-fund (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [8] Ec.europa.eu. (2015). *Ταμείο Συνοχής*. Διαθέσιμο σε: http://ec.europa.eu/regional_policy/el/funding/cohesion-fund/ (Ανακτήθηκε 12 Δεκεμβρίου, 2016).

- [9] Ec.europa.eu. (2016). *Γλωσσάριο*. Διαθέσιμο σε: http://ec.europa.eu/regional_policy/el/policy/what/glossary/c/cohesion-fund (Ανακτήθηκε 12 Δεκεμβρίου, 2016).
- [10] Ec.europa.eu. (2016). *Γλωσσάριο*. Διαθέσιμο σε: http://ec.europa.eu/regional_policy/el/policy/what/glossary/e/european-agricultural-fund-for-rural-development (Ανακτήθηκε 12 Δεκεμβρίου, 2016).
- [11] Αλιεία - European Commission. (2016). *Ευρωπαϊκό Ταμείο Θάλασσας και Αλιείας (ΕΤΘΑ) - Αλιεία - European Commission*. Διαθέσιμο σε: https://ec.europa.eu/fisheries/cfp/emff_el (Ανακτήθηκε 12 Δεκεμβρίου, 2016).
- [12] Ec.europa.eu. (2015). *Προγραμματισμός και υλοποίηση*. Διαθέσιμο σε: http://ec.europa.eu/regional_policy/el/policy/how/stages-step-by-step/ (Ανακτήθηκε 12 Δεκεμβρίου, 2016).
- [13] www.espa.gr. (2016). *ΕΣΠΑ 2014-2020*. Διαθέσιμο σε: <https://www.espa.gr/el/Pages/staticESPA2014-2020.aspx> (Ανακτήθηκε 12 Δεκεμβρίου, 2016).
- [14] www.espa.gr. (2016). *Στοιχεία και πληροφορίες*. Διαθέσιμο σε: <https://www.espa.gr/el/pages/Factsheet.aspx> (Ανακτήθηκε 12 Δεκεμβρίου, 2016).
- [15] www.espa.gr. (2015). *Επιχειρησιακά Προγράμματα*. Διαθέσιμο σε: <https://www.espa.gr/el/Pages/staticOperationalProgrammes.aspx> (Ανακτήθηκε 12 Δεκεμβρίου, 2016).
- [16] Applica, Ismeri Europa and Cambridge Economic Associates, (2016). *Task 3 Country Report Greec. Ex post evaluation of Cohesion Policy programmes 2007-2013, focusing on the European Regional Development Fund (ERDF) and the Cohesion Fund (CF)*. [online] Luxembourg: Publications Office of the European Union, 2016: EUROPEAN COMMISSION. Διαθέσιμο σε: http://ec.europa.eu/regional_policy/sources/docgener/evaluation/pdf/expost2013/wp1_el_report_en.pdf (Ανακτήθηκε 12 Δεκεμβρίου, 2016).
- [17] www.espa.gr. (2015). *Η Πολιτική Συνοχής της ΕΕ*. Διαθέσιμο σε: <https://www.espa.gr/el/pages/staticCohesionPolicy.aspx> (Ανακτήθηκε 12 Δεκεμβρίου, 2016).
- [18] Ec.europa.eu. (2014). *10 ερωτήσεις για την πολιτική συνοχής - Περιφερειακή Πολιτική της ΕΕ*. Διαθέσιμο σε: http://ec.europa.eu/regional_policy/el/faq/#1 (Ανακτήθηκε 12 Δεκεμβρίου, 2016).

- [19] www.espa.gr. (2015). *Πολιτική της Συνοχής*. Διαθέσιμο σε: <https://www.espa.gr/el/Pages/staticEUPolicies.aspx> (Ανακτήθηκε 12 Δεκεμβρίου, 2016).
- [20] Antoniou, G. and Van Harmelen, F. (2008). *A semantic Web primer*. 2nd ed. Cambridge, Mass.: MIT Press.
- [21] World Wide Web Foundation. (2016). *History of the Web*. Διαθέσιμο σε: <http://webfoundation.org/about/vision/history-of-the-web/> (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [22] Berners-Lee, T., Hendler, J. and Lassila, O. (2001). The Semantic Web. *Scientific American*, 284(5), pp.34-43.
- [23] W3.org. (2016). *Semantic Web - W3C*. Διαθέσιμο σε: <https://www.w3.org/standards/semanticweb/> (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [24] W3.org. (2016). *Semantic Web Standards*. Διαθέσιμο σε: https://www.w3.org/2001/sw/wiki/Main_Page (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [25] [En.wikipedia.org](http://en.wikipedia.org). (2016). *Semantic Web Stack*. Διαθέσιμο σε: https://en.wikipedia.org/wiki/Semantic_Web_Stack#Semantic_Web_technologies (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [26] W3.org. (2006). *Linked Data - Design Issues*. Διαθέσιμο σε: <https://www.w3.org/DesignIssues/LinkedData.html> (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [27] [En.wikipedia.org](http://en.wikipedia.org). (2016). *Ontology*. Διαθέσιμο σε: <https://en.wikipedia.org/wiki/Ontology> (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [28] [El.wikipedia.org](http://el.wikipedia.org). (2016). *Οντολογία (πληροφορική)*. Διαθέσιμο σε: [https://el.wikipedia.org/wiki/%CE%9F%CE%BD%CF%84%CE%BF%CE%BB%CE%BF%CE%B3%CE%AF%CE%B1_\(%CF%80%CE%BB%CE%B7%CF%81%CE%BF%CF%86%CE%BF%CF%81%CE%B9%CE%BA%CE%AE\)](https://el.wikipedia.org/wiki/%CE%9F%CE%BD%CF%84%CE%BF%CE%BB%CE%BF%CE%B3%CE%AF%CE%B1_(%CF%80%CE%BB%CE%B7%CF%81%CE%BF%CF%86%CE%BF%CF%81%CE%B9%CE%BA%CE%AE)) (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [29] [En.wikipedia.org](http://en.wikipedia.org). (2016). *Ontology (information science)*. Διαθέσιμο σε: [https://en.wikipedia.org/wiki/Ontology_\(information_science\)#Components](https://en.wikipedia.org/wiki/Ontology_(information_science)#Components) (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [30] [En.wikipedia.org](http://en.wikipedia.org). (2016). *Ontology engineering*. Διαθέσιμο σε: https://en.wikipedia.org/wiki/Ontology_engineering#Ontology_languages (Ανακτήθηκε 11 Δεκεμβρίου, 2016).

- [31] *ΤΙ ΕΙΝΑΙ ΤΟ ANAPTYXI.GOV.GR*. Διαθέσιμο σε: <http://anaptyxi.gov.gr/Default.aspx?tabid=249&language=el-GR> (Ανακτήθηκε 10 Δεκεμβρίου, 2016).
- [32] *API Ανοιχτών Δεδομένων*. Διαθέσιμο σε: <http://anaptyxi.gov.gr/ODHelp.aspx> (Ανακτήθηκε 10 Δεκεμβρίου, 2016).
- [33] *ΤΙ ΕΙΝΑΙ ΤΟ ANAPTYXI.GOV.GR - Χρήσιμοι ορισμοί - Έργα*. Διαθέσιμο σε: <http://anaptyxi.gov.gr/Default.aspx?tabid=249&language=el-GR#ergaTypesContainer-1> (Ανακτήθηκε 10 Δεκεμβρίου, 2016).
- [34] *ΤΙ ΕΙΝΑΙ ΤΟ ANAPTYXI.GOV.GR - Χρήσιμοι ορισμοί - Ενισχύσεις-Επιδότησεις*. Διαθέσιμο σε: <http://anaptyxi.gov.gr/Default.aspx?tabid=249&language=el-GR#ergaTypesContainer-3> (Ανακτήθηκε 10 Δεκεμβρίου, 2016).
- [35] *ΤΙ ΕΙΝΑΙ ΤΟ ANAPTYXI.GOV.GR - Χρήσιμοι ορισμοί - Έργα προτεραιότητας*. Διαθέσιμο σε: <http://anaptyxi.gov.gr/Default.aspx?tabid=249&language=el-GR#ergaTypesContainer-5> (Ανακτήθηκε 10 Δεκεμβρίου, 2016).
- [36] *ObeuBudgets.eu*. Διαθέσιμο σε: <http://openbudgets.eu/> (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [37] *ObeuBudgets.eu - Objectives*. Διαθέσιμο σε: <http://openbudgets.eu/about/objectives/> (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [38] Mynarz, J. (2016). *The OpenBudgets.eu data model*. <https://openbudgets.eu/>. Διαθέσιμο σε: https://openbudgets.eu/post/2016/11/11/openbudgets_eu_data_model/ (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [39] W3.org. (2016). *The RDF Data Cube Vocabulary*. Διαθέσιμο σε: <http://www.w3.org/TR/vocab-data-cube/> (Ανακτήθηκε 11 Δεκεμβρίου, 2016).
- [40] Musen, M. (2015). The protégé project. *AI Matters*, 1(4), pp.4-12.
- [41] Knap, T., Hanecak, P., Klimek, J., Mader, C., Necasky, M., Van Nuffelen, B. and Skoda, P. (2016). *UnifiedViews: An ETL Tool for RDF Data Management*. Semantic Web – Interoperability, Usability, Applicability an IOS Press Journal. Διαθέσιμο σε: <http://www.semantic-web-journal.net/system/files/swj1490.pdf> (Ανακτήθηκε 12 Δεκεμβρίου, 2016).
- [42] W3.org. (2016). *The RDF Data Cube Vocabulary*. Διαθέσιμο σε: <https://www.w3.org/TR/vocab-data-cube/#wf-rules> (Ανακτήθηκε 13 Δεκεμβρίου, 2016).

- [43] Klímek J., Kučera J., Mynarz J., Sedmihradská L., Zbranek J. (2015a): OpenBudgets.eu - Deliverable D1.2 - *Design of data structure definition for public budget data*, <http://openbudgets.eu/assets/deliverables/D1.2.pdf>.
- [44] Klímek J., Kučera J., Mynarz J., Sedmihradská L., Zbranek J. (2015b): OpenBudgets.eu - Deliverable D1.3 - *Design of data structure definition for public spending data*, <http://openbudgets.eu/assets/deliverables/D1.3.pdf>.
- [45] Mynarz, Jindřich; Svátek, Vojtěch; Karampatakis, Sotirios; Klímek, Jakub; Bratsas, Charalampos. Modeling fiscal data with the Data Cube Vocabulary. In Martin, Michael; Cuquet, Martí; Folmer, Erwin (eds.). *Joint Proceedings of the Posters and Demos Track of the 12th International Conference on Semantic Systems - SEMANTiCS2016 and the 1st International Workshop on Semantic Change & Evolving Semantics (SuCCESS'16) co-located with the 12th International Conference on Semantic Systems (SEMANTiCS 2016)*. Aachen: RWTH Aachen University, 2016. CEUR workshop proceedings, vol. 1695. Διαθέσιμο σε: <http://ceur-ws.org/Vol-1695/paper25.pdf>
- [46] Koupidis, K., Bratsas, C., Karampatakis, S., Martzopoulou, A. and Antoniou, I. (2016). Fiscal Knowledge discovery in Municipalities of Athens and Thessaloniki via Linked Open Data. 2016 *11th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)*.
- [47] Filippidis, P., Karampatakis, S., Koupidis, K., Ioannidis, L. and Bratsas, C. (2016). The code lists case: Identifying and linking the key parts of fiscal datasets. 2016 *11th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)*.
- [48] Red Flags. (2016). Red Flags. Διαθέσιμο σε: <http://www.redflags.eu/> (Ανακτήθηκε 15 Δεκεμβρίου, 2016).
- [49] Wensink, W. and Maarten de Vet, J. (2013). *Development of a methodology to estimate the direct costs of corruption and other elements for an EU-evaluation mechanism in the area of anti-corruption. Identifying and Reducing Corruption in Public Procurement in the EU*. Διαθέσιμο σε: http://ec.europa.eu/anti-fraud/sites/antifraud/files/docs/body/identifying_reducing_corruption_in_public_procurement_en.pdf (Ανακτήθηκε 15 Δεκεμβρίου, 2016).

- [50] Europe calling: Contribute to the EU procurement data standard. (2016). [Blog] <http://www.open-contracting.org/>. Διαθέσιμο σε: <http://www.open-contracting.org/2016/11/29/europe-calling-contribute-eu-procurement-data-standard/> (Ανακτήθηκε 15 Δεκεμβρίου, 2016).
- [51] Open Contracting Partnership. (2016). *Why Open Contracting - Open Contracting Partnership*. Διαθέσιμο σε: <http://www.open-contracting.org/why-open-contracting/> (Ανακτήθηκε 15 Δεκεμβρίου, 2016).
- [52] Anita, N. and Tünde, T. (2015). *Red Flags project - NEW WARNING SYSTEM FOR THE IDENTIFICATION OF RED FLAGS IN PUBLIC PROCUREMENTS*. Budapest: Transparency International Hungary and K-Monitor. Διαθέσιμο σε: <http://www.redflags.eu/files/redflags-summary-en.pdf> (Ανακτήθηκε 15 Δεκεμβρίου, 2016).
- [53] *Η ΠΡΟΛΗΨΗ ΤΗΣ ΑΠΑΤΗΣ στις ΔΙΑΠΘΡΩΤΙΚΕΣ ΔΡΑΣΕΙΣ*. (2014). ΓΕΝΙΚΗ ΓΡΑΜΜΑΤΕΙΑ ΔΗΜΟΣΙΩΝ ΕΠΕΝΔΥΣΕΩΝ - ΕΣΠΑ. Διαθέσιμο σε: <http://www.epdm.gr/Uploads/Files/KatapolemisiApatis/EYUYredflags.pdf> (Ανακτήθηκε 15 Δεκεμβρίου, 2016).
- [54] Εισαγωγή - Κόκκινες Σημαίες. (n.d.). ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΘΕΣΜΙΚΗΣ ΥΠΟΣΤΗΡΙΞΗΣ. Διαθέσιμο σε: <http://www.epdm.gr/Uploads/Files/KatapolemisiApatis/EYUYredflags.pdf> (Ανακτήθηκε 15 Δεκεμβρίου, 2016).