



ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ
ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΞΕΙΔΙΚΕΥΣΗ ΣΤΑ ΠΛΗΡΟΦΟΡΙΑΚΑ
ΣΥΣΤΗΜΑΤΑ

Μεταπτυχιακή Διπλωματική Εργασία

«Έλεγχος Συμμόρφωσης (Conformance Checking) Πειραματικών Διαδικασιών σε
Περιβάλλον Εικονικού Εργαστηρίου μέσω Εργαλείων Εξόρυξης Διεργασιών
(ProcessMining) ανοικτού κώδικα»

«ΚΟΣΜΑΔΑΚΗΣ ΝΙΚΟΛΑΟΣ»

Επιβλέπων καθηγητής: «Σύψας Αθανάσιος»

Πάτρα, Σεπτέμβριος 2026

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του/της φοιτητή/φοιτήτριας («συγγραφέας/δημιουργός») που την εκπόνησε. Στο πλαίσιο της πολιτικής ανοικτής πρόσβασης ο/η συγγραφέας/δημιουργός εκχωρεί στο ΕΑΠ, μη αποκλειστική άδεια χρήσης του δικαιώματος αναπαραγωγής, προσαρμογής, δημόσιου δανεισμού, παρουσίασης στο κοινό και ψηφιακής διάχυσής τους διεθνώς, σε ηλεκτρονική μορφή και σε οποιοδήποτε μέσο, για διδακτικούς και ερευνητικούς σκοπούς, άνευ ανταλλάγματος και για όλο το χρόνο διάρκειας των δικαιωμάτων πνευματικής ιδιοκτησίας. Η ανοικτή πρόσβαση στο πλήρες κείμενο για μελέτη και ανάγνωση δεν σημαίνει καθ' οιονδήποτε τρόπο παραχώρηση δικαιωμάτων διανοητικής ιδιοκτησίας του/της συγγραφέα/δημιουργού ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αντιγραφή, αποθήκευση, πώληση, εμπορική χρήση, μετάδοση, διανομή, έκδοση, εκτέλεση, «μεταφόρτωση» (downloading), «ανάρτηση» (uploading), μετάφραση, τροποποίηση με οποιονδήποτε τρόπο, τμηματικά ή περιληπτικά της εργασίας, χωρίς τη ρητή προηγούμενη έγγραφη συναίνεση του/της συγγραφέα/δημιουργού. Ο/Η συγγραφέας/δημιουργός διατηρεί το σύνολο των ηθικών και περιουσιακών του δικαιωμάτων.



«Έλεγχος Συμμόρφωσης (Conformance Checking) Πειραματικών
Διαδικασιών σε Περιβάλλον Εικονικού Εργαστηρίου μέσω
Εργαλείων Εξόρυξης Διεργασιών (ProcessMining) ανοικτού
κώδικα»

ΚΟΣΜΑΔΑΚΗΣ ΝΙΚΟΛΑΟΣ

Επιτροπή Επίβλεψης Διπλωματικής Εργασίας

Επιβλέπων Καθηγητής:
Σύψας Αθανάσιος
ΣΕΠ, Ελληνικό Ανοικτό Πανεπιστήμιο

Συν-Επιβλέπουσα Καθηγήτρια:
Τσώνη Ροδάνθη
ΣΕΠ, Ελληνικό Ανοικτό Πανεπιστήμιο

Πάτρα, Σεπτέμβριος 2026

«Θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου για τη συνολική καθοδήγηση και τη βοήθεια που μου παρείχε στην εκπόνηση της διπλωματικής μου εργασίας. Παράλληλα, εκφράζω τις ευχαριστίες μου στη συν-επιβλέπουσα καθηγήτριά μου για την πολύτιμη συμβολή και την υποστήριξή της. Πάνω απ' όλα, θα ήθελα να ευχαριστήσω την οικογένειά μου, για την απεριόριστη στήριξη και την πίστη που μου έδειξε σε κάθε μου βήμα.»

Περίληψη

Η Εξόρυξη Διεργασιών (Process Mining) είναι μια σύγχρονη τεχνολογία που συνδυάζει την ανάλυση επιχειρησιακών διαδικασιών με τα δεδομένα των πληροφοριακών συστημάτων. Μέσω της επεξεργασίας αρχείων καταγραφής συμβάντων, επιτρέπει την απεικόνιση των πραγματικών ροών εργασίας, αποκαλύπτοντας αποκλίσεις από τα θεωρητικά ή προκαθορισμένα μοντέλα. Μία από τις σημαντικότερες τεχνικές του Process Mining είναι ο Έλεγχος Συμμόρφωσης (Conformance Checking), το οποίο συγκρίνει τη διαδικασία όπως εκτελείται στην πράξη με ένα προκαθορισμένο πρότυπο ή μοντέλο αναφοράς. Στόχος είναι η διαπίστωση της συμμόρφωσης των διαδικασιών και η ανίχνευση αποκλίσεων που μπορεί να επηρεάζουν την αποδοτικότητα και την ποιότητα.

Το Conformance Checking παρέχει πολύτιμες πληροφορίες για την κατανόηση των αιτιών των αποκλίσεων και την πρόταση βελτιώσεων. Εφαρμόζεται σε διάφορους τομείς, όπως βιομηχανία, υγειονομική περίθαλψη και χρηματοοικονομικές υπηρεσίες, όπου η ακρίβεια και η τήρηση κανόνων είναι κρίσιμες. Μέσω αυτής της τεχνικής, οι οργανισμοί μπορούν να ενισχύσουν την αξιοπιστία των διαδικασιών τους και να επιτύχουν μεγαλύτερη αποτελεσματικότητα. Τα δεδομένα που παρέχει υποστηρίζουν τη λήψη στρατηγικών αποφάσεων και την ανάπτυξη βέλτιστων πρακτικών. Συνολικά, το Conformance Checking αποτελεί βασικό εργαλείο για τη σύνδεση της θεωρίας με την πραγματικότητα στην ανάλυση διαδικασιών.

Η παρούσα διπλωματική εργασία εξετάζει την εφαρμογή τεχνικών Process Mining και ειδικότερα Conformance Checking για την ανάλυση πειραματικών διαδικασιών που υλοποιούνται σε περιβάλλον εικονικού εργαστηρίου.

Η εργασία συνδυάζει τη μοντελοποίηση διαδικασιών μέσω της UML (κυρίως με χρήση Activity Diagrams) και τη χρήση εργαλείων ανοικτού κώδικα για ανάλυση ελέγχου συμμόρφωσης. Εξετάζεται κατά πόσο οι διαδικασίες που πραγματοποιούνται σε εικονικά εργαστήρια ακολουθούν τα καθορισμένα πρότυπα, βάσει των καταγεγραμμένων αρχείων καταγραφής συμβάντων που προκύπτουν από το περιβάλλον του εικονικού εργαστηρίου.

Λέξεις - Κλειδιά

Εξόρυξη Διεργασιών

Έλεγχος Συμμόρφωσης

Virtual Laboratories

UML activity diagrams

ProM

PM4Py

«Conformance Checking of Experimental Procedures in a Virtual Laboratory Environment Using Open-Source Process Mining Tools»

KOSMADAKIS NIKOLAOS

Abstract

Process Mining is a modern technology that connects business process analysis with data from information systems. By processing event logs, it enables the visualization of actual workflow execution, revealing deviations from theoretical or predefined models. One of the most important techniques in Process Mining is Conformance Checking, which compares the process as executed in practice with a predefined reference model. Its goal is to assess process compliance and detect deviations that may affect efficiency and quality.

Conformance Checking provides valuable insights into the causes of deviations and suggests potential improvements. It is applied across various domains such as industry, healthcare, and financial services, where accuracy and adherence to rules are critical. Through this technique, organizations can enhance the reliability of their processes and achieve greater effectiveness. The data it provides supports strategic decision-making and the development of best practices. Overall, Conformance Checking serves as a key tool for bridging the gap between theoretical models and practical reality in business process analysis.

This thesis investigates the application of Process Mining techniques, with a particular focus on Conformance Checking, for the analysis of experimental procedures conducted in a virtual laboratory environment.

The study combines process modeling through UML (primarily using Activity Diagrams)

with the use of open-source tools for conformance analysis. It examines the extent to which procedures carried out in virtual laboratories adhere to established standards, based on the event log files recorded within the virtual laboratory environment.

Keywords

Process Mining

Conformance Checking

Virtual Laboratories

UML activity diagrams

ProM

PM4Py

Περιεχόμενα

Περίληψη.....	v
Abstract.....	vii
Περιεχόμενα.....	ix
Κατάλογος Εικόνων / Σχημάτων	xi
Κατάλογος Πινάκων.....	xiii
Συντομογραφίες & Ακρωνύμια	xiv
1. Εξόρυξη Διεργασιών και Έλεγχος Συμμόρφωσης.....	1
1.1 Εξόρυξη Διεργασιών	1
1.1.1 Η έννοια της Εξόρυξης Διεργασιών.....	1
1.1.2 Τα βασικά σημεία της Εξόρυξης Διεργασιών	3
1.1.3 Αρχεία Γεγονότων (Event logs).....	5
1.1.4 Εξόρυξη Εκπαιδευτικών Διαδικασιών	9
1.1.5 Αλγόριθμοι για Εξόρυξη Διεργασιών	10
1.1.5.1 Συγκριτική αξιολόγηση και παραδείγματα χρήσης.....	17
1.2 Έλεγχος Συμμόρφωσης	21
1.2.1 Θεωρητικό υπόβαθρο	21
1.2.2 Μέθοδοι και Αλγόριθμοι για Έλεγχο Συμμόρφωσης	23
1.2.3 Μετρικές Ποιότητας.....	24
1.2.4 Προκλήσεις και Περιορισμοί.....	24
2. UML και UML Activity Diagram.....	26
2.1 Θεωρητικό υπόβαθρο (UML και UML Activity Diagram).....	26
2.2 Ρόλος στην Μοντελοποίηση Επιχειρησιακών Διαδικασιών.....	28
2.3 Το Λογισμικό Μοντελοποίησης Visual Paradigm	30
2.3.1 Τεχνικά χαρακτηριστικά και λειτουργικότητα του Visual Paradigm στο πλαίσιο της μοντελοποίησης συστημάτων	30
2.3.2 Η συμβολή του Visual Paradigm στη γέφυρωση μοντελοποίησης και εξόρυξης διεργασιών.....	31
2.4 Η χρήση των UML Activity Diagrams για την αναπαράσταση διαδικασιών	32
2.5 Σύνδεση των UML Activity Diagrams και της Εξόρυξης Διεργασιών.....	34
3. Εργαλεία Ανοικτού Κώδικα για Εξόρυξη Διεργασιών	45
3.1 Εισαγωγή	45
3.2 Το εργαλείο ProM	48
3.2.1 Περιγραφή του ProM	48
3.2.2 Αρχιτεκτονική και Λειτουργικότητα	52
3.2.3 Παραδείγματα εφαρμογής.....	54
3.3 Η βιβλιοθήκη PM4Py	70
3.3.1 Περιγραφή της βιβλιοθήκης	70

3.3.2 Βασικά Χαρακτηριστικά	72
3.3.3 Αρχιτεκτονική και Επεκτασιμότητα	74
3.4 Αξιολόγηση και κριτήρια επιλογής των εργαλείων εξόρυξης διεργασιών	76
4. Εικονικά Εργαστήρια	78
4.1 Εισαγωγή	78
4.2 Το εικονικό εργαστήριο OnLabs.....	79
4.2.1 Αρχιτεκτονική και Λειτουργικότητα του OnLabs	82
4.3 Η εφαρμογή της Εξόρυξης Διεργασιών και του Ελέγχου Συμμόρφωσης στην Εκπαιδευτική Διαδικασία	84
4.4 Event Logs στο OnLabs και Εφαρμογές Process Discovery	89
4.5 Διαδικασία Εξαγωγής Δεδομένων από το Εικονικό Εργαστήριο OnLabs.....	90
4.5.1 Εισαγωγή.....	90
4.5.2 Περιγραφή και Ροή της Πειραματικής Διαδικασίας.....	91
5. Μεθοδολογικό πλαίσιο, υλοποίηση και έλεγχος συμμόρφωσης της διεργασίας	106
5.1 Εισαγωγή.....	106
5.2 «Μικροσκοπική Παρατήρηση Δείγματος» στο ψηφιακό περιβάλλον OnLabs....	107
5.3: Αρχιτεκτονική και μεθοδολογική ροή υλοποίησης της πειραματικής διαδικασίας	111
5.4 Συλλογή δεδομένων και διαμόρφωση του ιστορικού συμβάντων (Event Log) ...	114
5.5 Ανακάλυψη διαδικασιών (ProcessDiscovery) στο περιβάλλον ProM	122
5.6 Σχεδιασμός του μοντέλου αναφοράς (Reference Model).....	128
5.7 Έλεγχος συμμόρφωσης (Conformance Checking) με χρήση της βιβλιοθήκης PM4Py	136
6. Τελικά συμπεράσματα και προοπτικές μελλοντικής έρευνας	146
Βιβλιογραφία	150
Παράρτημα Α: Κώδικας υλοποίησης Ελέγχου Συμμόρφωσης	155

Κατάλογος Εικόνων / Σχημάτων

Εικόνα 1 Επισκόπηση της Εξόρυξης Διεργασιών	1
Εικόνα 2 Εξόρυξη Διεργασιών = Επιστήμη Δεδομένων $\cap$ Επιστήμη Διεργασιών.....	2
Εικόνα 3 Τα βασικά σημεία του Process Mining	4
Εικόνα 4 Αρχείο καταγραφής συμβάντων (Event log).....	6
Εικόνα 5 Μοντέλο διεργασίας $\leftrightarrow$ Event log εικόνας 4.....	8
Εικόνα 6 Παράδειγμα αρχείου καταγραφής συμβάντων	13
Εικόνα 7 Παράδειγμα εφαρμογής του αλγόριθμου Alpha Miner.....	13
Εικόνα 8 Έλεγχος συμμόρφωσης: σύγκριση παρατηρούμενης συμπεριφοράς με μοντελοποιημένη συμπεριφορά.....	21
Εικόνα 9 UML AD για επαλήθευση κράτησης αεροπορικού εισιτηρίου	27
Εικόνα 10 Αρχικά στάδια του πειράματος μικροσκοπίας στο εικονικό εργαστήριο Onlabs.....	36
Εικόνα 11 Εισαγωγή του αρχείου καταγραφής συμβάντων στο ProM	39
Εικόνα 12 Μετατροπή .csv αρχείου σε μορφή .xes	40
Εικόνα 13 Εξαγωγή του μοντέλου διεργασίας με χρήση του αλγόριθμου Heuristics Miner	41
Εικόνα 14 Παραγωγή του μοντέλου διεργασίας με χρήση του Heuristics Miner	42
Εικόνα 15 Μετατροπή του Heuristics Net σε Petri Net	42
Εικόνα 16 Οπτικοποίηση του παραγόμενου Petri Net.....	43
Εικόνα 17 Δομή του Μοντέλου Αναφοράς σε Μορφή Petri Net	44
Εικόνα 18 Η ροή εργασιών της διπλωματικής εργασίας	48
Εικόνα 19 ProM: Αρχική σελίδα	49
Εικόνα 20 Συνοπτικά στοιχεία ενός αρχείου καταγραφής συμβάντων	50
Εικόνα 21 Επιλογή του αλγορίθμου Alpha Miner για την ανακάλυψη Μοντέλου Διεργασίας	51
Εικόνα 22 Μοντέλο διεργασιών με χρήση του αλγόριθμου Fuzzy Miner.....	53
Εικόνα 23 Αρχείο καταγραφής συμβάντων παραδείγματος	55
Εικόνα 24 Εισαγωγή του αρχείου καταγραφής συμβάντων στο ProM	56
Εικόνα 25 Δεδομένα των events με ‘Case ID’ = 2	57
Εικόνα 26 Συνοπτικά στοιχεία αρχείου καταγραφής συμβάντων	58
Εικόνα 27 Διεπαφή επιλογής του αλγορίθμου Heuristics Miner στο περιβάλλον ProM	61
Εικόνα 28 Παράθυρο ρύθμισης παραμέτρων του αλγορίθμου	62
Εικόνα 29 Heuristics Net – Μοντέλο διεργασίας από Heuristics Miner	63
Εικόνα 30 Επιλογή μετατροπής Heuristics net σε Petri net	64
Εικόνα 31 Τελικό μοντέλο διεργασίας σε μορφή δικτύου Petri net	65
Εικόνα 32 Επιλογή του αλγορίθμου Token-Based Replay για έλεγχο συμμόρφωσης.....	66
Εικόνα 33 Καθορισμός της τελικής σήμανσης του δικτύου Petri	67
Εικόνα 34 Παραμετροποίηση και επιλογή αλγορίθμου για την αναπαραγωγή των γεγονότων	68
Εικόνα 35 Καθορισμός κόστους αποκλίσεων για τη διαδικασία αναπαραγωγής	69
Εικόνα 36 Οπτικοποίηση των αποτελεσμάτων του Token-Based Replay επί του δικτύου Petri	70

Εικόνα 37 Παράδειγμα κώδικα με στόχο την ανακάλυψη διεργασιών (process discovery)	73
Εικόνα 38 Μοντέλο διεργασιών χρησιμοποιώντας τον αλγόριθμο Alpha Miner	73
Εικόνα 39 Εστίαση στο OnLabs – 3D Εικονικής Πραγματικότητας Εργαστήριο Βιολογίας.....	80
Εικόνα 40 Το Εργαστήριο OnLabs σε Λειτουργία Πειραματισμού.....	81
Εικόνα 41 Συνεδρία μέσω Skype για τη μελέτη μικροσκοπίας με χρήση περιβάλλοντος προσομοίωσης.....	83
Εικόνα 42 Η εφαρμογή της εξόρυξης διαδικασιών στον τομέα της εκπαίδευσης και ο κύκλος ζωής της.....	85
Εικόνα 43 Μελέτες Εξόρυξης Διαδικασιών ανά Τομέα στην Εκπαίδευση	88
Εικόνα 44 Εργαλεία για Εξόρυξη Διαδικασιών που χρησιμοποιούνται σε μελέτες Εκπαιδευτικής Εξόρυξης Διαδικασιών	89
Εικόνα 45 Ρύθμιση της έντασης του φωτισμού και προετοιμασία της οπτικής δέσμης.....	93
Εικόνα 46 Ρύθμιση του διαφράγματος της ίριδας για τον έλεγχο της αντίθεσης.....	94
Εικόνα 47 Τοποθέτηση του δείγματος στην αντικειμενοφόρο πλάκα.....	95
Εικόνα 48 Επικάλυψη του δείγματος με την καλυπτρίδα	96
Εικόνα 49 Τοποθέτηση και ευθυγράμμιση του παρασκευάσματος στην τράπεζα του μικροσκοπίου.....	97
Εικόνα 50 Μικρομετρική ρύθμιση και βελτιστοποίηση της εστίασης μέσω των προσοφθάλμιων φακών	98
Εικόνα 51 Εναλλαγή αντικειμενικών φακών μέσω της περιστρεφόμενης κεφαλής	99
Εικόνα 52 Μικρομετρική εστίαση του ειδώλου για τη μεγιστοποίηση της ευκρίνειας.....	100
Εικόνα 53 Έξοδος από το πεδίο μικροσκοπικής παρατήρησης και επιστροφή στον χώρο εργασίας.....	101
Εικόνα 54 Ολοκλήρωση πειραματικής διαδικασίας και απενεργοποίηση εργαστηριακού εξοπλισμού	102
Εικόνα 55 Εξαγωγή της δραστηριότητας χρήστη και αποθήκευση του αρχείου καταγραφής.....	103
Εικόνα 56 Απεικόνιση του παραγόμενου αρχείου καταγραφής (Event Log) στον φάκελο αποθήκευσης.....	104
Εικόνα 57 Στιγμιότυπο του αρχείου καταγραφής συμβάντων	105
Εικόνα 58 Η ροή εργασιών της διπλωματικής εργασίας	114
Εικόνα 59 Πρωτογενές αρχείο καταγραφής από το περιβάλλον OnLabs.....	116
Εικόνα 60 Προγραμματιστική επεξεργασία του αρχικού ιστορικού συμβάντων	119
Εικόνα 61 Δομή και περιεχόμενα του αρχείου αντιστοίχισης.....	120
Εικόνα 62 Δομή και περιεχόμενα του τελικού αρχείου καταγραφής συμβάντων.....	121
Εικόνα 63 Εισαγωγή του αρχείου καταγραφής συμβάντων στο περιβάλλον ProM	123
Εικόνα 64 Επιλογή του προσθέτου μετατροπής σε μορφή XES.....	124
Εικόνα 65 Επιλογή του αλγορίθμου Heuristic Miner για την ανακάλυψη μοντέλου.....	125
Εικόνα 66 Μερική απεικόνιση του παραγόμενου ευρετικού δικτύου	126
Εικόνα 67 Μετατροπή του ευρετικού δικτύου σε δίκτυο Petri	127
Εικόνα 68 Μερική απεικόνιση του παραγόμενου δικτύου Petri	128
Εικόνα 69 Ψηφιακή μοντελοποίηση της διεργασίας με χρήση του προτύπου BPMN	130
Εικόνα 70 Εισαγωγή του αρχείου BPMN στο λογισμικό ProM	131
Εικόνα 71 Επιτυχής επιβεβαίωση εισαγωγής του μοντέλου στο περιβάλλον εργασίας....	132

Εικόνα 72 Επιλογή του μοντέλου αναφοράς για την εκτέλεση ενέργειας.....	133
Εικόνα 73 Επιλογή του αλγορίθμου μετατροπής σε δίκτυο Petri	134
Εικόνα 74 Παραμετροποίηση των ιδιοτήτων μετάφρασης της διεργασίας	135
Εικόνα 75 Εξαγωγή και αποθήκευση του παραγόμενου δικτύου Petri	136
Εικόνα 76 Αποτελέσματα μετρικών και στατιστικών στοιχείων του ελέγχου συμμόρφωσης (Token-Based Replay).....	144

Κατάλογος Πινάκων

Πίνακας 1 Συγκριτικός Πίνακας Αξιολόγησης Αλγορίθμων Εξόρυξης Διεργασιών	19
--	----

Συντομογραφίες & Ακρωνύμια

XES	Extensible Event Stream
ProM	Process Mining Framework

PM4Py	Process Mining for Python
CSV	Comma-Separated Values
BPMN	Business Process Model and Notation
PNML	Petri Net Markup Language
XML	Extensible Markup Language
MXML	Mining Extensible Markup Language

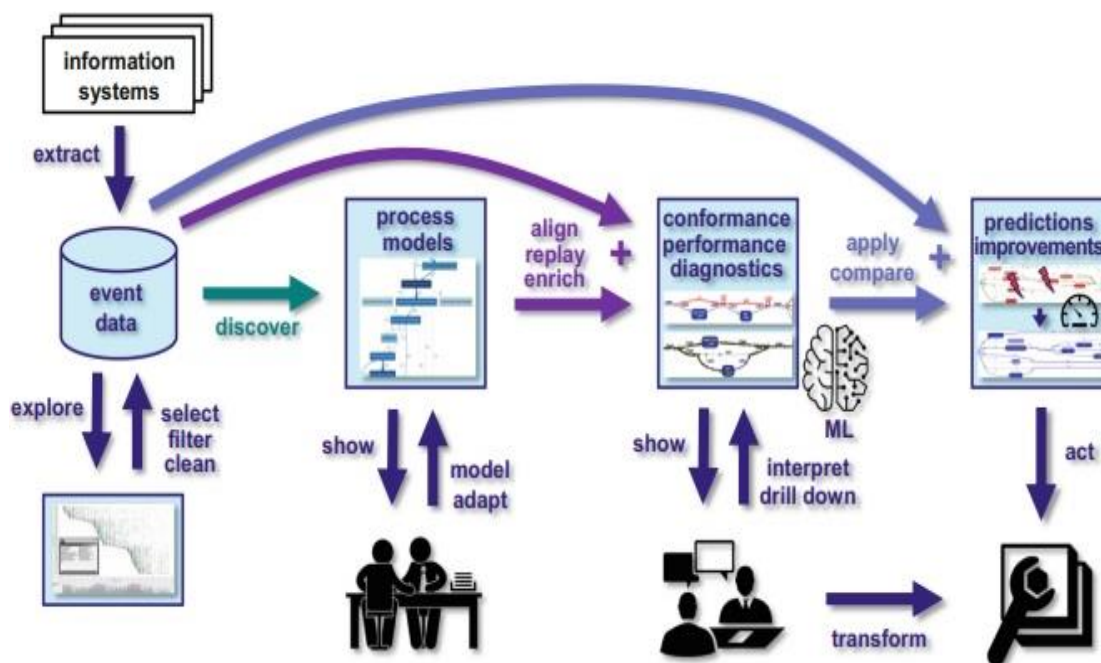
ΚΕΦΑΛΑΙΟ 1

Εξόρυξη Διεργασιών και Έλεγχος Συμμόρφωσης

1.1 Εξόρυξη Διεργασιών

1.1.1 Η έννοια της Εξόρυξης Διεργασιών

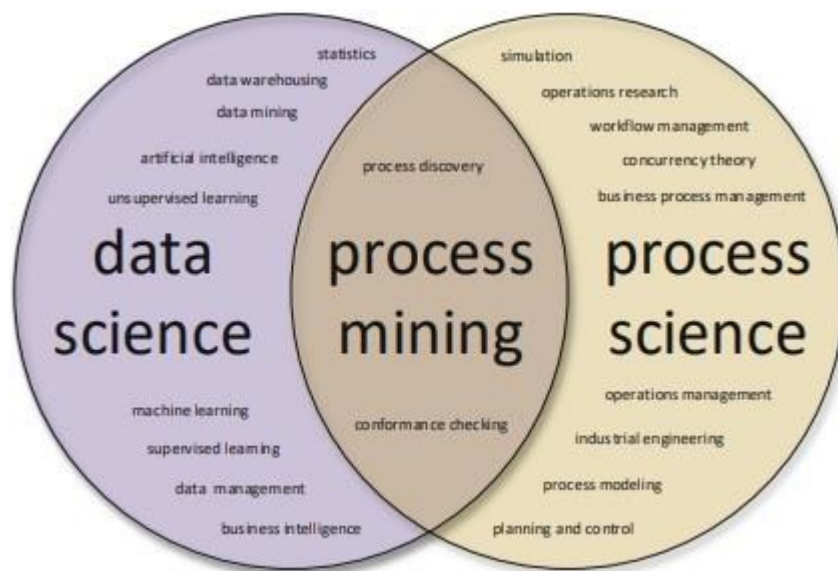
Η Εξόρυξη Διεργασιών (Process Mining) αποτελεί ένα από τα πλέον δυναμικά και αναπτυσσόμενα ερευνητικά πεδία στον χώρο της ανάλυσης δεδομένων και της διοίκησης πληροφοριακών συστημάτων (van der Aalst, 2011, 2016). Αναπτύχθηκε ως απάντηση στην αυξανόμενη ανάγκη των οργανισμών να κατανοήσουν σε βάθος τον τρόπο με τον οποίο εκτελούνται οι διαδικασίες τους στην πράξη και όχι απλώς όπως αυτές περιγράφονται σε θεωρητικά ή κανονιστικά μοντέλα. Η βασική καινοτομία της Εξόρυξης Διεργασιών έγκειται στη συστηματική αξιοποίηση των δεδομένων γεγονότων που καταγράφονται αυτόματα από πληροφοριακά συστήματα, όπως συστήματα ERP, CRM, LMS και συστήματα υγείας (Reinkemeyer, 2020; van der Aalst, 2016). Στην παρακάτω εικόνα [**Εικόνα 1**: van der Aalst (2016)], παρουσιάζεται μία επισκόπηση της Εξόρυξης Διεργασιών.



Εικόνα 1: Επισκόπηση της Εξόρυξης Διεργασιών

Σε αντίθεση με τις παραδοσιακές μεθόδους μοντελοποίησης επιχειρησιακών διαδικασιών, οι οποίες βασίζονται σε συνεντεύξεις με εμπλεκόμενους χρήστες, εργαστήρια μοντελοποίησης και εμπειρικές παρατηρήσεις, η Εξόρυξη Διεργασιών υιοθετεί μια δεδομενοκεντρική προσέγγιση (van der Aalst, 2011). Τα δεδομένα αυτά θεωρούνται αντικειμενικά, καθώς προκύπτουν από την πραγματική αλληλεπίδραση χρηστών και συστημάτων και όχι από υποκειμενικές εκτιμήσεις. Με τον τρόπο αυτό, μειώνεται ο κίνδυνος παραλείψεων ή στρεβλώσεων που συχνά χαρακτηρίζουν τις παραδοσιακές προσεγγίσεις.

Η έννοια της διαδικασίας στην Εξόρυξη Διεργασιών δεν περιορίζεται μόνο σε επιχειρησιακά περιβάλλοντα. Μπορεί να εφαρμοστεί σε κάθε επαναλαμβανόμενη ακολουθία δραστηριοτήτων που αφήνει ψηφιακά ίχνη. Παραδείγματα περιλαμβάνουν τη ροή μάθησης ενός φοιτητή σε ένα διαδικτυακό μάθημα, τη διαχείριση ενός ιατρικού περιστατικού σε νοσοκομείο ή τη διεκπεραίωση μιας διοικητικής αίτησης στο δημόσιο τομέα (Sypsas & Kalles, 2022; de Jong et al., 2013). Η ευελιξία αυτή καθιστά την Εξόρυξη Διεργασιών ιδιαίτερα ελκυστική για διεπιστημονική έρευνα. Η Εξόρυξη Διεργασιών, συνδυάζει την Επιστήμη των δεδομένων και την Επιστήμη των διεργασιών [**Εικόνα 2**: van der Aalst (2016)].



Εικόνα 2: Εξόρυξη Διεργασιών = Επιστήμη Δεδομένων $\cap$ Επιστήμη Διεργασιών

Κεντρικό ρόλο στην Εξόρυξη Διεργασιών διαδραματίζει η έννοια της περίπτωσης (case). Η

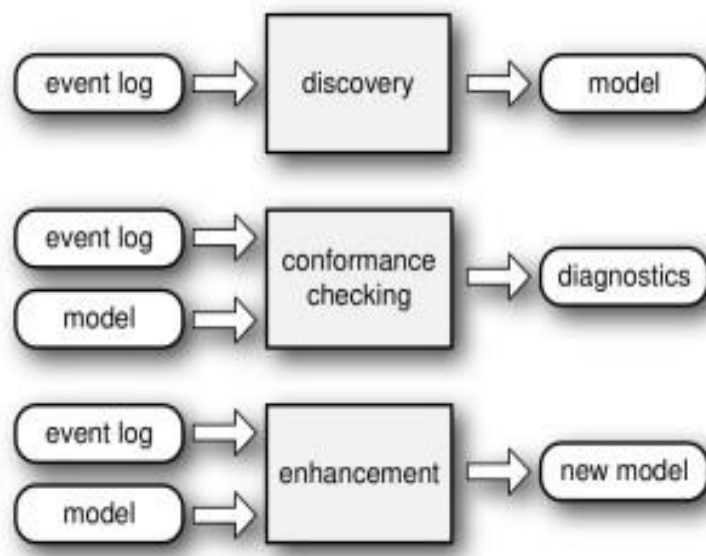
περίπτωση αντιπροσωπεύει μια μοναδική εκτέλεση της διαδικασίας και αποτελεί το σημείο αναφοράς για την οργάνωση των γεγονότων. Κάθε περίπτωση ακολουθεί μια συγκεκριμένη πορεία μέσα στη διαδικασία, η οποία μπορεί να διαφέρει σημαντικά από άλλες περιπτώσεις. Η ανάλυση αυτών των διαφοροποιήσεων επιτρέπει την κατανόηση της πολυπλοκότητας και της ευελιξίας των πραγματικών διαδικασιών (van der Aalst, 2011).

Επιπλέον, η Εξόρυξη Διεργασιών δεν περιορίζεται στην απλή απεικόνιση της ροής δραστηριοτήτων. Παρέχει τη δυνατότητα ποσοτικής ανάλυσης, όπως ο υπολογισμός χρόνων εκτέλεσης, καθυστερήσεων, συχνοτήτων και επιπέδων απόδοσης. Με τον τρόπο αυτό, συνδέει τη δομική ανάλυση των διαδικασιών με δείκτες απόδοσης (Key Performance Indicators – KPIs), ενισχύοντας τη στρατηγική λήψη αποφάσεων (Augusto et al., 2019).

Σε θεωρητικό επίπεδο, η Εξόρυξη Διεργασιών τοποθετείται ανάμεσα στην Εξόρυξη Δεδομένων (Data Mining) και τη Μοντελοποίηση Διαδικασιών. Ενώ η Εξόρυξη Δεδομένων επικεντρώνεται στην ανακάλυψη προτύπων χωρίς απαραίτητα χρονική ή ακολουθιακή διάσταση, η Εξόρυξη Διεργασιών δίνει έμφαση στη σειρά και τη χρονική αλληλουχία των γεγονότων. Αυτή η διάσταση του χρόνου είναι κρίσιμη για την κατανόηση της λειτουργικής συμπεριφοράς των διαδικασιών (van der Aalst, 2016).

1.1.2 Τα βασικά σημεία της Εξόρυξης Διεργασιών

Η θεωρητική και πρακτική βάση της Εξόρυξης Διεργασιών δομείται γύρω από τρεις βασικούς πυλώνες: την ανακάλυψη διαδικασιών, τον έλεγχο συμμόρφωσης και τη βελτίωση διαδικασιών (van der Aalst, 2011, 2016) [**Εικόνα 3**: van der Aalst, 2011, 2016]. Οι πυλώνες αυτοί δεν λειτουργούν ανεξάρτητα, αλλά αλληλοσυμπληρώνονται, δημιουργώντας ένα ολοκληρωμένο πλαίσιο ανάλυσης και βελτιστοποίησης.



Εικόνα 3: Τα βασικά σημεία του Process Mining

Η ανακάλυψη διαδικασιών αποτελεί συνήθως το πρώτο στάδιο εφαρμογής της Εξόρυξης Διεργασιών. Στο στάδιο αυτό, ο αναλυτής δεν διαθέτει προϋπάρχον μοντέλο ή επιθυμεί να ελέγξει αν το υφιστάμενο μοντέλο ανταποκρίνεται στην πραγματικότητα. Μέσω ειδικών αλγορίθμων, παράγεται ένα μοντέλο που απεικονίζει τη ροή των δραστηριοτήτων όπως αυτή προκύπτει από τα δεδομένα (Augusto et al., 2019). Το μοντέλο αυτό μπορεί να λάβει διάφορες μορφές, όπως Petri Nets ή BPMN διαγράμματα.

Ένα από τα βασικά ζητήματα στην ανακάλυψη διαδικασιών είναι η ισορροπία μεταξύ ακρίβειας και απλότητας. Ένα υπερβολικά λεπτομερές μοντέλο μπορεί να είναι δύσκολο στην κατανόηση και τη χρήση, ενώ ένα υπεραπλουστευμένο μοντέλο ενδέχεται να αποκρύπτει κρίσιμες πληροφορίες. Η επιλογή του κατάλληλου επιπέδου αφαίρεσης αποτελεί σημαντική πρόκληση και συχνά εξαρτάται από τον σκοπό της ανάλυσης (Reinkemeyer, 2020).

Ο έλεγχος συμμόρφωσης έρχεται να απαντήσει στο ερώτημα κατά πόσο η πραγματική εκτέλεση μιας διαδικασίας ευθυγραμμίζεται με ένα προκαθορισμένο μοντέλο ή σύνολο κανόνων. Στο πλαίσιο αυτό, εντοπίζονται αποκλίσεις, όπως παραλείψεις δραστηριοτήτων, αλλαγές στη σειρά εκτέλεσης ή μη επιτρεπόμενες επαναλήψεις (Rozinat & van der Aalst, 2008; Carmona et al., 2018). Οι αποκλίσεις αυτές δεν θεωρούνται απαραίτητα αρνητικές,

καθώς σε ορισμένες περιπτώσεις μπορεί να υποδηλώνουν αναγκαία ευελιξία ή προσαρμογή.

Η βελτίωση διαδικασιών αξιοποιεί τα ευρήματα των προηγούμενων σταδίων για την ενίσχυση της αποδοτικότητας και της ποιότητας. Μέσω της ανάλυσης χρόνων αναμονής, σημείων συμφόρησης και κατανομής πόρων, οι οργανισμοί μπορούν να προτείνουν στοχευμένες παρεμβάσεις (van der Aalst, 2016). Η Εξόρυξη Διεργασιών, επομένως, δεν αποτελεί απλώς εργαλείο ανάλυσης, αλλά και στρατηγικό μηχανισμό υποστήριξης αλλαγών.

Ιδιαίτερη σημασία έχει και η έννοια της συνεχούς βελτίωσης. Σε αντίθεση με στατικές μελέτες διαδικασιών, η Εξόρυξη Διεργασιών μπορεί να εφαρμοστεί επαναλαμβανόμενα, επιτρέποντας τη συνεχή παρακολούθηση της απόδοσης και την έγκαιρη ανίχνευση προβλημάτων. Αυτό καθιστά την προσέγγιση ιδιαίτερα συμβατή με σύγχρονες μεθοδολογίες διοίκησης ποιότητας (Reinkemeyer, 2020).

1.1.3 Αρχεία Γεγονότων (Event logs)

Τα αρχεία γεγονότων (event logs) αποτελούν τον ακρογωνιαίο λίθο της Εξόρυξης Διεργασιών, καθώς συνιστούν την πρωτογενή πηγή δεδομένων πάνω στην οποία βασίζονται όλες οι αναλύσεις. Χωρίς τη διαθεσιμότητα αξιόπιστων και επαρκώς δομημένων αρχείων γεγονότων, η εφαρμογή τεχνικών Εξόρυξης Διεργασιών καθίσταται είτε αδύνατη είτε οδηγεί σε αποτελέσματα χαμηλής εγκυρότητας (Reinkemeyer, 2020). Τα αρχεία αυτά προκύπτουν από την αυτόματη καταγραφή δραστηριοτήτων σε πληροφοριακά συστήματα και αποτυπώνουν τη δυναμική εκτέλεση των διαδικασιών σε πραγματικό χρόνο.

Ένα αρχείο γεγονότων αποτελείται από ένα σύνολο περιπτώσεων (cases), όπου κάθε περίπτωση αντιπροσωπεύει μία μοναδική εκτέλεση της διαδικασίας. Για παράδειγμα, σε ένα εκπαιδευτικό περιβάλλον, μία περίπτωση μπορεί να αντιστοιχεί σε έναν φοιτητή και την πορεία του μέσα σε ένα μάθημα, ενώ σε ένα επιχειρησιακό περιβάλλον μπορεί να αφορά την επεξεργασία μιας παραγγελίας (Sypsas & Kalles, 2022). Κάθε περίπτωση περιλαμβάνει μία ακολουθία γεγονότων, τα οποία είναι χρονικά διατεταγμένα.

Κάθε γεγονός χαρακτηρίζεται από ένα ελάχιστο σύνολο υποχρεωτικών γνωρισμάτων. Το σημαντικότερο είναι το αναγνωριστικό της περίπτωσης, το οποίο επιτρέπει την ομαδοποίηση των γεγονότων. Εξίσου κρίσιμη είναι η δραστηριότητα (activity), η οποία

περιγράφει τι ακριβώς εκτελέστηκε, καθώς και η χρονική σήμανση (timestamp), η οποία καθορίζει πότε έλαβε χώρα το γεγονός. Η χρονική πληροφορία είναι θεμελιώδης, καθώς επιτρέπει την ανάλυση της ακολουθίας, της διάρκειας και των καθυστερήσεων μεταξύ δραστηριοτήτων (van der Aalst, 2016). Στην παρακάτω εικόνα, απεικονίζεται ένα παράδειγμα ενός αρχείου γεγονότων [**Εικόνα 4**: DataMiningApps, n.d.]



Case	Time	Activity	Extra attributes ...
1	2015-04-01 14:21:42	Form Filling	Handling analyst: Tom
1	2015-04-02 12:15:54	Risk Analysis	
2	2015-04-02 12:32:11	Form Filling	Handling analyst: Susan
1	2015-04-04 15:46:06	Append Additional Details	Field changed: F203
2	2015-04-05 10:06:56	Financial History Check	
1	2015-04-05 11:14:16	Append Additional Details	Field changed: F101
1	2015-04-11 09:11:03	Financial History Check	
1	2015-04-11 16:08:13	Append Additional Details	Field changed: F204
2	2015-04-14 14:36:33	Risk Analysis	
2	2015-04-16 13:10:12	Managerial Decision	Outcome: approved
1	2015-04-20 11:56:41	Managerial Decision	Outcome: rejected

Εικόνα 4: Αρχείο καταγραφής συμβάντων (Event log)

Πέρα από τα βασικά αυτά γνωρίσματα, τα αρχεία γεγονότων συχνά περιλαμβάνουν και πρόσθετες πληροφορίες, γνωστές ως επεκτάσεις ή χαρακτηριστικά (attributes). Τέτοια χαρακτηριστικά μπορεί να αφορούν αυτόν που υλοποιεί την δραστηριότητα (π.χ. χρήστης, εκπαιδευτικός, υπάλληλος), το κόστος, το χρησιμοποιούμενο σύστημα ή το αποτέλεσμα της δραστηριότητας. Η ύπαρξη αυτών των εμπλουτισμένων δεδομένων επιτρέπει την εφαρμογή πιο προηγμένων τεχνικών ανάλυσης, όπως η ανάλυση πόρων και η πολυδιάστατη αξιολόγηση απόδοσης (van der Aalst, 2016).

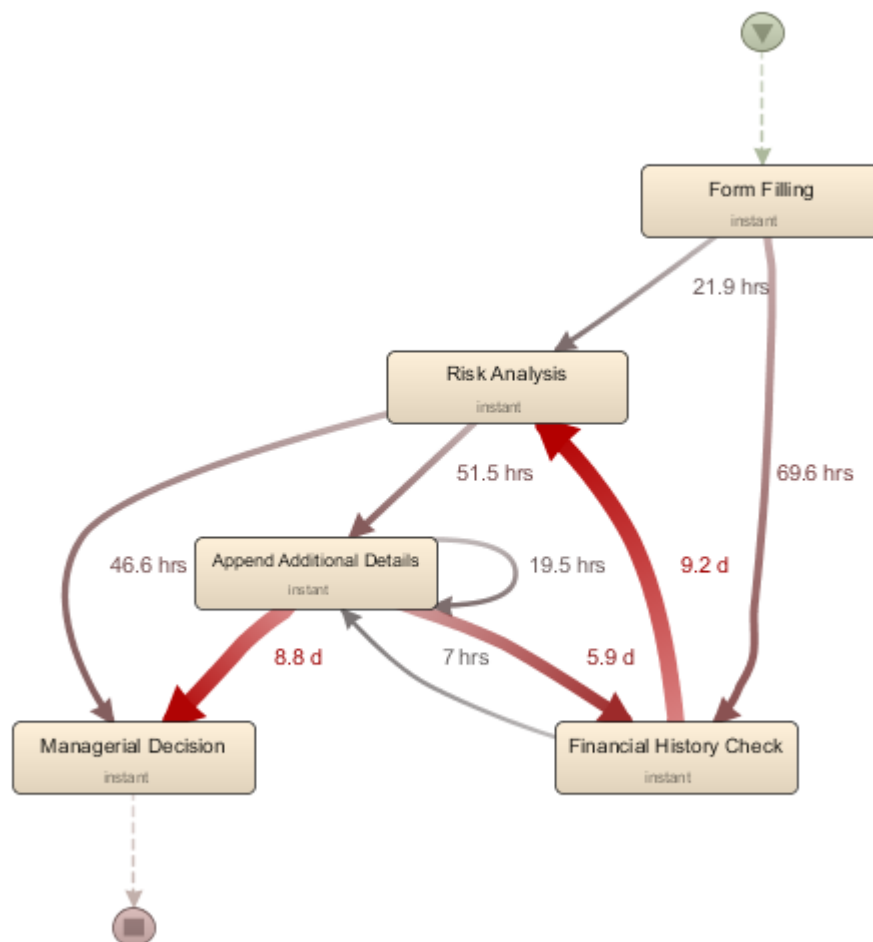
Η δομή και η διάρθρωση των αρχείων γεγονότων παίζουν καθοριστικό ρόλο στη

δυνατότητα επεξεργασίας τους. Ένα από τα πιο διαδεδομένα πρότυπα είναι το XES (eXtensible Event Stream), το οποίο έχει σχεδιαστεί ειδικά για εφαρμογές Εξόρυξης Διεργασιών. Το XES επιτρέπει την τυποποιημένη αναπαράσταση γεγονότων και περιπτώσεων, διασφαλίζοντας τη διαλειτουργικότητα μεταξύ διαφορετικών εργαλείων ανάλυσης (Berti et al., 2019). Παράλληλα, χρησιμοποιούνται και απλούστερες μορφές, όπως αρχεία CSV, τα οποία όμως απαιτούν πρόσθετη προεπεξεργασία.

Ένα από τα σημαντικότερα ζητήματα που σχετίζονται με τα αρχεία γεγονότων είναι η ποιότητα των δεδομένων. Στην πράξη, τα δεδομένα που καταγράφονται από πληροφοριακά συστήματα συχνά περιέχουν θόρυβο, ελλείψεις ή ασυνέπειες. Για παράδειγμα, ενδέχεται να λείπουν χρονικές σημάνσεις, να υπάρχουν διπλότυπα γεγονότα ή να καταγράφονται δραστηριότητες με διαφορετικές ονομασίες παρότι αναφέρονται στην ίδια ενέργεια. Τα προβλήματα αυτά μπορούν να επηρεάσουν σημαντικά τα παραγόμενα μοντέλα διαδικασιών (Augusto et al., 2019).

Για τον λόγο αυτό, η προεπεξεργασία των αρχείων γεγονότων αποτελεί κρίσιμο στάδιο στην Εξόρυξη Διεργασιών. Η προεπεξεργασία μπορεί να περιλαμβάνει τον καθαρισμό δεδομένων, την ενοποίηση δραστηριοτήτων, την αφαίρεση ακραίων περιπτώσεων και το φιλτράρισμα σπάνιων συμπεριφορών. Μέσω αυτών των ενεργειών, επιδιώκεται η δημιουργία ενός αντιπροσωπευτικού και αξιόπιστου συνόλου δεδομένων (Reinkemeyer, 2020; van der Aalst, 2016).

Ιδιαίτερη πρόκληση αποτελεί και η επιλογή του κατάλληλου επιπέδου ανάλυσης. Σε ορισμένες περιπτώσεις, η λεπτομερής καταγραφή όλων των γεγονότων μπορεί να οδηγήσει σε υπερβολικά πολύπλοκα μοντέλα, τα οποία είναι δύσκολο να ερμηνευθούν. Αντίθετα, η υπερβολική αφαίρεση ενδέχεται να αποκρύψει κρίσιμες πληροφορίες (van der Aalst, 2016). Η ισορροπία μεταξύ λεπτομέρειας και αναγνωσιμότητας αποτελεί βασικό αντικείμενο μελέτης στο πεδίο. Η παρακάτω εικόνα, απεικονίζει το μοντέλο διεργασίας που προκύπτει με βάση το παραπάνω αρχείο καταγραφής συμβάντων [**Εικόνα 5**: DataMiningApps, n.d.]



Εικόνα 5: Μοντέλο διεργασίας ↔ Eventlog εικόνας 4

Τέλος, πρέπει να ληφθούν υπόψη και ζητήματα δεοντολογίας και προστασίας προσωπικών δεδομένων. Τα αρχεία γεγονότων συχνά περιέχουν ευαίσθητες πληροφορίες, ιδιαίτερα σε τομείς όπως η εκπαίδευση και η υγεία. Η ανωνυμοποίηση δεδομένων και η συμμόρφωση με κανονισμούς προστασίας προσωπικών δεδομένων αποτελούν απαραίτητες προϋποθέσεις για την υπεύθυνη εφαρμογή της Εξόρυξης Διεργασιών (Sypsas & Kalles, 2022).

Στο πλαίσιο της παρούσας εργασίας, η θεωρητική ανάλυση των αρχείων γεγονότων που προηγήθηκε αποτελεί το αναγκαίο υπόβαθρο για τη μετάβαση στο εφαρμοσμένο σκέλος της έρευνας. Ενώ τα event logs εξετάστηκαν εδώ ως στατικές δομές δεδομένων, στα επόμενα κεφάλαια αποκτούν λειτουργική οντότητα, καθώς συνιστούν το ψηφιακό αποτύπωμα των χρηστών στο εικονικό εργαστήριο OnLabs. Η ορθή δόμηση αυτών των αρχείων, με έμφαση

στην ακρίβεια των χρονικών σημάνσεων και τη σαφήνεια των δραστηριοτήτων, είναι αυτή που επιτρέπει την αποτελεσματική εφαρμογή των αλγορίθμων ανακάλυψης (process discovery).

Η κρισιμότητα των logs αναδεικνύεται περαιτέρω κατά τη διαδικασία του ελέγχου συμμόρφωσης (conformance checking), όπου η πραγματική συμπεριφορά, όπως αυτή καταγράφεται στα αρχεία γεγονότων, αντιπαραβάλλεται με τα κανονιστικά μοντέλα UML που έχουν σχεδιαστεί. Η ανάλυση αυτή δεν στοχεύει απλώς στην επιβεβαίωση της εκτέλεσης των πειραμάτων, αλλά στην ανάδειξη των ποιοτικών χαρακτηριστικών της μάθησης, όπως η πιστή τήρηση του πρωτοκόλλου ή η εμφάνιση αποκλίνουσας συμπεριφοράς. Συνεπώς, η διασφάλιση της ποιότητας των δεδομένων και η επιλογή του κατάλληλου επιπέδου αφαίρεσης, ζητήματα που συζητήθηκαν ανωτέρω, αποτελούν καθοριστικούς παράγοντες για την εγκυρότητα των συμπερασμάτων που θα ακολουθήσουν σχετικά με την εκπαιδευτική αξία του OnLabs.

1.1.4 Εξόρυξη Εκπαιδευτικών Διαδικασιών

Η Εξόρυξη Εκπαιδευτικών Διαδικασιών (Educational Process Mining) αποτελεί εξειδικευμένη εφαρμογή της Εξόρυξης Διεργασιών στον τομέα της εκπαίδευσης και εντάσσεται στο ευρύτερο πεδίο της Εκπαιδευτικής Ανάλυσης Δεδομένων (Learning Analytics) (Sypsas & Kalles, 2022; Baker & Inventado, 2014). Στόχος της είναι η κατανόηση, ανάλυση και βελτίωση των μαθησιακών και διοικητικών διαδικασιών που λαμβάνουν χώρα σε εκπαιδευτικά περιβάλλοντα, τόσο δια ζώσης όσο και εξ αποστάσεως. Η ραγδαία ανάπτυξη των ψηφιακών εκπαιδευτικών πλατφορμών έχει δημιουργήσει τεράστιους όγκους δεδομένων, οι οποίοι μπορούν να αξιοποιηθούν για την εξαγωγή χρήσιμων συμπερασμάτων σχετικά με τη μαθησιακή συμπεριφορά (Tsoni, Stavropoulos, & Verykios, 2019).

Τα συστήματα διαχείρισης μάθησης (Learning Management Systems – LMS), όπως το Moodle και το Blackboard, καταγράφουν λεπτομερώς τις ενέργειες των εκπαιδευόμενων και των εκπαιδευτικών. Κάθε σύνδεση, προβολή εκπαιδευτικού υλικού, υποβολή εργασίας ή συμμετοχή σε δραστηριότητα δημιουργεί γεγονότα που μπορούν να αναλυθούν με τεχνικές Εξόρυξης Διεργασιών (Sypsas & Kalles, 2023). Μέσω αυτών των δεδομένων, είναι δυνατή

η ανακατασκευή των πραγματικών μαθησιακών διαδρομών που ακολουθούν οι φοιτητές.

Ένα από τα βασικά πλεονεκτήματα της Εξόρυξης Εκπαιδευτικών Διαδικασιών είναι η δυνατότητα εντοπισμού προτύπων μάθησης. Οι εκπαιδευόμενοι δεν ακολουθούν πάντα τη γραμμική πορεία που έχει σχεδιαστεί από τον εκπαιδευτικό. Αντίθετα, συχνά παρατηρούνται παρακάμψεις, επαναλήψεις ή εγκατάλειψη δραστηριοτήτων. Η ανάλυση αυτών των αποκλίσεων επιτρέπει την καλύτερη κατανόηση των μαθησιακών δυσκολιών και των σημείων όπου απαιτείται παιδαγωγική παρέμβαση (Sypsas & Kalles, 2022; Sypsas & Kalles, 2020).

Ιδιαίτερα σημαντική είναι η συμβολή της Εξόρυξης Εκπαιδευτικών Διαδικασιών στην πρόβλεψη της μαθησιακής επίδοσης και της εγκατάλειψης σπουδών. Μέσω της ανάλυσης των αρχείων γεγονότων, μπορούν να εντοπιστούν πρώιμα σημάδια μειωμένης εμπλοκής, όπως καθυστερήσεις στην υποβολή εργασιών ή μειωμένη αλληλεπίδραση με το εκπαιδευτικό υλικό (Wang et al., 2019). Τα ευρήματα αυτά μπορούν να αξιοποιηθούν για την παροχή εξατομικευμένης υποστήριξης.

Πέρα από τη μαθησιακή διάσταση, η Εξόρυξη Εκπαιδευτικών Διαδικασιών εφαρμόζεται και στη διοικητική λειτουργία των εκπαιδευτικών ιδρυμάτων. Διαδικασίες όπως οι εγγραφές, η αξιολόγηση μαθημάτων και η διαχείριση εξετάσεων μπορούν να αναλυθούν προκειμένου να εντοπιστούν καθυστερήσεις και αναποτελεσματικότητες. Με τον τρόπο αυτό, συμβάλλει στη βελτίωση της συνολικής ποιότητας των εκπαιδευτικών υπηρεσιών (Cairns et al., 2015).

Ωστόσο, η εφαρμογή της Εξόρυξης Διεργασιών στην εκπαίδευση συνοδεύεται από σημαντικές προκλήσεις. Ζητήματα προστασίας προσωπικών δεδομένων, διαφάνειας και ηθικής χρήσης των δεδομένων αποτελούν κρίσιμα σημεία προβληματισμού (Bogarín et al., 2018). Η ανάλυση μαθησιακών δεδομένων πρέπει να γίνεται με σεβασμό στην ιδιωτικότητα των εκπαιδευόμενων και σύμφωνα με το ισχύον θεσμικό πλαίσιο.

1.1.5 Αλγόριθμοι για Εξόρυξη Διεργασιών

Οι αλγόριθμοι Εξόρυξης Διεργασιών αποτελούν τον πυρήνα της αυτόματης ανάλυσης διαδικασιών, καθώς είναι υπεύθυνοι για τη μετατροπή των αρχείων γεγονότων σε κατανοητά και αναλύσιμα μοντέλα (van der Aalst, 2016). Η επιλογή του κατάλληλου αλγορίθμου επηρεάζει καθοριστικά την ποιότητα, την ακρίβεια και τη χρηστικότητα των

παραγόμενων αποτελεσμάτων. Κατά την εξέλιξη του πεδίου, έχουν προταθεί πολυάριθμοι αλγόριθμοι, καθένας με διαφορετικά χαρακτηριστικά και περιορισμούς.

Ο αλγόριθμος AlphaMiner, αποτελεί έναν από τους πρώτους και πιο θεμελιώδεις αλγορίθμους στον χώρο της Εξόρυξης Διεργασιών. Εισήγαγε βασικές έννοιες, όπως η αιτιότητα, η παράλληλη εκτέλεση και η επιλογή μεταξύ δραστηριοτήτων. Παρότι η συμβολή του υπήρξε καθοριστική για τη θεμελίωση του πεδίου, παρουσιάζει σημαντικούς περιορισμούς, ιδιαίτερα στη διαχείριση θορύβου και πολύπλοκων δομών (van der Aalst et al., 2011).

Για την αντιμετώπιση αυτών των περιορισμών, αναπτύχθηκε ο αλγόριθμος Heuristic Miner, ο οποίος λαμβάνει υπόψη τη συχνότητα εμφάνισης των ακολουθιών δραστηριοτήτων. Με τον τρόπο αυτό, είναι σε θέση να φιλτράρει σπάνιες ή θορυβώδεις συμπεριφορές, παράγοντας πιο σταθερά και αντιπροσωπευτικά μοντέλα (Weijters & Ribeiro, 2011). Ο Heuristic Miner είναι ιδιαίτερα χρήσιμος σε πραγματικά δεδομένα, τα οποία σπάνια είναι πλήρως καθαρά.

Ο αλγόριθμος Inductive Miner αποτελεί μια πιο σύγχρονη προσέγγιση, η οποία εγγυάται τη δημιουργία ορθών και κατανοητών μοντέλων. Βασίζεται σε μια αναδρομική διαδικασία διαχωρισμού των δεδομένων και παράγει μοντέλα που είναι πάντα δομημένα και χωρίς αδιέξοδα (Leemans et al., 2013). Η ιδιότητα αυτή καθιστά τον Inductive Miner ιδιαίτερα δημοφιλή σε ακαδημαϊκές και πρακτικές εφαρμογές.

Πέρα από τους αλγορίθμους ανακάλυψης, σημαντική θέση κατέχουν και οι αλγόριθμοι ελέγχου συμμόρφωσης, οι οποίοι συγκρίνουν τα δεδομένα γεγονότων με υπάρχοντα μοντέλα. Μέσω αυτών, μπορούν να εντοπιστούν αποκλίσεις και να ποσοτικοποιηθεί ο βαθμός συμμόρφωσης της διαδικασίας (van der Aalst, 2016). Οι τεχνικές αυτές είναι ιδιαίτερα χρήσιμες σε περιβάλλοντα με αυστηρές κανονιστικές απαιτήσεις.

Η επιλογή του κατάλληλου αλγορίθμου δεν αποτελεί απλή διαδικασία και εξαρτάται από πολλούς παράγοντες, όπως το μέγεθος του αρχείου γεγονότων, ο βαθμός θορύβου, το επιθυμητό επίπεδο λεπτομέρειας και ο σκοπός της ανάλυσης (van der Aalst et al., 2011). Στο πλαίσιο μιας διπλωματικής εργασίας, η συγκριτική αξιολόγηση διαφορετικών αλγορίθμων μπορεί να αποτελέσει σημαντικό στοιχείο επιστημονικής συμβολής.

Παρακάτω, αναλύονται μερικοί από τους αλγόριθμους για Εξόρυξη Διεργασιών.

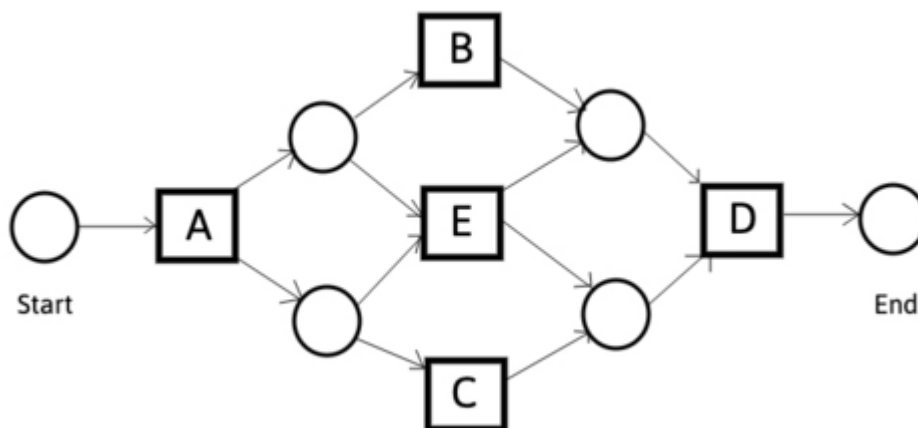
Alpha Miner

Ο αλγόριθμος Alpha Miner αποτελεί έναν από τους πρώτους και πλέον γνωστούς αλγόριθμους στο πεδίο του Process Mining, ο οποίος σχεδιάστηκε για την ανακάλυψη επιχειρησιακών διαδικασιών από αρχεία καταγραφής γεγονότων (event logs) (van der Aalst, 2011). Ο βασικός σκοπός του Alpha Miner είναι η κατασκευή ενός τυπικού μοντέλου διαδικασίας, συνήθως σε μορφή Petri net, το οποίο να αναπαριστά τις σχέσεις μεταξύ των δραστηριοτήτων που παρατηρούνται στα δεδομένα. Ο αλγόριθμος ξεκινά από το δεδομένο event log, όπου κάθε εγγραφή αντιστοιχεί σε μια εκτέλεση μιας διαδικασίας (case) και περιλαμβάνει τη σειρά των δραστηριοτήτων που εκτελέστηκαν. Αρχικά, ο Alpha Miner εντοπίζει τις μοναδικές δραστηριότητες και εξετάζει τις ζεύξεις δραστηριοτήτων, δηλαδή ποια ενέργεια ακολουθεί ποια σε ένα case, προκειμένου να αποκαλύψει άμεσες διαδοχικές σχέσεις. Στη συνέχεια, καθορίζει τις σχέσεις αιτιακής εξάρτησης, αλληλοαποκλεισμού και παράλληλης εκτέλεσης, διακρίνοντας έτσι μεταξύ των δραστηριοτήτων που μπορούν να εκτελεστούν ταυτόχρονα και εκείνων που έχουν σειριακή εξάρτηση (van der Aalst, 1998). Ο αλγόριθμος κατασκευάζει επίσης σύνολα εισόδου και εξόδου (input/output places) για το Petri net, τα οποία εξασφαλίζουν ότι όλες οι ροές ξεκινούν από το αρχικό σημείο της διαδικασίας και καταλήγουν στο τελικό, διατηρώντας τη συνέπεια και την ορθότητα του μοντέλου. Η δύναμη του Alpha Miner έγκειται στην ικανότητά του να παράγει ένα πλήρες και συνεπές μοντέλο με βάση μόνο τα δεδομένα, χωρίς να απαιτεί προηγούμενη γνώση της διαδικασίας. Ο Alpha Miner παραμένει ένα σημαντικό εργαλείο στην ιστορική και θεωρητική εξέλιξη του Process Mining, αποτελώντας τη βάση για πιο εξελιγμένους αλγόριθμους, όπως οι Heuristic Miner και Fuzzy Miner [van der Aalst et al., 2007], οι οποίοι προσπαθούν να αντιμετωπίσουν τα όρια του αρχικού αλγορίθμου και να διαχειριστούν καλύτερα τα θορυβώδη και μη δομημένα δεδομένα.

Στην παρακάτω εικόνα (**Εικόνα 7**) [Wikipedia, n.d.], βλέπουμε το αποτέλεσμα (δημιουργία μοντέλου διεργασίας) από την εφαρμογή του αλγόριθμου με βάση το αρχείο καταγραφής συμβάντων (**Εικόνα 6**) [Wikipedia, n.d.].

Case id	Activity	Timestamp
1	A	2019/10/09 14:50:17.000
1	B	2019/10/09 15:50:17.000
1	C	2019/10/09 15:56:17.000
1	D	2019/10/10 13:50:17.000
2	A	2019/10/10 14:30:17.000
2	C	2019/10/10 14:50:14.000
2	B	2019/10/11 09:50:17.000
2	D	2019/10/11 10:50:17.000
3	A	2019/10/11 12:50:17.000
3	E	2019/10/21 14:50:17.000
3	D	2019/10/21 15:50:17.000

Εικόνα 6: Παράδειγμα αρχείου καταγραφής συμβάντων



Εικόνα 7: Παράδειγμα εφαρμογής του αλγόριθμου AlphaMiner

Heuristic Miner

Ο αλγόριθμος Heuristic Miner αποτελεί μια εξέλιξη του κλασικού Alpha Miner και έχει σχεδιαστεί για να αντιμετωπίζει πρακτικές προκλήσεις που προκύπτουν κατά την εξόρυξη διαδικασιών από πραγματικά δεδομένα, τα οποία συχνά είναι ασυνεπή, ατελή ή περιέχουν

θόρυβο [Weijters & Ribeiro, 2011]. Σε αντίθεση με τον Alpha Miner, ο οποίος βασίζεται σε αυστηρούς λογικούς κανόνες για την αναγνώριση των σχέσεων μεταξύ δραστηριοτήτων, ο Heuristic Miner χρησιμοποιεί στατιστικά κριτήρια και πιθανότητες συσχέτισης ώστε να καθορίσει τις συνδέσεις μεταξύ γεγονότων. Ουσιαστικά, εφαρμόζει μια ευρετική προσέγγιση, λαμβάνοντας υπόψη τη συχνότητα εμφάνισης και τη σχετική βαρύτητα των ακολουθιών δραστηριοτήτων μέσα στα event logs. Έτσι, αγνοεί σπάνιες ή μη αντιπροσωπευτικές περιπτώσεις, δημιουργώντας μοντέλα που αντανακλούν καλύτερα τη συνηθισμένη συμπεριφορά των διαδικασιών. Ο αλγόριθμος κατασκευάζει ένα γράφημα εξαρτήσεων (dependency graph), όπου κάθε κόμβος αντιπροσωπεύει μια δραστηριότητα και κάθε ακμή φέρει έναν δείκτη ισχύος της σχέσης μεταξύ δύο δραστηριοτήτων. Ανάλογα με την τιμή αυτού του δείκτη, αποφασίζεται αν η σχέση θα περιληφθεί στο τελικό μοντέλο. Στη συνέχεια, το γράφημα αυτό μετατρέπεται σε μοντέλο ροής εργασίας (workflow model), το οποίο μπορεί να απεικονιστεί με τη μορφή Petri net ή Causal net [Weijters et al., 2006]. Ο Heuristic Miner είναι ιδιαίτερα χρήσιμος για την ανάλυση πραγματικών επιχειρησιακών διαδικασιών, καθώς μπορεί να χειριστεί πολύπλοκα μονοπάτια, παράλληλες δραστηριότητες και επαναλήψεις. Ωστόσο, η ακρίβειά του εξαρτάται από την ποιότητα και τον όγκο των δεδομένων: σε περιπτώσεις με περιορισμένα ή μη αντιπροσωπευτικά logs, το αποτέλεσμα ενδέχεται να αποκλίνει από την πραγματική διαδικασία. Παρά τους περιορισμούς του, ο Heuristic Miner έχει καθιερωθεί ως πρότυπο αλγόριθμο για πρακτικές εφαρμογές Process Mining, καθώς συνδυάζει αποδοτικότητα, ευελιξία και ρεαλιστική απεικόνιση των επιχειρησιακών ροών.

Fuzzy Miner

Ο αλγόριθμος Fuzzy Miner αποτελεί έναν σύγχρονο αλγόριθμο Process Mining που σχεδιάστηκε για την εξαγωγή και απεικόνιση πολύπλοκων και θορυβωδών διαδικασιών από πραγματικά event logs [Günther & van der Aalst, 2007]. Σε αντίθεση με τους Alpha και Heuristic Miner, οι οποίοι επικεντρώνονται σε καθαρές ή στατιστικά επαναλαμβανόμενες ακολουθίες, ο Fuzzy Miner χρησιμοποιεί τεχνικές συμπύκνωσης και αφαίρεσης ασήμαντων λεπτομερειών, ώστε να δημιουργεί ευανάγνωστα μοντέλα. Η βασική ιδέα του είναι η αξιολόγηση των δραστηριοτήτων και των σχέσεων τους με βάση δύο δείκτες: τη σπουδαιότητα (significance) και τη συνάφεια (correlation). Οι δραστηριότητες ή οι σχέσεις

που έχουν χαμηλή σπουδαιότητα ή συνάφεια μπορούν να συγχωνευθούν ή να απλοποιηθούν, μειώνοντας την πολυπλοκότητα του μοντέλου. Ο αλγόριθμος εφαρμόζει επίσης φιλτράρισμα και ομαδοποίηση κόμβων, δημιουργώντας ένα ιεραρχικό γράφημα που αποκαλύπτει τις κύριες ροές και τις υπο-ροές της διαδικασίας. Επιπλέον, μπορεί να αναγνωρίσει παράλληλες, επαναλαμβανόμενες και μη γραμμικές ακολουθίες, γεγονός που τον καθιστά ιδιαίτερα κατάλληλο για πραγματικά, μη δομημένα περιβάλλοντα. Ο Fuzzy Miner παρέχει εποπτικά μοντέλα, που βοηθούν τους αναλυτές να κατανοήσουν την κύρια συμπεριφορά της διαδικασίας και να εντοπίσουν σημεία συμφόρησης ή βελτίωσης. Η μέθοδος είναι ιδιαίτερα ανεκτική σε θόρυβο και σπάνιες περιπτώσεις, καθώς αυτές δεν επηρεάζουν σημαντικά το τελικό μοντέλο. Παρά την ευελιξία του, η ακριβής ρύθμιση των παραμέτρων σπουδαιότητας και συνάφειας απαιτεί εμπειρία και πειραματισμό. Ο αλγόριθμος είναι επίσης επεκτάσιμος και μπορεί να ενσωματωθεί με άλλα εργαλεία ανάλυσης για προχωρημένες εφαρμογές. Συνολικά, ο Fuzzy Miner προσφέρει μια ρεαλιστική και κατανοητή απεικόνιση των σύνθετων διαδικασιών, καθιστώντας τον πολύτιμο για οργανισμούς που επιθυμούν να αποκτήσουν γρήγορη και ουσιαστική γνώση από τα δεδομένα τους.

Token-Based Replay

Ο αλγόριθμος Token-Based Replay είναι μια τεχνική conformance checking στο πλαίσιο του Process Mining, που χρησιμοποιείται για να αξιολογήσει κατά πόσο ένα μοντέλο διαδικασίας (συνήθως σε μορφή Petri-net) συμφωνεί με τα πραγματικά δεδομένα που προέρχονται από event logs [Rozinat & van der Aalst, 2008]. Η βασική ιδέα του αλγορίθμου είναι η αναπαράσταση της εκτέλεσης κάθε case με tokens, τα οποία κινούνται μέσα στο Petri-net σύμφωνα με τις δραστηριότητες που παρατηρούνται στο log. Κατά την αναπαραγωγή της διαδικασίας, κάθε δραστηριότητα που εκτελείται «καίει» ή μετακινεί ένα token από μια θέση (place) σε άλλη, ακολουθώντας τους κανόνες του μοντέλου. Ο αλγόριθμος μετράει επίσης σφάλματα και αποκλίσεις, όπως tokens που λείπουν ή tokens που δεν μπορούν να τοποθετηθούν λόγω ασυμφωνίας με το μοντέλο, επιτρέποντας την ποσοτικοποίηση της συμμόρφωσης (fitness) του μοντέλου με τα δεδομένα. Η προσέγγιση αυτή είναι ιδιαίτερα χρήσιμη για την ανάλυση ρεαλιστικών και θορυβωδών logs, καθώς παρέχει λεπτομερή πληροφόρηση για τα σημεία όπου το μοντέλο δεν αντικατοπτρίζει

σωστά την εκτέλεση της διαδικασίας. Παράλληλα, επιτρέπει την κατηγοριοποίηση των αποκλίσεων σε λειτουργικές ή μη λειτουργικές, βοηθώντας τους αναλυτές να εντοπίσουν αδύναμα σημεία ή σημεία βελτίωσης στις διαδικασίες. Ο Token-Based Replay είναι επίσης σημαντικός για την αξιολόγηση της ποιότητας και της πληρότητας ενός μοντέλου, καθώς δείχνει με ακρίβεια ποσοτικά ποσοστά συμμόρφωσης για κάθε περίπτωση. Επιπλέον, μπορεί να χρησιμοποιηθεί σε συνδυασμό με άλλες τεχνικές Process Mining, όπως μέθοδοι ανακάλυψης μοντέλων και στατιστικές αναλύσεις, για μια ολιστική προσέγγιση της διαδικασίας. Η ευελιξία του αλγορίθμου τον καθιστά κατάλληλο για εφαρμογές σε πολύπλοκες επιχειρησιακές διαδικασίες, όπου οι παραδοσιακοί αλγόριθμοι ανακάλυψης ενδέχεται να μην αποδίδουν πλήρως. Επιπλέον, η απεικόνιση των tokens και των αποκλίσεων διευκολύνει την οπτική κατανόηση της εκτέλεσης της διαδικασίας. Η μέθοδος έχει βρει ευρεία χρήση σε τομείς όπως η υγειονομική περίθαλψη, οι τράπεζες και η παραγωγή, όπου η συμμόρφωση είναι κρίσιμη. Τέλος, ο Token-Based Replay παρέχει ένα ακριβές και επεξηγηματικό εργαλείο για την αξιολόγηση και βελτίωση των επιχειρησιακών διαδικασιών, συνδυάζοντας την ανάλυση δεδομένων με την επαλήθευση του μοντέλου.

Performance Mining

Ο αλγόριθμος Performance Mining αποτελεί μια τεχνική Process Mining που εστιάζει στην αξιολόγηση της απόδοσης των διαδικασιών και στην κατανόηση των χρονικών χαρακτηριστικών τους [van der Aalst et al., 2016]. Σε αντίθεση με τους αλγορίθμους ανακάλυψης διαδικασιών, που επικεντρώνονται στη δομή και τη ροή των δραστηριοτήτων, ο Performance Mining αναλύει στοιχεία όπως χρόνοι εκτέλεσης, καθυστερήσεις, συμφόρηση και χρονικά μοτίβα που εμφανίζονται στα event logs. Η βασική ιδέα είναι η συσχέτιση των χρονικών δεδομένων με τις διαδικασίες, ώστε να εντοπιστούν σημεία συμφόρησης, υπερβολικά μεγάλοι χρόνοι αναμονής ή καθυστερημένες ακολουθίες. Ο αλγόριθμος επιτρέπει τη δημιουργία μοντέλων χρονικής απόδοσης, τα οποία μπορούν να εμφανιστούν ως χρωματικοί χάρτες (heatmaps), χρονικά διαγράμματα ή στατιστικά γραφήματα πάνω στο μοντέλο της διαδικασίας. Μέσω αυτής της ανάλυσης, οι οργανισμοί μπορούν να εντοπίσουν κρίσιμες διαδρομές, δραστηριότητες που περιορίζουν την ταχύτητα της διαδικασίας και σημεία όπου η απόδοση δεν ανταποκρίνεται στα επιχειρησιακά πρότυπα. Ο Performance Mining μπορεί επίσης να συνδυαστεί με τεχνικές conformance

checking, επιτρέποντας την ταυτόχρονη αξιολόγηση της συμμόρφωσης και της απόδοσης. Η μέθοδος είναι ιδιαίτερα χρήσιμη για τη βελτιστοποίηση διαδικασιών, καθώς παρέχει δεδομένα για τη λήψη αποφάσεων σχετικά με επανασχεδιασμό ή αυτοματοποίηση. Επιπλέον, ο αλγόριθμος μπορεί να εφαρμόσει στατιστική ανάλυση και εξόρυξη μοτίβων, ώστε να προβλέψει πιθανά σημεία συμφόρησης στο μέλλον. Ο Performance Mining είναι επίσης κατάλληλος για την παρακολούθηση υπηρεσιών σε πραγματικό χρόνο, δίνοντας τη δυνατότητα άμεσης παρέμβασης σε προβληματικά σημεία. Τέλος, η εφαρμογή του σε διάφορους τομείς, όπως η παραγωγή, οι τράπεζες και η υγεία, επιτρέπει την συνεχή βελτίωση και τον έλεγχο της απόδοσης των επιχειρησιακών διαδικασιών, συμβάλλοντας στη λήψη τεκμηριωμένων αποφάσεων.

1.1.5.1 Συγκριτική αξιολόγηση και παραδείγματα χρήσης

Η ολοκλήρωση της θεωρητικής παρουσίασης των αλγορίθμων οδηγεί στο κρίσιμο ερώτημα της επιλογής της κατάλληλης μεθοδολογίας για την ανάλυση των δεδομένων. Στο πεδίο της Εξόρυξης Διεργασιών (Process Mining), δεν υφίσταται ένας καθολικός αλγόριθμος που να ανταποκρίνεται με την ίδια επιτυχία σε κάθε είδους αρχείο γεγονότων. Αντίθετα, η επιλογή εξαρτάται από την ποιότητα των δεδομένων, το επίπεδο του «θορύβου» που εμπεριέχεται σε αυτά και το αν ο αναλυτής στοχεύει στην απλή οπτικοποίηση ή στον αυστηρό μαθηματικό έλεγχο. Ξεκινώντας από τον Alpha Miner, αξιολογείται ως ένας αλγόριθμος με μεγάλη θεωρητική σημασία, καθώς ήταν ο πρώτος που κατάφερε να μετατρέψει τις απλές καταγραφές γεγονότων σε δομημένα μοντέλα Petri Nets. Ωστόσο, στην πράξη, ο Alpha Miner αποδεικνύεται ιδιαίτερα «εύθραυστος». Η κύρια αδυναμία του έγκειται στο γεγονός ότι θεωρεί τα δεδομένα τέλεια και πλήρη. Αν το αρχείο γεγονότων περιέχει έστω και ελάχιστα λάθη ή τυχαίες ενέργειες (θόρυβος), ο αλγόριθμος συχνά αποτυγχάνει να παράγει ένα σωστό μοντέλο ή δημιουργεί μια αναπαράσταση που δεν αντιστοιχεί στην πραγματικότητα. Επιπλέον, δυσκολεύεται να αναγνωρίσει σύνθετες λογικές σχέσεις, όπως όταν πολλές δραστηριότητες συμβαίνουν ταυτόχρονα ή όταν υπάρχουν εξαρτήσεις που δεν είναι άμεσα ορατές στην αλληλουχία των γεγονότων.

Σε αντίθεση με την αυστηρότητα του Alpha Miner, ο Heuristic Miner προσφέρει μια πολύ πιο ρεαλιστική και «έξυπνη» προσέγγιση. Ο αλγόριθμος αυτός δεν εξετάζει κάθε γεγονός ως απόλυτο, αλλά χρησιμοποιεί στατιστικά κριτήρια και συχνότητες για να καταλάβει ποιες

διαδρομές είναι οι πιο σημαντικές. Με αυτόν τον τρόπο, καταφέρνει να φιλτράρει τις σπάνιες ή τυχαίες συμπεριφορές που θα μπορούσαν να μπερδέψουν την ανάλυση, αναδεικνύοντας τον βασικό κορμό της διεργασίας. Αυτό το χαρακτηριστικό τον καθιστά εξαιρετικά αξιόπιστο για την ανακάλυψη μοντέλων από πραγματικά δεδομένα, καθώς παράγει μια καθαρή εικόνα της διαδικασίας (μέσω των Causal Nets), η οποία παραμένει πιστή στην πλειονότητα των περιπτώσεων που καταγράφηκαν. Η δυνατότητά του να εξισορροπεί την πολυπλοκότητα με την αναγνωσιμότητα είναι ο λόγος που θεωρείται ένας από τους πιο αποτελεσματικούς αλγορίθμους για την ανακάλυψη της ροής μιας διαδικασίας.

Μια εντελώς διαφορετική φιλοσοφία ακολουθεί ο Fuzzy Miner, ο οποίος έχει σχεδιαστεί για να αντιμετωπίζει το πρόβλημα των εξαιρετικά μη δομημένων και περίπλοκων μοντέλων. Σε πολλές περιπτώσεις, οι διαδικασίες είναι τόσο περίπλοκες που η προσπάθεια δημιουργίας ενός αυστηρού μοντέλου οδηγεί σε ένα χάος από γραμμές και κόμβους που είναι αδύνατο να ερμηνευτούν. Ο Fuzzy Miner λύνει αυτό το πρόβλημα χρησιμοποιώντας τεχνικές αφαίρεσης: ομαδοποιεί τις λιγότερο σημαντικές δραστηριότητες και αναδεικνύει μόνο τους ισχυρούς δεσμούς. Παρόλο που προσφέρει την καλύτερη οπτικοποίηση για πολύπλοκα δεδομένα, υστερεί στο γεγονός ότι το μοντέλο που παράγει δεν έχει την απαιτούμενη μαθηματική δομή για να χρησιμοποιηθεί σε επόμενα στάδια ελέγχου συμμόρφωσης. Είναι, λοιπόν, ένα εξαιρετικό εργαλείο για μια πρώτη «ανάγνωση» της κατάστασης, αλλά δεν επαρκεί για μια εις βάθος τεχνική ανάλυση.

Στην αξιολόγηση πρέπει να συμπεριληφθεί και το Performance Mining, το οποίο, αν και συχνά λειτουργεί συμπληρωματικά, προσφέρει κάτι που οι άλλοι αλγόριθμοι παραλείπουν: τη διάσταση του χρόνου. Ενώ οι υπόλοιποι επικεντρώνονται στο «τι έγινε» και με «ποια σειρά», το Performance Mining εστιάζει στο «πόσο διήρκεσε». Η ικανότητά του να εντοπίζει καθυστερήσεις και σημεία όπου η διαδικασία «κολλάει» (bottlenecks), προβάλλοντας αυτές τις πληροφορίες πάνω στο μοντέλο, το καθιστά απαραίτητο για την αξιολόγηση της αποδοτικότητας. Τέλος, ο αλγόριθμος Token-Based Replay (TBR) έρχεται να δώσει απάντηση στο ερώτημα της εγκυρότητας. Η μέθοδος αυτή παίρνει τα πραγματικά ίχνη των γεγονότων και προσπαθεί να τα «τρέξει» πάνω σε ένα υπάρχον μοντέλο. Είναι ο πιο κατανοητός τρόπος για να δούμε πού ακριβώς υπάρχει απόκλιση, καθώς καταγράφει κάθε φορά που μια ενέργεια δεν ταιριάζει με το προβλεπόμενο σχέδιο. Ο TBR δεν ανακαλύπτει μια νέα διαδικασία, αλλά ελέγχει την πιστότητα της υπάρχουσας, προσφέροντας πολύτιμα

διαγνωστικά στοιχεία.

Παρακάτω ακολουθεί ο συγκριτικός πίνακας της αξιολόγησης των παραπάνω αλγορίθμων Εξόρυξης Διεργασιών και η ανάλυση των κριτηρίων αξιολόγησης,

Κριτήριο Αξιολόγησης	Alpha Miner	Heuristic Miner	Fuzzy Miner	Token-Base d Replay	Performanc e Mining
Ανοχή σε Θόρυβο & Λάθη	Πολύ Χαμηλή	Υψηλή	Πολύ Υψηλή	Μέτρια	Μέτρια
Ανίχνευση Παραλληλίας	Ναι	Ναι	Μερικώς	Ναι	Ναι
Επίπεδο Αφαίρεσης/Απλότητα	Χαμηλό	Μέτριο	Πολύ Υψηλό	Χαμηλό	Μέτριο
Μαθηματική Αυστηρότητα	Υψηλή	Υψηλή	Χαμηλή	Πολύ Υψηλή	Μέτρια
Διαγνωστική Ικανότητα	Χαμηλή	Μέτρια	Υψηλή	Κορυφαία	Υψηλή
Ανάλυση Χρονικών Δεδομένων	Όχι	Όχι	Όχι	Όχι	Κυρίαρχη

Πίνακας 1: Συγκριτικός Πίνακας Αξιολόγησης Αλγορίθμων Εξόρυξης Διεργασιών

Για την πληρέστερη κατανόηση του παραπάνω πίνακα, είναι απαραίτητο να αναλυθεί η σημασία των κριτηρίων που χρησιμοποιήθηκαν, καθώς αυτά καθορίζουν την επιτυχία της εξαγωγής συμπερασμάτων από τα δεδομένα:

1. Ανοχή σε Θόρυβο και Λάθη (Noise Robustness): Το κριτήριο αυτό αναφέρεται στην ικανότητα του αλγορίθμου να αγνοεί μεμονωμένα, τυχαία περιστατικά που δεν αντιπροσωπεύουν την πραγματική διαδικασία. Όπως παρατηρούμε, ο Alpha Miner υστερεί σημαντικά εδώ, καθώς αντιμετωπίζει κάθε καταγραφή ως εξίσου σημαντική. Αντίθετα, ο Heuristic Miner και ο Fuzzy Miner «φιλτράρουν» την πληροφορία, εστιάζοντας στις συχνές επαναλήψεις, γεγονός που τους καθιστά πιο αξιόπιστους σε πραγματικά σενάρια όπου τα ανθρώπινα λάθη ή οι τεχνικές δυσλειτουργίες δημιουργούν «θόρυβο» στα αρχεία γεγονότων.

2. Ανίχνευση Παραλληλίας (Concurrency): Πρόκειται για την ικανότητα του αλγορίθμου να καταλαβαίνει πότε δύο δραστηριότητες συμβαίνουν ταυτόχρονα ή με τυχαία σειρά, χωρίς η μία να εξαρτάται από την άλλη. Η σωστή αναγνώριση της παραλληλίας αποτρέπει τη

δημιουργία υπερβολικά πολύπλοκων μοντέλων που μοιάζουν με «λαβύρινθο», επιτρέποντας μια πιο καθαρή και λογική αναπαράσταση της διεργασίας.

3. Επίπεδο Αφαίρεσης και Απλότητα: Το κριτήριο αυτό αξιολογεί πόσο εύκολα μπορεί ένας άνθρωπος να διαβάσει και να κατανοήσει το παραγόμενο μοντέλο. Ο Fuzzy Miner υπερέχει σε αυτόν τον τομέα, καθώς «κρύβει» τις λεπτομέρειες που μπερδεύουν το μάτι. Ωστόσο, η υπερβολική απλούστευση μπορεί μερικές φορές να οδηγήσει στην απώλεια κρίσιμων πληροφοριών, γι' αυτό και ο Heuristic Miner θεωρείται ότι προσφέρει την καλύτερη ισορροπία μεταξύ λεπτομέρειας και ευκρίνειας.

4. Μαθηματική Αυστηρότητα και Σημασιολογία: Αναφέρεται στο αν το μοντέλο που προκύπτει ακολουθεί συγκεκριμένους κανόνες λογικής (όπως τα Petri Nets). Αυτό είναι κρίσιμο για τη διπλωματική εργασία, καθώς μόνο μοντέλα με υψηλή αυστηρότητα μπορούν να χρησιμοποιηθούν για τον έλεγχο συμμόρφωσης. Εδώ, ο Heuristic Miner και ο Token-Based Replay αποτελούν τις κορυφαίες επιλογές, καθώς επιτρέπουν τη διεξαγωγή ασφαλών μαθηματικών υπολογισμών πάνω στη δομή της διαδικασίας.

5. Διαγνωστική Ικανότητα: Το κριτήριο αυτό μετρά πόσο καλά μπορεί ένας αλγόριθμος να μας υποδείξει πού ακριβώς εντοπίζεται ένα πρόβλημα ή μια απόκλιση. Ο Token-Based Replay είναι ο αδιαμφισβήτητος ηγέτης εδώ, καθώς «τρέχει» τα δεδομένα πάνω στο μοντέλο και μας δείχνει με ακρίβεια το σημείο όπου η πραγματικότητα διαφέρει από το σχεδιασμό.

6. Ανάλυση Χρονικών Δεδομένων: Ενώ οι περισσότεροι αλγόριθμοι εστιάζουν στη δομή (ποια ενέργεια ακολουθεί ποια), το Performance Mining είναι η μοναδική τεχνική που αξιοποιεί πλήρως τις χρονικές σημάνσεις (timestamps) για να υπολογίσει καθυστερήσεις, χρόνους αναμονής και συνολική διάρκεια διεκπεραίωσης.

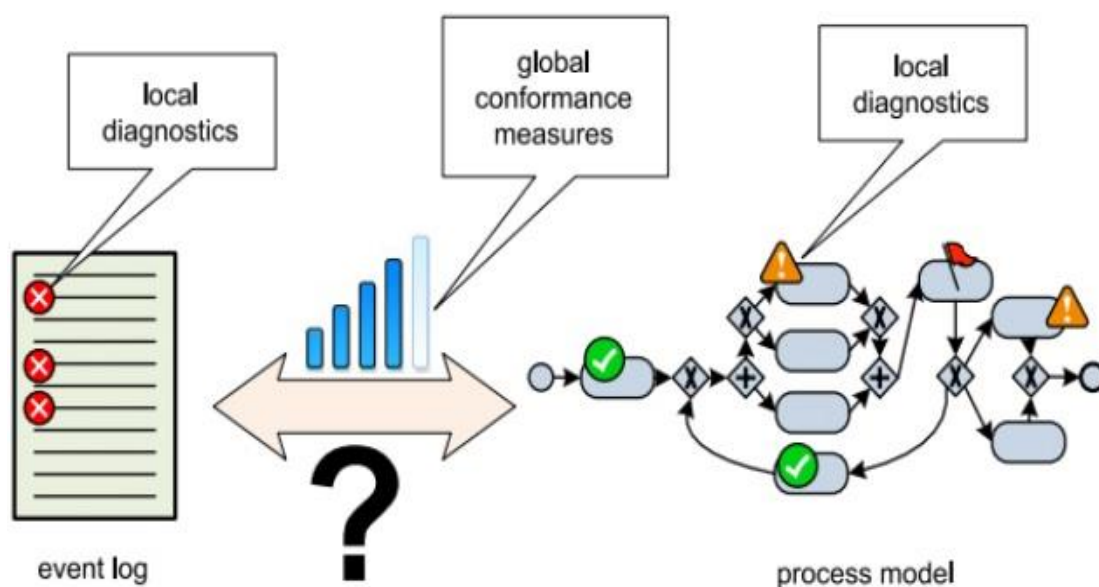
Λαμβάνοντας υπόψη όλα τα παραπάνω, για τις ανάγκες της παρούσας εργασίας προκρίνεται ο συνδυασμός του **Heuristic Miner** και του **Token-Based Replay**. Ο πρώτος επιλέγεται ως το βασικό εργαλείο ανακάλυψης (Discovery), καθώς είναι ο πλέον ικανός να διαχειριστεί τον θόρυβο και να παράγει ένα μοντέλο που είναι ταυτόχρονα απλό και ακριβές. Ο δεύτερος επιλέγεται ως το εργαλείο ελέγχου (Conformance), διότι επιτρέπει τη σύγκριση της πραγματικότητας με το μοντέλο με τρόπο που είναι μετρήσιμος και εύκολα αναγνώσιμος.

Αυτή η συνδυαστική προσέγγιση εξασφαλίζει ότι η ανάλυση θα βασίζεται σε ένα στέρεο μοντέλο, το οποίο στη συνέχεια θα ελεγχθεί εξονυχιστικά για τυχόν αποκλίσεις, προσφέροντας μια ολοκληρωμένη εικόνα της διαδικασίας που μελετάται.

1.2 Έλεγχος Συμμόρφωσης

1.2.1 Θεωρητικό υπόβαθρο

Ο Έλεγχος Συμμόρφωσης (Conformance Checking) αποτελεί θεμελιώδη συνιστώσα της Εξόρυξης Διεργασιών και εστιάζει στη συστηματική σύγκριση μεταξύ της πραγματικής συμπεριφοράς μιας διεργασίας και της αναμενόμενης ή σχεδιασμένης συμπεριφοράς, όπως αυτή αποτυπώνεται σε ένα μοντέλο διεργασίας (van der Aalst, 2016) (**Εικόνα 8**) [Mukhiya & Ingvaldsen, 2018]. Η ανάγκη για Έλεγχο Συμμόρφωσης προκύπτει από το γεγονός ότι οι οργανισμοί λειτουργούν σε δυναμικά και συχνά πολύπλοκα περιβάλλοντα, όπου οι πραγματικές εκτελέσεις αποκλίνουν συστηματικά από τις τυπικές προδιαγραφές, είτε λόγω ανθρώπινων παραγόντων, είτε λόγω οργανωσιακών περιορισμών, είτε λόγω εξωτερικών συνθηκών (Rozinat & van der Aalst, 2008).



Εικόνα 8: Έλεγχος συμμόρφωσης: σύγκριση παρατηρούμενης συμπεριφοράς με μοντελοποιημένη συμπεριφορά

Στο θεωρητικό πλαίσιο του Conformance Checking, κεντρικό ρόλο διαδραματίζει η έννοια

του αρχείου καταγραφής συμβάντων. Ένα αρχείο καταγραφής συμβάντων, αποτελείται από ένα σύνολο ιχνών εκτέλεσης (traces), όπου κάθε ίχνος αντιστοιχεί σε μία συγκεκριμένη περίπτωση της διεργασίας και περιλαμβάνει μία διατεταγμένη ακολουθία γεγονότων. Κάθε γεγονός περιγράφεται από ένα σύνολο χαρακτηριστικών, όπως η δραστηριότητα που εκτελέστηκε, ο χρόνος εκτέλεσης, ο πόρος που την εκτέλεσε και επιπλέον δεδομένα που σχετίζονται με την κατάσταση της διεργασίας. Το event log, επομένως, αποτυπώνει την πραγματική επιχειρησιακή συμπεριφορά με τρόπο λεπτομερή και αντικειμενικό (van der Aalst et al., 2011).

Το μοντέλο διεργασίας, από την άλλη πλευρά, εκφράζει τη θεωρητικά επιθυμητή ή κανονιστικά αποδεκτή συμπεριφορά της διεργασίας. Τα μοντέλα αυτά μπορεί να έχουν διαφορετικές μορφές, με τα Petri Nets να αποτελούν την επικρατέστερη τυπική αναπαράσταση λόγω της εκφραστικότητάς τους και της ισχυρής μαθηματικής τους βάσης (Murata, 1989). Ωστόσο, συχνά χρησιμοποιούνται και άλλες αναπαραστάσεις, όπως BPMN διαγράμματα, process trees ή δηλωτικά μοντέλα, τα οποία εστιάζουν περισσότερο σε περιορισμούς παρά σε ακριβείς ακολουθίες δραστηριοτήτων (Dumas et al., 2018).

Θεωρητικά, ο Έλεγχος Συμμόρφωσης μπορεί να ιδωθεί ως πρόβλημα σύγκρισης δύο γλωσσών: της γλώσσας που παράγεται από το μοντέλο διεργασίας και της γλώσσας που προκύπτει από τα event logs. Η σύγκριση αυτή επιτρέπει την αναγνώριση αποκλίσεων, οι οποίες μπορεί να εκφράζονται ως παραλείψεις δραστηριοτήτων, επιπλέον εκτελέσεις, αλλαγές στη σειρά εκτέλεσης ή παραβιάσεις παράλληλων ροών (van der Aalst, 2016). Στο πλαίσιο αυτό, η συμμόρφωση δεν αντιμετωπίζεται ως απόλυτη έννοια, αλλά ως βαθμός εγγύτητας μεταξύ των δύο αναπαραστάσεων.

Η θεωρητική σημασία του Conformance Checking ενισχύεται περαιτέρω από τη σύνδεσή του με την επιχειρησιακή διακυβέρνηση και τη συμμόρφωση με κανονιστικά πλαίσια. Σε τομείς όπως η υγεία, οι χρηματοοικονομικές υπηρεσίες και η δημόσια διοίκηση, η απόκλιση από καθορισμένες διαδικασίες μπορεί να έχει σοβαρές νομικές και λειτουργικές συνέπειες (van der Aalst, 2016; Dumas et al., 2018). Κατά συνέπεια, ο Έλεγχος Συμμόρφωσης λειτουργεί όχι μόνο ως εργαλείο ανάλυσης, αλλά και ως μηχανισμός ελέγχου, διαφάνειας και διασφάλισης ποιότητας.

1.2.2 Μέθοδοι και Αλγόριθμοι για Έλεγχο Συμμόρφωσης

Οι μέθοδοι που έχουν αναπτυχθεί για τον Έλεγχο Συμμόρφωσης παρουσιάζουν σημαντική ποικιλομορφία, αντανακλώντας τόσο τη διαφορετικότητα των μοντέλων διεργασίας όσο και την πολυπλοκότητα των πραγματικών δεδομένων [van der Aalst, 2016; van Dongen et al., 2005]. Οι πρώτες προσεγγίσεις βασίστηκαν στην ιδέα της αναπαραγωγής των event logs πάνω στο μοντέλο διεργασίας, με χαρακτηριστικό παράδειγμα τη μέθοδο token-based replay. Στην προσέγγιση αυτή, κάθε γεγονός του log επιχειρείται να «προσπελάσει» το αντίστοιχο transition του μοντέλου, με την παρακολούθηση της ροής των tokens να αποκαλύπτει αποκλίσεις και ασυνέπειες [Rozinat & van der Aalst, 2008].

Παρά τη σχετική απλότητά της, η token-based προσέγγιση παρουσιάζει περιορισμούς ως προς την ακρίβεια, ιδιαίτερα σε μοντέλα με έντονη παράλληλη συμπεριφορά ή εναλλακτικές διαδρομές. Για την αντιμετώπιση αυτών των περιορισμών, αναπτύχθηκε η μέθοδος των ευθυγραμμίσεων (alignments), η οποία θεωρείται σήμερα η πλέον θεωρητικά θεμελιωμένη προσέγγιση στον Έλεγχο Συμμόρφωσης [Adriansyah et al., 2011]. Η βασική ιδέα των alignments είναι η εύρεση της βέλτιστης αντιστοίχισης μεταξύ ενός trace του log και μιας επιτρεπτής ακολουθίας του μοντέλου, λαμβάνοντας υπόψη ένα προκαθορισμένο σύστημα κόστους.

Οι αλγόριθμοι ευθυγράμμισης βασίζονται συνήθως σε τεχνικές αναζήτησης καταστάσεων, όπως ο αλγόριθμος A^* , και επιτρέπουν την ακριβή ταυτοποίηση των σημείων απόκλισης [van Dongen et al., 2012]. Κάθε απόκλιση κατηγοριοποιείται ως κίνηση μόνο στο log ή μόνο στο μοντέλο, προσφέροντας σαφή ερμηνεία των διαφορών. Η προσέγγιση αυτή παρέχει πλούσια διαγνωστική πληροφορία, καθιστώντας την ιδιαίτερα χρήσιμη σε περιβάλλοντα όπου η κατανόηση των αιτιών των αποκλίσεων είναι εξίσου σημαντική με την ανίχνευσή τους [van der Aalst, 2016].

Πέρα από τα Petri Nets, σημαντική ερευνητική δραστηριότητα έχει αναπτυχθεί γύρω από τον Έλεγχο Συμμόρφωσης σε δηλωτικά μοντέλα διεργασιών. Σε αυτά τα μοντέλα, η έμφαση μετατοπίζεται από την ακριβή ακολουθία δραστηριοτήτων στους περιορισμούς που πρέπει να ικανοποιούνται. Ο έλεγχος συμμόρφωσης, συνεπώς, επικεντρώνεται στον εντοπισμό παραβιάσεων περιορισμών, γεγονός που προσφέρει μεγαλύτερη ευελιξία αλλά και αυξημένη υπολογιστική πολυπλοκότητα [Montali et al., 2010].

Τα τελευταία χρόνια, παρατηρείται αυξανόμενο ενδιαφέρον για data-aware και time-aware Conformance Checking, όπου η συμμόρφωση δεν αφορά μόνο τη δομή της διεργασίας αλλά και τις τιμές των δεδομένων και τους χρονικούς περιορισμούς. Οι προσεγγίσεις αυτές διευρύνουν σημαντικά το πεδίο εφαρμογής του Conformance Checking, καθιστώντας το πιο κοντά στις πραγματικές ανάγκες των οργανισμών [De Leoni et al., 2012].

1.2.3 Μετρικές Ποιότητας

Η ποσοτικοποίηση της συμμόρφωσης απαιτεί τη χρήση μετρικών ποιότητας που αποτυπώνουν διαφορετικές πτυχές της σχέσης μεταξύ μοντέλου και event log. Η μετρική fitness αποτελεί τον βασικό δείκτη [van der Aalst, 2011; Adriansyah et al., 2011], καθώς εκφράζει τον βαθμό στον οποίο το μοντέλο μπορεί να εξηγήσει τις παρατηρούμενες εκτελέσεις. Ένα υψηλό fitness υποδηλώνει ότι οι περισσότερες περιπτώσεις του log μπορούν να αναπαραχθούν από το μοντέλο χωρίς σημαντικές αποκλίσεις.

Ωστόσο, το fitness από μόνο του δεν επαρκεί για την αξιολόγηση της ποιότητας της συμμόρφωσης. Ένα υπερβολικά γενικό μοντέλο μπορεί να επιτυγχάνει υψηλό fitness, επιτρέποντας όμως συμπεριφορές που δεν παρατηρούνται ποτέ στην πράξη [Rozinat & van der Aalst, 2008; Adriansyah, 2014]. Για τον λόγο αυτό εισάγεται η μετρική precision, η οποία αξιολογεί το κατά πόσο το μοντέλο περιορίζει τη συμπεριφορά του στις παρατηρούμενες εκτελέσεις. Η ισορροπία μεταξύ fitness και precision αποτελεί κεντρικό ζήτημα στη βιβλιογραφία της Εξόρυξης Διεργασιών.

Η γενίκευση συμπληρώνει τις παραπάνω μετρικές, αξιολογώντας την ικανότητα του μοντέλου να αποδέχεται εύλογες, αλλά μη παρατηρημένες, συμπεριφορές. Τέλος, η απλότητα σχετίζεται με τη δομική πολυπλοκότητα του μοντέλου και την κατανοησιμότητά του από ανθρώπινους αναλυτές [van der Aalst et al., 2004]. Στο πλαίσιο του Conformance Checking, οι μετρικές αυτές χρησιμοποιούνται συνδυαστικά, προκειμένου να παραχθεί μια σφαιρική εικόνα της ποιότητας και της χρησιμότητας του μοντέλου.

1.2.4 Προκλήσεις και Περιορισμοί

Παρά τη θεωρητική ωριμότητα και τη συνεχή εξέλιξη των μεθόδων Conformance Checking, εξακολουθούν να υφίστανται σημαντικές προκλήσεις. Η διαχείριση μεγάλου όγκου δεδομένων αποτελεί μία από τις βασικότερες, καθώς τα σύγχρονα πληροφοριακά

συστήματα παράγουν event logs με εκατομμύρια γεγονότα [van der Aalst, 2011]. Η κλιμάκωση των αλγορίθμων και η διατήρηση αποδεκτών χρόνων εκτέλεσης αποτελούν ενεργά ερευνητικά πεδία.

Ένας ακόμη σημαντικός περιορισμός αφορά την ποιότητα των δεδομένων. Τα event logs συχνά περιέχουν θόρυβο, ελλειπείς εγγραφές ή ασυνέπειες, οι οποίες επηρεάζουν άμεσα τα αποτελέσματα του ελέγχου συμμόρφωσης [Rozinat & van der Aalst, 2008]. Επιπλέον, η ερμηνεία των αποκλίσεων απαιτεί βαθιά γνώση του επιχειρησιακού πλαισίου, καθώς δεν είναι όλες οι αποκλίσεις ανεπιθύμητες ή προβληματικές.

Τέλος, η μετάβαση από τον παραδοσιακό, δομικό Έλεγχο Συμμόρφωσης σε πιο σύνθετες μορφές που ενσωματώνουν δεδομένα, χρόνο και πόρους παραμένει πρόκληση [van der Aalst et al., 2012]. Η ανάγκη για ολοκληρωμένες προσεγγίσεις που θα υποστηρίζουν τη συνεχή βελτίωση διεργασιών καθιστά το Conformance Checking ένα από τα πλέον δυναμικά και ερευνητικά ενεργά πεδία της σύγχρονης Εξόρυξης Διεργασιών.

ΚΕΦΑΛΑΙΟ 2

UML και UML Activity Diagram

2.1 Θεωρητικό υπόβαθρο (UML και UML Activity Diagram)

Η Ενοποιημένη Γλώσσα Μοντελοποίησης (Unified Modeling Language – UML) αποτελεί ένα διεθνώς αναγνωρισμένο πρότυπο οπτικής μοντελοποίησης, το οποίο χρησιμοποιείται ευρέως για την ανάλυση, τον σχεδιασμό και την τεκμηρίωση σύνθετων συστημάτων λογισμικού και επιχειρησιακών διαδικασιών (Ahmad et al., 2019). Η UML αναπτύχθηκε με στόχο την ενοποίηση διαφορετικών αντικειμενοστραφών μεθοδολογιών και καθιερώθηκε ως πρότυπο από τον οργανισμό Object Management Group (OMG). Η βασική φιλοσοφία της UML έγκειται στη δημιουργία μιας κοινής, τυποποιημένης γλώσσας επικοινωνίας μεταξύ όλων των εμπλεκόμενων μερών ενός έργου, ανεξαρτήτως τεχνικού υπόβαθρου.

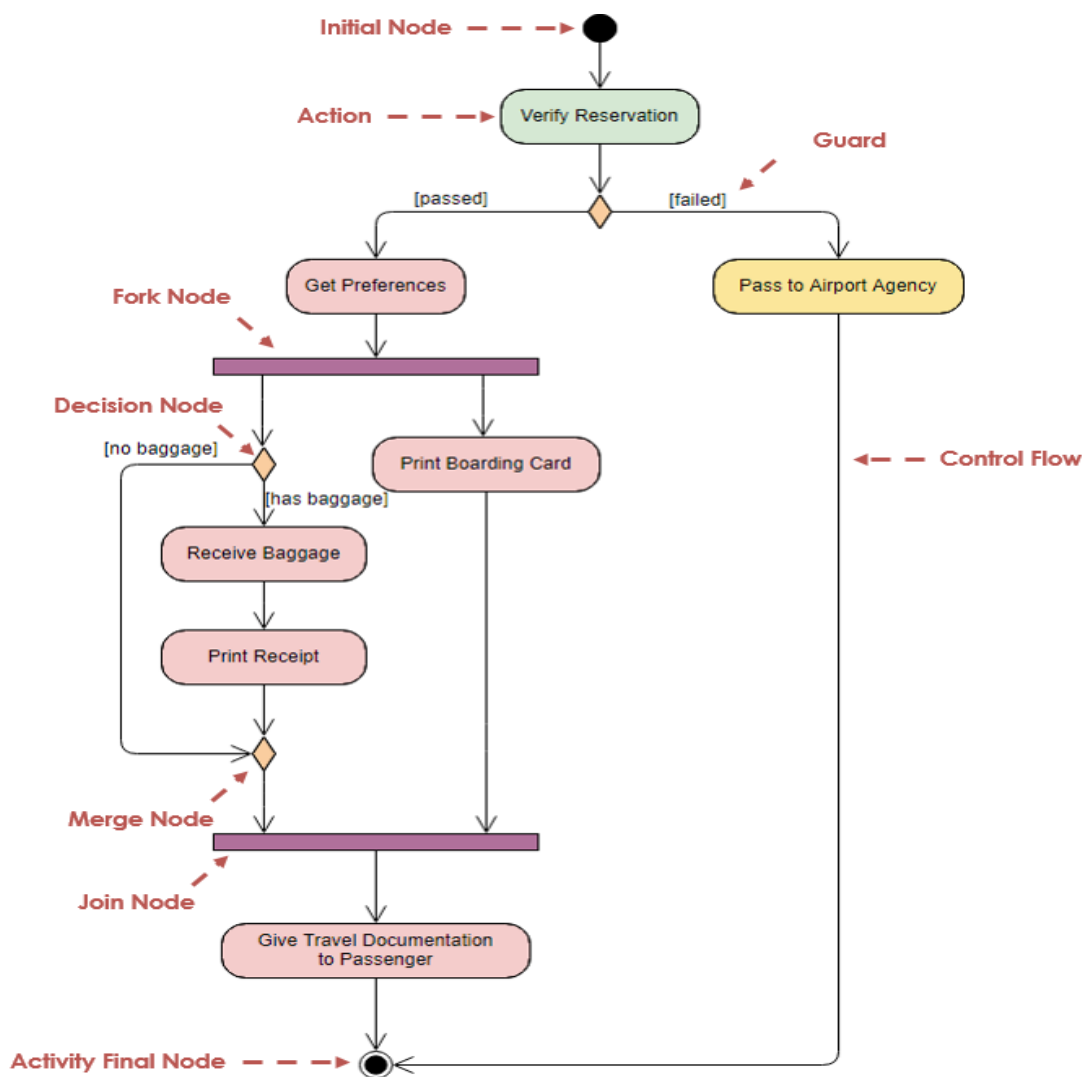
Η UML δεν αποτελεί μεθοδολογία ανάπτυξης λογισμικού, αλλά ένα σύνολο συμβολισμών και κανόνων που επιτρέπουν τη σαφή και ακριβή απεικόνιση τόσο της δομής όσο και της συμπεριφοράς ενός συστήματος (Ahmad et al., 2019). Το πρότυπο περιλαμβάνει πληθώρα διαγραμμάτων, τα οποία κατηγοριοποιούνται σε δομικά και συμπεριφορικά. Τα δομικά διαγράμματα επικεντρώνονται στη στατική απεικόνιση των στοιχείων ενός συστήματος, ενώ τα συμπεριφορικά διαγράμματα αποτυπώνουν τη δυναμική αλληλεπίδραση και τη ροή εκτέλεσης των λειτουργιών (Ahmad et al., 2019).

Τα UML Activity Diagrams ανήκουν στην κατηγορία των συμπεριφορικών διαγραμμάτων και χρησιμοποιούνται για τη μοντελοποίηση ροών εργασίας και ακολουθιών ενεργειών (Ahmad et al., 2019). Αποτελούν εξέλιξη των κλασικών διαγραμμάτων ροής, ενσωματώνοντας έννοιες όπως ο παραλληλισμός, ο συγχρονισμός και οι εναλλακτικές διαδρομές εκτέλεσης. Η δυνατότητα αυτή καθιστά τα Activity Diagrams ιδιαίτερα κατάλληλα για την αναπαράσταση σύνθετων διαδικασιών, τόσο σε τεχνικό όσο και σε επιχειρησιακό επίπεδο.

Στα βασικά στοιχεία ενός UML Activity Diagram περιλαμβάνονται οι δραστηριότητες και οι ενέργειες, οι οποίες αναπαριστούν τα επιμέρους βήματα μιας διαδικασίας. Η ροή ελέγχου καθορίζεται μέσω ακμών που συνδέουν τις δραστηριότητες, ενώ ειδικοί κόμβοι

χρησιμοποιούνται για την αναπαράσταση της έναρξης και της ολοκλήρωσης της διαδικασίας. Οι κόμβοι απόφασης επιτρέπουν τη μοντελοποίηση συνθηκών και διακλαδώσεων, ενώ οι κόμβοι συγχώνευσης επαναφέρουν τη ροή σε κοινό σημείο (Ahmad et al., 2019).

Ιδιαίτερη σημασία στα Activity Diagrams έχουν οι μηχανισμοί παραλληλισμού. Μέσω των κόμβων διακλάδωσης (fork) και συγχρονισμού (join), είναι δυνατή η ταυτόχρονη εκτέλεση πολλαπλών δραστηριοτήτων και η επανασύνδεσή τους σε μεταγενέστερο στάδιο. Η δυνατότητα αυτή αντανακλά με ακρίβεια τον τρόπο λειτουργίας πολλών πραγματικών συστημάτων και επιχειρησιακών διαδικασιών, στις οποίες διάφορες εργασίες εκτελούνται ταυτόχρονα από διαφορετικούς ρόλους ή συστήματα.



Εικόνα 9: UML AD για επαλήθευση κράτησης αεροπορικού εισιτηρίου

Επιπλέον, τα UML Activity Diagrams υποστηρίζουν τη χρήση λωρίδων ευθύνης (swimlanes), οι οποίες επιτρέπουν την κατανομή των δραστηριοτήτων σε διαφορετικούς εκτελεστές, όπως χρήστες, οργανωτικές μονάδες ή πληροφοριακά συστήματα. Η χρήση των swimlanes ενισχύει τη σαφήνεια του διαγράμματος και διευκολύνει την κατανόηση των ρόλων και των ευθυνών εντός μιας διαδικασίας (Hettab et al., 2015).

Η εξέλιξη της UML στην έκδοση 2.x οδήγησε σε σημαντική ενίσχυση της εκφραστικής δύναμης των Activity Diagrams. Προστέθηκαν έννοιες όπως η ροή αντικειμένων, η ιεραρχική αποσύνθεση δραστηριοτήτων και η δυνατότητα μοντελοποίησης σύνθετων καταστάσεων συγχρονισμού. Ως αποτέλεσμα, τα Activity Diagrams μπορούν πλέον να χρησιμοποιηθούν όχι μόνο για απλή περιγραφή διαδικασιών, αλλά και ως εργαλείο ανάλυσης και σχεδιασμού πολύπλοκων συστημάτων (Hettab et al., 2015).

Συνοψίζοντας, τα UML Activity Diagrams αποτελούν ένα θεμελιώδες εργαλείο της UML για την αναπαράσταση της συμπεριφοράς και της λογικής ροής ενός συστήματος. Η ευελιξία, η σαφήνεια και η τυποποίησή τους τα καθιστούν ιδιαίτερα χρήσιμα σε περιβάλλοντα όπου απαιτείται ακριβής επικοινωνία και τεκμηρίωση σύνθετων διαδικασιών (Hettab et al., 2015).

2.2 Ρόλος στην Μοντελοποίηση Επιχειρησιακών Διαδικασιών

Η μοντελοποίηση επιχειρησιακών διαδικασιών αποτελεί βασικό στοιχείο της σύγχρονης οργανωσιακής διαχείρισης και του ψηφιακού μετασχηματισμού. Στόχος της είναι η συστηματική καταγραφή, ανάλυση και βελτίωση των διαδικασιών που διέπουν τη λειτουργία ενός οργανισμού. Στο πλαίσιο αυτό, τα UML Activity Diagrams διαδραματίζουν σημαντικό ρόλο, προσφέροντας ένα τυποποιημένο και κατανοητό μέσο απεικόνισης των επιχειρησιακών ροών.

Τα Activity Diagrams επιτρέπουν την αποτύπωση της αλληλουχίας των δραστηριοτήτων που συνθέτουν μια επιχειρησιακή διαδικασία, καθιστώντας εμφανή τα σημεία λήψης αποφάσεων, τις εναλλακτικές ροές και τις εξαρτήσεις μεταξύ επιμέρους ενεργειών (Jena et al., 2014). Η οπτική αυτή αναπαράσταση διευκολύνει την κατανόηση της διαδικασίας από όλα τα εμπλεκόμενα μέρη και συμβάλλει στη μείωση ασάφειας και παρερμηνειών.

Ένα από τα κύρια πλεονεκτήματα των UML Activity Diagrams είναι η δυνατότητα

μοντελοποίησης παράλληλων δραστηριοτήτων. Στις σύγχρονες επιχειρήσεις, πολλές διαδικασίες δεν εκτελούνται σειριακά, αλλά περιλαμβάνουν ταυτόχρονες ενέργειες από διαφορετικούς ρόλους ή συστήματα. Τα Activity Diagrams μπορούν να αποτυπώσουν αυτή την πολυπλοκότητα με σαφή και κατανοητό τρόπο, γεγονός που τα καθιστά ιδιαίτερα χρήσιμα σε περιβάλλοντα υψηλής οργανωσιακής πολυπλοκότητας (Jena et al., 2014).

Η χρήση swimlanes ενισχύει περαιτέρω τη χρησιμότητα των Activity Diagrams στη μοντελοποίηση επιχειρησιακών διαδικασιών. Μέσω αυτών, είναι δυνατή η απεικόνιση της κατανομής αρμοδιοτήτων και της αλληλεπίδρασης μεταξύ διαφορετικών οργανωτικών μονάδων ή πληροφοριακών συστημάτων (Ding et al., 2016). Έτσι, τα διαγράμματα δεν περιορίζονται μόνο στη λειτουργική διάσταση της διαδικασίας, αλλά ενσωματώνουν και την οργανωτική της δομή (Ding et al., 2016).

Στην πράξη, τα UML Activity Diagrams χρησιμοποιούνται συχνά για την τεκμηρίωση της υφιστάμενης κατάστασης μιας διαδικασίας (as-is). Η καταγραφή αυτή αποτελεί απαραίτητη προϋπόθεση για την κατανόηση των υφιστάμενων προβλημάτων και την αναγνώριση ευκαιριών βελτίωσης. Στη συνέχεια, τα ίδια διαγράμματα μπορούν να αξιοποιηθούν για τον σχεδιασμό της επιθυμητής μελλοντικής κατάστασης (to-be), ενσωματώνοντας αλλαγές που στοχεύουν στη βελτίωση της αποδοτικότητας, της ποιότητας ή της ευελιξίας (Ding et al., 2016).

Σε σύγκριση με άλλες γλώσσες μοντελοποίησης επιχειρησιακών διαδικασιών, όπως η BPMN, τα UML Activity Diagrams χαρακτηρίζονται από μεγαλύτερη γενικότητα και λιγότερη εξειδίκευση. Παρότι αυτό μπορεί να θεωρηθεί περιορισμός σε ορισμένα σενάρια, αποτελεί ταυτόχρονα πλεονέκτημα σε περιβάλλοντα όπου απαιτείται στενή σύνδεση της επιχειρησιακής ανάλυσης με τον σχεδιασμό λογισμικού. Η ενσωμάτωση των Activity Diagrams στο οικοσύστημα της UML διευκολύνει τη διασύνδεσή τους με άλλα διαγράμματα και την υιοθέτηση μιας ενιαίας προσέγγισης μοντελοποίησης (Hettab et al., 2015).

Συνολικά, τα UML Activity Diagrams αποτελούν ένα ευέλικτο και ισχυρό εργαλείο για τη μοντελοποίηση επιχειρησιακών διαδικασιών. Η δυνατότητα συνδυασμού λειτουργικής και οργανωτικής πληροφορίας, σε συνδυασμό με την απλότητα και την τυποποίησή τους, τα καθιστά ιδιαίτερα χρήσιμα τόσο σε ακαδημαϊκό όσο και σε επαγγελματικό επίπεδο (Hettab

et al., 2015).

2.3 Το Λογισμικό Μοντελοποίησης Visual Paradigm

2.3.1 Τεχνικά χαρακτηριστικά και λειτουργικότητα του Visual Paradigm στο πλαίσιο της μοντελοποίησης συστημάτων

Η επιλογή του Visual Paradigm ως το κεντρικό εργαλείο μοντελοποίησης για την παρούσα διπλωματική εργασία βασίστηκε στην ικανότητά του να προσφέρει ένα ολοκληρωμένο και αυστηρά δομημένο περιβάλλον σχεδίασης. Στο πλαίσιο του εικονικού εργαστηρίου OnLabs, η ανάγκη για μια ακριβή αναπαράσταση της πειραματικής διαδικασίας καθιστά απαραίτητη τη χρήση ενός προηγμένου εργαλείου. Το Visual Paradigm ξεπερνά τα όρια μιας απλής ζωγραφικής απεικόνισης, καθώς επιτρέπει στον ερευνητή να ενσωματώσει κανόνες και περιορισμούς απευθείας πάνω στο μοντέλο. Με τον τρόπο αυτό, οι λεκτικές οδηγίες ενός εργαστηριακού πρωτοκόλλου μετατρέπονται σε ένα ψηφιακό μοντέλο αναφοράς, όπου κάθε ενέργεια (Action) και κάθε ροή ελέγχου (Control Flow) ακολουθεί πιστά τους κανόνες της UML 2.5.

Ένα από τα πιο χρήσιμα χαρακτηριστικά του λογισμικού είναι η δυνατότητα οργάνωσης των δραστηριοτήτων μέσω της χρήσης Swimlanes (Partitioning). Στην εργασία αυτή, η συγκεκριμένη λειτουργία αξιοποιείται για να διαχωρίσουμε ξεκάθαρα τις οντότητες που συμμετέχουν στο πείραμα. Έτσι, οι ενέργειες κατηγοριοποιούνται ανάλογα με το αν αφορούν τον εργαστηριακό εξοπλισμό (όπως οι ρυθμίσεις του μικροσκοπίου), τα αναλώσιμα υλικά (όπως η αντικειμενοφόρος πλάκα και το δείγμα) ή τις κινήσεις του ίδιου του χρήστη. Αυτός ο διαχωρισμός βοηθά σημαντικά στη μετέπειτα ανάλυση, καθώς επιτρέπει την εύκολη σύνδεση των ψηφιακών ιχνών που παράγει το OnLabs με τις αντίστοιχες λογικές ενότητες του μοντέλου μας.

Επιπλέον, το Visual Paradigm διαθέτει ενσωματωμένους μηχανισμούς ελέγχου (Model Validation), οι οποίοι διασφαλίζουν ότι το τελικό διάγραμμα δεν περιέχει λογικά σφάλματα, όπως αδιέξοδα (deadlocks) ή αποκομμένα βήματα που δεν οδηγούν πουθενά. Παράλληλα, η δυνατότητα να ορίζονται προϋποθέσεις (Pre-conditions) και αποτελέσματα (Post-conditions) για κάθε δραστηριότητα, μας επιτρέπει να χτίσουμε ένα μοντέλο που περιγράφει την ορθή αλληλουχία του πειράματος. Για παράδειγμα, μπορούμε να ορίσουμε

ότι η εστίαση του δείγματος είναι εφικτή μόνο αφού έχει ενεργοποιηθεί πρώτα η φωτεινή πηγή. Αυτή η αυστηρότητα στη σχεδίαση είναι απαραίτητη για τη φάση του Conformance Checking, καθώς μας δίνει έναν αδιαμφισβήτητο «οδικό χάρτη» για να κρίνουμε αν οι χρήστες λειτούργησαν σωστά στο εικονικό περιβάλλον.

2.3.2 Η συμβολή του Visual Paradigm στη γέφυρωση μοντελοποίησης και εξόρυξης διεργασιών

Η συνεισφορά του Visual Paradigm εκτείνεται πέρα από τη σχεδίαση, καθώς αποτελεί τη γέφυρα που ενώνει τη θεωρητική μοντελοποίηση με την εφαρμοσμένη Εξόρυξη Διεργασιών. Το μεγάλο πλεονέκτημα του λογισμικού σε αυτή τη μεθοδολογία είναι η εξαιρετική διαλειτουργικότητα που προσφέρει. Δίνει τη δυνατότητα να εξάγουμε τα μοντέλα που σχεδιάσαμε σε διεθνή ψηφιακά πρότυπα, όπως το BPMN 2.0 και το format XML (XMI). Η λειτουργία αυτή είναι κομβικής σημασίας, επειδή τα εξειδικευμένα εργαλεία Process Mining (όπως το ProM και το Disco) χρειάζονται τέτοιου είδους δομημένα αρχεία για να εκτελέσουν τους αλγορίθμους τους. Με απλά λόγια, το Visual Paradigm αναλαμβάνει να «μεταφράσει» τις εικόνες του διαγράμματος σε μηχαναγνώσιμη μορφή, έτοιμη να συγκριθεί με τα event logs του OnLabs.

Κατά τον σχεδιασμό του πειράματος, το εργαλείο επιτρέπει τον ακριβή προσδιορισμό των σημείων απόφασης (Decision Nodes) και των επαναλαμβανόμενων διαδικασιών (Iterative Loops). Αυτά τα στοιχεία αποτελούν τον πυρήνα της ανάλυσης στο Process Mining. Επιπλέον, η δυνατότητα του Visual Paradigm να προσθέτει metadata (συμπληρωματικές πληροφορίες) σε κάθε στοιχείο του διαγράμματος, βοηθά στην ακριβή αντιστοίχιση των βημάτων. Έτσι, μειώνεται η ασάφεια κατά τη σύγκριση των δεδομένων, αφού κάθε ψηφιακό ίχνος στο log (για παράδειγμα η εντολή `rotated mic_fine_focus_knob`) συνδέεται αυτόματα και μονοσήμαντα με έναν συγκεκριμένο κόμβο στο μοντέλο αναφοράς. Αυτό επιτρέπει στους αλγορίθμους να υπολογίσουν με ακρίβεια δείκτες όπως η πιστότητα (fitness) της ροής.

Τέλος, η χρήση του Visual Paradigm ενισχύει την επιστημονική τεκμηρίωση της εργασίας μέσα από τη δυνατότητα αυτόματης παραγωγής αναλυτικών αναφορών (Report Generation). Οι αναφορές αυτές καταγράφουν όλες τις ιδιότητες και τους περιορισμούς που έχουν τεθεί στο μοντέλο, δίνοντας μια καθαρή εικόνα για το ποιο είναι το «ιδανικό

μονοπάτι» του πειράματος. Όταν, λοιπόν, οι αλγόριθμοι του Process Mining εντοπίσουν μια απόκλιση στο event log, μπορούμε να ανατρέξουμε στην τεκμηρίωση του Visual Paradigm για να κατανοήσουμε ποιος κανόνας παραβιάστηκε. Συμπερασματικά, το Visual Paradigm δεν είναι απλώς ένα πρόγραμμα σχεδίασης, αλλά ένα πολύτιμο αναλυτικό εργαλείο που υποστηρίζει ολόκληρη τη μεθοδολογική δομή της διπλωματικής εργασίας.

2.4 Η χρήση των UML Activity Diagrams για την αναπαράσταση διαδικασιών

Η αναπαράσταση διαδικασιών αποτελεί θεμελιώδες στοιχείο στην ανάλυση και κατανόηση σύνθετων συστημάτων, ανεξαρτήτως του πεδίου εφαρμογής τους. Διαδικασίες μπορούν να εντοπιστούν σε πλήθος περιβαλλόντων, όπως πληροφοριακά συστήματα, υπολογιστικές ροές, τεχνικά συστήματα, οργανωσιακές λειτουργίες και επιστημονικά πρωτόκολλα. Η ανάγκη για τυποποιημένη, κατανοητή και επεκτάσιμη αναπαράσταση αυτών των διαδικασιών έχει οδηγήσει στην ανάπτυξη διαφόρων γλωσσών και μεθοδολογιών μοντελοποίησης. Στο πλαίσιο αυτό, η Unified Modeling Language (UML) προσφέρει ένα ευρέως αποδεκτό σύνολο διαγραμμάτων, μεταξύ των οποίων τα Activity Diagrams κατέχουν κεντρική θέση για την απεικόνιση της ροής ελέγχου και της δυναμικής συμπεριφοράς των διαδικασιών (Ahmad et al., 2019).

Τα UML Activity Diagrams έχουν σχεδιαστεί για να αναπαριστούν την ακολουθία ενεργειών, τις συνθήκες απόφασης και την παράλληλη εκτέλεση δραστηριοτήτων μέσα σε μια διαδικασία. Σε αντίθεση με διαγράμματα που εστιάζουν στη δομή ή στις στατικές σχέσεις μεταξύ οντοτήτων, τα Activity Diagrams αποτυπώνουν τη χρονική εξέλιξη και τη λογική ροή μιας διαδικασίας, καθιστώντας τα κατάλληλα για την περιγραφή σύνθετων συμπεριφορών. Η γενικότητα της προσέγγισης αυτής επιτρέπει την εφαρμογή τους σε ποικίλα πεδία, από την ανάλυση αλγορίθμων και ροών εργασίας έως τη μοντελοποίηση αλληλεπιδράσεων μεταξύ συστημάτων και χρηστών.

Κεντρικό στοιχείο των UML Activity Diagrams είναι οι δραστηριότητες (activities) και οι ενέργειες (actions), οι οποίες αναπαριστούν τα επιμέρους βήματα μιας διαδικασίας. Η σύνδεση μεταξύ αυτών επιτυγχάνεται μέσω ροών ελέγχου (control flows), οι οποίες καθορίζουν τη σειρά εκτέλεσης. Επιπλέον, οι κόμβοι απόφασης (decision nodes) επιτρέπουν τη μοντελοποίηση εναλλακτικών μονοπατιών εκτέλεσης βάσει λογικών συνθηκών, ενώ οι

κόμβοι συγχώνευσης (merge nodes) χρησιμοποιούνται για την επανένωση διαφορετικών ροών. Η ύπαρξη κόμβων διακλάδωσης και συγχρονισμού (fork και join nodes) επιτρέπει την αναπαράσταση παράλληλων ή ταυτόχρονων δραστηριοτήτων, χαρακτηριστικό που απαντάται συχνά σε πραγματικές διαδικασίες ανεξαρτήτως τομέα εφαρμογής (Ahmad et al., 2019).

Ένα επιπλέον χαρακτηριστικό που ενισχύει τη χρησιμότητα των Activity Diagrams είναι η δυνατότητα χρήσης partitions ή swimlanes. Μέσω αυτών, οι δραστηριότητες μπορούν να ομαδοποιηθούν βάσει υπεύθυνων οντοτήτων, όπως ρόλοι, υποσυστήματα ή τεχνικές συνιστώσες. Η δυνατότητα αυτή διευκολύνει την ανάλυση της κατανομής ευθυνών και την κατανόηση της αλληλεπίδρασης μεταξύ διαφορετικών στοιχείων ενός συστήματος. Στο πλαίσιο της γενικής αναπαράστασης διαδικασιών, τα swimlanes δεν περιορίζονται σε οργανωσιακούς ρόλους, αλλά μπορούν να χρησιμοποιηθούν για τη διάκριση μεταξύ λογικών ή φυσικών συνιστωσών ενός συστήματος.

Τα UML Activity Diagrams χρησιμοποιούνται συχνά ως εργαλείο εννοιολογικής μοντελοποίησης, ιδίως στα αρχικά στάδια ανάλυσης μιας διαδικασίας. Το υψηλό επίπεδο αφαίρεσης και η γραφική φύση τους επιτρέπουν την αποτύπωση της βασικής λογικής μιας διαδικασίας χωρίς να απαιτείται λεπτομερής τεχνική εξειδίκευση. Η ιδιότητα αυτή καθιστά τα Activity Diagrams ιδιαίτερα χρήσιμα για την κατανόηση και τεκμηρίωση διαδικασιών, καθώς και για την επικοινωνία μεταξύ διαφορετικών εμπλεκόμενων μερών, όπως αναλυτές, σχεδιαστές και τελικούς χρήστες (Ding et al., 2016).

Παράλληλα, τα Activity Diagrams μπορούν να λειτουργήσουν ως ενδιάμεση αναπαράσταση μεταξύ άτυπων περιγραφών διαδικασιών και πιο αυστηρών τυπικών μοντέλων. Σε περιπτώσεις όπου απαιτείται περαιτέρω ανάλυση ή αυτοματοποιημένη επεξεργασία, τα Activity Diagrams μπορούν να μετατραπούν ή να χαρτογραφηθούν σε φορμαλισμούς με αυστηρή σημασιολογία, όπως τα Petri nets. Η δυνατότητα αυτή είναι ιδιαίτερα σημαντική σε περιβάλλοντα όπου η ανάλυση της πραγματικής εκτέλεσης των διαδικασιών βασίζεται σε δεδομένα, όπως στο process mining (Ding et al., 2016).

Στο πλαίσιο του process mining, οι διαδικασίες αναλύονται με βάση αρχεία καταγραφής γεγονότων, τα οποία αποτυπώνουν την πραγματική ακολουθία εκτέλεσης ενεργειών σε ένα σύστημα. Τα UML Activity Diagrams μπορούν να χρησιμοποιηθούν για την αναπαράσταση

του αναμενόμενου ή προβλεπόμενου τρόπου εκτέλεσης μιας διαδικασίας, παρέχοντας ένα εννοιολογικό μοντέλο αναφοράς. Αν και τα Activity Diagrams δεν διαθέτουν εγγενώς την τυπική σημασιολογία που απαιτείται για την απευθείας εφαρμογή αλγορίθμων process mining, συμβάλλουν ουσιαστικά στην κατανόηση της δομής και της λογικής της διαδικασίας πριν από την εφαρμογή πιο αυστηρών αναλυτικών τεχνικών.

Η συμβολή των UML Activity Diagrams είναι επίσης εμφανής στο πλαίσιο του conformance checking, το οποίο αφορά τον έλεγχο της συμμόρφωσης μεταξύ ενός μοντέλου διαδικασίας και των πραγματικών εκτελέσεών της. Τα Activity Diagrams μπορούν να αποτυπώσουν τις επιτρεπόμενες ροές, τις συνθήκες και τους περιορισμούς μιας διαδικασίας, λειτουργώντας ως βάση για τον εντοπισμό αποκλίσεων. Παρότι ο έλεγχος συμμόρφωσης πραγματοποιείται συνήθως πάνω σε τυπικά μοντέλα, η χρήση Activity Diagrams διευκολύνει την εννοιολογική ερμηνεία των αποτελεσμάτων και την κατανόηση των αποκλίσεων από τη σκοπιά του σχεδιασμού της διαδικασίας (Jena et al., 2014).

Παρά τα πλεονεκτήματά τους, τα UML Activity Diagrams παρουσιάζουν και ορισμένους περιορισμούς. Η έλλειψη αυστηρής τυπικής σημασιολογίας μπορεί να οδηγήσει σε ασάφειες, ιδιαίτερα σε πολύπλοκες διαδικασίες με πολλαπλές εξαιρέσεις ή συγχρονισμούς. Επιπλέον, διαφορετικοί αναλυτές ενδέχεται να αποτυπώσουν την ίδια διαδικασία με διαφορετικό τρόπο, γεγονός που επηρεάζει τη συγκρισιμότητα των μοντέλων. Για τον λόγο αυτό, τα Activity Diagrams χρησιμοποιούνται συχνά συμπληρωματικά με άλλες τεχνικές και εργαλεία μοντελοποίησης.

Συμπερασματικά, τα UML Activity Diagrams αποτελούν ένα ευέλικτο και εκφραστικό μέσο για τη γενική αναπαράσταση διαδικασιών. Η δυνατότητά τους να αποτυπώνουν τη ροή ελέγχου, τις εναλλακτικές και την παράλληλη εκτέλεση δραστηριοτήτων τα καθιστά κατάλληλα για ποικίλα πεδία εφαρμογής. Στο πλαίσιο της παρούσας διπλωματικής εργασίας, τα Activity Diagrams λειτουργούν ως εννοιολογικό εργαλείο που γεφυρώνει τη θεωρητική περιγραφή των διαδικασιών με την ανάλυση της πραγματικής τους εκτέλεσης μέσω τεχνικών process mining και conformance checking.

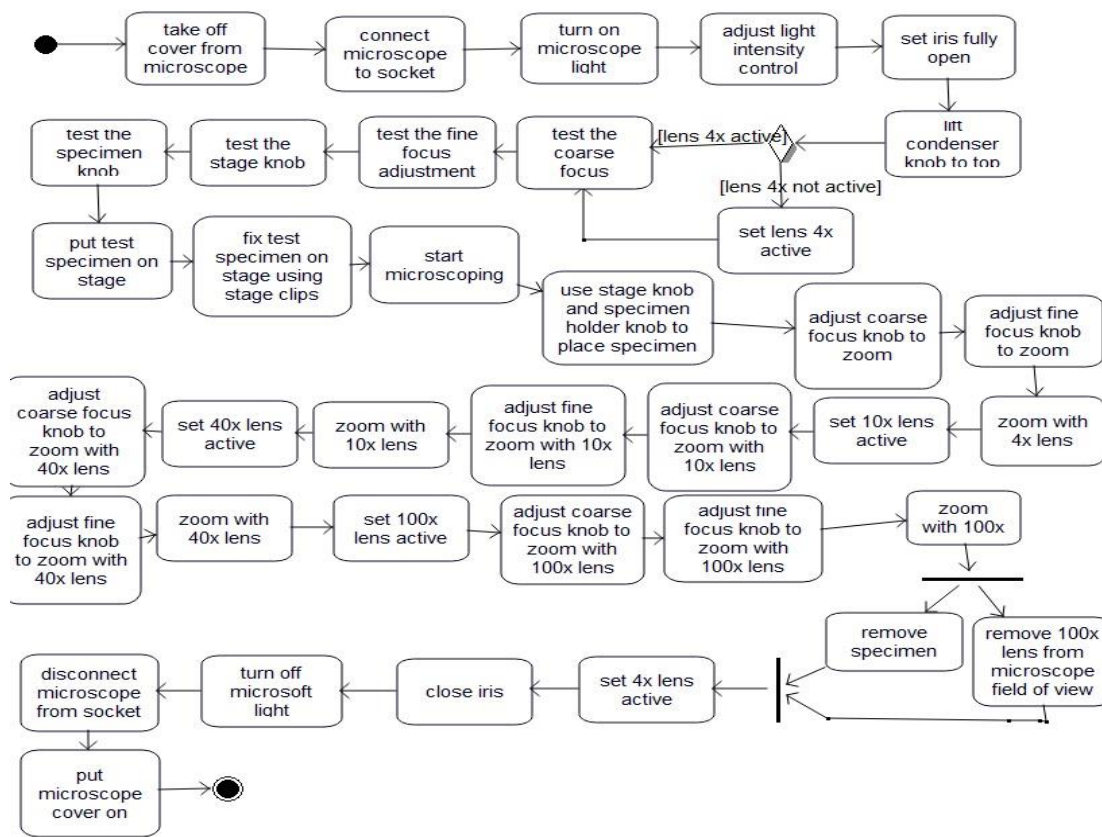
2.5 Σύνδεση των UML Activity Diagrams και της Εξόρυξης Διεργασιών

Η εξόρυξη διεργασιών (Process Mining) αποτελεί ένα σύγχρονο επιστημονικό πεδίο που

συνδυάζει στοιχεία από τη μοντελοποίηση διαδικασιών, την ανάλυση δεδομένων και τα πληροφοριακά συστήματα (Hettab et al., 2015). Στόχος της είναι η εξαγωγή γνώσης σχετικά με τη λειτουργία των επιχειρησιακών διαδικασιών, βασιζόμενη σε δεδομένα που προκύπτουν από την πραγματική εκτέλεσή τους. Τα δεδομένα αυτά καταγράφονται σε αρχεία γεγονότων, τα οποία περιέχουν πληροφορίες σχετικά με τις δραστηριότητες, τη χρονική τους ακολουθία και τους εμπλεκόμενους πόρους (Hettab et al., 2015).

Η σύνδεση των UML Activity Diagrams με την εξόρυξη διεργασιών προκύπτει από το κοινό τους αντικείμενο μελέτης, δηλαδή τη ροή των δραστηριοτήτων (Ahmad et al., 2019). Ενώ τα Activity Diagrams χρησιμοποιούνται κυρίως για τον εννοιολογικό σχεδιασμό και την τεκμηρίωση διαδικασιών, η εξόρυξη διεργασιών επιτρέπει την ανακάλυψη και ανάλυση μοντέλων που αντικατοπτρίζουν την πραγματική συμπεριφορά των συστημάτων (Ahmad et al., 2019). Τα UML Activity Diagrams μπορούν να χρησιμοποιηθούν ως αρχικά μοντέλα διαδικασιών, τα οποία στη συνέχεια συγκρίνονται με τα πραγματικά event logs μέσω τεχνικών Conformance Checking.

Το UML Activity Diagram (AD) που παρουσιάζεται στην παρακάτω εικόνα (**Εικόνα 10**) [Sypsas, 2023] αποτυπώνει τα αρχικά στάδια του πειράματος μικροσκοπικής που πραγματοποιείται στο Ελληνικό Ανοικτό Πανεπιστήμιο, στο πλαίσιο του εικονικού εργαστηρίου Onlabs.



Εικόνα 10: Αρχικά στάδια του πειράματος μικροσκοπίας στο εικονικό εργαστήριο Onlabs

Ένας βασικός τρόπος αξιοποίησης των UML Activity Diagrams στο πλαίσιο της εξόρυξης διεργασιών είναι η χρήση τους ως μοντέλων αναφοράς. Τα μοντέλα αυτά μπορούν να συγκριθούν με τα μοντέλα που προκύπτουν από την ανάλυση των event logs, με σκοπό τον εντοπισμό αποκλίσεων. Η διαδικασία αυτή, γνωστή ως έλεγχος συμμόρφωσης, επιτρέπει την αξιολόγηση του βαθμού στον οποίο η πραγματική εκτέλεση μιας διαδικασίας ευθυγραμμίζεται με τον αρχικό σχεδιασμό της (Ding et al., 2016).

Οι αποκλίσεις που εντοπίζονται μέσω του ελέγχου συμμόρφωσης μπορεί να οφείλονται σε διάφορους παράγοντες, όπως οργανωτικές παρεκκλίσεις, εξαιρέσεις, ή ανεπαρκή ορισμό της διαδικασίας. Η ανάλυση αυτών των αποκλίσεων παρέχει πολύτιμες πληροφορίες για τη βελτίωση τόσο των ίδιων των διαδικασιών όσο και των μοντέλων που τις περιγράφουν.

Παράλληλα, τα αποτελέσματα της εξόρυξης διεργασιών μπορούν να αξιοποιηθούν για τον εμπλουτισμό των UML Activity Diagrams. Μέσω της ανάλυσης πραγματικών δεδομένων, είναι δυνατόν να εντοπιστούν εναλλακτικές ροές, επαναλήψεις ή σπάνια σεναρία, τα οποία

δεν είχαν προβλεφθεί κατά τον αρχικό σχεδιασμό. Η ενσωμάτωση αυτών των ευρημάτων οδηγεί στη δημιουργία πιο ρεαλιστικών και ολοκληρωμένων μοντέλων (Ding et al., 2016).

Σε ερευνητικό επίπεδο, έχουν αναπτυχθεί διάφορες τεχνικές χαρτογράφησης μοντέλων εξόρυξης διεργασιών, όπως τα Petri Nets, σε UML Activity Diagrams. Η χαρτογράφηση αυτή διευκολύνει τη μετάβαση από τα μαθηματικά ή φορμαλιστικά μοντέλα της εξόρυξης διεργασιών σε πιο κατανοητές οπτικές αναπαραστάσεις, ενισχύοντας τη χρησιμότητά τους σε επιχειρησιακά περιβάλλοντα (Jena et al., 2014).

Η συνέργεια μεταξύ UML Activity Diagrams και εξόρυξης διεργασιών αναδεικνύει τη σημασία της συνδυαστικής προσέγγισης στη μελέτη και βελτίωση επιχειρησιακών διαδικασιών. Ενώ τα Activity Diagrams παρέχουν μια δομημένη και κατανοητή απεικόνιση σε επίπεδο σχεδιασμού, η εξόρυξη διεργασιών προσφέρει εμπειρικά δεδομένα και αναλυτικά εργαλεία που επιτρέπουν τη συνεχή αξιολόγηση και προσαρμογή των διαδικασιών στις πραγματικές συνθήκες λειτουργίας (Jena et al., 2014).

Στο πλαίσιο της παρούσας διπλωματικής εργασίας, η παραπάνω θεωρητική συνέργεια αποκτά μια απόλυτα πρακτική διάσταση. Μεταφέρεται, συγκεκριμένα, από το γενικό επίπεδο των επιχειρησιακών διαδικασιών στο εξειδικευμένο περιβάλλον της ψηφιακής εκπαίδευσης. Η ανάγκη για τη σύνδεση των UML Activity Diagrams με την εξόρυξη διεργασιών γίνεται εύκολα κατανοητή αν αναλογιστούμε τη φύση του εικονικού εργαστηρίου OnLabs. Το OnLabs δεν είναι ένα στατικό εκπαιδευτικό λογισμικό, αλλά ένα δυναμικό περιβάλλον αλληλεπίδρασης. Μέσα σε αυτό, κάθε ενέργεια του φοιτητή —όπως η περιστροφή ενός κοχλία ή η τοποθέτηση μιας αντικειμενοφόρου πλάκας— καταγράφεται και αφήνει ένα μοναδικό ψηφιακό αποτύπωμα. Επομένως, το Activity Diagram της Εικόνας 10 δεν αποτελεί μια απλή, εκ των υστέρων ζωγραφική αναπαράσταση. Αντίθετα, λειτουργεί ως ο επίσημος «οδικός χάρτης» του πειράματος, ο οποίος ορίζει εκ των προτέρων ποια είναι η επιστημονικά ορθή αλληλουχία των βημάτων.

Αυτή η καθοδηγητική ισχύς του UML Activity Diagram αποτελεί τη βάση για τον έλεγχο συμμόρφωσης (Conformance Checking) που εξετάζει η εργασία. Στην πράξη, η σύνδεση αυτή υλοποιείται μέσα από μια διαδικασία αντιστοίχισης (mapping). Με τον τρόπο αυτό, οι θεωρητικοί κόμβοι του διαγράμματος UML συνδέονται απευθείας με τις πραγματικές καταγραφές (Activities) του event log του OnLabs. Για παράδειγμα, το διάγραμμα ορίζει ότι

η ενεργοποίηση της φωτεινής πηγής πρέπει υποχρεωτικά να προηγείται της εστίασης. Η εξόρυξη διεργασιών έρχεται τότε να ελέγξει αν στο αρχείο .csv η δραστηριότητα `set mic_AC_switch on` εμφανίζεται όντως πριν από τη δραστηριότητα `rotated mic_fine_focus_knob`. Έτσι, το Activity Diagram μετατρέπεται από ένα στατικό σχέδιο σε ένα δυναμικό κριτήριο αξιολόγησης, με το οποίο το λογισμικό Process Mining (όπως το ProM) ελέγχει αυτόματα αν οι φοιτητές λειτούργησαν σωστά.

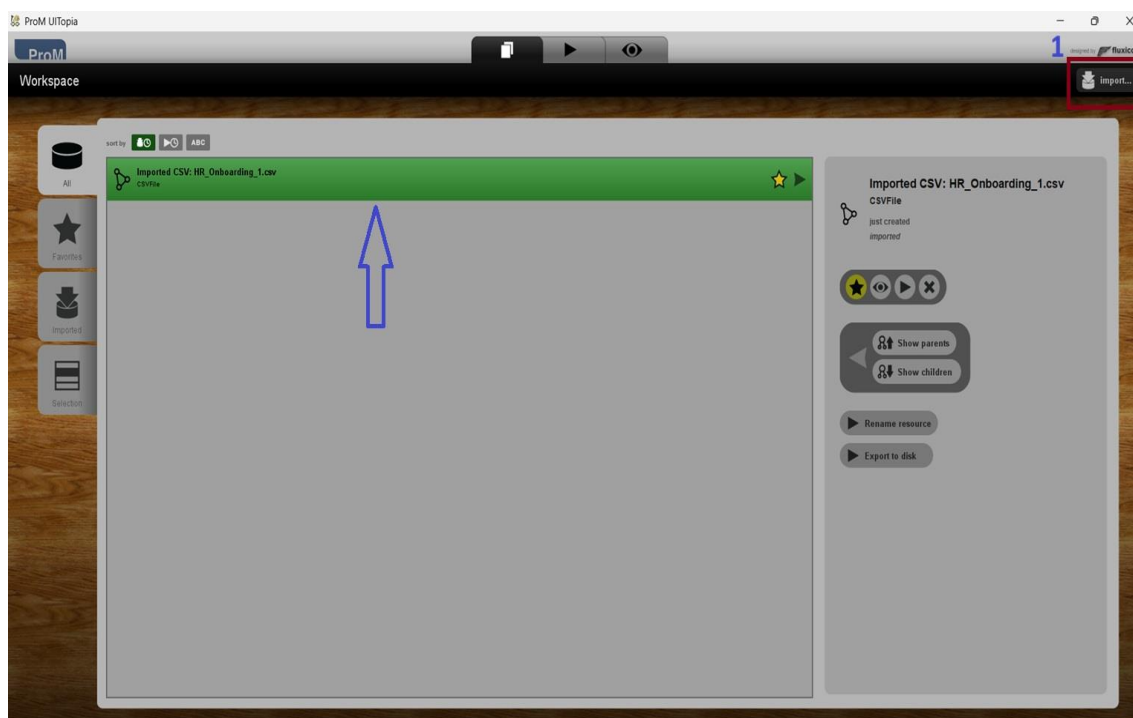
Η ουσιαστική σύνδεση των δύο εννοιών στην παρούσα μελέτη κλειδώνει πάνω στην ανάγκη ποσοτικής αξιολόγησης του πειράματος της μικροσκοπίας. Ενώ το Activity Diagram της Εικόνας 10 ορίζει τη σωστή δομή, δεν μπορεί να αποκαλύψει πώς συμπεριφέρονται οι φοιτητές σε πραγματικό χρόνο. Εδώ ακριβώς εισέρχεται η εξόρυξη διεργασιών, η οποία χρησιμοποιεί το διάγραμμα UML ως σταθερό άξονα σύγκρισης για να μετρήσει την πιστότητα (fitness) των φοιτητών. Η διασύνδεση αυτή επιτρέπει στους αλγορίθμους του ProM να επικάθονται πάνω στις ροές του Visual Paradigm και να αναδεικνύουν με ακρίβεια τις επαναλαμβανόμενες κινήσεις (loops) των χρηστών —όπως η άσκοπη εναλλαγή μεταξύ μακρομετρικού και μικρομετρικού κοχλία— καθώς και τα βήματα που προκαλούν τις μεγαλύτερες χρονικές καθυστερήσεις (bottlenecks).

Συνεπώς, η σύνδεση του UML Activity Diagram με το Process Mining αποτελεί τον κεντρικό μεθοδολογικό πυλώνα αυτής της εργασίας για την εξαγωγή εκπαιδευτικών συμπερασμάτων. Το διάγραμμα UML παρέχει τους απαραίτητους κανόνες και τους περιορισμούς του μικροσκοπίου, ενώ η εξόρυξη διεργασιών επεξεργάζεται τα δεδομένα του OnLabs για να αποδείξει σε ποια συγκεκριμένα βήματα του πρωτοκόλλου οι φοιτητές αποκλίνουν από το ορθό μονοπάτι. Αυτή η άμεση σύγκριση του σχεδιασμού με την πράξη επιτρέπει τον αυτόματο εντοπισμό των σφαλμάτων των χρηστών, προσφέροντας στην εργασία μια αντικειμενική μέθοδο για την αξιολόγηση και τη μελλοντική αναβάθμιση των ασκήσεων του εικονικού εργαστηρίου.

Προκειμένου να γίνει απόλυτα σαφής η παραπάνω θεωρητική ανάλυση, κρίνεται σκόπιμη η παρουσίαση ενός συγκεκριμένου και πρακτικού παραδείγματος με τη χρήση ενός αρχείου καταγραφής συμβάντων (event log). Το παράδειγμα αυτό αφορά τη Διαδικασία Υποδοχής και Ένταξης Νέου Προσωπικού (HR Onboarding Process) σε μια επιχείρηση. Η συγκεκριμένη διεργασία περιλαμβάνει όλες τις τυπικές διοικητικές ενέργειες που απαιτούνται για την ενσωμάτωση ενός υπαλλήλου στο νέο του εργασιακό περιβάλλον. Πιο

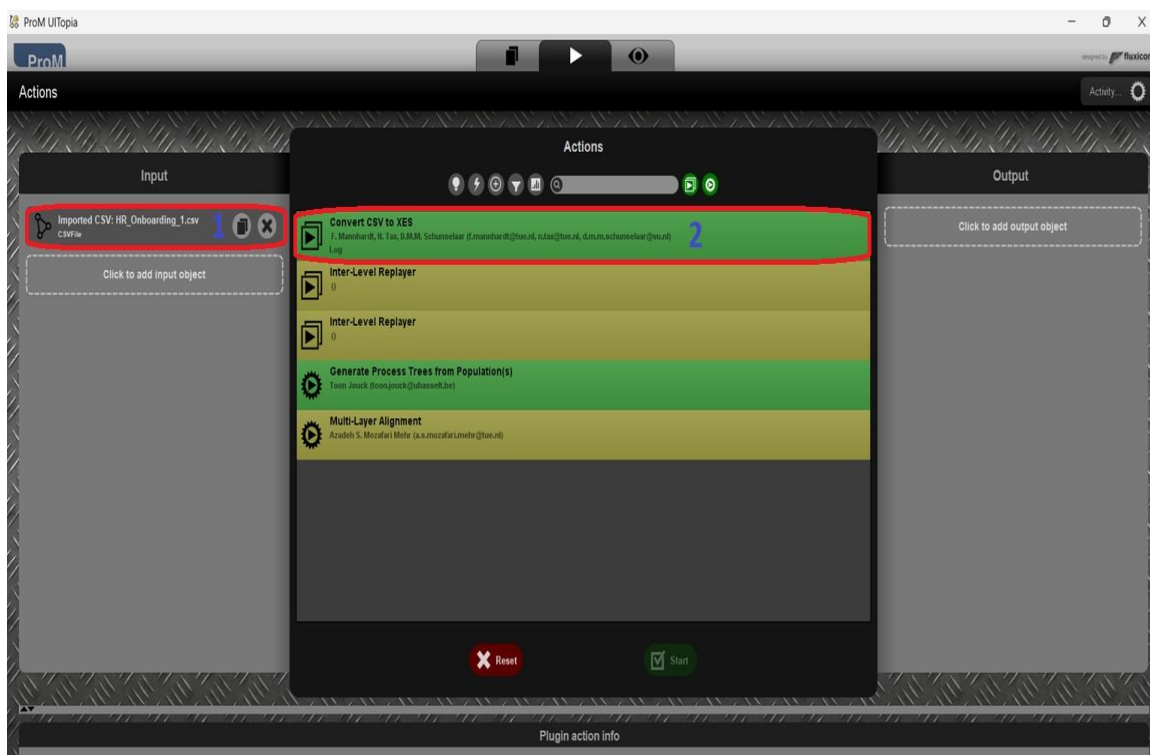
αναλυτικά, η ροή περιγράφει βήματα όπως ο έλεγχος και η επαλήθευση των απαραίτητων εγγράφων του υπαλλήλου, η προετοιμασία και η υπογραφή της σύμβασης εργασίας, καθώς και οι τελικές εγκρίσεις που απαιτούνται από το τμήμα Ανθρωπίνων Πόρων για την ολοκλήρωση της πρόσληψης.

Στη συνέχεια της ενότητας αυτής παρουσιάζεται η τεχνική ροή επεξεργασίας αυτών των δεδομένων, προκειμένου ναδειχθεί στην πράξη πώς οι καταγεγραμμένες ενέργειες μετατρέπονται σε οπτικά μοντέλα. Αρχικά, εισάγεται το αρχείο καταγραφής συμβάντων (event log) της διαδικασίας στο λογισμικό ProM, από το οποίο εξάγεται αυτόματα το πραγματικό μοντέλο της διεργασίας (Process Discovery). Αμέσως μετά, παρουσιάζεται ο τρόπος με τον οποίο το μοντέλο αυτό μετατρέπεται, μέσω του ίδιου εργαλείου, σε ένα τυποποιημένο Petri net, δίνοντας στη ροή τη μαθηματική δομή που απαιτείται για την ανάλυση συστημάτων. Τέλος, παρατίθεται το μοντέλο αναφοράς για τη συγκεκριμένη διαδικασία, προσφέροντας μια ολοκληρωμένη και καθαρή εικόνα για το πώς μια ροή δραστηριοτήτων μεταφράζεται από τη μορφή ενός log αρχείου σε ένα λειτουργικό μοντέλο.



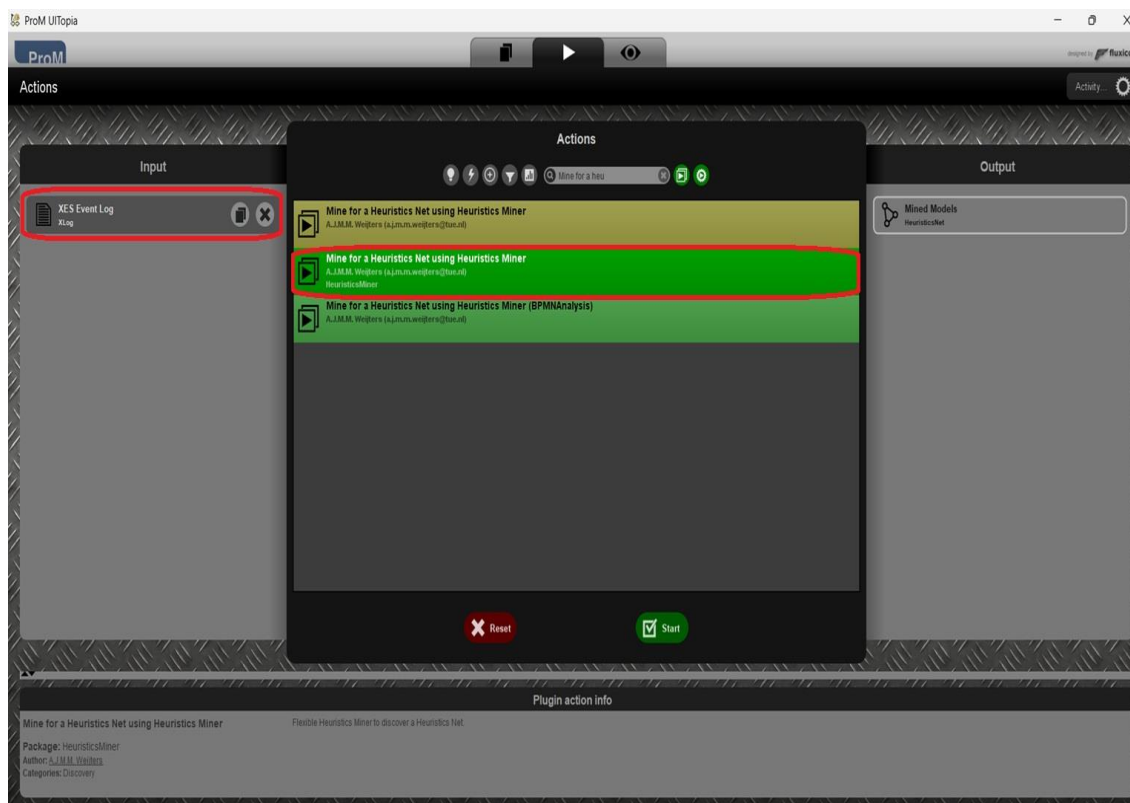
Εικόνα 11: Εισαγωγή του αρχείου καταγραφής συμβάντων στο ProM

Στο συγκεκριμένο στιγμιότυπο (**Εικόνα 11**) απεικονίζεται η επιτυχής εισαγωγή του αρχείου καταγραφής συμβάντων (event log) στο λογισμικό ProM. Το σύστημα επιβεβαιώνει την ανάγνωση των δεδομένων από το αρχείο .csv, το οποίο είναι πλέον διαθέσιμο ως αντικείμενο στο Workspace του προγράμματος. Αυτό το βήμα αποτελεί την αφετηρία της ανάλυσης, καθώς μετατρέπει τα ακατέργαστα δεδομένα της διεργασίας σε κατάλληλη μορφή για την εφαρμογή αλγορίθμων εξόρυξης.



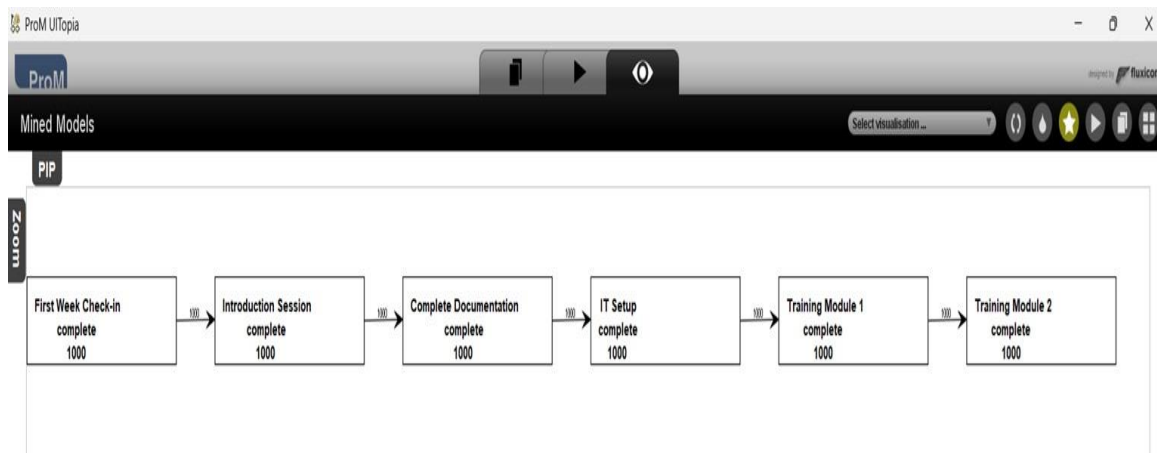
Εικόνα 12: Μετατροπή .csv αρχείου σε μορφή .xes

Στο συγκεκριμένο στάδιο (**Εικόνα 12**) πραγματοποιείται η μετατροπή του αρχείου .csv στο πρότυπο σχήμα .xes (eXtensible Event Stream), το οποίο αποτελεί το επίσημο αρχιτεκτονικό πρότυπο για το Process Mining. Η διαδικασία αυτή εξασφαλίζει τη συμβατότητα των δεδομένων με το σύνολο των αλγορίθμων του ProM, καθώς οργανώνει τις πληροφορίες σε μια αυστηρή ιεραρχική δομή υποθέσεων και γεγονότων. Με την ολοκλήρωση της μετατροπής, το log αποκτά την απαραίτητη τυποποίηση για την εξαγωγή ακριβών μοντέλων διεργασιών.

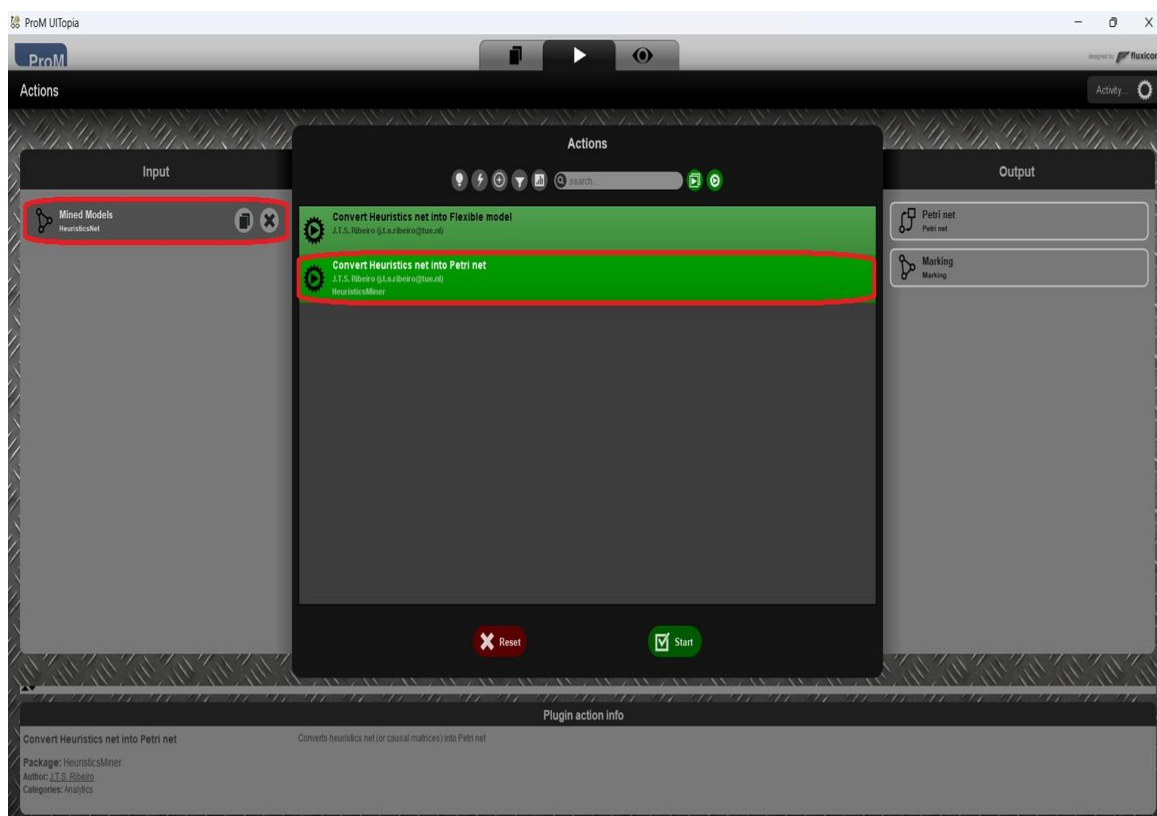


Εικόνα 13: Εξαγωγή του μοντέλου διεργασίας με χρήση του αλγόριθμου Heuristics Miner

Στο συγκεκριμένο στιγμιότυπο (**Εικόνα 13**) επιλέγεται ο αλγόριθμος Heuristics Miner για την ανακάλυψη του μοντέλου διεργασίας από το αρχείο .xes. Ο συγκεκριμένος αλγόριθμος είναι ιδανικός για την εξαγωγή μοντέλων που περιλαμβάνουν θόρυβο ή εξαιρέσεις, καθώς εστιάζει στις πιο συχνές διαδρομές και στις αιτιακές σχέσεις μεταξύ των δραστηριοτήτων. Με την εφαρμογή του, τα ακατέργαστα δεδομένα μετατρέπονται σε μια δομημένη οπτική αναπαράσταση της πραγματικής ροής της διεργασίας (**Εικόνα 14**).



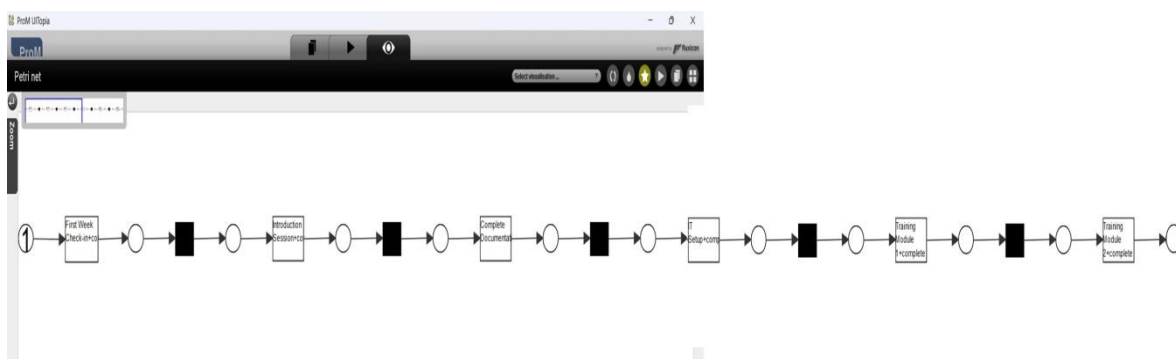
Εικόνα 14: Παραγωγή του μοντέλου διεργασίας με χρήση του Heuristics Miner



Εικόνα 15: Μετατροπή του Heuristics Net σε Petri Net

Στο συγκεκριμένο στιγμιότυπο (Εικόνα 15) επιλέγεται η ενέργεια μετατροπής του παραχθέντος Heuristics net σε δίκτυο Petri (Petri net). Η διαδικασία αυτή είναι καθοριστική,

καθώς μεταφράζει το αρχικό μοντέλο σε μια αυστηρή, μαθηματικά θεμελιωμένη γραφική αναπαράσταση, η οποία αποτελείται από θέσεις (places), μεταβάσεις (actions) και βέλη. Με τη συγκεκριμένη μετατροπή, η διεργασία HR Onboarding αποκτά μια τυποποιημένη μορφή που επιτρέπει τον έλεγχο της ορθότητας της ροής, καθώς και τη μετέπειτα διενέργεια αναλύσεων συμμόρφωσης (conformance checking).

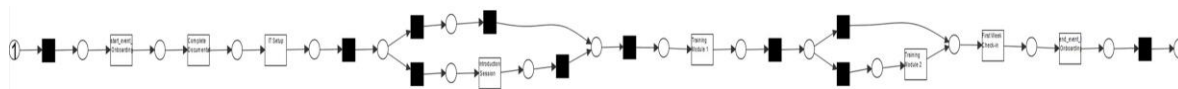


Εικόνα 16: Οπτικοποίηση του παραγόμενου Petri Net

Στο συγκεκριμένο στιγμιότυπο (**Εικόνα 16**) απεικονίζεται το τελικό μοντέλο διεργασίας σε μορφή Petri net, όπως αυτό προέκυψε μετά τη μετατροπή από το Heuristics net. Η οθόνη προβολής (View) του ProM εμφανίζει τη μαθηματική δομή του δικτύου, όπου οι κύκλοι αντιπροσωπεύουν τις θέσεις (places) και τα ορθογώνια κουτιά τις μεταβάσεις (activities/actions). Η γραφική αυτή αναπαράσταση επιτρέπει τη λεπτομερή μελέτη της δυναμικής συμπεριφοράς της διεργασίας HR Onboarding, καθιστώντας σαφή τη ροή, τις εξαρτήσεις και τη διαδοχή των δραστηριοτήτων από την αρχή έως το τέλος της.

Η συνδυαστική μελέτη του αρχικού event log (.csv) και του τελικά παραχθέντος Petri net επιτρέπει την εξαγωγή ασφαλών συμπερασμάτων για τη δομή και τη συμπεριφορά της διεργασίας HR Onboarding. Το Petri net επιβεβαιώνει ότι η ροή ακολουθεί μια εξαιρετικά δομημένη και γραμμική αλληλουχία 6.000 εγγραφών, η οποία ξεκινά σταθερά με το *First Week Check-in* και ολοκληρώνεται χωρίς αποκλίσεις στο *Training Module 2*. Η απουσία παράλληλων διακλαδώσεων ή βρόχων ανατροφοδότησης (loops) στο μοντέλο υποδεικνύει ότι η διαδικασία ένταξης των νέων υπαλλήλων είναι πλήρως τυποποιημένη και εκτελείται με αυστηρή πειθαρχία, διασφαλίζοντας τη σταθερότητα και την προβλεψιμότητα των

επιχειρησιακών ροών του οργανισμού.



Εικόνα 17: Δομή του Μοντέλου Αναφοράς σε Μορφή Petri Net

Στο συγκεκριμένο στιγμιότυπο (**Εικόνα 17**) απεικονίζεται το Petri net που αποτελεί το επίσημο μοντέλο αναφοράς (reference model) για τη διεργασία HR Onboarding του οργανισμού. Η γραφική αυτή αναπαράσταση αποτυπώνει τον ιδανικό σχεδιασμό της διαδικασίας, ενσωματώνοντας συγκεκριμένους επιχειρησιακούς κανόνες και πύλες ελέγχου. Αυτό γίνεται εμφανές από τις διακλαδώσεις και τις εναλλακτικές διαδρομές που επιτρέπει το δίκτυο σε σημεία όπως το *Introduction Session* και το *Training Module 2*, προσφέροντας μια ολοκληρωμένη και τυποποιημένη βάση για τη μετέπειτα αξιολόγηση της πραγματικής ροής.

ΚΕΦΑΛΑΙΟ 3

Εργαλεία Ανοικτού Κώδικα για Εξόρυξη Διεργασιών

3.1 Εισαγωγή

Η Εξόρυξη Διεργασιών (Process Mining) αποτελεί έναν σύγχρονο και ταχέως αναπτυσσόμενο ερευνητικό και εφαρμοστικό τομέα που γεφυρώνει το χάσμα μεταξύ της διαχείρισης επιχειρησιακών διεργασιών (Business Process Management – BPM), της επιστήμης δεδομένων και της υπολογιστικής νοημοσύνης (van der Aalst, 2016; Reinkemeyer, 2020; van der Aalst & Carmona, 2022). Στόχος της είναι η ανακάλυψη, ανάλυση και βελτίωση διεργασιών μέσω της αξιοποίησης καταγεγραμμένων δεδομένων γεγονότων (event logs), τα οποία παράγονται από πληροφοριακά συστήματα όπως ERP, CRM και συστήματα διαχείρισης ροών εργασίας. Σε αντίθεση με τις παραδοσιακές προσεγγίσεις μοντελοποίησης διεργασιών, οι οποίες βασίζονται κυρίως σε συνεντεύξεις και υποκειμενικές περιγραφές, η Εξόρυξη Διεργασιών προσφέρει μια αντικειμενική, δεδομενοκεντρική προσέγγιση που αντανάκλα την πραγματική εκτέλεση των διεργασιών (van der Aalst, 2011).

Στο πλαίσιο αυτό, ιδιαίτερη σημασία έχει η διαθεσιμότητα αξιόπιστων και επεκτάσιμων εργαλείων λογισμικού που υποστηρίζουν τις βασικές κατηγορίες της Εξόρυξης Διεργασιών, δηλαδή την ανακάλυψη διεργασιών (process discovery), τον έλεγχο συμμόρφωσης (conformance checking) και τη βελτίωση ή επέκταση μοντέλων (process enhancement) (van der Aalst, 2016; Reinkemeyer, 2020). Τα εργαλεία ανοικτού κώδικα διαδραματίζουν καθοριστικό ρόλο τόσο στην ακαδημαϊκή έρευνα όσο και στην πρακτική εφαρμογή, καθώς προάγουν τη διαφάνεια, την αναπαραγωγικότητα των πειραμάτων και τη συνεχή εξέλιξη μέσω της συνεισφοράς της ερευνητικής κοινότητας. Ανάμεσα στα πιο διαδεδομένα και ώριμα εργαλεία ανοικτού κώδικα στον χώρο της Εξόρυξης Διεργασιών συγκαταλέγονται το ProM (<https://promtools.org/>) και το PM4Py (<https://processintelligence.solutions/pm4py>).

Το ProM αποτελεί ένα από τα πρώτα και πλέον καθιερωμένα εργαλεία Εξόρυξης Διεργασιών και έχει αναπτυχθεί κυρίως στο ακαδημαϊκό περιβάλλον. Πρόκειται για ένα πλαίσιο (framework) βασισμένο στη γλώσσα προγραμματισμού Java, το οποίο υιοθετεί μια

αρθρωτή αρχιτεκτονική μέσω της χρήσης προσθέτων (plugins). Η αρχιτεκτονική αυτή επιτρέπει την ενσωμάτωση και τον πειραματισμό με μεγάλο αριθμό αλγορίθμων και τεχνικών, καλύπτοντας ένα ευρύ φάσμα λειτουργιών, όπως η ανακάλυψη μοντέλων Petri Nets, BPMN και Heuristic Nets, ο έλεγχος συμμόρφωσης μέσω ευθυγραμμίσεων (alignments) και η ανάλυση επιδόσεων διεργασιών. Το ProM έχει καθιερωθεί ως σημείο αναφοράς στην έρευνα, καθώς χρησιμοποιείται ευρέως για την υλοποίηση και αξιολόγηση νέων αλγορίθμων, ενώ παράλληλα προσφέρει γραφικό περιβάλλον χρήστη που διευκολύνει την πειραματική ανάλυση (van der Aalst & Carmona, 2022).

Ωστόσο, παρά τα πλεονεκτήματά του, το ProM παρουσιάζει και ορισμένους περιορισμούς, ιδίως σε ότι αφορά την ευχρηστία και την ενσωμάτωσή του σε σύγχρονες ροές ανάλυσης δεδομένων. Η ανάγκη κατανόησης της εσωτερικής δομής των plugins και η σχετικά υψηλή καμπύλη εκμάθησης μπορεί να αποτελέσουν εμπόδιο για χρήστες χωρίς ισχυρό τεχνικό υπόβαθρο. Επιπλέον, η διαχείριση μεγάλων συνόλων δεδομένων ενδέχεται να οδηγήσει σε αυξημένες απαιτήσεις σε υπολογιστικούς πόρους, γεγονός που καθιστά αναγκαία την αναζήτηση πιο ευέλικτων λύσεων.

Σε αυτό το πλαίσιο, το PM4Py έχει αναδειχθεί ως ένα σύγχρονο εργαλείο ανοικτού κώδικα που στοχεύει στην κάλυψη αυτών των αναγκών, αξιοποιώντας τη δυναμική της γλώσσας προγραμματισμού Python (Berti, van Zelst, & van der Aalst, 2019). Το PM4Py έχει σχεδιαστεί ως βιβλιοθήκη (library) και όχι ως αυτόνομο γραφικό εργαλείο, γεγονός που το καθιστά ιδιαίτερα κατάλληλο για ενσωμάτωση σε pipelines ανάλυσης δεδομένων, εφαρμογές μηχανικής μάθησης και περιβάλλοντα όπως τα Jupyter Notebooks. Παρέχει υποστήριξη για τα βασικά πρότυπα καταγραφής γεγονότων, όπως το XES, καθώς και για πληθώρα αλγορίθμων ανακάλυψης διεργασιών, συμπεριλαμβανομένων των Alpha Miner, Inductive Miner και Heuristics Miner.

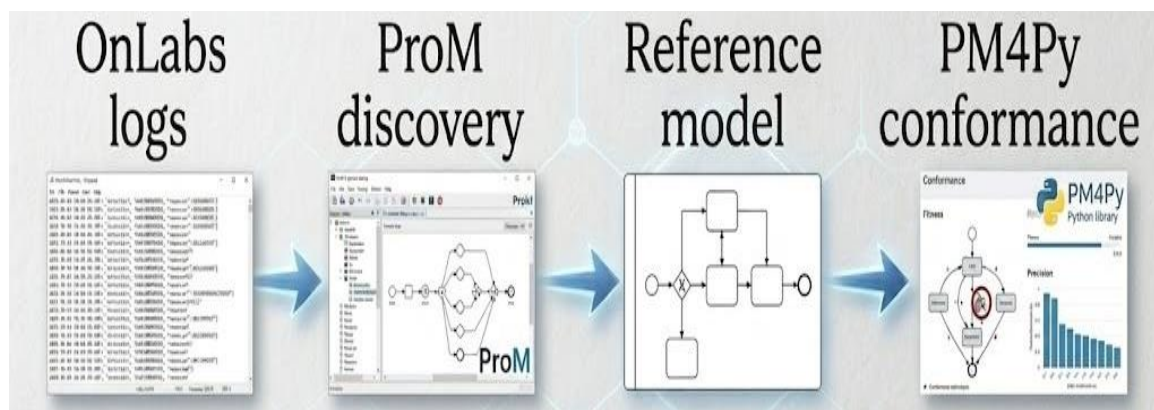
Ένα από τα βασικά πλεονεκτήματα του PM4Py είναι η ευελιξία και η επεκτασιμότητά του, καθώς επιτρέπει στον χρήστη να προσαρμόζει εύκολα τη λειτουργικότητα στις ανάγκες της εκάστοτε ανάλυσης. Η στενή του διασύνδεση με δημοφιλείς βιβλιοθήκες ανάλυσης δεδομένων της Python, όπως οι pandas και numpy, διευκολύνει την προεπεξεργασία, τον καθαρισμό και τη στατιστική ανάλυση των event logs. Παράλληλα, η δυνατότητα αυτοματοποίησης επαναλαμβανόμενων πειραμάτων το καθιστά ιδιαίτερα ελκυστικό για ερευνητικές και βιομηχανικές εφαρμογές μεγάλης κλίμακας (Berti, van Zelst, & van der

Aalst, 2019).

Συγκρίνοντας τα δύο εργαλεία, προκύπτει ότι το ProM και το PM4Py εξυπηρετούν συμπληρωματικούς ρόλους στον χώρο της Εξόρυξης Διεργασιών. Το ProM υπερέχει ως ερευνητική πλατφόρμα για την ανάπτυξη και αξιολόγηση νέων μεθόδων, προσφέροντας πλούσιο φάσμα έτοιμων αλγορίθμων και οπτικοποιήσεων. Αντίθετα, το PM4Py δίνει έμφαση στην πρακτική εφαρμογή και στην ενσωμάτωση με σύγχρονες τεχνολογίες ανάλυσης δεδομένων, καθιστώντας το ιδανικό για αυτοματοποιημένες και επαναλήψιμες αναλύσεις. Η επιλογή του κατάλληλου εργαλείου εξαρτάται από το εκάστοτε ερευνητικό ή επιχειρησιακό πλαίσιο, καθώς και από τις απαιτήσεις σε επίπεδο ευχρηστίας, επεκτασιμότητας και υπολογιστικής απόδοσης.

Συνοψίζοντας, τα εργαλεία ανοικτού κώδικα ProM και PM4Py διαδραματίζουν κομβικό ρόλο στην εξέλιξη και τη διάδοση της Εξόρυξης Διεργασιών. Η συνδυαστική τους χρήση μπορεί να προσφέρει μια ολοκληρωμένη προσέγγιση στην ανάλυση επιχειρησιακών διεργασιών, αξιοποιώντας τα πλεονεκτήματα κάθε πλατφόρμας (van der Aalst, 2016; Berti, van Zelst, & van der Aalst, 2019). Στο πλαίσιο της παρούσας διπλωματικής εργασίας, η μελέτη και αξιολόγηση των εργαλείων αυτών συμβάλλει στην κατανόηση των δυνατοτήτων και των περιορισμών τους, καθώς και στη διερεύνηση της καταλληλότητάς τους για την υποστήριξη σύγχρονων εφαρμογών Εξόρυξης Διεργασιών.

Η παρούσα εργασία εστιάζει στην υλοποίηση μιας ολοκληρωμένης ροής εργασιών (pipeline) για την ανάλυση των δεδομένων της πλατφόρμας OnLabs. Η διαδικασία εκκινεί με την εξαγωγή του αρχείου καταγραφής συμβάντων από το OnLabs και τη χρήση του εργαλείου ProM για την αυτόματη ανακάλυψη της πραγματικής διεργασίας (process discovery). Παράλληλα, αξιοποιείται ένα καθορισμένο μοντέλο αναφοράς (reference model), το οποίο λειτουργεί ως το πρότυπο σύγκρισης για την αξιολόγηση της ορθότητας των ενεργειών. Η ροή ολοκληρώνεται με τη χρήση της βιβλιοθήκης PM4Py, όπου πραγματοποιείται ο έλεγχος συμμόρφωσης (conformance checking) μεταξύ του αρχείου καταγραφής συμβάντων και του πρότυπου μοντέλου. Η διαδοχή των σταδίων αυτών και η αρχιτεκτονική της προσέγγισης απεικονίζονται στην παρακάτω εικόνα (**Εικόνα 18**).

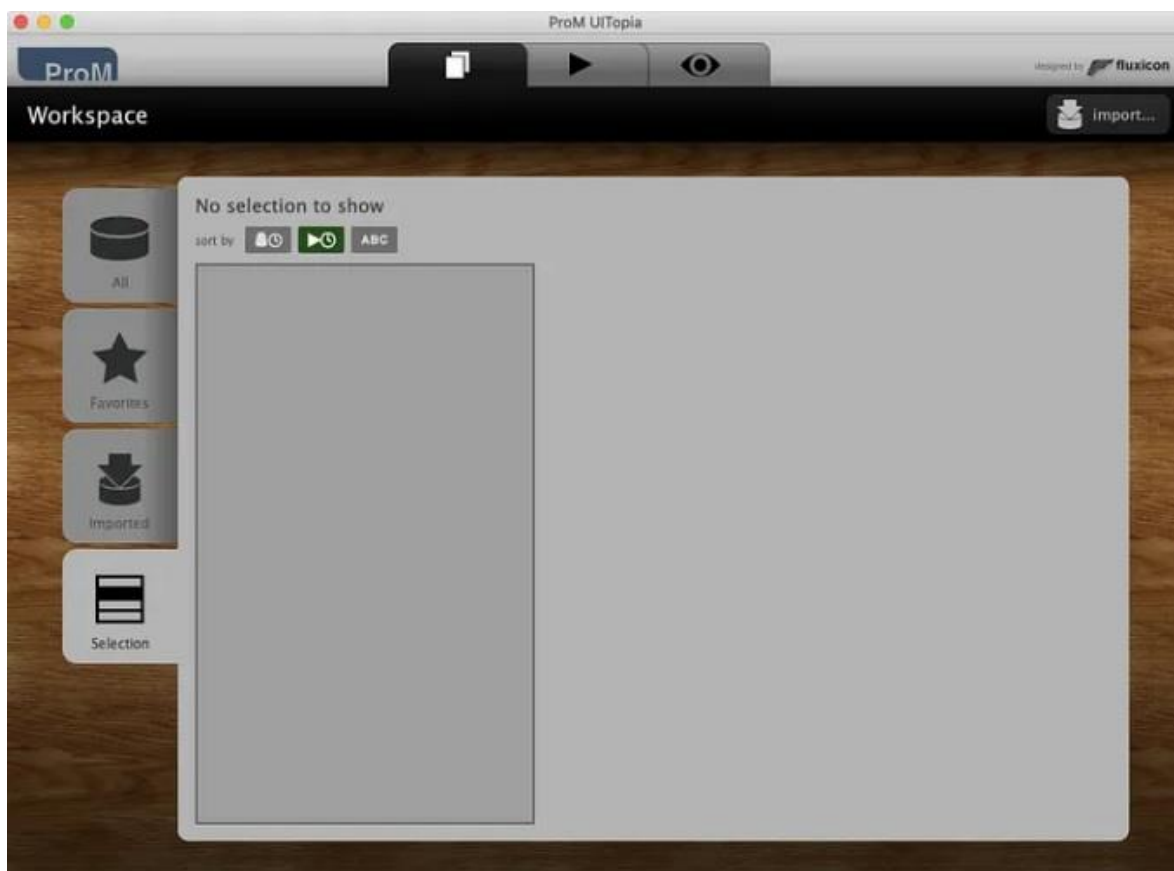


Εικόνα 18: Η ροή εργασιών της διπλωματικής εργασίας

3.2 Το εργαλείο ProM

3.2.1 Περιγραφή του ProM

Το ProM αποτελεί ένα από τα πιο αναγνωρισμένα και καθιερωμένα λογισμικά εργαλεία στον τομέα της εξόρυξης διεργασιών (process mining). Αναπτύχθηκε αρχικά σε ακαδημαϊκό περιβάλλον και συγκεκριμένα στο Τεχνολογικό Πανεπιστήμιο του Eindhoven, με στόχο την παροχή ενός ολοκληρωμένου πλαισίου ανάλυσης επιχειρησιακών διεργασιών βασισμένου σε πραγματικά δεδομένα εκτέλεσης. Με την πάροδο του χρόνου, το ProM εξελίχθηκε σε πλατφόρμα ανοικτού κώδικα, η οποία χρησιμοποιείται ευρέως τόσο στην ακαδημαϊκή έρευνα όσο και σε πιλοτικές ή πειραματικές εφαρμογές σε οργανισμούς (van der Aalst, 2011).



Εικόνα 19: ProM: Αρχική σελίδα

Η εξόρυξη διεργασιών αποτελεί έναν ταχέως αναπτυσσόμενο επιστημονικό κλάδο, ο οποίος συνδυάζει έννοιες από τη διαχείριση επιχειρησιακών διεργασιών, τη μηχανική λογισμικού και την ανάλυση δεδομένων. Το ProM ενσωματώνει αυτή τη διεπιστημονική προσέγγιση, προσφέροντας ένα περιβάλλον στο οποίο οι χρήστες μπορούν να μελετήσουν τη συμπεριφορά των πληροφοριακών συστημάτων μέσα από αρχεία καταγραφής γεγονότων. Τα αρχεία αυτά περιέχουν λεπτομερή στοιχεία σχετικά με τις εκτελέσεις των διεργασιών, όπως δραστηριότητες, χρονικές σφραγίδες και συσχετίσεις μεταξύ ενεργειών (Reinkemeyer, 2020).

Ένα από τα βασικά πλεονεκτήματα του ProM είναι η δυνατότητα ανακάλυψης μοντέλων διεργασιών χωρίς την ανάγκη ύπαρξης προκαθορισμένης τεκμηρίωσης. Μέσω εξειδικευμένων αλγορίθμων, το εργαλείο μπορεί να εξάγει μοντέλα που απεικονίζουν τη λογική ροή των δραστηριοτήτων, βασιζόμενο αποκλειστικά στα δεδομένα εκτέλεσης. Τα μοντέλα αυτά αντικατοπτρίζουν την πραγματική λειτουργία των συστημάτων και συχνά

αποκαλύπτουν αποκλίσεις από τις επίσημα σχεδιασμένες διαδικασίες (van der Aalst, 2016; van der Aalst & Carmona, n.d.).



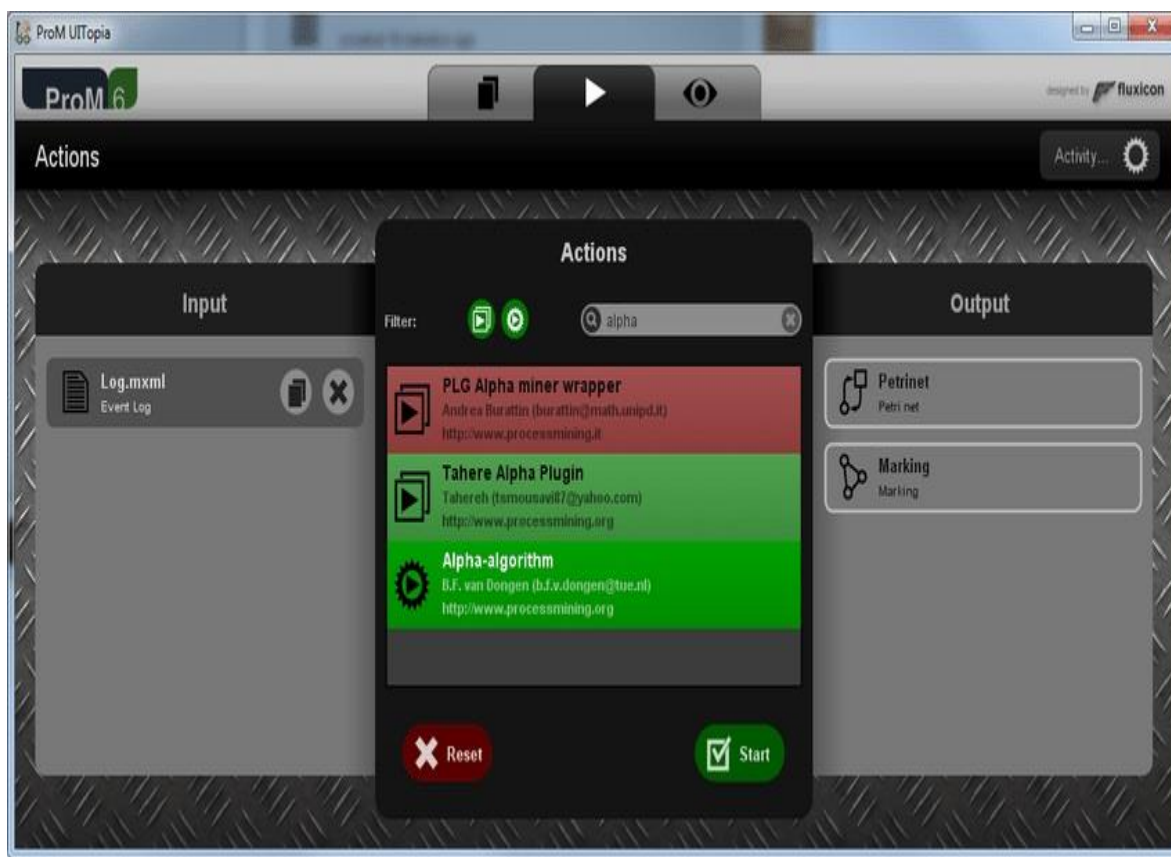
Εικόνα 20: Συνοπτικά στοιχεία ενός αρχείου καταγραφής συμβάντων

Πέρα από την ανακάλυψη διεργασιών, το ProM υποστηρίζει και την ανάλυση συμμόρφωσης, επιτρέποντας τη σύγκριση μεταξύ ενός υφιστάμενου μοντέλου και των πραγματικών δεδομένων. Μέσω αυτής της σύγκρισης, καθίσταται δυνατός ο εντοπισμός παραβιάσεων, αποκλίσεων και μη αναμενόμενων συμπεριφορών. Η λειτουργία αυτή είναι ιδιαίτερα σημαντική σε οργανισμούς που λειτουργούν υπό κανονιστικά πλαίσια, όπου η συμμόρφωση με πρότυπα και διαδικασίες αποτελεί κρίσιμο παράγοντα (Adriansyah et al., 2011).

Επιπλέον, το ProM παρέχει δυνατότητες ενίσχυσης των μοντέλων διεργασιών με πληροφορίες απόδοσης. Μέσα από την ανάλυση χρονικών δεδομένων, το εργαλείο μπορεί να αποκαλύψει χρόνους εκτέλεσης, καθυστερήσεις και σημεία συμφόρησης. Οι

πληροφορίες αυτές επιτρέπουν τη βαθύτερη κατανόηση της λειτουργίας των διεργασιών και υποστηρίζουν τη λήψη τεκμηριωμένων αποφάσεων για τη βελτίωσή τους (Berti et al., 2019).

Η ανοιχτή αρχιτεκτονική του ProM ενισχύει σημαντικά τη χρηστικότητά του. Ως λογισμικό ανοικτού κώδικα, επιτρέπει την ανάπτυξη νέων προσθέτων και αλγορίθμων από την ερευνητική κοινότητα. Αυτό έχει οδηγήσει στη δημιουργία ενός πλούσιου οικοσυστήματος λειτουργιών, καθιστώντας το ProM ιδιαίτερα ευέλικτο και προσαρμόσιμο σε διαφορετικά σενάρια ανάλυσης (Reinkemeyer, 2020).



Εικόνα 21: Επιλογή του αλγορίθμου Alpha Miner για την ανακάλυψη Μοντέλου Διεργασίας

Συνολικά, το ProM αποτελεί ένα ισχυρό και ολοκληρωμένο εργαλείο εξόρυξης διεργασιών, το οποίο υποστηρίζει σε βάθος την ανάλυση, κατανόηση και βελτίωση επιχειρησιακών διαδικασιών. Η επιστημονική του τεκμηρίωση, σε συνδυασμό με την πρακτική του εφαρμοσιμότητα, το καθιστούν σημείο αναφοράς στον χώρο του process mining (van der Aalst, 2011; van der Aalst, 2016).

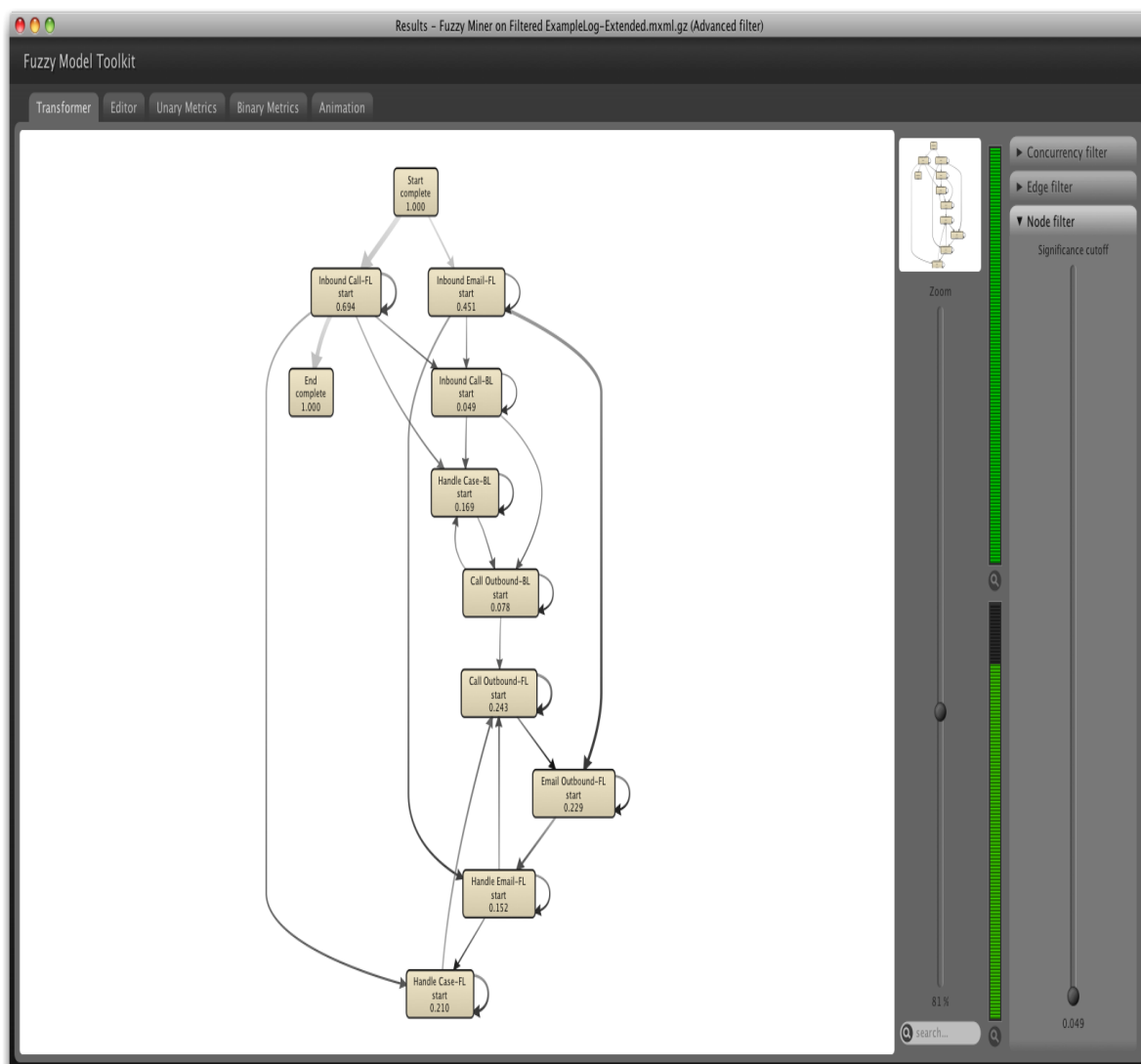
3.2.2 Αρχιτεκτονική και Λειτουργικότητα

Η αρχιτεκτονική του ProM έχει σχεδιαστεί με κύριο στόχο την επεκτασιμότητα και την ευελιξία. Το σύστημα βασίζεται σε μια αρθρωτή δομή, όπου κάθε λειτουργία υλοποιείται μέσω ανεξάρτητων προσθέτων. Τα πρόσθετα αυτά αντιστοιχούν σε συγκεκριμένους αλγορίθμους ή τεχνικές ανάλυσης και μπορούν να λειτουργούν είτε αυτόνομα είτε σε συνδυασμό με άλλα πρόσθετα, δημιουργώντας σύνθετες ροές εργασίας (Aalst et al., 2012).

Στον πυρήνα του ProM βρίσκεται ένας μηχανισμός διαχείρισης αντικειμένων, ο οποίος είναι υπεύθυνος για την αναπαράσταση των δεδομένων και των μοντέλων. Τα αρχεία καταγραφής γεγονότων, τα μοντέλα διεργασιών και τα αποτελέσματα των αναλύσεων αποθηκεύονται ως αντικείμενα, τα οποία μπορούν να ανταλλάσσονται μεταξύ διαφορετικών προσθέτων. Αυτή η κοινή δομή δεδομένων διευκολύνει τη συνεργασία μεταξύ αλγορίθμων και επιτρέπει την επαναχρησιμοποίηση των αποτελεσμάτων (Aalst, 2016).

Η λειτουργικότητα του ProM ξεκινά με την εισαγωγή και προετοιμασία των δεδομένων. Το εργαλείο υποστηρίζει πρότυπες μορφές αρχείων, όπως XES και MXML, οι οποίες είναι ευρέως αποδεκτές στον χώρο της εξόρυξης διεργασιών. Παρέχονται επίσης δυνατότητες καθαρισμού και φιλτραρίσματος των δεδομένων, επιτρέποντας την απομάκρυνση θορύβου και μη σχετικών γεγονότων πριν από την ανάλυση (Dongen et al., 2013).

Στη συνέχεια, ο χρήστης μπορεί να επιλέξει ανάμεσα σε πλήθος αλγορίθμων ανακάλυψης διεργασιών. Οι αλγόριθμοι αυτοί διαφέρουν ως προς τη φιλοσοφία και τα χαρακτηριστικά τους, προσφέροντας διαφορετικά επίπεδα ακρίβειας, γενίκευσης και ανθεκτικότητας στον θόρυβο. Η δυνατότητα επιλογής καθιστά το ProM κατάλληλο για διαφορετικά είδη δεδομένων και αναλυτικών στόχων. Στην παρακάτω εικόνα (**Εικόνα 22**) [Rozinat, 2010], απεικονίζεται η εξαγωγή ενός μοντέλου διεργασιών (process model) εισάγοντας ένα αρχείο καταγραφής συμβάντων (log file) κάνοντας χρήση του αλγόριθμου Fuzzy Miner.



Εικόνα 22: Μοντέλο διεργασιών με χρήση του αλγόριθμου Fuzzy Miner

Ιδιαίτερη σημασία έχει η υποστήριξη της ανάλυσης συμμόρφωσης, η οποία υλοποιείται μέσω εξειδικευμένων προσθέτων. Τα πρόσθετα αυτά συγκρίνουν ένα μοντέλο διεργασίας με τα δεδομένα εκτέλεσης και υπολογίζουν μετρικές αξιολόγησης, όπως η καταλληλότητα και η ακρίβεια. Οι μετρικές αυτές παρέχουν ποσοτική βάση για την αξιολόγηση της ποιότητας των μοντέλων και τη βελτίωσή τους (van der Aalst et al., 2011).

Η αρχιτεκτονική του ProM υποστηρίζει επίσης προηγμένες τεχνικές οπτικοποίησης. Τα μοντέλα διεργασιών μπορούν να παρουσιαστούν γραφικά και να ενισχυθούν με πληροφορίες απόδοσης, όπως συχνότητες δραστηριοτήτων και χρονικές καθυστερήσεις. Η οπτικοποίηση αυτή διευκολύνει την κατανόηση πολύπλοκων διεργασιών και καθιστά τα αποτελέσματα της ανάλυσης πιο προσιτά σε ευρύτερο κοινό (Berti et al., 2019).

Τέλος, το ProM επιτρέπει τη δημιουργία σύνθετων σεναρίων ανάλυσης μέσω του συνδυασμού πολλαπλών προσθέτων. Τα αποτελέσματα μιας ανάλυσης μπορούν να χρησιμοποιηθούν ως είσοδος για επόμενα στάδια, υποστηρίζοντας μια ολοκληρωμένη προσέγγιση στη μελέτη και βελτίωση διεργασιών. Η δυνατότητα αυτή αναδεικνύει τη δύναμη της αρθρωτής αρχιτεκτονικής του εργαλείου (Aalst, 2016).

Συμπερασματικά, η αρχιτεκτονική και η λειτουργικότητα του ProM αντανακλούν έναν ώριμο και επιστημονικά τεκμηριωμένο σχεδιασμό. Η ευελιξία, η επεκτασιμότητα και η πλούσια λειτουργικότητα καθιστούν το ProM ένα από τα σημαντικότερα εργαλεία στον τομέα της εξόρυξης διεργασιών, με ουσιαστική συμβολή τόσο στην έρευνα όσο και στην πρακτική εφαρμογή.

3.2.3 Παραδείγματα εφαρμογής

Στην ενότητα αυτή, θα παρουσιαστεί ένα παράδειγμα της εξόρυξης διεργασιών (Process Mining) χρησιμοποιώντας το εργαλείο ProM.

Για το παράδειγμα, θα χρησιμοποιηθεί ένα αρχείο καταγραφής συμβάντων (event log), τα δεδομένα του οποίου έχουν εξαχθεί από ένα σύστημα διαχείρισης αιτήσεων πελατών. Στην παρακάτω εικόνα (**Εικόνα 23**) [<https://processmining.org>], βλέπουμε τα δεδομένα του αρχείου σε μορφή .xls:

	A	B	C	D	E	F
1	Case ID	Event ID	dd-MM-yyyy:HH.mm	Activity	Resource	Costs
2	1	35654423	30-12-2010:11.02	register request	Pete	50
3	1	35654424	31-12-2010:10.06	examine thoroughly	Sue	400
4	1	35654425	05-01-2011:15.12	check ticket	Mike	100
5	1	35654426	06-01-2011:11.18	decide	Sara	200
6	1	35654427	07-01-2011:14.24	reject request	Pete	200
7	2	35654483	30-12-2010:11.32	register request	Mike	50
8	2	35654485	30-12-2010:12.12	check ticket	Mike	100
9	2	35654487	30-12-2010:14.16	examine casually	Sean	400
10	2	35654488	05-01-2011:11.22	decide	Sara	200
11	2	35654489	08-01-2011:12.05	pay compensation	Ellen	200
12	3	35654521	30-12-2010:14.32	register request	Pete	50
13	3	35654522	30-12-2010:15.06	examine casually	Mike	400
14	3	35654524	30-12-2010:16.34	check ticket	Ellen	100
15	3	35654525	06-01-2011:09.18	decide	Sara	200
16	3	35654526	06-01-2011:12.18	reinitiate request	Sara	200
17	3	35654527	06-01-2011:13.06	examine thoroughly	Sean	400
18	3	35654530	08-01-2011:11.43	check ticket	Pete	100
19	3	35654531	09-01-2011:09.55	decide	Sara	200
20	3	35654533	15-01-2011:10.45	pay compensation	Ellen	200
21	4	35654641	06-01-2011:15.02	register request	Pete	50
22	4	35654643	07-01-2011:12.06	check ticket	Mike	100
23	4	35654644	08-01-2011:14.43	examine thoroughly	Sean	400
24	4	35654645	09-01-2011:12.02	decide	Sara	200
25	4	35654647	12-01-2011:15.44	reject request	Ellen	200
26	5	35654711	06-01-2011:09.02	register request	Ellen	50
27	5	35654712	07-01-2011:10.16	examine casually	Mike	400
28	5	35654714	08-01-2011:11.22	check ticket	Pete	100
29	5	35654715	10-01-2011:13.28	decide	Sara	200
30	5	35654716	11-01-2011:16.18	reinitiate request	Sara	200
31	5	35654718	14-01-2011:14.33	check ticket	Ellen	100
32	5	35654719	16-01-2011:15.50	examine casually	Mike	400
33	5	35654720	19-01-2011:11.18	decide	Sara	200
34	5	35654721	20-01-2011:12.48	reinitiate request	Sara	200
35	5	35654722	21-01-2011:09.06	examine casually	Sue	400
36	5	35654724	21-01-2011:11.34	check ticket	Pete	100
37	5	35654725	23-01-2011:13.12	decide	Sara	200
38	5	35654726	24-01-2011:14.56	reject request	Mike	200
39	6	35654871	06-01-2011:15.02	register request	Mike	50
40	6	35654873	06-01-2011:16.06	examine casually	Ellen	400
41	6	35654874	07-01-2011:16.22	check ticket	Mike	100
42	6	35654875	07-01-2011:16.52	decide	Sara	200
43	6	35654877	16-01-2011:11.47	pay compensation	Mike	200

Εικόνα 23: Αρχείο καταγραφής συμβάντων παραδείγματος

Στο αρχείο αυτό, διακρίνουμε τα παρακάτω πεδία:

CaseID: Αντιπροσωπεύει το αναγνωριστικό μιας συγκεκριμένης ροής εργασίας.

EventID: Μοναδικός αναγνωριστικός αριθμός για κάθε γεγονός/ενέργεια μέσα στο αρχείο καταγραφής συμβάντων.

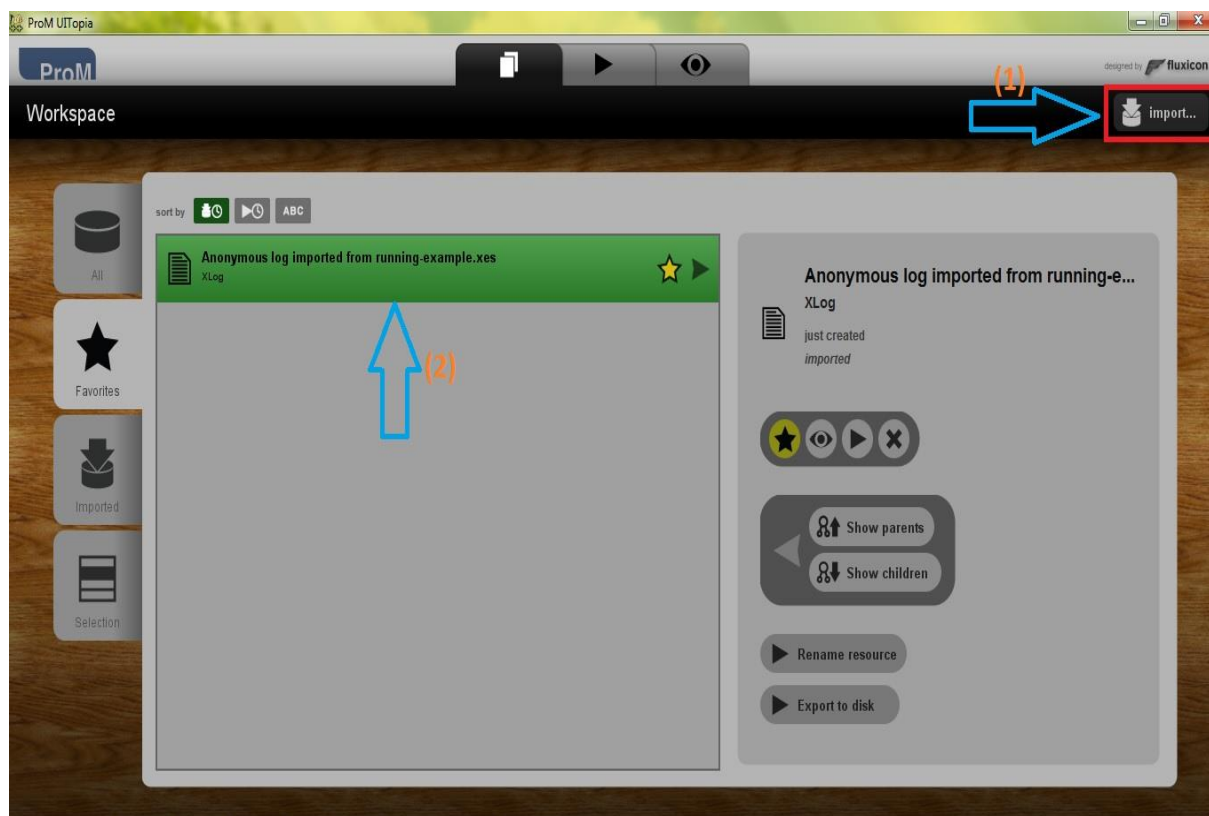
dd-MM-yyyy:HH.mm: Η ημερομηνία και η ώρα που συνέβη το γεγονός. Χρησιμοποιείται για να καθοριστεί η σειρά των γεγονότων και για χρονική ανάλυση.

Activity: Η ενέργεια ή δραστηριότητα που πραγματοποιήθηκε. Δείχνει τι ακριβώς έγινε σε κάθε γεγονός.

Resource: Το άτομο ή το σύστημα που εκτέλεσε την ενέργεια. Χρησιμεύει για να εντοπιστεί ποιος ή τι ολοκλήρωσε την δραστηριότητα.

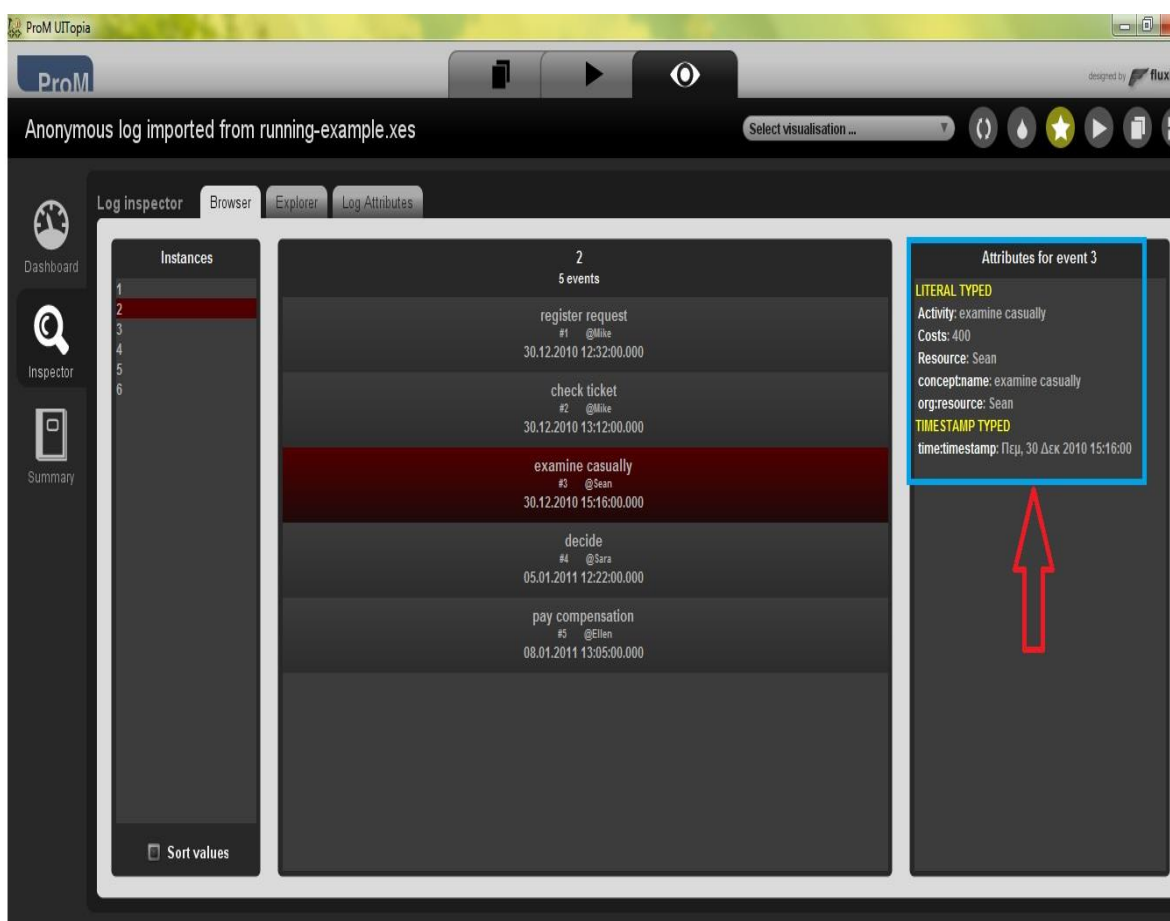
Costs: Το κόστος που συνδέεται με την εκτέλεση της δραστηριότητας.

Έπειτα, εισάγεται το αρχείο καταγραφής συμβάντων (.xes file) στο εργαλείο εξόρυξης διεργασιών ProM όπως δείχνει η παρακάτω εικόνα (**Εικόνα 24**):



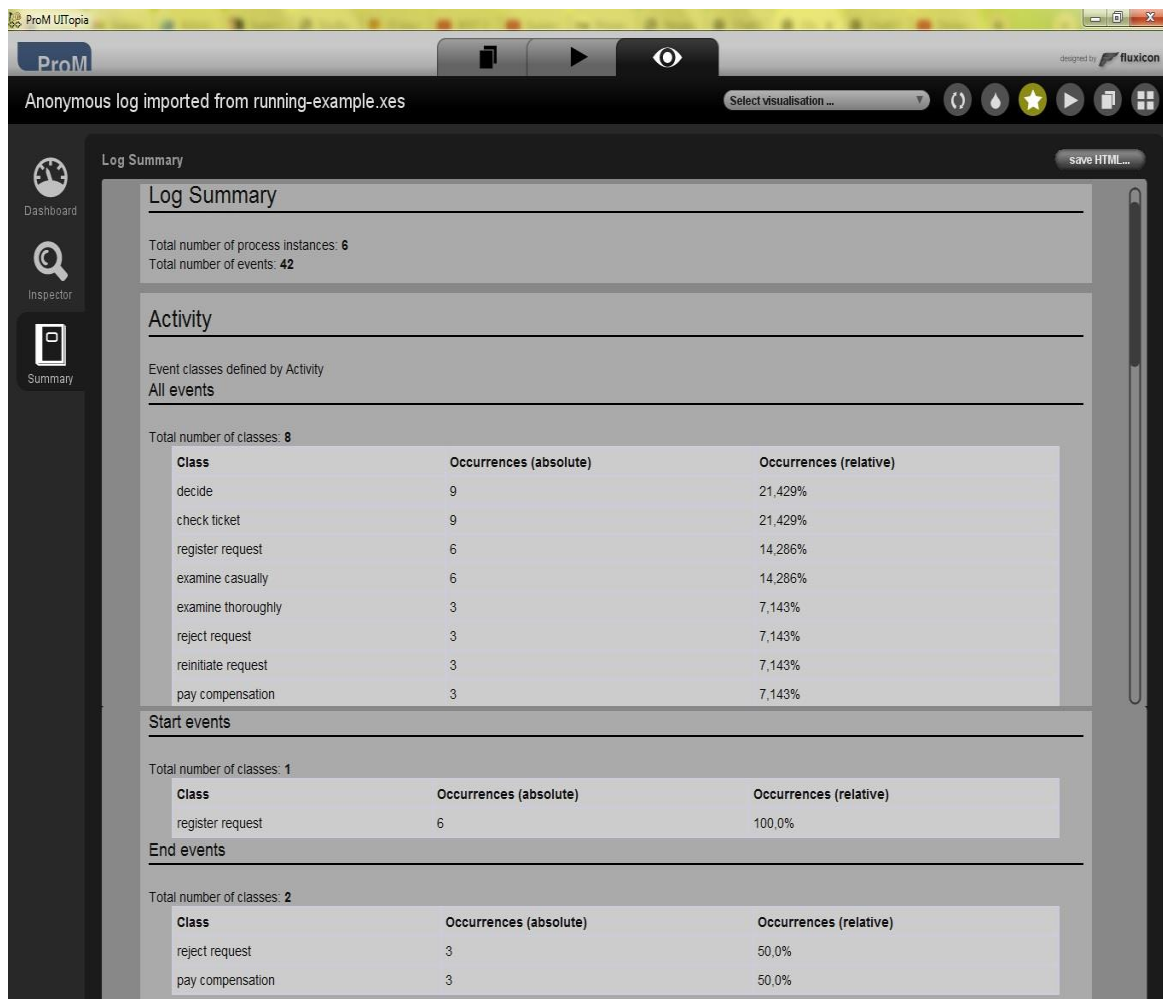
Εικόνα 24: Εισαγωγή του αρχείου καταγραφής συμβάντων στο ProM

Αρχικά, επιλέγοντας από το μενού: <viewresource>, μπορούμε να δούμε τόσο τα περιεχόμενα του αρχείου καταγραφής συμβάντων όσο και κάποια ενδιαφέροντα στατιστικά στοιχεία. Στην παρακάτω εικόνα (**Εικόνα 25**), παρουσιάζεται για κάθε ‘CaseID’ του event log αρχείου, αναλυτικά τα δεδομένα του κάθε γεγονότος (event) που αντιστοιχεί στο εκάστοτε ‘CaseID’. Για παράδειγμα, στην εικόνα βλέπουμε τα δεδομένα του κάθε event που αντιστοιχούν στο ‘CaseID’ 2:



Εικόνα 25: Δεδομένα των events με ‘CaseID’ = 2

Στην παρακάτω εικόνα (**Εικόνα 26**), βλέπουμε μία σύνοψη του αρχείου καταγραφής συμβάντων.



Εικόνα 26: Συνοπτικά στοιχεία αρχείου καταγραφής συμβάντων

Πιο συγκεκριμένα, τα δεδομένα τα οποία παίρνουμε είναι τα εξής:

Συνολικός αριθμός περιπτώσεων (cases): 6

Συνολικός αριθμός γεγονότων (events): 42

Συνολικός αριθμός κλάσεων γεγονότων (πόσες διαφορετικές κλάσεις παρουσιάζονται στο αρχείο): 8

Έπειτα, για κάθε κλάση εμφανίζονται στατιστικά στοιχεία και πιο συγκεκριμένα πόσες φορές εμφανίζεται στο αρχείο καταγραφής συμβάντων και το αντίστοιχο ποσοστό (%).

Start events: Αναφέρονται στα γεγονότα (activities) με τα οποία ξεκινούν τα cases ενός event log αρχείου. Συγκεκριμένα, για κάθε case εξετάζεται ποιο event εμφανίζεται πρώτο στην χρονική σειρά και αυτά τα events καταγράφονται ως “start events”. Έτσι, μπορούμε να δούμε ποιες δραστηριότητες αποτελούν την πραγματική αρχή της διαδικασίας και πόσο συχνά εμφανίζονται, δηλαδή σε πόσα cases ξεκινά η διαδικασία με κάθε μία από αυτές.

Στο παράδειγμα μας, η δραστηριότητα η οποία θεωρείται ως “start event” και με την οποία ξεκινούν και τα 6 cases (ποσοστό 100%) είναι η: «**register request**».

End events: Αντίστοιχα, τα “End events” είναι τα γεγονότα που εμφανίζονται τελευταία σε κάθε case και σηματοδοτούν την ολοκλήρωση της διαδικασίας. Για κάθε case, καταγράφεται το τελευταίο event στην χρονική ακολουθία και αυτά τα events παρουσιάζονται στο Log Summary ως “End events” μαζί με την συχνότητά τους. Με αυτό τον τρόπο, γίνεται σαφές πως τερματίζεται η διαδικασία στην πράξη.

Στο παράδειγμα μας, οι δραστηριότητες οι οποίες θεωρούνται ως “End events” και με τις οποίες τα 3 από τα 6 cases (ποσοστό 50%) ολοκληρώνουν τις διαδικασίες τους είναι οι: «reject request» και «pay compensation».

Η πληροφορία για τα “Start events” και “End events” είναι ιδιαίτερα σημαντική, καθώς επιτρέπει την κατανόηση της πραγματικής δομής της διαδικασίας. Βοηθάει στον έλεγχο της ποιότητας του event log αρχείου, στον εντοπισμό ασυνήθιστων ή ελλειπόν περιπτώσεων, καθώς και στην σωστή ανακάλυψη μοντέλων διαδικασιών, αφού καθορίζονται με ακρίβεια τα σημεία έναρξης και λήξης της διαδικασίας.

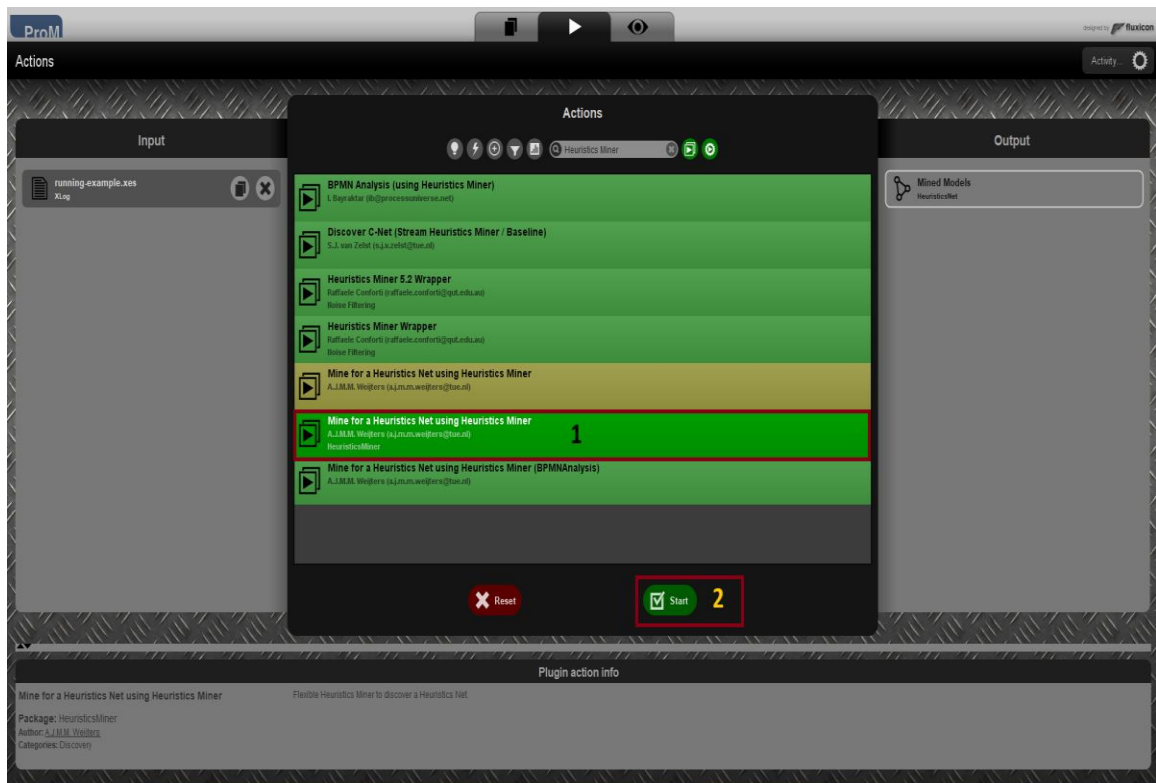
Στη συνέχεια της ανάλυσης που προηγήθηκε στην Ενότητα 1.1.5.1, και κατόπιν συγκριτικής αξιολόγησης διαφόρων αλγορίθμων εξόρυξης, επιλέχθηκαν για τις ανάγκες της παρούσας εργασίας ο **Heuristics Miner** και η τεχνική **Token-Based Replay**. Η επιλογή του Heuristics Miner βασίστηκε στην ικανότητά του να επεξεργάζεται ατελή δεδομένα (noisy logs) και να εξάγει συμπαγή και μεθοδολογικά συνεπή μοντέλα, εξασφαλίζοντας μια αξιόπιστη βάση για τον μετέπειτα έλεγχο συμμόρφωσης. Αντίστοιχα, το Token-Based Replay επιλέχθηκε ως η καταλληλότερη μέθοδος για τον έλεγχο συμμόρφωσης, καθώς επιτρέπει την ακριβή αποτύπωση της ροής των “τεκμηρίων” (tokens) πάνω στο μοντέλο, προσφέροντας σαφή εικόνα για το πού και γιατί προκύπτουν αποκλίσεις από την προβλεπόμενη συμπεριφορά.

Παρακάτω, παρουσιάζεται η πρακτική εφαρμογή των δύο αυτών μεθόδων μέσα από συγκεκριμένα παραδείγματα.

Heuristics Miner

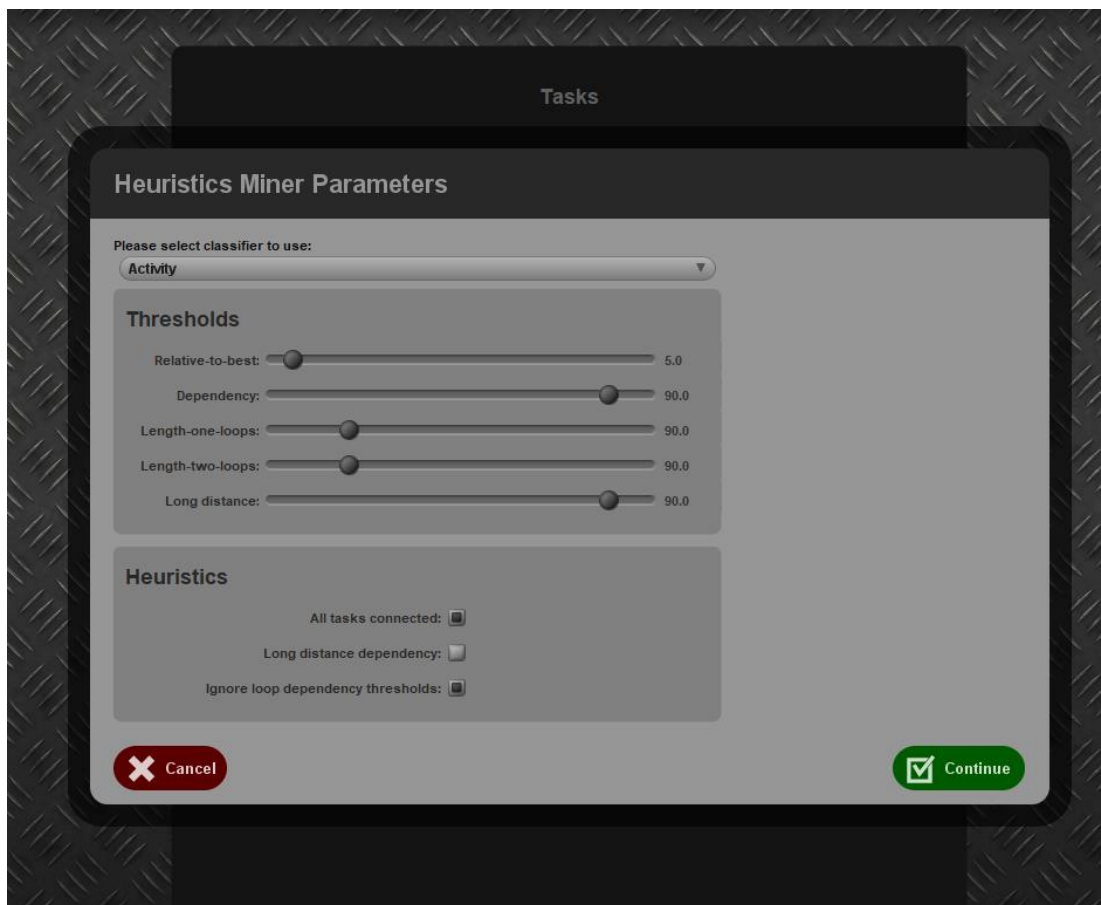
Στο σημείο αυτό, θα παρουσιαστεί μια ενδεικτική εφαρμογή του αλγορίθμου Heuristics Miner, χρησιμοποιώντας ένα αντιπροσωπευτικό αρχείο καταγραφής συμβάντων. Στόχος είναι η πρακτική επίδειξη των δυνατοτήτων του αλγορίθμου στην ανακάλυψη μοντέλων διεργασιών, καθώς και η παρουσίαση των επιμέρους βημάτων της διαδικασίας, από την εισαγωγή των δεδομένων έως την εξαγωγή του τελικού Heuristics Net και τη μετατροπή του σε Petri Net.

Για την έναρξη της διαδικασίας ανακάλυψης του μοντέλου, είναι απαραίτητη η επιλογή του κατάλληλου αλγορίθμου στο περιβάλλον του ProM. Όπως παρουσιάζεται στην **Εικόνα 27**, έχοντας ήδη εισαγάγει το αρχείο καταγραφής συμβάντων ως είσοδο, χρησιμοποιείται η λειτουργία αναζήτησης για τον εντοπισμό των διαθέσιμων plugins για τον Heuristics Miner. Από τη λίστα που προκύπτει, επιλέγεται η βασική υλοποίηση (Mine for a Heuristics Net using Heuristics Miner) (1), και στη συνέχεια η διαδικασία εκκινείται με το πάτημα του κουμπιού "Start" (2), οδηγώντας στο επόμενο στάδιο της παραμετροποίησης.



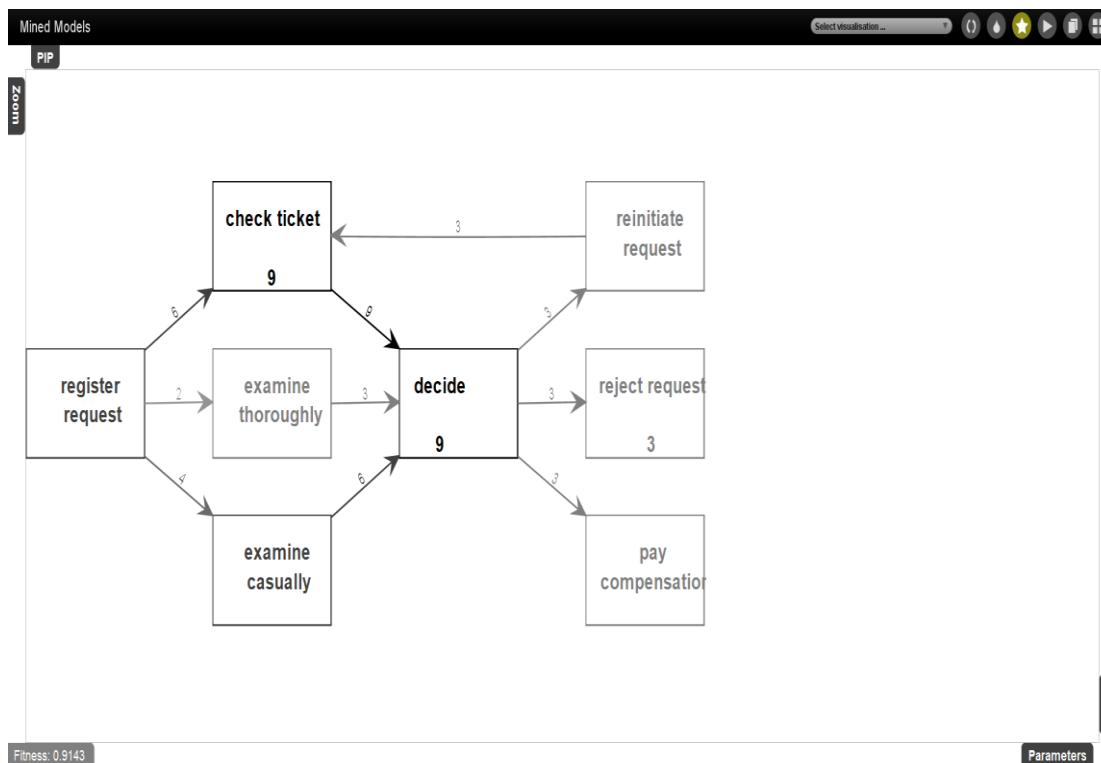
Εικόνα 27: Διεπαφή επιλογής του αλγορίθμου Heuristics Miner στο περιβάλλον ProM

Μετά την εκκίνηση του αλγορίθμου, εμφανίζεται το παράθυρο ρυθμίσεων του Heuristics Miner, όπως βλέπουμε στην **Εικόνα 28**. Στόχος αυτού του σταδίου είναι να "εκπαιδεύσουμε" τον αλγόριθμο πώς να ερμηνεύει τις εξαρτήσεις μεταξύ των δραστηριοτήτων. Αυτό επιτυγχάνεται ρυθμίζοντας τα κατώφλια (thresholds), τα οποία καθορίζουν πόσο συχνά πρέπει να εμφανίζεται μια ακολουθία δραστηριοτήτων στο αρχείο καταγραφής συμβάντων ώστε να θεωρηθεί έγκυρη εξάρτηση και να συμπεριληφθεί στο τελικό Heuristics Net. Παράλληλα, μπορούμε να επιλέξουμε τον "Activity" ταξινομητή και να εφαρμόσουμε επιπλέον ευρετικούς κανόνες, όπως ο περιορισμός όλες οι εργασίες να είναι συνδεδεμένες, διασφαλίζοντας έτσι τη συνεκτικότητα του μοντέλου.



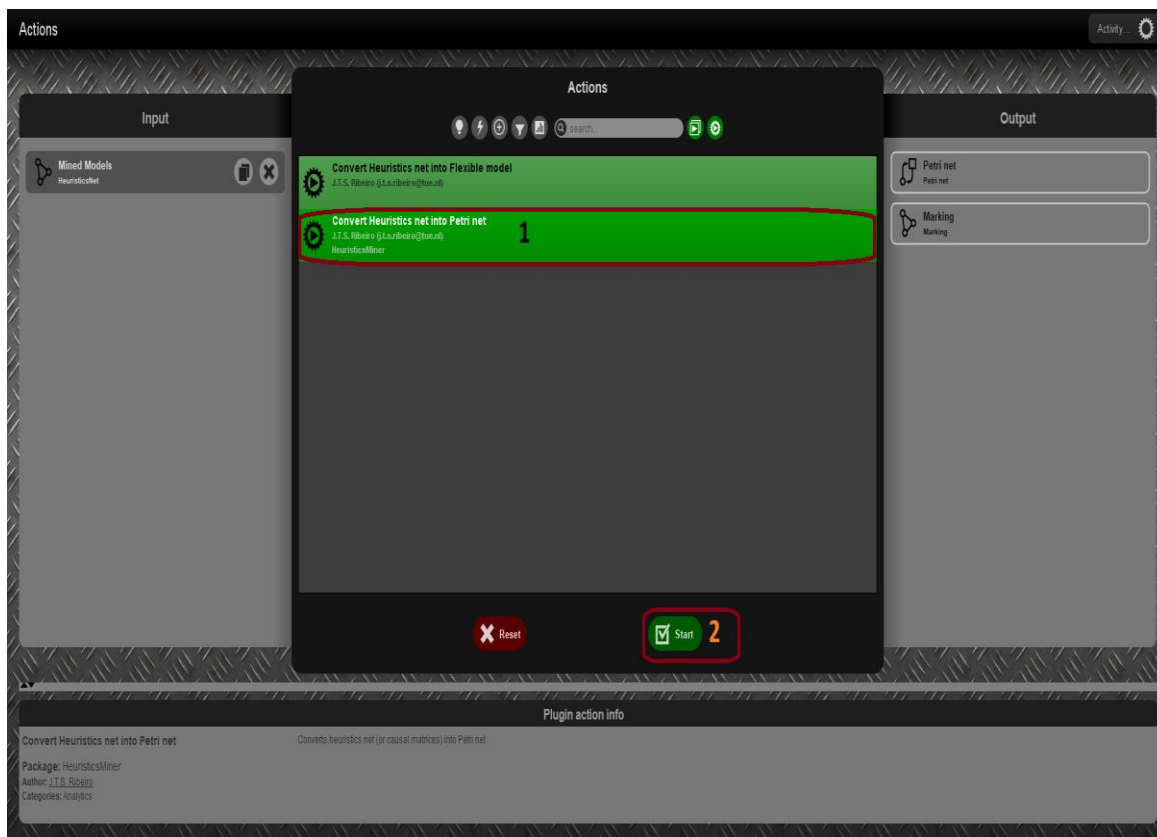
Εικόνα 28: Παράθυρο ρύθμισης παραμέτρων του αλγορίθμου

Ως αποτέλεσμα της διαδικασίας εξόρυξης, ο αλγόριθμος παρήγαγε το Heuristics Net που απεικονίζεται στην **Εικόνα 29**. Αυτό το μοντέλο μάς προσφέρει μια πρώτη, καθαρή εικόνα της ροής εργασιών. Κάθε κουτάκι αντιπροσωπεύει μια δραστηριότητα (π.χ. "register request", "check ticket"), ενώ τα βέλη δείχνουν τη σειρά με την οποία αυτές συμβαίνουν. Οι αριθμοί πάνω στα βέλη και μέσα στα κουτάκια είναι ιδιαίτερα σημαντικοί, καθώς υποδεικνύουν τη συχνότητα εμφάνισης. Για παράδειγμα, βλέπουμε ότι η διαδρομή από το "register request" στο "check ticket" ακολουθήθηκε 6 φορές, ενώ η δραστηριότητα "check ticket" εκτελέστηκε συνολικά 9 φορές. Αυτή η οπτικοποίηση μας επιτρέπει να εντοπίσουμε άμεσα τις πιο συχνές διαδρομές μέσα στη διεργασία.



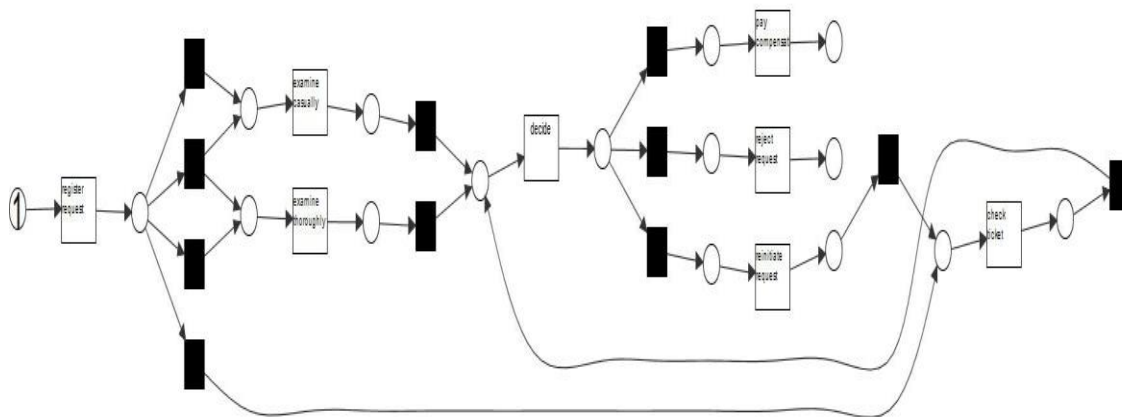
Εικόνα 29: Heuristics Net – Μοντέλο διεργασίας από Heuristics Miner

Στο στάδιο αυτό, το λογισμικό ProM χρησιμοποιείται για να μετατρέψει το αρχικό μοντέλο που προέκυψε από τον Heuristics Miner σε μορφή Petri net. Στο κεντρικό μενού των ενεργειών, επιλέγεται η λειτουργία "Convert Heuristics net into Petri net" (σημείο 1), η οποία αναλαμβάνει να αλλάξει τη δομή του μοντέλου ώστε να είναι συμβατή με επόμενες αναλύσεις. Η μετατροπή αυτή είναι αναγκαία, καθώς το Token Based Replay απαιτεί ένα Petri net μοντέλο ως βάση για να εκτελέσει τον έλεγχο συμμόρφωσης. Στο δεξιό τμήμα της οθόνης φαίνονται τα αποτελέσματα που θα παραχθούν, δηλαδή το μοντέλο Petri net και η αρχική του σήμανση (Marking). Η διαδικασία ξεκινά πατώντας το κουμπί "Start" (σημείο 2), ολοκληρώνοντας έτσι τη μετάβαση μεταξύ των δύο διαφορετικών τύπων μοντέλων (Εικόνα 30).



Εικόνα 30: Επιλογή μετατροπής Heuristics net σε Petri net

Η παρακάτω εικόνα (**Εικόνα 31**), παρουσιάζει το τελικό μοντέλο διεργασίας σε μορφή Petri net που παράχθηκε από το ProM, ακολουθώντας τη διαδικασία μετατροπής από το Heuristics Net. Το μοντέλο αυτό αποτελείται από θέσεις (κύκλοι) και μεταβάσεις (ορθογώνια), που αναπαριστούν τις καταστάσεις και τις δραστηριότητες της διεργασίας αντίστοιχα. Στο αριστερό άκρο, ο κύκλος με τον αριθμό '1' υποδηλώνει την αρχική θέση (input place) της διεργασίας, από όπου ξεκινά η ροή. Είναι εμφανές ότι οι μεταβάσεις φέρουν τις ονομασίες των πραγματικών δραστηριοτήτων, όπως "register request", "examine casually", "decide" και "check ticket". Το δίκτυο Petri αποτελεί πλέον ένα τυπικό μοντέλο διεργασίας, έτοιμο να χρησιμοποιηθεί ως βάση για την εφαρμογή του αλγορίθμου Token Based Replay και τον έλεγχο συμμόρφωσης με τα δεδομένα του log καταγραφής.



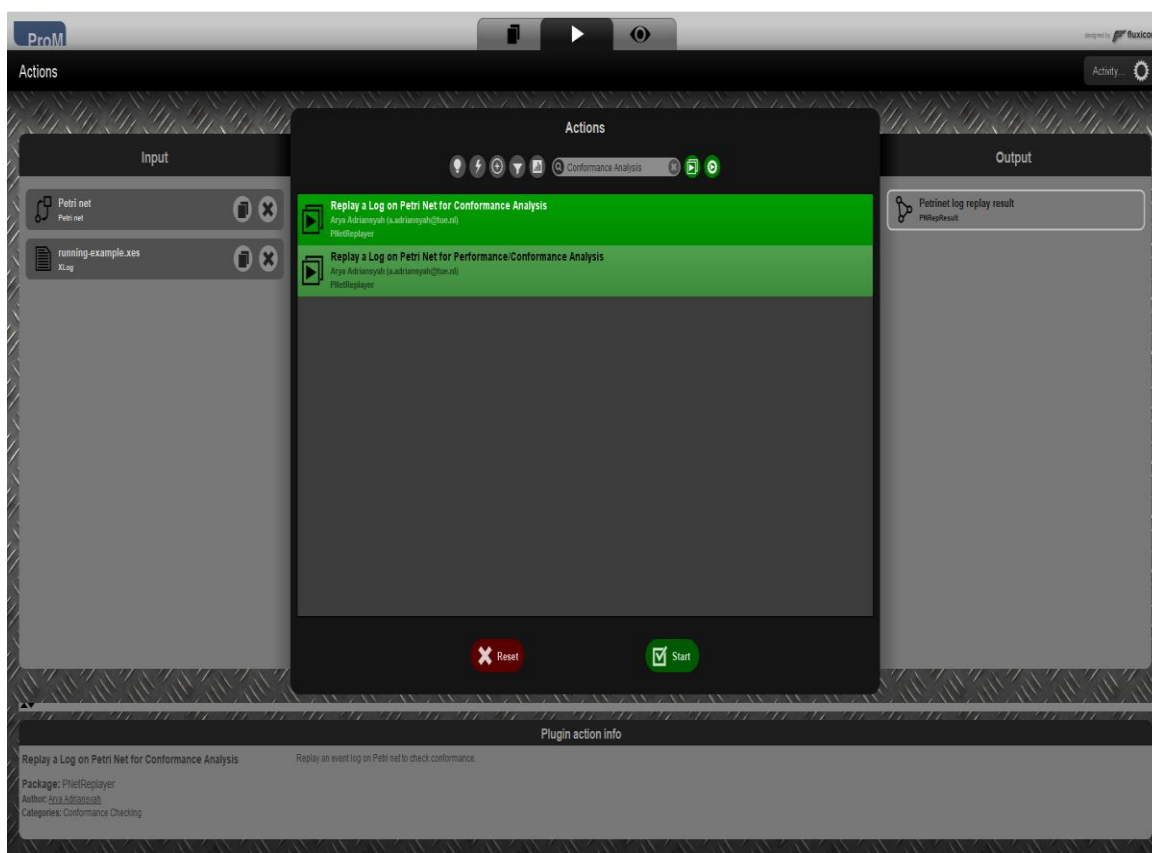
Εικόνα 31: Τελικό μοντέλο διεργασίας σε μορφή δικτύου Petri net

Token-Based Replay

Έχοντας ολοκληρώσει τη διαδικασία εξόρυξης του μοντέλου διεργασίας με τη χρήση του αλγορίθμου Heuristics Miner και τη μετατροπή του στο τελικό δίκτυο Petri, το επόμενο κρίσιμο βήμα στην ανάλυση είναι ο έλεγχος συμμόρφωσης (conformance checking). Στην ενότητα αυτή, θα παρουσιαστεί η εφαρμογή της τεχνικής Token-Based Replay στο λογισμικό ProM. Η συγκεκριμένη τεχνική επιτρέπει την αναπαράσταση της ροής των καταγραφών (events) από το αρχείο καταγραφής (event log) επάνω στο τελικό δίκτυο Petri, χρησιμοποιώντας διακριτικά (tokens). Μέσω αυτής της διαδικασίας, καθίσταται δυνατός ο εντοπισμός αποκλίσεων μεταξύ του μοντέλου και της πραγματικής συμπεριφοράς, καθώς και ο υπολογισμός μετρικών καταλληλότητας (fitness), αξιολογώντας έτσι την ακρίβεια του μοντέλου που προέκυψε.

Σε αυτό το στάδιο, προετοιμάζουμε την εκτέλεση του ελέγχου συμμόρφωσης στο ProM, συνδυάζοντας το δίκτυο Petri που φτιάξαμε με το αρχικό αρχείο καταγραφής γεγονότων (event log). Στην αριστερή στήλη ("Input") φαίνονται τα δύο απαραίτητα στοιχεία εισόδου: το δίκτυο Petri και το αρχείο "running-example.xes". Από τη λίστα των διαθέσιμων ενεργειών στο κέντρο, επιλέγουμε το plugin **"Replay a Log on Petri Net for Conformance"**

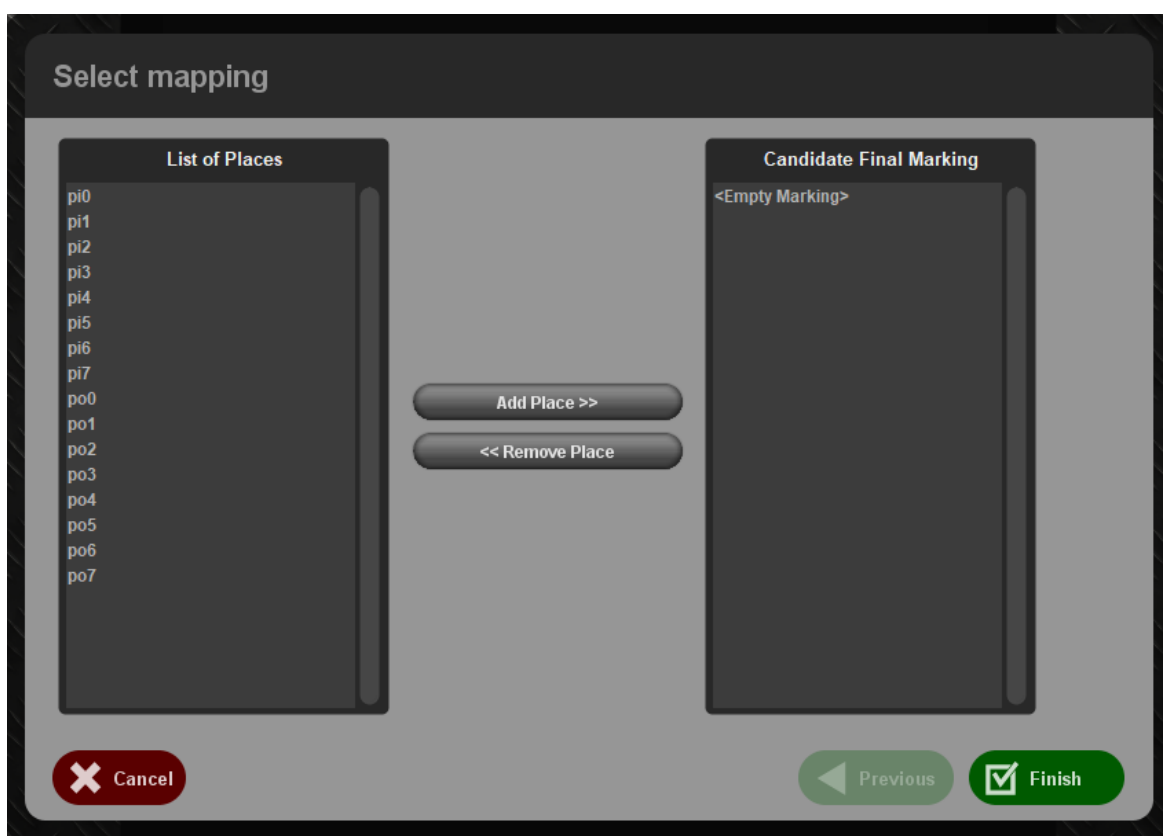
Analysis", το οποίο υλοποιεί την τεχνική του Token-Based Replay. Στη δεξιά πλευρά ("Output") εμφανίζεται το αναμενόμενο αποτέλεσμα της διαδικασίας, το οποίο θα περιλαμβάνει τα στατιστικά στοιχεία της αναπαραγωγής των γεγονότων πάνω στο μοντέλο. Πατώντας το κουμπί "Start", ξεκινά η διαδικασία που θα μας επιτρέψει να αξιολογήσουμε πόσο πιστά ακολουθεί το μοντέλο την πραγματική ροή της διεργασίας (**Εικόνα 32**).



Εικόνα 32: Επιλογή του αλγορίθμου Token-Based Replay για έλεγχο συμμόρφωσης

Σε αυτό το παράθυρο ρυθμίσεων του ProM (**Εικόνα 33**), ο χρήστης καλείται να ορίσει την τελική σήμανση (Final Marking) για το δίκτυο Petri, δηλαδή την κατάσταση στην οποία πρέπει να βρίσκεται το μοντέλο όταν ολοκληρώνεται μια διεργασία. Στην αριστερή λίστα ("List of Places") εμφανίζονται όλες οι διαθέσιμες θέσεις του δικτύου, ενώ στη δεξιά πλευρά ("Candidate Final Marking") επιλέγονται οι θέσεις που θα περιέχουν διακριτικά (tokens) στο τέλος της εκτέλεσης. Ο σωστός ορισμός της τελικής σήμανσης είναι θεμελιώδης για τον

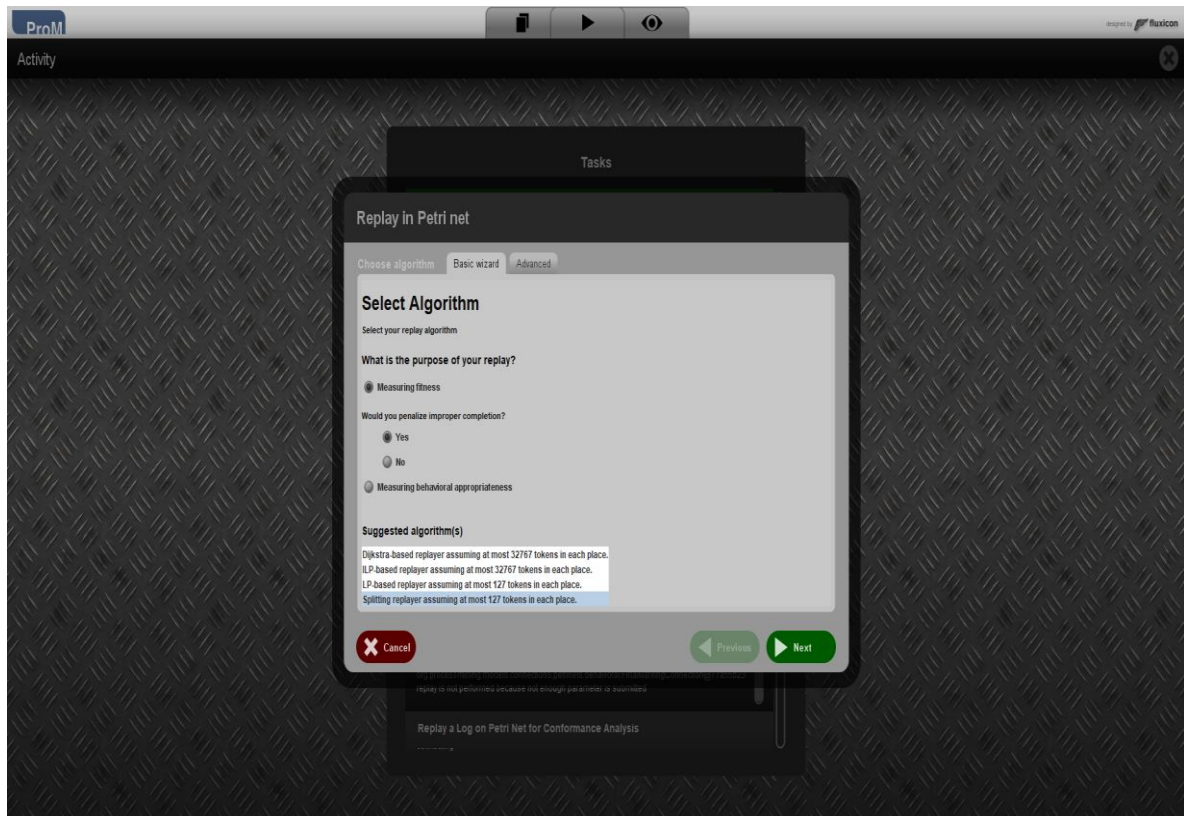
αλγόριθμο Token-Based Replay, καθώς επιτρέπει στον έλεγχο συμμόρφωσης να αναγνωρίζει αν μια υπόθεση (case) ολοκληρώθηκε επιτυχώς ή αν παρέμειναν εκκρεμότητες στο μοντέλο. Η διαδικασία οριστικοποιείται πατώντας το κουμπί "Finish", ώστε να προχωρήσει η ανάλυση στα επόμενα στάδια.



Εικόνα 33: Καθορισμός της τελικής σήμανσης του δικτύου Petri

Σε αυτό το βήμα, το λογισμικό ProM ζητά τον καθορισμό των παραμέτρων για τη διαδικασία της αναπαραγωγής (replay). Αρχικά, ο χρήστης επιλέγει ως κύριο σκοπό τη μέτρηση της καταλληλότητας (Measuring fitness), ενώ παράλληλα ορίζεται αν θα υπάρξει ποινή για περιπτώσεις λανθασμένης ολοκλήρωσης των διαδρομών. Στο κάτω μέρος του παραθύρου, το σύστημα προτείνει μια λίστα από εξειδικευμένους αλγορίθμους, όπως ο "Dijkstra-based replayer", οι οποίοι διαχειρίζονται τον τρόπο με τον οποίο τα διακριτικά (tokens) θα κινηθούν μέσα στο δίκτυο Petri. Η σωστή επιλογή των ρυθμίσεων αυτών είναι

κρίσιμη, καθώς καθορίζει την ακρίβεια των αποτελεσμάτων που θα προκύψουν από τον έλεγχο συμμόρφωσης. Πατώντας το κουμπί "Next", προχωράμε στο τελικό στάδιο πριν την εκτέλεση της ανάλυσης (**Εικόνα 34**).



Εικόνα 34: Παραμετροποίηση και επιλογή αλγορίθμου για την αναπαραγωγή των γεγονότων

Στο συγκεκριμένο παράθυρο ρυθμίσεων, ορίζονται οι παράμετροι κόστους που σχετίζονται με τις αποκλίσεις ανάμεσα στο μοντέλο και το αρχείο καταγραφής. Ο χρήστης μπορεί να καθορίσει το "κόστος μετακίνησης στο μοντέλο" (Move on Model Cost), που αφορά ενέργειες που προβλέπονται από το δίκτυο Petri αλλά λείπουν από το log, καθώς και το "κόστος μετακίνησης στο log" (Move on Log Cost), για ενέργειες που καταγράφηκαν αλλά δεν προβλέπονταν από το μοντέλο. Αυτές οι τιμές επηρεάζουν άμεσα τον τελικό υπολογισμό της καταλληλότητας (fitness), καθώς ο αλγόριθμος προσπαθεί να βρει τη διαδρομή με το μικρότερο δυνατό συνολικό κόστος. Επιπλέον, μέσω της μπάρας κύλισης στο επάνω μέρος,

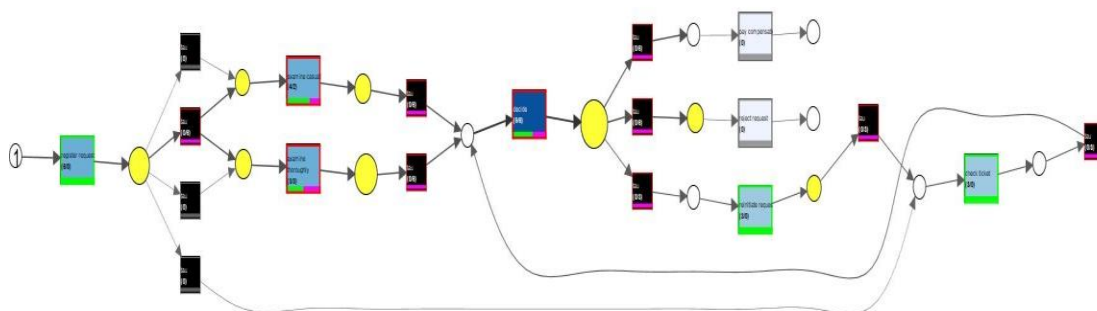
ρυθμίζεται ο μέγιστος αριθμός καταστάσεων που θα εξεταστούν, διασφαλίζοντας την αποδοτικότητα της ανάλυσης. Πατώντας το κουμπί "Finish", ολοκληρώνεται η παραμετροποίηση και ξεκινά η εκτέλεση του Token-Based Replay (Εικόνα 35).



Εικόνα 35 : Καθορισμός κόστους αποκλίσεων για τη διαδικασία αναπαραγωγής

Η εικόνα αυτή παρουσιάζει το τελικό αποτέλεσμα της διαδικασίας ελέγχου συμμόρφωσης, όπου το αρχείο καταγραφής έχει αναπαραχθεί πάνω στο δίκτυο Petri. Οι χρωματικοί δείκτες στις μεταβάσεις παρέχουν άμεση πληροφόρηση για τον τρόπο με τον οποίο εκτελέστηκαν οι δραστηριότητες: το πράσινο χρώμα υποδεικνύει πλήρη συμμόρφωση, ενώ το κόκκινο και το μοβ φανερώνουν σημεία όπου υπήρξαν προβλήματα στην κίνηση των διακριτικών (tokens). Συγκεκριμένα, απεικονίζονται περιπτώσεις όπου έλειπαν διακριτικά για να εκτελεστεί μια ενέργεια ή περιπτώσεις όπου περίσσεψαν διακριτικά μετά την ολοκλήρωση της διεργασίας.

Αυτή η οπτική αναπαράσταση επιτρέπει τον εύκολο εντοπισμό των αδύναμων σημείων του μοντέλου και των αποκλίσεων της πραγματικής ροής, προσφέροντας μια συνολική εικόνα της αξιοπιστίας της διεργασίας (**Εικόνα 36**).



Εικόνα 36: Οπτικοποίηση των αποτελεσμάτων του Token-Based Replay επί του δικτύου Petri

3.3 Η βιβλιοθήκη PM4Py

3.3.1 Περιγραφή της βιβλιοθήκης

Η PM4Py (Process Mining for Python) αποτελεί μία από τις πιο διαδεδομένες και επιστημονικά τεκμηριωμένες βιβλιοθήκες ανοικτού κώδικα για την ανάλυση επιχειρησιακών διαδικασιών μέσω τεχνικών process mining (Berti, van Zelst, & van der Aalst, 2019). Η βιβλιοθήκη έχει αναπτυχθεί με στόχο να προσφέρει ένα ολοκληρωμένο σύνολο εργαλείων για την ανακάλυψη, ανάλυση, αξιολόγηση και βελτιστοποίηση διαδικασιών, αξιοποιώντας δεδομένα καταγραφής γεγονότων (event logs) που προέρχονται από πληροφοριακά συστήματα (van der Aalst, 2016). Η PM4Py είναι υλοποιημένη στη γλώσσα προγραμματισμού Python, γεγονός που την καθιστά ιδιαίτερα ελκυστική για ερευνητές, φοιτητές και επαγγελματίες, καθώς μπορεί να ενσωματωθεί εύκολα σε επιστημονικές ροές εργασίας, περιβάλλοντα ανάλυσης δεδομένων και εφαρμογές μηχανικής μάθησης (Berti et al., 2019).

Η έννοια του process mining, την οποία υποστηρίζει η PM4Py, γεφυρώνει το χάσμα μεταξύ

της κλασικής μοντελοποίησης επιχειρησιακών διαδικασιών και της ανάλυσης δεδομένων. Αντί οι διαδικασίες να περιγράφονται αποκλειστικά με βάση συνεντεύξεις ή θεωρητικές προδιαγραφές, το process mining βασίζεται σε πραγματικά δεδομένα εκτέλεσης, προσφέροντας μια αντικειμενική και αποδείξιμη εικόνα της λειτουργίας ενός οργανισμού (van der Aalst, 2011· van der Aalst, 2016). Σε αυτό το πλαίσιο, η PM4Py λειτουργεί ως το τεχνολογικό υπόβαθρο που επιτρέπει την εφαρμογή αλγορίθμων ανακάλυψης διαδικασιών, ελέγχου συμμόρφωσης και βελτίωσης απόδοσης (Reinkemeyer, 2020).

Η βιβλιοθήκη υποστηρίζει ευρέως διαδεδομένα πρότυπα αρχείων καταγραφής γεγονότων, όπως το XES (eXtensible Event Stream) και το CSV, επιτρέποντας την εύκολη εισαγωγή δεδομένων από συστήματα ERP, CRM, BPM και άλλες επιχειρησιακές πλατφόρμες (Berti et al., 2019). Παράλληλα, προσφέρει δυνατότητες προεπεξεργασίας δεδομένων, όπως φιλτράρισμα θορύβου, επιλογή περιπτώσεων (cases), χρονική ανάλυση και εμπλουτισμό δεδομένων, οι οποίες είναι κρίσιμες για την αξιοπιστία των αποτελεσμάτων (van der Aalst, 2016).

Η PM4Py έχει έντονο ερευνητικό προσανατολισμό και βασίζεται σε θεωρητικά θεμέλια που έχουν αναπτυχθεί στον χώρο του process mining τις τελευταίες δύο δεκαετίες. Πολλοί από τους αλγορίθμους που περιλαμβάνονται στη βιβλιοθήκη έχουν δημοσιευθεί σε διεθνή επιστημονικά περιοδικά και συνέδρια, γεγονός που ενισχύει την εγκυρότητα και την αξιοπιστία της. Ταυτόχρονα, η βιβλιοθήκη δεν περιορίζεται σε ακαδημαϊκή χρήση, αλλά χρησιμοποιείται και σε βιομηχανικά περιβάλλοντα, όπου απαιτείται ανάλυση πολύπλοκων και μεγάλου όγκου δεδομένων (Reinkemeyer, 2020).

Ένα σημαντικό πλεονέκτημα της PM4Py είναι η ευελιξία που προσφέρει ως προς τη χρήση της. Μπορεί να αξιοποιηθεί τόσο μέσω απλών σεναρίων Python όσο και σε διαδραστικά περιβάλλοντα, όπως τα Jupyter Notebooks, διευκολύνοντας την εξερευνητική ανάλυση και την παρουσίαση αποτελεσμάτων (Berti et al., 2019). Επιπλέον, η βιβλιοθήκη επιτρέπει την οπτικοποίηση των ανακαλυφθέντων μοντέλων διαδικασιών, χρησιμοποιώντας γραφήματα Petri nets, BPMN διαγράμματα και άλλες αναπαραστάσεις, καθιστώντας τα αποτελέσματα κατανοητά ακόμη και σε μη τεχνικούς χρήστες (van der Aalst, 2016).

Συνολικά, η PM4Py μπορεί να θεωρηθεί ως ένα πλήρες οικοσύστημα εργαλείων process mining, το οποίο συνδυάζει θεωρητική αυστηρότητα, πρακτική χρησιμότητα και

τεχνολογική ευελιξία. Η σημασία της για την επιστημονική έρευνα αλλά και για την εφαρμοσμένη ανάλυση διαδικασιών την καθιστά βασικό σημείο αναφοράς στον συγκεκριμένο τομέα (Berti et al., 2019; van der Aalst, 2016).

3.3.2 Βασικά Χαρακτηριστικά

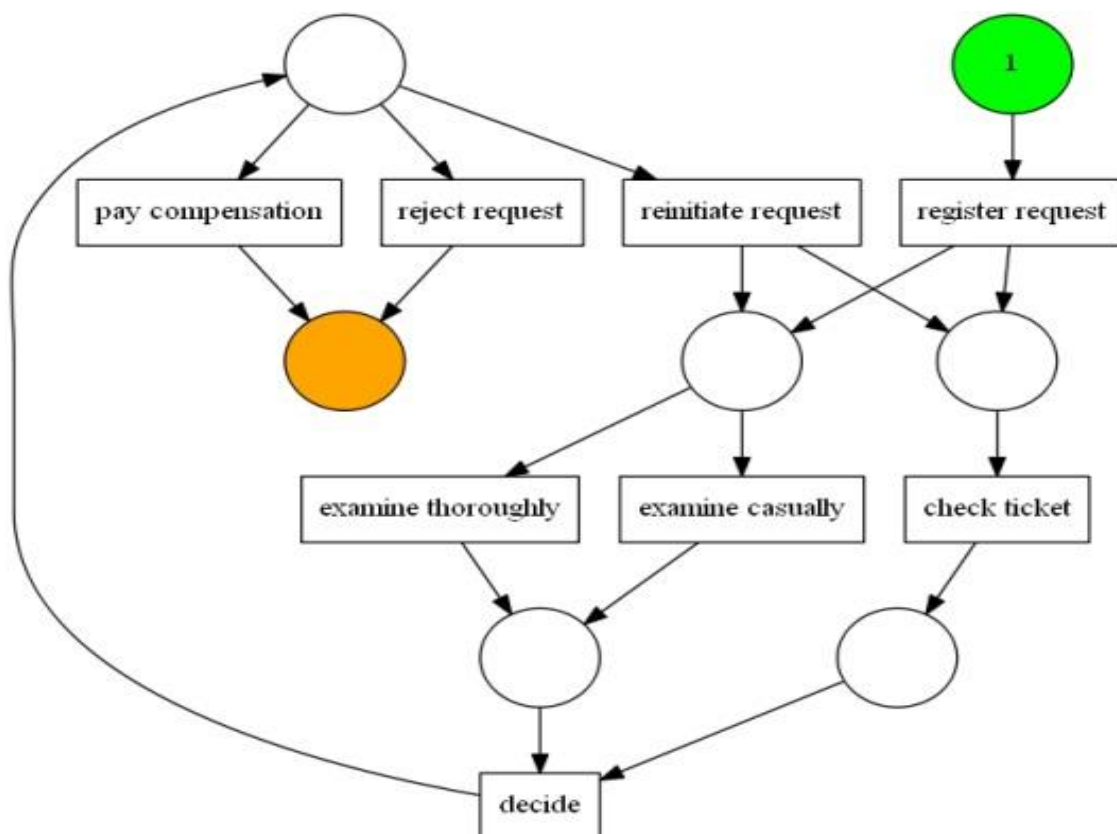
Ένα από τα βασικότερα χαρακτηριστικά της PM4Py είναι η εκτενής υποστήριξη αλγορίθμων ανακάλυψης διαδικασιών. Η βιβλιοθήκη περιλαμβάνει κλασικούς αλλά και σύγχρονους αλγορίθμους, όπως τον Alpha Miner, τον Heuristic Miner, τον Inductive Miner και παραλλαγές αυτών. Οι αλγόριθμοι αυτοί επιτρέπουν τη δημιουργία μοντέλων διαδικασιών με διαφορετικά επίπεδα αυστηρότητας και γενίκευσης, ανάλογα με τις ανάγκες της ανάλυσης και την ποιότητα των δεδομένων (van der Aalst, 2016). Ιδιαίτερη σημασία έχει ο Inductive Miner, ο οποίος παράγει πάντα δομημένα και ορθά μοντέλα, γεγονός που τον καθιστά ιδιαίτερα χρήσιμο σε ακαδημαϊκά και βιομηχανικά σενάρια (van der Aalst, 2016).

Εξίσου σημαντικός είναι ο τομέας του ελέγχου συμμόρφωσης (conformance checking), στον οποίο η PM4Py παρέχει εργαλεία για τη σύγκριση μεταξύ ενός θεωρητικού μοντέλου διαδικασίας και των πραγματικών εκτελέσεων που καταγράφονται στα event logs. Μέσω τεχνικών όπως τα alignments και οι μετρικές fitness και precision, είναι δυνατή η ποσοτική αξιολόγηση του βαθμού στον οποίο μια διαδικασία ακολουθεί τις προκαθορισμένες προδιαγραφές της. Αυτό είναι ιδιαίτερα χρήσιμο σε περιβάλλοντα όπου η συμμόρφωση με κανονιστικά πλαίσια και εσωτερικούς κανόνες είναι κρίσιμη (Carmona et al., 2018).

Στην **Εικόνα 37** [<https://www.researchgate.net>], βλέπουμε ένα παράδειγμα κώδικα χρησιμοποιώντας την βιβλιοθήκη PM4Py, και πιο συγκεκριμένα τον αλγόριθμο Alpha Miner, με στόχο την ανακάλυψη διεργασιών (process discovery), ενώ στην **Εικόνα 38** [<https://www.researchgate.net>], βλέπουμε την απεικόνιση του μοντέλου διεργασιών χρησιμοποιώντας στο συγκεκριμένο παράδειγμα τον αλγόριθμο Alpha Miner.

```
from pm4py.objects.log.importer.xes import factory as xes_importer
from pm4py.algo.discovery.alpha import factory as alpha_miner
from pm4py.visualization.petrinet import factory as pn_vis_factory
log = xes_importer.apply("C:\\receipt.xes")
# discovers a Petri net along with an initial (im)
# and a final marking (fm)
net, im, fm = alpha_miner.apply(log)
gviz = pn_vis_factory.apply(net, im, fm)
pn_vis_factory.view(gviz)
```

Εικόνα 37: Παράδειγμα κώδικα με στόχο την ανακάλυψη διεργασιών (process discovery)



Εικόνα 38: Μοντέλο διεργασιών χρησιμοποιώντας τον αλγόριθμο Alpha Miner

Η PM4Py υποστηρίζει επίσης αναλυτικές λειτουργίες απόδοσης (performance analysis), επιτρέποντας τη μελέτη χρονικών καθυστερήσεων, χρόνων διεκπεραίωσης και σημείων συμφόρησης (bottlenecks). Μέσω αυτών των δυνατοτήτων, οι χρήστες μπορούν να εντοπίσουν προβλήματα αποδοτικότητας και να προτείνουν στοχευμένες βελτιώσεις στις

επιχειρησιακές διαδικασίες. Η ανάλυση αυτή μπορεί να πραγματοποιηθεί τόσο σε επίπεδο συνολικής διαδικασίας όσο και σε επίπεδο μεμονωμένων δραστηριοτήτων (van der Aalst et al., 2012).

Ένα ακόμη βασικό χαρακτηριστικό της βιβλιοθήκης είναι η ισχυρή υποστήριξη προεπεξεργασίας και καθαρισμού δεδομένων. Στην πράξη, τα δεδομένα καταγραφής γεγονότων συχνά περιέχουν θόρυβο, ελλείψεις ή ασυνέπειες. Η PM4Py παρέχει μηχανισμούς φιλτραρίσματος βάσει χρόνου, δραστηριοτήτων, πόρων και περιπτώσεων, επιτρέποντας τη δημιουργία καθαρών και αντιπροσωπευτικών συνόλων δεδομένων. Αυτή η λειτουργικότητα είναι κρίσιμη για την εγκυρότητα των αποτελεσμάτων του process mining (van der Aalst, 2016).

Η διαλειτουργικότητα της PM4Py με άλλες βιβλιοθήκες του οικοσυστήματος Python αποτελεί επίσης σημαντικό πλεονέκτημα. Η δυνατότητα συνδυασμού της με βιβλιοθήκες όπως η pandas, η numpy και εργαλεία μηχανικής μάθησης επιτρέπει την υλοποίηση σύνθετων αναλύσεων και πειραματικών μεθοδολογιών. Έτσι, η PM4Py δεν λειτουργεί απομονωμένα, αλλά εντάσσεται σε ένα ευρύτερο πλαίσιο ανάλυσης δεδομένων (Reinkemeyer, 2020).

3.3.3 Αρχιτεκτονική και Επεκτασιμότητα

Η αρχιτεκτονική της βιβλιοθήκης PM4Py έχει σχεδιαστεί με βασική αρχή τη διαχωρισμένη λειτουργικότητα και τη modular ανάπτυξη, ώστε να υποστηρίζει τόσο απλές όσο και ιδιαίτερα σύνθετες αναλύσεις process mining. Η βιβλιοθήκη οργανώνεται σε επιμέρους υποσυστήματα, τα οποία καλύπτουν διαφορετικά στάδια της διαδικασίας ανάλυσης, όπως η εισαγωγή και αναπαράσταση δεδομένων, η προεπεξεργασία event logs, η ανακάλυψη μοντέλων διαδικασιών, ο έλεγχος συμμόρφωσης, η ανάλυση απόδοσης και η οπτικοποίηση αποτελεσμάτων. Αυτή η αρχιτεκτονική επιτρέπει στον χρήστη να αξιοποιεί μόνο τα τμήματα που είναι απαραίτητα για το εκάστοτε ερευνητικό ή εφαρμοσμένο σενάριο (Berti et al., 2019).

Κεντρικό ρόλο στην αρχιτεκτονική της PM4Py κατέχει η εσωτερική αναπαράσταση των event logs. Τα δεδομένα οργανώνονται σε δομές που βασίζονται σε αντικείμενα, επιτρέποντας τη σαφή μοντελοποίηση εννοιών όπως τα γεγονότα, οι περιπτώσεις (cases), οι

δραστηριότητες και οι χρονικές σφραγίδες. Αυτή η προσέγγιση διευκολύνει τόσο την ανάλυση όσο και την επέκταση της βιβλιοθήκης, καθώς νέες ιδιότητες ή διαστάσεις δεδομένων μπορούν να προστεθούν χωρίς να απαιτείται ριζική αναδιάρθρωση του υπάρχοντος κώδικα (Berti et al., 2019).

Η επεκτασιμότητα της PM4Py αποτελεί ένα από τα πιο σημαντικά χαρακτηριστικά της, ιδιαίτερα στον ακαδημαϊκό χώρο. Η βιβλιοθήκη έχει σχεδιαστεί έτσι ώστε να επιτρέπει την εύκολη ενσωμάτωση νέων αλγορίθμων process mining, είτε πρόκειται για νέες μεθόδους ανακάλυψης διαδικασιών είτε για εναλλακτικές τεχνικές ελέγχου συμμόρφωσης. Οι ερευνητές μπορούν να υλοποιήσουν τις προτεινόμενες μεθόδους τους αξιοποιώντας τις υπάρχουσες διεπαφές (interfaces) και να τις συγκρίνουν πειραματικά με καθιερωμένους αλγορίθμους της βιβλιοθήκης, διασφαλίζοντας έτσι την επιστημονική εγκυρότητα των αποτελεσμάτων (Berti et al., 2019· van der Aalst, 2016).

Ένα ακόμη σημαντικό στοιχείο της αρχιτεκτονικής είναι η υποστήριξη πολλαπλών αναπαραστάσεων μοντέλων διαδικασιών. Η PM4Py επιτρέπει τη μετάβαση μεταξύ διαφορετικών μορφών, όπως Petri nets, process trees και BPMN διαγράμματα. Αυτή η δυνατότητα δεν είναι απλώς ζήτημα οπτικοποίησης, αλλά επηρεάζει άμεσα την ανάλυση και την επεκτασιμότητα, καθώς διαφορετικοί αλγόριθμοι και τεχνικές βασίζονται σε διαφορετικές μαθηματικές αναπαραστάσεις. Η ύπαρξη μηχανισμών μετατροπής μεταξύ αυτών των μοντέλων καθιστά τη βιβλιοθήκη ιδιαίτερα ευέλικτη (van der Aalst, 2011· van der Aalst, 2016).

Η επιλογή της Python ως γλώσσας υλοποίησης επηρεάζει καθοριστικά την αρχιτεκτονική και την επεκτασιμότητα της PM4Py. Η Python προσφέρει υψηλό επίπεδο αφαίρεσης, γεγονός που επιτρέπει την ταχεία ανάπτυξη και πειραματισμό με νέες ιδέες. Παράλληλα, η συμβατότητα με βιβλιοθήκες επιστημονικών υπολογισμών και ανάλυσης δεδομένων ενισχύει τη δυνατότητα ενσωμάτωσης της PM4Py σε σύνθετες ροές εργασίας, όπως αυτές που συνδυάζουν process mining με μηχανική μάθηση ή στατιστική ανάλυση (Berti et al., 2019).

Επιπλέον, η αρχιτεκτονική της PM4Py υποστηρίζει τη σταδιακή κλιμάκωση (scalability) της ανάλυσης. Παρότι η βιβλιοθήκη χρησιμοποιείται συχνά σε ερευνητικά περιβάλλοντα, έχει σχεδιαστεί έτσι ώστε να μπορεί να διαχειριστεί μεγάλα σύνολα δεδομένων, αξιοποιώντας

αποδοτικούς αλγορίθμους και δομές δεδομένων. Αυτό επιτρέπει την εφαρμογή της σε πραγματικά επιχειρησιακά σενάρια, όπου οι διαδικασίες παράγουν μεγάλο όγκο γεγονότων σε συνεχή βάση (Berti et al., 2019· van der Aalst, 2016).

Συνοψίζοντας, η αρχιτεκτονική και η επεκτασιμότητα της PM4Py την καθιστούν ένα ιδιαίτερα ισχυρό και ευέλικτο εργαλείο για το process mining. Η αρθρωτή σχεδίαση, η υποστήριξη πολλαπλών μοντέλων, η ευκολία ενσωμάτωσης νέων αλγορίθμων και η συμβατότητα με το ευρύτερο οικοσύστημα της Python συμβάλλουν στη μακροχρόνια βιωσιμότητα και επιστημονική αξία της βιβλιοθήκης (Berti et al., 2019).

3.4 Αξιολόγηση και κριτήρια επιλογής των εργαλείων εξόρυξης διεργασιών

Η επιλογή της κατάλληλης εργαλειοθήκης για μια διπλωματική εργασία που εστιάζει στο Process Mining και το Conformance Checking αποτελεί μια κρίσιμη απόφαση, καθώς καθορίζει την εγκυρότητα της ανάλυσης και την ποιότητα των συμπερασμάτων. Ο συνδυασμός των ProM, PM4Py και Visual Paradigm επιλέχθηκε με γνώμονα τη δημιουργία μιας ολοκληρωμένης μεθοδολογικής ροής, η οποία καλύπτει το κενό ανάμεσα στον θεωρητικό σχεδιασμό και την προηγμένη υπολογιστική ανάλυση δεδομένων.

Το ProM (Process Mining framework) αποτελεί την πλέον αναγνωρισμένη και καθιερωμένη πλατφόρμα στην ακαδημαϊκή έρευνα. Η επιλογή του έναντι εμπορικών λογισμικών, όπως το Disco ή το Celonis, βασίζεται στον ανοιχτό και ερευνητικό του χαρακτήρα. Ενώ τα εμπορικά εργαλεία προσφέρουν ταχύτητα και φιλικό περιβάλλον, συχνά προσφέρουν τυποποιημένες λειτουργίες χωρίς πρόσβαση στη λογική των αλγορίθμων τους, περιορίζοντας τον χρήστη σε προκαθορισμένες διαδικασίες. Το ProM, αντίθετα, παρέχει πρόσβαση σε εκατοντάδες plugins που αναπτύσσονται από την παγκόσμια επιστημονική κοινότητα, επιτρέποντας στον ερευνητή να πειραματιστεί με state-of-the-art τεχνικές ανακάλυψης διεργασιών. Με αυτόν τον τρόπο, η διπλωματική εργασία αποκτά επιστημονικό βάθος, καθώς η ανάλυση δεν περιορίζεται σε μια απλή απεικόνιση, αλλά εμβαθύνει στις λεπτομέρειες των αλγορίθμων και της δομής των event logs.

Συμπληρωματικά προς το ProM, η βιβλιοθήκη PM4Py (Process Mining for Python) εισάγει στην έρευνα την απαραίτητη υπολογιστική ευελιξία και τον προγραμματιστικό έλεγχο. Σε

σύγκριση με άλλες βιβλιοθήκες, όπως η bupaR (της γλώσσας R), η PM4Py υπερέχει λόγω της κυριαρχίας της Python στο σύγχρονο οικοσύστημα της επιστήμης δεδομένων (Data Science). Η επιλογή της είναι καθοριστική για το κομμάτι του Conformance Checking, καθώς προσφέρει εξαιρετικά αποδοτικούς αλγορίθμους για τον υπολογισμό μετρικών όπως το fitness και το precision. Η χρήση της PM4Py επιτρέπει την αυτοματοποίηση της ανάλυσης και τη διασύνδεση των αποτελεσμάτων με εργαλεία μηχανικής μάθησης, διασφαλίζοντας ότι η μελέτη είναι αναπαραγώγιμη και συμβατή με τις τρέχουσες τεχνολογικές τάσεις.

Για τη δημιουργία του μοντέλου αναφοράς (reference model), επιλέχθηκε το Visual Paradigm, ένα εργαλείο που ξεφεύγει από τα όρια μιας απλής εφαρμογής σχεδίασης διαγραμμάτων. Σε αντίθεση με γενικά εργαλεία όπως το Visio ή το Lucidchart, το Visual Paradigm είναι ένα ολοκληρωμένο περιβάλλον CASE που προσφέρει αυστηρή υποστήριξη του προτύπου BPMN 2.0. Αυτό είναι κρίσιμο για τη διαδικασία του ελέγχου συμμόρφωσης, καθώς διασφαλίζει ότι η «ιδανική» διαδικασία που σχεδιάζουμε είναι συντακτικά ορθή και απαλλαγμένη από λογικά σφάλματα. Η δυνατότητα εξαγωγής των μοντέλων σε τυποποιημένες μορφές (όπως .bpnm ή .xml) επιτρέπει την άμεση εισαγωγή τους στα αναλυτικά εργαλεία (ProM και PM4Py), ολοκληρώνοντας έτσι μια ενιαία ροή εργασίας χωρίς προβλήματα ασυμβατότητας.

Συνοψίζοντας, η συνέργεια αυτών των τριών εργαλείων προσφέρει μια ισχυρή μεθοδολογική βάση. Το Visual Paradigm ορίζει το θεωρητικό πλαίσιο (τι πρέπει να συμβαίνει), το ProM ανακαλύπτει την πραγματική κατάσταση από τα δεδομένα (τι πραγματικά συμβαίνει) και η PM4Py γεφυρώνει τα δύο μέσω του ελέγχου συμμόρφωσης (πού υπάρχουν αποκλίσεις). Ο συνδυασμός αυτός εξασφαλίζει ότι η διπλωματική εργασία δεν είναι απλώς μια περιγραφή δεδομένων, αλλά μια τεκμηριωμένη επιστημονική μελέτη που αξιοποιεί τα πιο εξελιγμένα εργαλεία που διαθέτει σήμερα η ακαδημαϊκή και προγραμματιστική κοινότητα.

ΚΕΦΑΛΑΙΟ 4

Εικονικά Εργαστήρια

4.1 Εισαγωγή

Η έννοια των εικονικών εργαστηρίων έχει αναδειχθεί ως βασικό στοιχείο της σύγχρονης εκπαιδευτικής τεχνολογίας, ιδιαίτερα στο πλαίσιο της ανώτατης εκπαίδευσης και της δια βίου μάθησης. Τα εικονικά εργαστήρια ορίζονται ως ψηφιακά περιβάλλοντα τα οποία επιτρέπουν την προσομοίωση εργαστηριακών διαδικασιών, πειραμάτων και πρακτικών δραστηριοτήτων μέσω υπολογιστικών συστημάτων. Στόχος τους είναι η υποστήριξη της ενεργητικής μάθησης, δίνοντας τη δυνατότητα στους εκπαιδευόμενους να αλληλεπιδρούν με εικονικά αντικείμενα και συστήματα σε ένα ελεγχόμενο και ασφαλές περιβάλλον (de Jong, 2013· De Jong, Linn, & Zacharia, 2013).

Η ανάγκη για ανάπτυξη εικονικών εργαστηρίων προέκυψε κυρίως από τους περιορισμούς των παραδοσιακών φυσικών εργαστηρίων. Παράγοντες όπως το υψηλό κόστος εξοπλισμού, οι αυξημένες απαιτήσεις συντήρησης, οι περιορισμοί χωρητικότητας και η ανάγκη φυσικής παρουσίας καθιστούν συχνά δύσκολη την ισότιμη πρόσβαση όλων των φοιτητών σε εργαστηριακές υποδομές. Τα εικονικά εργαστήρια έρχονται να αντιμετωπίσουν αυτά τα ζητήματα, προσφέροντας ευέλικτες και επεκτάσιμες λύσεις (Sypsas et al., 2024).

Ιδιαίτερη σημασία αποκτούν τα εικονικά εργαστήρια στο πλαίσιο της εξ αποστάσεως εκπαίδευσης. Σε εκπαιδευτικά ιδρύματα όπως το Ελληνικό Ανοικτό Πανεπιστήμιο, όπου η φυσική παρουσία δεν αποτελεί βασική προϋπόθεση φοίτησης, τα εικονικά εργαστήρια διασφαλίζουν τη σύνδεση της θεωρητικής γνώσης με την πρακτική εφαρμογή. Με τον τρόπο αυτό, οι φοιτητές δεν περιορίζονται στη μελέτη θεωρητικού υλικού, αλλά αποκτούν βιωματική εμπειρία, η οποία ενισχύει την κατανόηση σύνθετων εννοιών (Sypsas et al., 2024).

Επιπλέον, τα εικονικά εργαστήρια υποστηρίζουν σύγχρονες παιδαγωγικές προσεγγίσεις, όπως η διερευνητική μάθηση και η μάθηση μέσω προβλημάτων. Οι φοιτητές ενθαρρύνονται να πειραματιστούν, να κάνουν λάθη και να αναστοχαστούν πάνω στα αποτελέσματα των ενεργειών τους. Η διαδικασία αυτή συμβάλλει ουσιαστικά στην ανάπτυξη κριτικής σκέψης

και δεξιοτήτων επίλυσης προβλημάτων (Dalgarno & Lee, 2010· De Jong et al., 2013).

4.2 Το εικονικό εργαστήριο OnLabs

Το OnLabs αποτελεί ένα σύγχρονο εικονικό εργαστήριο που έχει αναπτυχθεί στο πλαίσιο του Ελληνικού Ανοικτού Πανεπιστημίου με στόχο την υποστήριξη της εργαστηριακής εκπαίδευσης σε περιβάλλον εξ αποστάσεως μάθησης. Η ανάπτυξη του βασίστηκε στην ανάγκη παροχής ενός αξιόπιστου, ευέλικτου και παιδαγωγικά τεκμηριωμένου συστήματος, ικανού να καλύψει τις απαιτήσεις διαφορετικών γνωστικών αντικειμένων και μεγάλου αριθμού φοιτητών (Paxinou et al., 2018).

Η φιλοσοφία σχεδίασης του OnLabs δίνει ιδιαίτερη έμφαση στην προσβασιμότητα και τη χρηστικότητα. Η πρόσβαση στο σύστημα πραγματοποιείται μέσω ενός απλού διαδικτυακού φυλλομετρητή, χωρίς την ανάγκη εγκατάστασης εξειδικευμένου λογισμικού ή τη χρήση ειδικού εξοπλισμού. Η επιλογή αυτή είναι ιδιαίτερα σημαντική στο πλαίσιο του ΕΑΠ, καθώς διασφαλίζει την ισότιμη συμμετοχή φοιτητών με διαφορετικό τεχνολογικό υπόβαθρο και διαφορετικές συνθήκες μελέτης (Sypsas et al., 2024). Η **Εικόνα 39** [Karpouzas et al., 2025], απεικονίζει το OnLabs, ένα 3D εργαστήριο βιολογίας σε εικονική πραγματικότητα. Το περιβάλλον αυτό προσφέρει μια εμπειρία μάθησης, επιτρέποντας στους χρήστες να κατανοήσουν και να εφαρμόσουν πειραματικές τεχνικές.

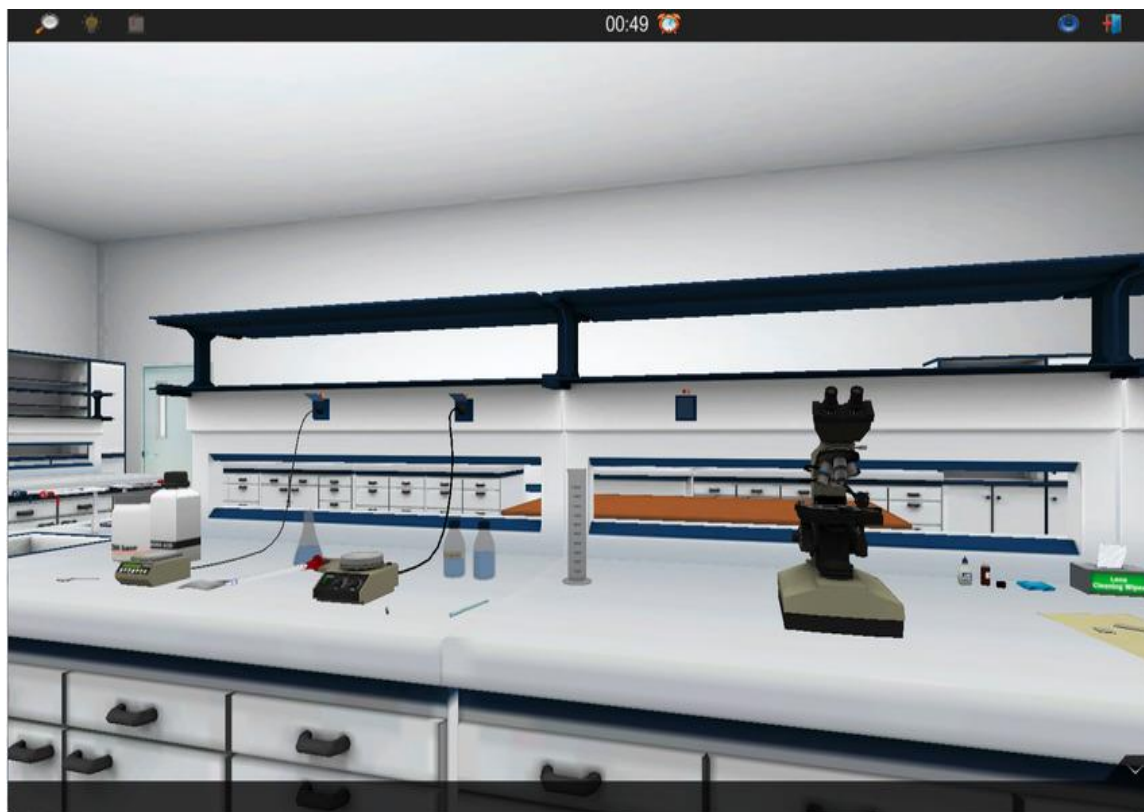


Εικόνα 39: Εστίαση στο OnLabs – 3D Εικονικής Πραγματικότητας Εργαστήριο Βιολογίας

Το OnLabs λειτουργεί ως ένα ολοκληρωμένο ψηφιακό οικοσύστημα μάθησης και όχι απλώς ως εργαλείο εκτέλεσης εργαστηριακών ασκήσεων. Ενσωματώνει παιδαγωγικές αρχές που προάγουν τη βιωματική, διερευνητική και ενεργητική μάθηση. Οι φοιτητές καλούνται να αλληλεπιδρούν με εικονικά συστήματα, να λαμβάνουν αποφάσεις και να παρατηρούν τις συνέπειες των ενεργειών τους, γεγονός που ενισχύει την κατανόηση σύνθετων και αφηρημένων εννοιών.

Ιδιαίτερη σημασία έχει η δυνατότητα επαναληψιμότητας των εργαστηριακών δραστηριοτήτων. Οι φοιτητές μπορούν να επαναλαμβάνουν τα σενάρια όσες φορές επιθυμούν, να δοκιμάζουν εναλλακτικές προσεγγίσεις και να μαθαίνουν μέσα από τα λάθη τους, χωρίς τον φόβο αποτυχίας. Το στοιχείο αυτό συμβάλλει ουσιαστικά στην εμβάθυνση της γνώσης και στην ανάπτυξη δεξιοτήτων επίλυσης προβλημάτων. Η **Εικόνα 40** [<https://www.researchgate.net>], απεικονίζει το OnLabs σε λειτουργία πειραματισμού, προσφέροντας μια ρεαλιστική εμπειρία εργαστηριακής πρακτικής. Το περιβάλλον αυτό βοηθά τους χρήστες να κατανοήσουν και να εφαρμόσουν πειραματικές διαδικασίες με

ασφαλή τρόπο.



Εικόνα 40: Το Εργαστήριο OnLabs σε Λειτουργία Πειραματισμού

Παράλληλα, το OnLabs γεφυρώνει αποτελεσματικά το χάσμα μεταξύ θεωρίας και πράξης, το οποίο συχνά παρατηρείται στην εξ αποστάσεως εκπαίδευση. Οι εργαστηριακές ασκήσεις είναι άμεσα συνδεδεμένες με το θεωρητικό περιεχόμενο των μαθημάτων, επιτρέποντας στους φοιτητές να εφαρμόζουν στην πράξη όσα έχουν μελετήσει. Η σύνδεση αυτή ενισχύει τη μεταφορά γνώσης και την ουσιαστική κατανόηση του γνωστικού αντικειμένου.

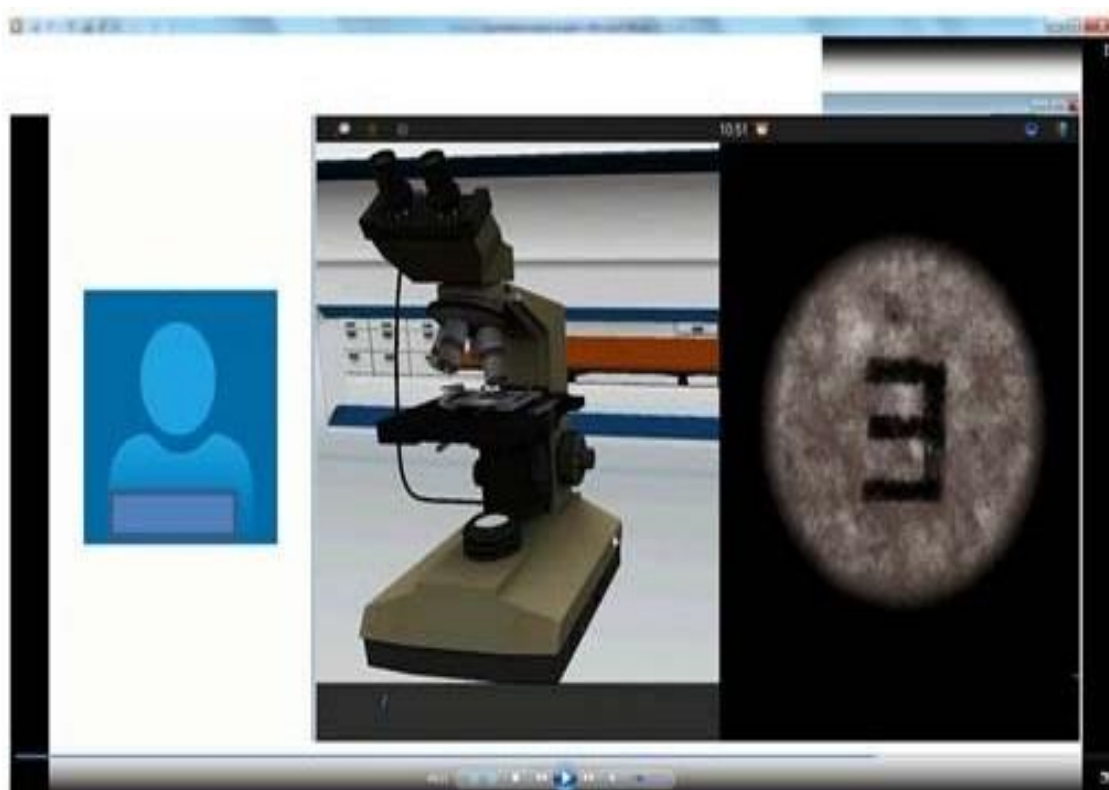
Τέλος, το OnLabs ενισχύει την αυτονομία και την αυτορρύθμιση των φοιτητών. Οι εκπαιδευόμενοι έχουν τη δυνατότητα να οργανώνουν τον χρόνο μελέτης τους, να παρακολουθούν την πρόοδό τους και να αναστοχάζονται πάνω στη μαθησιακή τους πορεία. Το χαρακτηριστικό αυτό συνάδει πλήρως με τη φιλοσοφία του ΕΑΠ και καθιστά το OnLabs βασικό πυλώνα της εξ αποστάσεως εργαστηριακής εκπαίδευσης.

4.2.1 Αρχιτεκτονική και Λειτουργικότητα του OnLabs

Η αρχιτεκτονική του OnLabs βασίζεται σε μια σύγχρονη πολυεπίπεδη προσέγγιση, η οποία διασφαλίζει την ευελιξία, την επεκτασιμότητα και τη μακροχρόνια βιωσιμότητα του συστήματος. Ο σαφής διαχωρισμός μεταξύ διεπαφής χρήστη, επιχειρησιακής λογικής και διαχείρισης δεδομένων επιτρέπει την ανεξάρτητη εξέλιξη κάθε επιπέδου, χωρίς να επηρεάζεται η συνολική λειτουργία του εικονικού εργαστηρίου (Sypsas & Kalles, 2023; Sypsas & Kalles, 2022a).

Η διεπαφή χρήστη αποτελεί το βασικό σημείο αλληλεπίδρασης των φοιτητών με το OnLabs. Έχει σχεδιαστεί με γνώμονα τη χρηστικότητα και τη μείωση της γνωστικής επιβάρυνσης, επιτρέποντας στους χρήστες να επικεντρωθούν στις εργαστηριακές δραστηριότητες. Η σαφής οργάνωση των λειτουργιών και η οπτική ανατροφοδότηση διευκολύνουν την πλοήγηση και ενισχύουν την εμπειρία χρήσης (Parastatidis et al., 2022).

Η **Εικόνα 41** [Paxinou et al., 2018], παρουσιάζει μια συνεδρία Skype για μικροσκοπηση στο πλαίσιο του εικονικού εργαστηρίου OnLabs, αξιοποιώντας το περιβάλλον προσομοίωσης. Η διαδικασία αυτή επιτρέπει στους συμμετέχοντες να παρακολουθήσουν και να αλληλεπιδράσουν με πειράματα μικροσκοπίας σε πραγματικό χρόνο, προσφέροντας μια ασφαλή και διαδραστική μαθησιακή εμπειρία.



Εικόνα 41: Συνεδρία μέσω Skype για τη μελέτη μικροσκοπίας με χρήση περιβάλλοντος προσομοίωσης

Το επίπεδο της επιχειρησιακής λογικής είναι υπεύθυνο για τη διαχείριση των εκπαιδευτικών σεναρίων και τη ροή των εργαστηριακών δραστηριοτήτων. Σε αυτό το επίπεδο υλοποιούνται οι κανόνες που καθορίζουν τις επιτρεπόμενες ενέργειες του χρήστη, οι συνθήκες μετάβασης μεταξύ των σταδίων μιας άσκησης και οι μηχανισμοί ανατροφοδότησης. Η ύπαρξη προκαθορισμένων ροών εργασίας διασφαλίζει τη συνοχή και την παιδαγωγική εγκυρότητα των σεναρίων (Sypsas & Kalles, 2023; Sypsas & Kalles, 2022a).

Η λειτουργικότητα του OnLabs έχει σχεδιαστεί ώστε να υποστηρίζει τόσο τη μαζική όσο και την εξατομικευμένη εκπαίδευση. Το σύστημα μπορεί να εξυπηρετήσει μεγάλο αριθμό φοιτητών ταυτόχρονα, διατηρώντας σταθερή απόδοση, ενώ παράλληλα προσφέρει τη δυνατότητα προσαρμογής των εργαστηριακών σεναρίων ανάλογα με τις ανάγκες του μαθήματος ή το επίπεδο των φοιτητών (Sypsas et al., 2024).

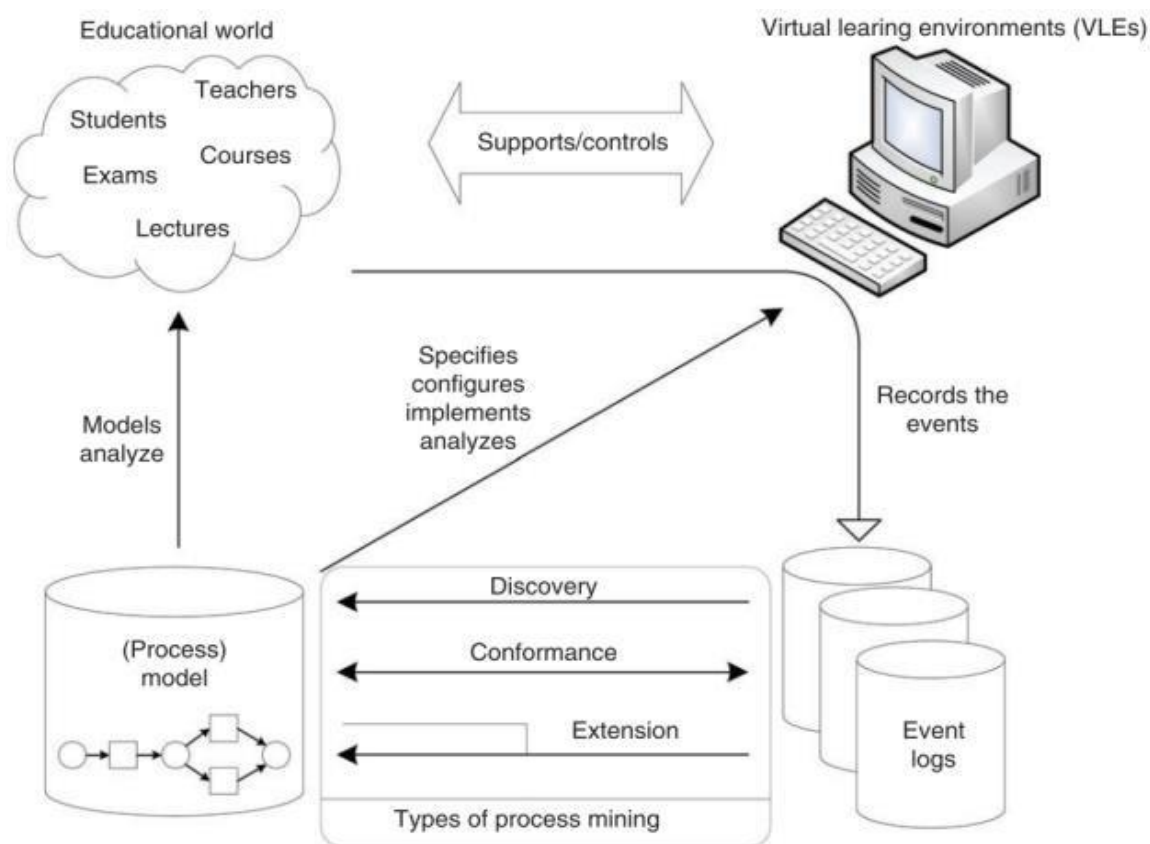
Ιδιαίτερη έμφαση δίνεται στη λειτουργία της ανατροφοδότησης. Το OnLabs μπορεί να

παρέχει άμεση ανατροφοδότηση κατά την εκτέλεση μιας δραστηριότητας, επιτρέποντας στους φοιτητές να εντοπίζουν και να διορθώνουν λάθη σε πραγματικό χρόνο. Παράλληλα, υποστηρίζεται και η έμμεση ανατροφοδότηση μέσω αναφορών και συγκεντρωτικών αποτελεσμάτων, τα οποία ενισχύουν τον αναστοχασμό και την αυτοαξιολόγηση (Sypsas & Kalles, 2022b).

Στο επίπεδο των δεδομένων, το OnLabs αποθηκεύει με οργανωμένο και ασφαλή τρόπο πληροφορίες που σχετίζονται με τους χρήστες, τις δραστηριότητες και τα αποτελέσματα. Η ύπαρξη αυτής της υποδομής δεδομένων αποτελεί τη βάση για την ανάλυση της μαθησιακής συμπεριφοράς και την περαιτέρω αξιοποίηση των δεδομένων σε ερευνητικό επίπεδο (Sypsas & Kalles, 2023).

4.3 Η εφαρμογή της Εξόρυξης Διεργασιών και του Ελέγχου Συμμόρφωσης στην Εκπαιδευτική Διαδικασία

Η εκπαιδευτική διαδικασία, όπως και άλλοι τομείς της κοινωνίας, βιώνει την επιρροή της ψηφιοποίησης και της τεχνολογικής εξέλιξης. Η τεχνολογία προσφέρει νέες δυνατότητες και εργαλεία που ενδέχεται να επηρεάσουν σημαντικά την ποιότητα, την αποτελεσματικότητα και την εξατομίκευση της μάθησης. Σε αυτό το πλαίσιο, η εφαρμογή της εξόρυξης διεργασιών (process mining) και του ελέγχου συμμόρφωσης (conformance checking) στην εκπαίδευση αναδύεται ως μια καινοτόμος προσέγγιση, η οποία έχει τη δυνατότητα να μεταμορφώσει την εκπαιδευτική διαδικασία, βελτιώνοντας τη μαθησιακή εμπειρία, την παρακολούθηση και τη διαχείριση των μαθητών, και τη διασφάλιση της ποιότητας και της δικαιοσύνης (Sypsas & Kalles, 2022c; Sypsas & Kalles, 2022b).



Εικόνα 42: Η εφαρμογή της εξόρυξης διαδικασιών στον τομέα της εκπαίδευσης και ο κύκλος ζωής της

Η εξόρυξη διεργασιών αποτελεί μια τεχνολογική προσέγγιση που χρησιμοποιεί δεδομένα από διάφορες πηγές για να αναλύσει και να κατανοήσει τις διεργασίες που εκτυλίσσονται μέσα σε ένα σύστημα. Στον τομέα της εκπαίδευσης, η εξόρυξη διεργασιών αναφέρεται στη διαδικασία ανάλυσης των δεδομένων που παράγονται από τις μαθησιακές δραστηριότητες των μαθητών μέσα από πλατφόρμες διαχείρισης μάθησης (LMS) ή άλλες ψηφιακές εκπαιδευτικές πλατφόρμες. Μέσω αυτής της διαδικασίας, οι εκπαιδευτικοί και οι διαχειριστές εκπαίδευσης αποκτούν τη δυνατότητα να κατανοήσουν τις πραγματικές μαθησιακές διαδικασίες που ακολουθούν οι μαθητές, να αναγνωρίσουν τις αλληλεπιδράσεις τους με το περιεχόμενο, και να εντοπίσουν τυχόν προβλήματα ή δυσκολίες στην πορεία της μάθησης (Sypsas & Kalles, 2022c).

Η εξόρυξη διεργασιών δίνει στους δασκάλους και τους εκπαιδευτικούς τη δυνατότητα να παρακολουθούν τη συμπεριφορά των μαθητών με μεγαλύτερη ακρίβεια και σε πραγματικό χρόνο. Με την ανάλυση των δεδομένων από την αλληλεπίδραση των μαθητών με τις

πλατφόρμες μάθησης, οι εκπαιδευτικοί μπορούν να εντοπίσουν μοτίβα συμπεριφοράς, να παρακολουθήσουν την πρόοδο των μαθητών και να προβλέψουν πιθανούς κινδύνους αποτυχίας ή καθυστέρησης (Sypsas & Kalles, 2022c; Sypsas & Kalles, 2023). Αυτή η συνεχής ανάλυση μπορεί να αποκαλύψει τι λειτουργεί και τι όχι, παρέχοντας χρήσιμες πληροφορίες για την προσαρμογή του εκπαιδευτικού περιεχομένου και την προώθηση της εξατομικευμένης μάθησης. Για παράδειγμα, ένας δάσκαλος μπορεί να παρατηρήσει ότι οι μαθητές που αφιερώνουν περισσότερη ώρα στην επίλυση συγκεκριμένων τύπων ασκήσεων έχουν υψηλότερη απόδοση στις εξετάσεις, επιτρέποντας έτσι την τροποποίηση του περιεχομένου και την προσθήκη δραστηριοτήτων που ανταποκρίνονται καλύτερα στις ανάγκες των μαθητών.

Η εξόρυξη διεργασιών μπορεί επίσης να ενισχύσει τη διάγνωση των μαθησιακών αναγκών των μαθητών. Αντί να βασίζονται αποκλειστικά σε παραδοσιακές αξιολογήσεις ή εξετάσεις, οι δάσκαλοι μπορούν να αναλύσουν τη συμπεριφορά των μαθητών και να προσαρμόσουν τις μεθόδους διδασκαλίας για να εντοπίσουν συγκεκριμένα σημεία όπου απαιτείται ενίσχυση. Αυτή η εξατομίκευση της μάθησης είναι ιδιαίτερα χρήσιμη σε μεγάλες τάξεις, όπου οι παραδοσιακές μέθοδοι μπορεί να μην είναι αρκετές για να καλύψουν τις ανάγκες όλων των μαθητών (Sypsas & Kalles, 2022c).

Από την άλλη πλευρά, ο έλεγχος συμμόρφωσης αναφέρεται στη διαδικασία που διασφαλίζει ότι οι εκπαιδευτικές διαδικασίες και οι μαθησιακές δραστηριότητες συμμορφώνονται με τα καθορισμένα πρότυπα, κανονισμούς και απαιτήσεις που έχουν τεθεί από τους φορείς εκπαίδευσης. Ο έλεγχος συμμόρφωσης στηρίζεται στην εφαρμογή αυστηρών κανόνων και διαδικασιών για την παρακολούθηση της απόδοσης των μαθητών και την αξιολόγηση της εκπαιδευτικής διαδικασίας, προκειμένου να διασφαλιστεί ότι οι διαδικασίες μάθησης είναι δίκαιες, ακριβείς και συμβατές με τα εκπαιδευτικά πρότυπα (Sypsas & Kalles, 2022c; Carmona et al., 2018). Η εφαρμογή του ελέγχου συμμόρφωσης είναι κρίσιμη, καθώς εξασφαλίζει τη διαφάνεια, την ακεραιότητα και την αξιοπιστία του εκπαιδευτικού συστήματος, αποφεύγοντας αδικίες και αποκλίσεις στη διαδικασία της μάθησης.

Ένα βασικό παράδειγμα εφαρμογής του ελέγχου συμμόρφωσης στην εκπαίδευση είναι η παρακολούθηση της συμμόρφωσης των μαθητών με το πρόγραμμα σπουδών και τις ακαδημαϊκές απαιτήσεις. Μέσω αυτής της διαδικασίας, οι εκπαιδευτικοί μπορούν να εντοπίσουν αν οι μαθητές ακολουθούν το σύστημα αξιολόγησης και αν οι διδασκαλίες τους

συμμορφώνονται με τα προγράμματα σπουδών που έχουν θεσπιστεί. Επίσης, ο έλεγχος συμμόρφωσης μπορεί να αφορά την εφαρμογή του σωστού τύπου αξιολόγησης για κάθε δραστηριότητα, διασφαλίζοντας ότι όλοι οι μαθητές υποβάλλονται στις ίδιες απαιτήσεις και ότι οι αξιολογήσεις είναι αντικειμενικές και αμερόληπτες (Sypsas & Kalles, 2022c; Sypsas & Kalles, 2022b).

Η συνδυασμένη χρήση της εξόρυξης διεργασιών και του ελέγχου συμμόρφωσης παρέχει ένα ισχυρό εργαλείο για την ενίσχυση της ποιότητας της εκπαίδευσης και την παρακολούθηση της διαδικασίας μάθησης. Η εξόρυξη διεργασιών επιτρέπει στους δασκάλους να αποκτούν μια πληρέστερη εικόνα για τη συμπεριφορά των μαθητών και την πρόοδό τους, ενώ ο έλεγχος συμμόρφωσης εξασφαλίζει ότι αυτή η διαδικασία τηρεί τα απαιτούμενα πρότυπα και κανονισμούς. Μέσω της ολοκληρωμένης αυτής προσέγγισης, οι δάσκαλοι μπορούν να εντοπίσουν τις αδυναμίες του εκπαιδευτικού συστήματος και να αναπτύξουν στρατηγικές για να το βελτιώσουν (Sypsas & Kalles, 2022c; Sypsas & Kalles, 2022b).

Παρά τα αναμφισβήτητα πλεονεκτήματα της εφαρμογής αυτών των εργαλείων στην εκπαίδευση, υπάρχουν και αρκετές προκλήσεις. Πρώτα απ' όλα, η προστασία των προσωπικών δεδομένων των μαθητών αποτελεί σημαντικό ζήτημα. Η ανάλυση των δεδομένων απαιτεί τη συλλογή και αποθήκευση ευαίσθητων πληροφοριών, γεγονός που απαιτεί αυστηρούς κανόνες προστασίας της ιδιωτικότητας και της ασφάλειας των δεδομένων. Οι εκπαιδευτικοί και οι οργανισμοί εκπαίδευσης πρέπει να διασφαλίσουν ότι τηρούν τις αυστηρές ρυθμίσεις που προβλέπουν οι νόμοι περί προστασίας προσωπικών δεδομένων (Sypsas & Kalles, 2022c; Sypsas & Kalles, 2023).

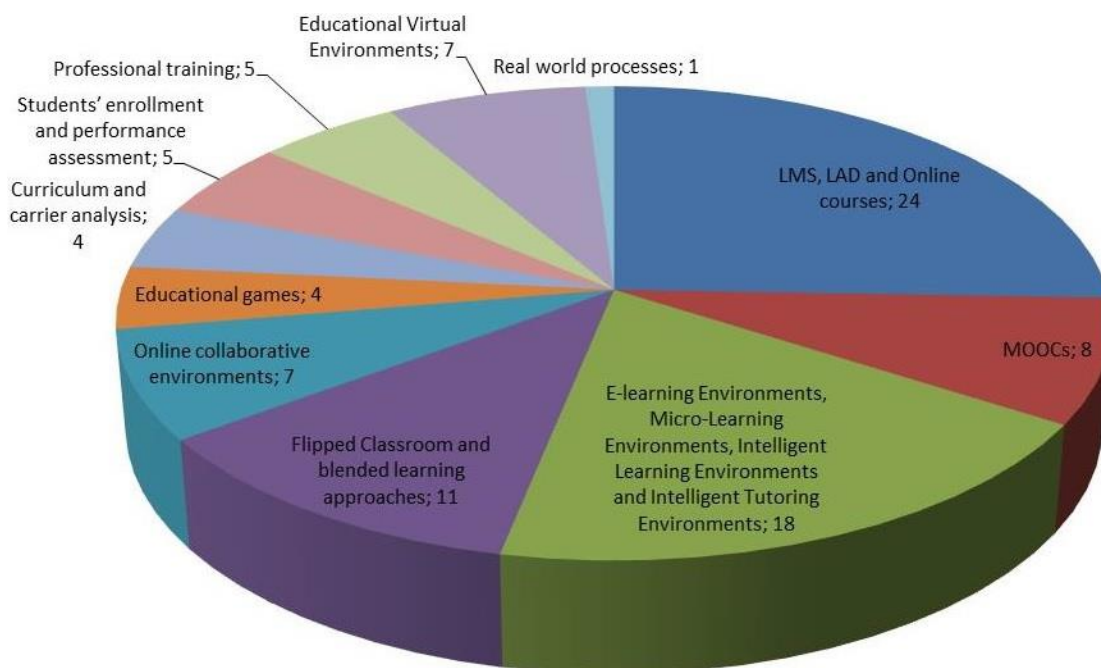
Επιπλέον, η εφαρμογή αυτών των εργαλείων απαιτεί την εκπαίδευση των δασκάλων και των άλλων φορέων εκπαίδευσης, κάτι που μπορεί να συναντήσει αντιστάσεις ή να αποδειχθεί δύσκολο σε παραδοσιακά εκπαιδευτικά περιβάλλοντα. Η επιτυχία της εφαρμογής εξαρτάται από την αποδοχή και την εξοικείωση των δασκάλων με τις νέες τεχνολογίες και τη συνεχιζόμενη εκπαίδευση για την αξιοποίηση των δεδομένων και των αναλύσεων που προκύπτουν.

Ωστόσο, όταν χρησιμοποιούνται σωστά, η εξόρυξη διεργασιών και ο έλεγχος συμμόρφωσης προσφέρουν έναν ισχυρό μηχανισμό για τη βελτίωση της εκπαιδευτικής διαδικασίας, δίνοντας τη δυνατότητα για μια πιο αποτελεσματική και εξατομικευμένη μάθηση, καθώς και

για την εξασφάλιση της διαφάνειας και της δικαιοσύνης στο εκπαιδευτικό σύστημα. Με την εφαρμογή αυτών των εργαλείων, η εκπαίδευση γίνεται πιο προσαρμοσμένη στις ανάγκες των μαθητών και μπορεί να ανταποκριθεί στις συνεχώς μεταβαλλόμενες απαιτήσεις του σύγχρονου κόσμου (Sypsas & Kalles, 2022c; Carmona et al., 2018).

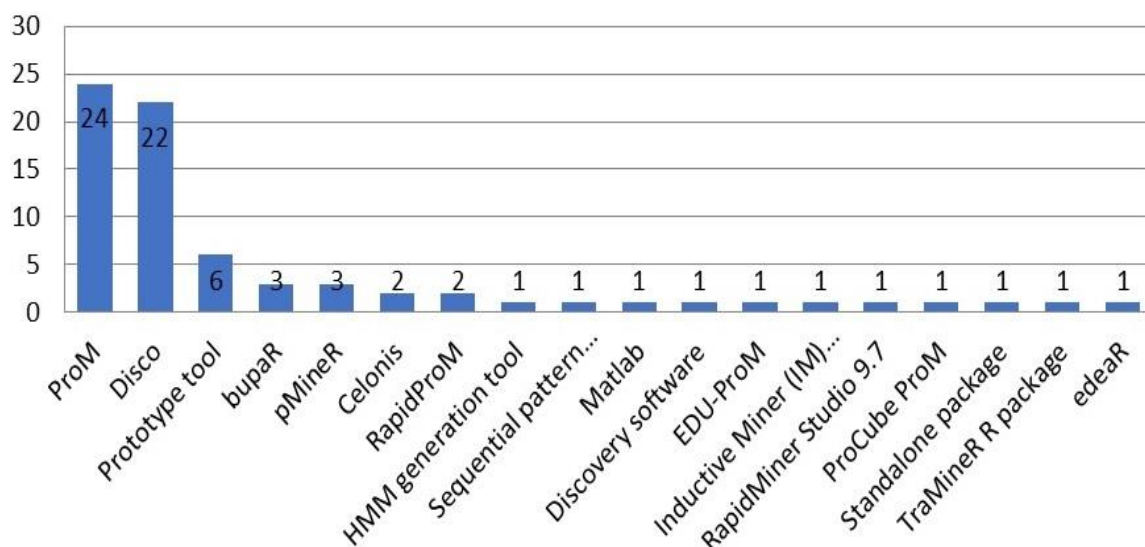
Η **Εικόνα 43** [Sypsas & Kalles, 2022], παρουσιάζει τον αριθμό και την κατανομή των μελετών εκπαιδευτικής εξόρυξης διαδικασιών σε διάφορους τομείς. Δείχνει ποιοι τομείς έχουν εστιαστεί περισσότερο στην έρευνα και ποιοι λιγότερο.

Τα δεδομένα βοηθούν στην κατανόηση των ερευνητικών τάσεων στον χώρο της Εκπαιδευτικής Εξόρυξης Διαδικασιών.



Εικόνα 43: Μελέτες Εξόρυξης Διαδικασιών ανά Τομέα στην Εκπαίδευση

Η **Εικόνα 44** [Sypsas & Kalles, 2022], απεικονίζει τα βασικά εργαλεία που χρησιμοποιούνται σε μελέτες Εκπαιδευτικής Εξόρυξης Διαδικασιών (EDM). Αναδεικνύει τη συχνότητα χρήσης τους και τις προτιμήσεις των ερευνητών στον συγκεκριμένο επιστημονικό χώρο. Τα δεδομένα παρέχουν μια συνοπτική εικόνα των διαθέσιμων εργαλείων και της εφαρμοσιμότητάς τους στην εκπαιδευτική έρευνα.



Εικόνα 44: Εργαλεία για Εξόρυξη Διαδικασιών που χρησιμοποιούνται σε μελέτες Εκπαιδευτικής Εξόρυξης Διαδικασιών

4.4 Event Logs στο OnLabs και Εφαρμογές Process Discovery

Η καταγραφή Event Logs στο OnLabs αποτελεί θεμελιώδες στοιχείο της λειτουργίας του συστήματος και προσδίδει σε αυτό έναν έντονα αναλυτικό και ερευνητικό χαρακτήρα. Τα Event Logs καταγράφουν με δομημένο τρόπο κάθε ενέργεια που πραγματοποιείται από τους χρήστες κατά τη διάρκεια της αλληλεπίδρασής τους με το εικονικό εργαστήριο. Κάθε καταγραφή συνοδεύεται από πληροφορίες όπως χρονική σήμανση, τύπος δραστηριότητας και ταυτότητα χρήστη, δημιουργώντας ένα πλήρες ιστορικό ενεργειών (Sypsas & Kalles, 2022c; Sypsas & Kalles, 2023).

Η ύπαρξη αυτών των δεδομένων καθιστά δυνατή την εφαρμογή τεχνικών Process Discovery, οι οποίες εντάσσονται στο πεδίο του Process Mining. Μέσω των τεχνικών αυτών, τα Event Logs μετατρέπονται σε μοντέλα διαδικασιών που απεικονίζουν τον πραγματικό τρόπο με τον οποίο οι φοιτητές αλληλεπιδρούν με το OnLabs. Τα μοντέλα αυτά αντικατοπτρίζουν την πραγματική μαθησιακή συμπεριφορά και συχνά διαφοροποιούνται από τα αρχικά θεωρητικά σενάρια (Sypsas & Kalles, 2022c).

Στο εκπαιδευτικό πλαίσιο του OnLabs, η ανάλυση των Event Logs επιτρέπει τον εντοπισμό κρίσιμων σημείων της μαθησιακής διαδικασίας. Μπορούν να αναδειχθούν δραστηριότητες που προκαλούν δυσκολίες, σημεία καθυστέρησης ή επαναλαμβανόμενα μοτίβα

συμπεριφοράς. Τα ευρήματα αυτά παρέχουν πολύτιμες πληροφορίες στους διδάσκοντες για τη βελτίωση της σχεδίασης των εργαστηριακών ασκήσεων (Sypsas & Kalles, 2022b; Sypsas & Kalles, 2023).

Η αξιοποίηση των τεχνικών Process Discovery δεν περιορίζεται μόνο σε διαγνωστική ανάλυση, αλλά μπορεί να υποστηρίξει και προγνωστικές εφαρμογές. Μέσα από τη συστηματική μελέτη των δεδομένων, είναι δυνατόν να εντοπιστούν ενδείξεις μειωμένης εμπλοκής ή πιθανής εγκατάλειψης, επιτρέποντας την έγκαιρη εκπαιδευτική παρέμβαση (Sypsas & Kalles, 2023).

Τέλος, τα αποτελέσματα της ανάλυσης των Event Logs μπορούν να αποτελέσουν τη βάση για την ανάπτυξη προσαρμοστικών μηχανισμών μάθησης. Το OnLabs μπορεί να εξελιχθεί σε ένα ευφύες εκπαιδευτικό περιβάλλον, το οποίο προσαρμόζει τη ροή των δραστηριοτήτων ή παρέχει εξατομικευμένη υποστήριξη με βάση τη συμπεριφορά των φοιτητών. Με τον τρόπο αυτό, το σύστημα συμβάλλει ουσιαστικά τόσο στη βελτίωση της μαθησιακής εμπειρίας όσο και στην εκπαιδευτική έρευνα (Sypsas & Kalles, 2022c; Sypsas & Kalles, 2023).

4.5 Διαδικασία Εξαγωγής Δεδομένων από το Εικονικό Εργαστήριο OnLabs

4.5.1 Εισαγωγή

Στην παρούσα ενότητα αναλύεται η διαδικασία παραγωγής και εξαγωγής των δεδομένων από το ψηφιακό περιβάλλον του εικονικού εργαστηρίου OnLabs. Η κατανόηση αυτής της διαδικασίας είναι θεμελιώδους σημασίας για τη διπλωματική εργασία, καθώς αποτελεί το πρώτο και κρίσιμότερο στάδιο της αλυσίδας ανάλυσης. Στόχος είναι η μετατροπή των αλληλεπιδράσεων του χρήστη σε δομημένη πληροφορία, η οποία θα αποτελέσει το πρωτογενές υλικό για την εφαρμογή των μεθοδολογιών ελέγχου συμμόρφωσης (Conformance Checking).

Για τις ανάγκες της ανάλυσης, επιλέχθηκε η υλοποίηση του πειράματος «**Μικροσκοπική Παρατήρηση Δείγματος**». Το συγκεκριμένο πείραμα αποτελεί ιδανική περίπτωση μελέτης, καθώς περιλαμβάνει μια σειρά από διαδοχικά και αυστηρά ορισμένα βήματα — από την τοποθέτηση της αντικειμενοφόρου πλάκας και την εναλλαγή των αντικειμενικών φακών,

μέχρι τη λεπτομερή ρύθμιση της εστίασης και την τελική απενεργοποίηση του εξοπλισμού. Κάθε μία από αυτές τις ενέργειες απαιτεί τη λήψη αποφάσεων από τον χρήστη, γεγονός που επιτρέπει την καταγραφή αποκλίσεων ή σφαλμάτων σε σχέση με το προβλεπόμενο εργαστηριακό πρωτόκολλο.

Η παραγωγή του αρχείου καταγραφής (event log) ξεκινά με την εκτέλεση του πειράματος από τον χρήστη σε πραγματικό χρόνο. Καθώς ο χρήστης αλληλεπιδρά με τα ψηφιακά αντικείμενα — όπως η περιστροφή της κεφαλής των φακών ή η ρύθμιση του μικρομετρικού κοχλία — το σύστημα OnLabs ιχνηλατεί κάθε κίνηση στο παρασκήνιο. Η ιχνηλασιμότητα αυτή δεν περιορίζεται μόνο στην ταυτοποίηση της ενέργειας, αλλά επεκτείνεται και στη χρονική σήμανση της κάθε δραστηριότητας, προσφέροντας μια πλήρη εικόνα της δυναμικής του πειράματος.

Η διαδικασία ολοκληρώνεται με τη χρήση της λειτουργίας αποθήκευσης που προσφέρει η εφαρμογή. Με την επιλογή του σχετικού εικονιδίου, ο χρήστης ενεργοποιεί τη διαδικασία εξαγωγής των δεδομένων σε ένα ψηφιακό αρχείο. Το αρχείο αυτό δεν είναι απλώς μια αναπαράσταση της κατάστασης του πειράματος, αλλά λειτουργεί ως ένα ιστορικό αρχείο συμβάντων. Η αποθήκευση σε συγκεκριμένο φάκελο του συστήματος επιτρέπει την αρχειοθέτηση πολλαπλών δοκιμών, οι οποίες στη συνέχεια μπορούν να εισαχθούν σε εξειδικευμένα λογισμικά Process Mining.

Στις υποενότητες που ακολουθούν, παρουσιάζεται μέσω οπτικού υλικού (screenshots) η διαδρομή από την εκτέλεση των εργαστηριακών πράξεων έως την τελική μορφή του αρχείου καταγραφής. Η λεπτομερής αυτή παρουσίαση αποσκοπεί στο να καταδείξει πώς μια σειρά από χειροκίνητες ενέργειες σε ένα εικονικό περιβάλλον μετατρέπονται σε αντικειμενικά δεδομένα, έτοιμα να αξιολογηθούν ως προς την ορθότητά τους. Με τον τρόπο αυτό, τεκμηριώνεται η αξιοπιστία της εισόδου των δεδομένων στο σύστημα ελέγχου, διασφαλίζοντας την εγκυρότητα των συμπερασμάτων της έρευνας.

4.5.2: Περιγραφή και Ροή της Πειραματικής Διαδικασίας

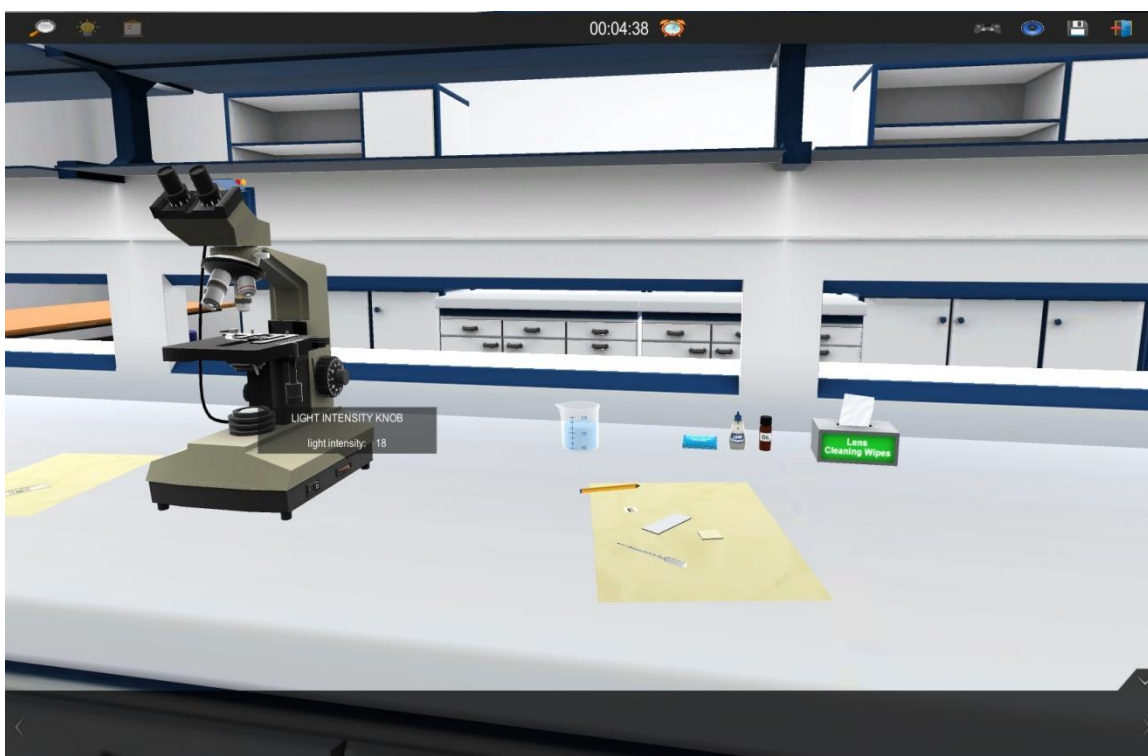
Για την εξαγωγή ενός αντιπροσωπευτικού αρχείου καταγραφής γεγονότων, κρίθηκε απαραίτητη η εκτέλεση ενός ολοκληρωμένου σεναρίου χρήσης, το οποίο καλύπτει όλο το φάσμα των λειτουργιών του εικονικού μικροσκοπίου. Το πείραμα της «**Μικροσκοπικής**

Παρατήρησης Δείγματος» ακολουθεί ένα αυστηρό πρωτόκολλο, η ορθή τήρηση του οποίου αποτελεί το αντικείμενο του ελέγχου συμμόρφωσης. Οι επιμέρους φάσεις και οι αντίστοιχες ψηφιακές καταγραφές (actions) περιγράφονται αναλυτικά παρακάτω:

- **Προετοιμασία Συστήματος:** Η διαδικασία εκκινεί με τη διασφάλιση της ηλεκτρικής τροφοδοσίας μέσω της σύνδεσης του καλωδίου στην αντίστοιχη υποδοχή (used mic_plug with socket1). Ακολουθεί η ενεργοποίηση του κύριου διακόπτη (set mic_AC_switch on) και η αρχική ρύθμιση της έντασης της φωτεινής πηγής μέσω του ροοστάτη (rotated mic_light_intensity_knob), καθώς και των μηχανισμών ελέγχου του φωτός (rotated mic_aperture_knob και rotated mic_condenser_knob).
- **Προετοιμασία Δείγματος και Αντικειμενοφόρου Πλάκας:** Ο χρήστης προετοιμάζει τον χώρο εργασίας μεταφέροντας την πλάκα στον πάγκο (moved slide to bench) και προβαίνει στον καθαρισμό της χρησιμοποιώντας το ειδικό χαρτί (used paper_scrap with slide). Στη συνέχεια, γίνεται η λήψη του δείγματος από το δοχείο με τη χρήση πιπέτας (used pasteur_body with beaker), η τοποθέτησή του στην πλάκα (used pasteur_body with slide) και η κάλυψή του (used slide_cover with slide).
- **Τοποθέτηση και Αρχική Εστίαση:** Η έτοιμη πλάκα τοποθετείται στην τράπεζα του μικροσκοπίου (used slide with mic_stage) και ο χρήστης εισέρχεται στη διεπαφή παρατήρησης (focused on microscope_instrument). Ακολουθεί η επιλογή του κατάλληλου φακού (rotated mic_revolving_nosepiece) και η χρήση του μακρομετρικού κοχλίου (rotated mic_coarse_focus_knob) για την ανύψωση της τράπεζας και την εύρεση του ειδώλου.
- **Βελτιστοποίηση Παρατήρησης:** Σε αυτό το στάδιο, πραγματοποιείται η ρύθμιση των προσοφθάλμιων φακών (rotated mic_ocular_right και rotated mic_ocular_left) και ο συνδυασμός κινήσεων του μικρομετρικού κοχλίου (rotated mic_fine_focus_knob) για τη μεγιστοποίηση της ευκρίνειας. Παράλληλα, ο χρήστης μετακινεί το δείγμα στους άξονες X και Y (rotated mic_x_axis_knob και rotated mic_y_axis_knob) για τη σάρωση των περιοχών ενδιαφέροντος.
- **Ολοκλήρωση Διαδικασίας:** Μετά το πέρας της παρατήρησης, ο χρήστης εξέρχεται από το όργανο (unfocused from microscope_instrument) και απομακρύνει το δείγμα (picked up slide to the inventory). Η διαδικασία ολοκληρώνεται με την επαναφορά

του διαφράγματος (rotated mic_aperture_knob), την απενεργοποίηση του διακόπτη (set mic_AC_switch off) και την πλήρη αποσύνδεση της συσκευής (pulled mic_plug).

Προκειμένου να γίνει πλήρως κατανοητός ο τρόπος με τον οποίο οι παραπάνω ενέργειες μεταφράζονται σε ψηφιακά δεδομένα εντός του συστήματος, κρίνεται σκόπιμη η οπτική αναπαράσταση των κρίσιμων σταδίων του πειράματος. Στο σημείο αυτό, παρατίθεται μια λεπτομερής σειρά στιγμιότυπων (screenshots), τα οποία αποτυπώνουν την αλληλεπίδραση του χρήστη με τη διεπαφή του OnLabs. Η παρουσίαση αυτή δεν περιορίζεται μόνο στην απεικόνιση των εργαστηριακών χειρισμών, αλλά επεκτείνεται μέχρι την τελική φάση της αποθήκευσης και της προβολής του αρχείου καταγραφής (event log), αναδεικνύοντας τη διασύνδεση της εικονικής πράξης με την πληροφοριακή της καταγραφή.



Εικόνα 45: Ρύθμιση της έντασης του φωτισμού και προετοιμασία της οπτικής δέσμης

Σε αυτό το βήμα, ο χρήστης ρυθμίζει την ένταση του φωτός μέσω του ροοστάτη, ώστε το δείγμα να φωτίζεται σωστά και να είναι ευδιάκριτο. Η ενέργεια αυτή είναι απαραίτητη για

να πετύχουμε την κατάλληλη αντίθεση, αποφεύγοντας παράλληλα τον υπερβολικό φωτισμό που θα κούραζε τα μάτια ή θα «έκρυβε» λεπτομέρειες του παρασκευάσματος. Στη ροή της εργασίας, η ρύθμιση του φωτός καταγράφεται ως ένα σημαντικό προπαρασκευαστικό βήμα πριν την εστίαση. Η σωστή εκτέλεσή του επιβεβαιώνει ότι ο χρήστης ακολουθεί το ορθό πρωτόκολλο για την προετοιμασία του μικροσκοπίου (**Εικόνα 45**).



Εικόνα 46: Ρύθμιση του διαφράγματος της ίριδας για τον έλεγχο της αντίθεσης

Σε αυτό το σημείο, ο χρήστης μετακινεί τον μοχλό του διαφράγματος προς τα αριστερά, προκειμένου να ελέγξει τη δέσμη φωτός που φτάνει στο δείγμα. Η κίνηση αυτή μεταβάλλει το άνοιγμα του διαφράγματος, επηρεάζοντας άμεσα την αντίθεση και το βάθος πεδίου της εικόνας. Στο πλαίσιο του πειράματος, η συγκεκριμένη ενέργεια είναι καθοριστική για τη βελτίωση της ευκρίνειας, ειδικά σε διαφανή παρασκευάσματα. Για τον έλεγχο συμμόρφωσης, η σωστή χρήση του διαφράγματος καταδεικνύει ότι ο χρήστης γνωρίζει πώς να βελτιστοποιεί την ποιότητα της εικόνας πέρα από την απλή ρύθμιση της φωτεινότητας (**Εικόνα 46**).



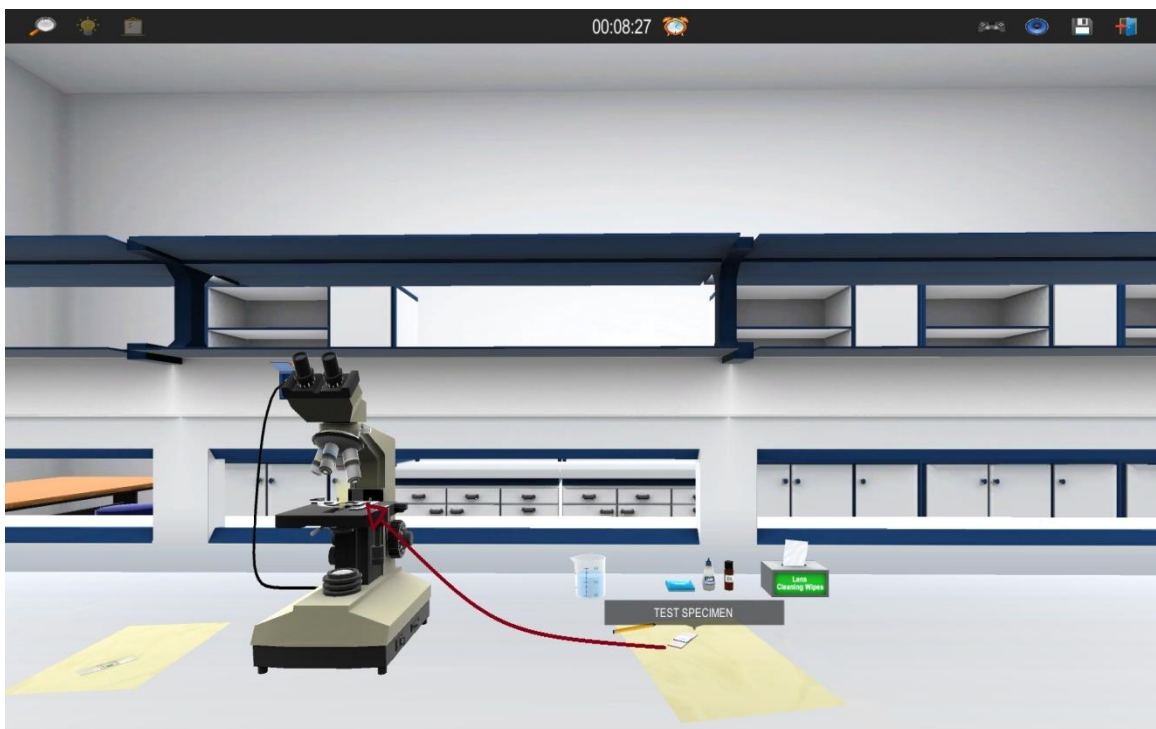
Εικόνα 47: Τοποθέτηση του δείγματος στην αντικειμενοφόρο πλάκα

Σε αυτό το στάδιο, ο χρήστης τοποθετεί το τμήμα του χαρτιού πάνω στην αντικειμενοφόρο πλάκα, προετοιμάζοντας το προς εξέταση δείγμα. Η ενέργεια αυτή αποτελεί την αφετηρία της δημιουργίας ενός νωπού παρασκευάσματος, όπου το αντικείμενο πρέπει να τοποθετηθεί προσεκτικά στο κέντρο της πλάκας για να είναι εφικτή η μετέπειτα μικροσκόπηση του. Από πλευράς ανάλυσης διεργασιών, η δραστηριότητα αυτή σηματοδοτεί την έναρξη της πειραματικής διαδικασίας στο φυσικό επίπεδο. Η ορθή τοποθέτηση είναι απαραίτητη, καθώς καθορίζει αν το δείγμα θα βρίσκεται εντός του οπτικού πεδίου όταν ο χρήστης ξεκινήσει την παρατήρηση (**Εικόνα 47**).



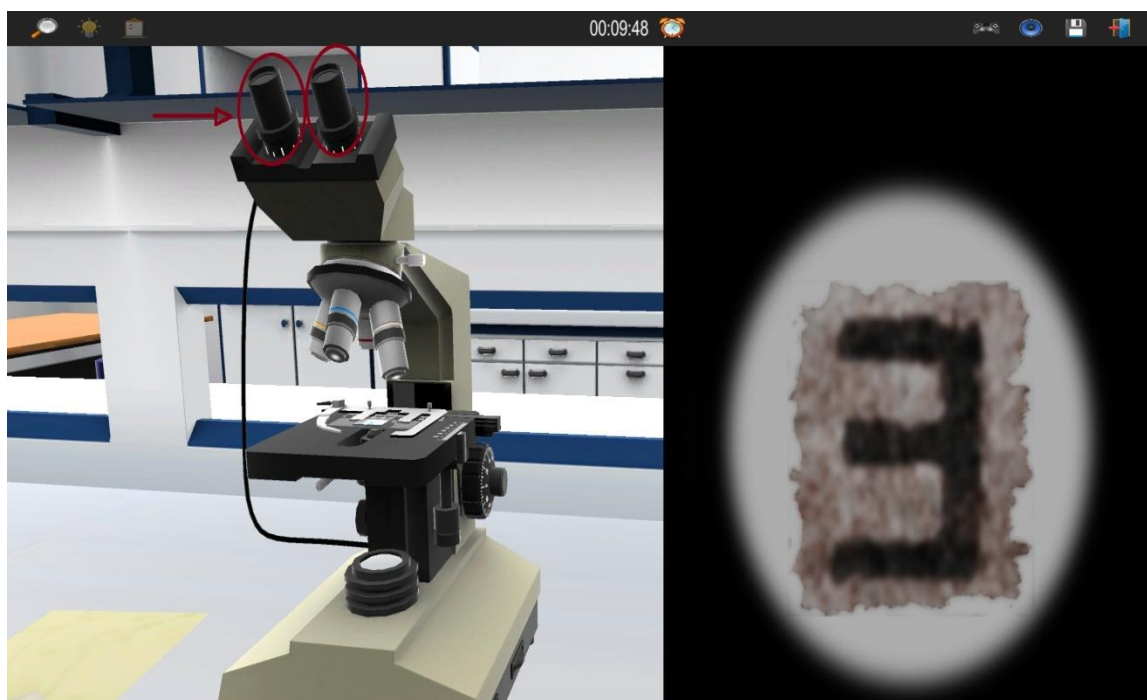
Εικόνα 48: Επικάλυψη του δείγματος με την καλυπτρίδα

Σε αυτό το βήμα, ο χρήστης τοποθετεί την καλυπτρίδα (slide cover) πάνω από το δείγμα που βρίσκεται στην αντικειμενοφόρο πλάκα. Η ενέργεια αυτή είναι θεμελιώδης, καθώς προστατεύει τους φακούς του μικροσκοπίου από τυχόν επαφή με το δείγμα και κρατά το παρασκεύασμα επίπεδο και σταθερό. Στη ροή της διαδικασίας, η τοποθέτηση της καλυπτρίδας ολοκληρώνει τη φάση προετοιμασίας του παρασκευάσματος. Η καταγραφή αυτού του βήματος στο Process Mining επιβεβαιώνει ότι τηρούνται οι κανόνες ασφάλειας και ορθής παρασκευής πριν το δείγμα τοποθετηθεί για τελική παρατήρηση (**Εικόνα 48**).



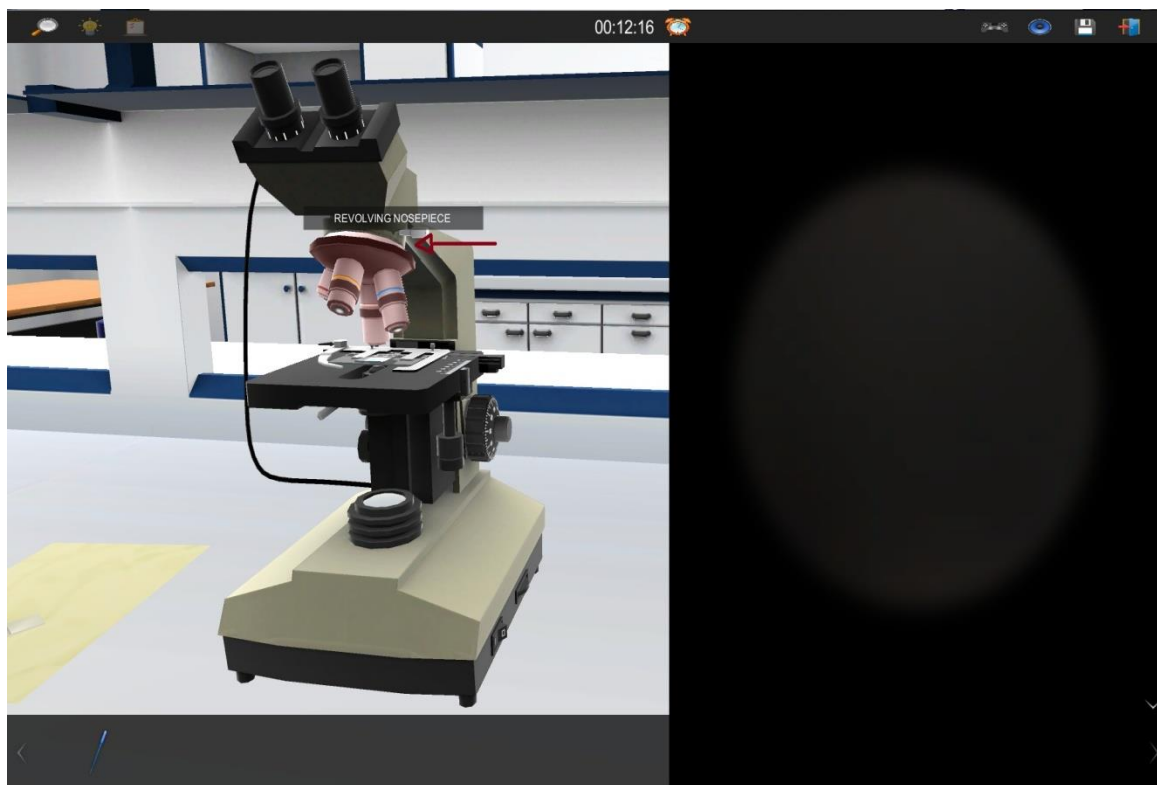
Εικόνα 49: Τοποθέτηση και ευθυγράμμιση του παρασκευάσματος στην τράπεζα του μικροσκοπίου

Στο σημείο αυτό, ο χρήστης μεταφέρει την έτοιμη αντικειμενοφόρο πλάκα και την τοποθετεί πάνω στην τράπεζα του μικροσκοπίου. Η κίνηση αυτή στο OnLabs, προσομοιώνει την προσεκτική τοποθέτηση του δείγματος κάτω από τους αντικειμενικούς φακούς, ώστε να ξεκινήσει η διαδικασία της εστίασης. Για την παρούσα διπλωματική εργασία, αυτή η ενέργεια αποτελεί τη μετάβαση από τη φάση της «Προετοιμασίας» στη φάση της «Παρατήρησης». Η επιτυχής ολοκλήρωση αυτού του βήματος είναι κρίσιμη για το Conformance Checking, καθώς χωρίς τη σωστή τοποθέτηση του δείγματος στην τράπεζα, οποιαδήποτε επόμενη προσπάθεια εστίασης θα οδηγούσε σε σφάλμα της διαδικασίας (Εικόνα 49).



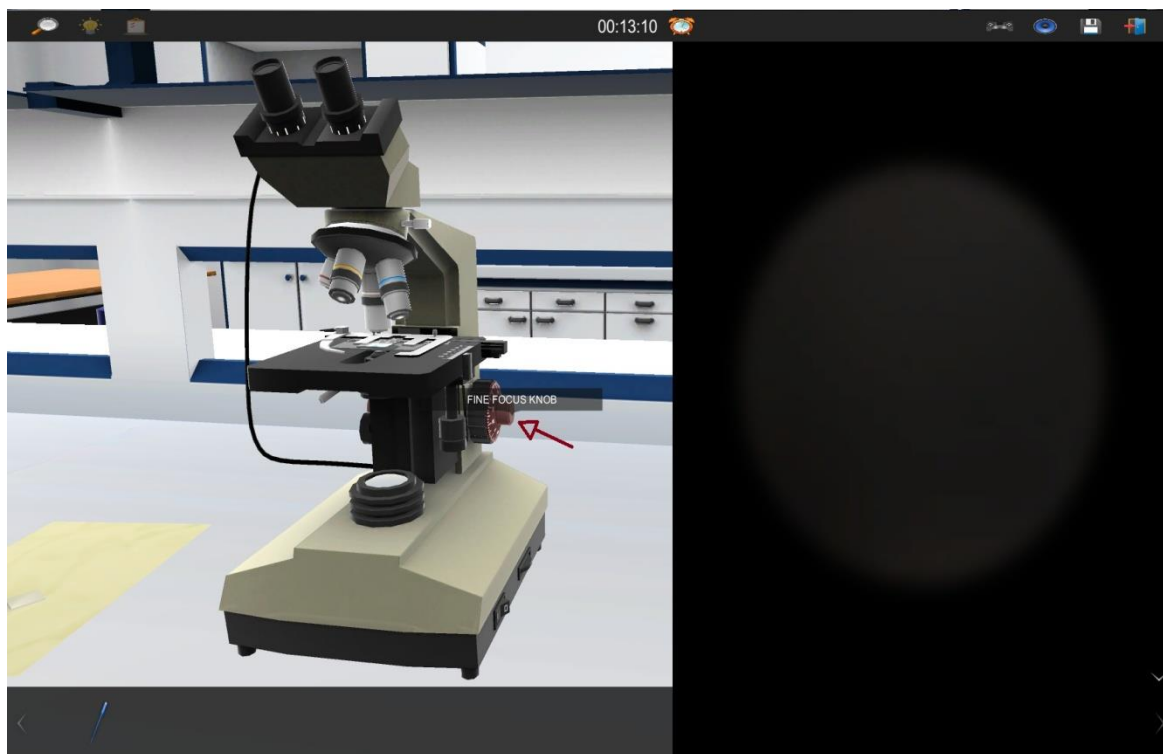
Εικόνα 50: Μικρομετρική ρύθμιση και βελτιστοποίηση της εστίασης μέσω των προσοφθάλμιων φακών

Σε αυτό το στάδιο, ο χρήστης χρησιμοποιεί τους προσοφθάλμιους φακούς για να επιτύχει την τελική, λεπτομερή εστίαση του ειδώλου. Η ενέργεια αυτή περιλαμβάνει τη μικρομετρική ρύθμιση των φακών, ώστε η εικόνα του δείγματος να γίνει απόλυτα καθαρή και ευκρινής για τον παρατηρητή. Στο πλαίσιο του Conformance Checking, η συγκεκριμένη δραστηριότητα αποτελεί το τελευταίο βήμα της διαδικασίας εστίασης πριν από την καταγραφή των αποτελεσμάτων. Η επιτυχής ολοκλήρωσή της επιβεβαιώνει ότι ο χρήστης έχει ρυθμίσει το όργανο με ακρίβεια, εξασφαλίζοντας ότι η οπτική παρατήρηση είναι αξιόπιστη και απαλλαγμένη από σφάλματα θολότητας (**Εικόνα 50**).



Εικόνα 51: Εναλλαγή αντικειμενικών φακών μέσω της περιστρεφόμενης κεφαλής

Σε αυτό το βήμα, ο χρήστης περιστρέφει την κεφαλή (revolving nosepiece) προς τα αριστερά για να επιλέξει τον κατάλληλο αντικειμενικό φακό. Η κίνηση αυτή αλλάζει τη μεγέθυνση του μικροσκοπίου, επιτρέποντας μια πιο λεπτομερή εξέταση του παρασκευάσματος. Στη διαδικασία του Conformance Checking, η ενέργεια αυτή είναι κομβική, καθώς η μετάβαση από μικρότερη σε μεγαλύτερη μεγέθυνση πρέπει να γίνεται σταδιακά και με τη σωστή σειρά. Η ορθή επιλογή του φακού επιβεβαιώνει ότι ο χρήστης ακολουθεί τη μεθοδολογία που απαιτείται για τη σωστή ανάλυση της μικροδομής του δείγματος (Εικόνα 51).



Εικόνα 52: Μικρομετρική εστίαση του ειδώλου για τη μεγιστοποίηση της ευκρίνειας

Στο στάδιο αυτό, ο χρήστης χειρίζεται τον κοχλία της μικρομετρικής εστίασης (fine focus knob) για να εξαλείψει τη θολότητα και να αναδείξει τις λεπτομέρειες του δείγματος. Σε αντίθεση με την αδρή εστίαση, η κίνηση αυτή πραγματοποιεί ανεπαίσθητες μετατοπίσεις στην τράπεζα του μικροσκοπίου, προσφέροντας την τελική καθαρότητα στην εικόνα. Στη μελέτη του Conformance Checking, η ενέργεια αυτή καταγράφεται ως το τελικό βήμα ακριβείας στην αλυσίδα των ρυθμίσεων. Η σωστή χρήση του μικρομετρικού κοχλία αποδεικνύει ότι ο χρήστης έχει τη δυνατότητα να φέρει το οπτικό αποτέλεσμα στο βέλτιστο επίπεδο ανάλυσης που επιτρέπει ο επιλεγμένος φακός (**Εικόνα 52**).



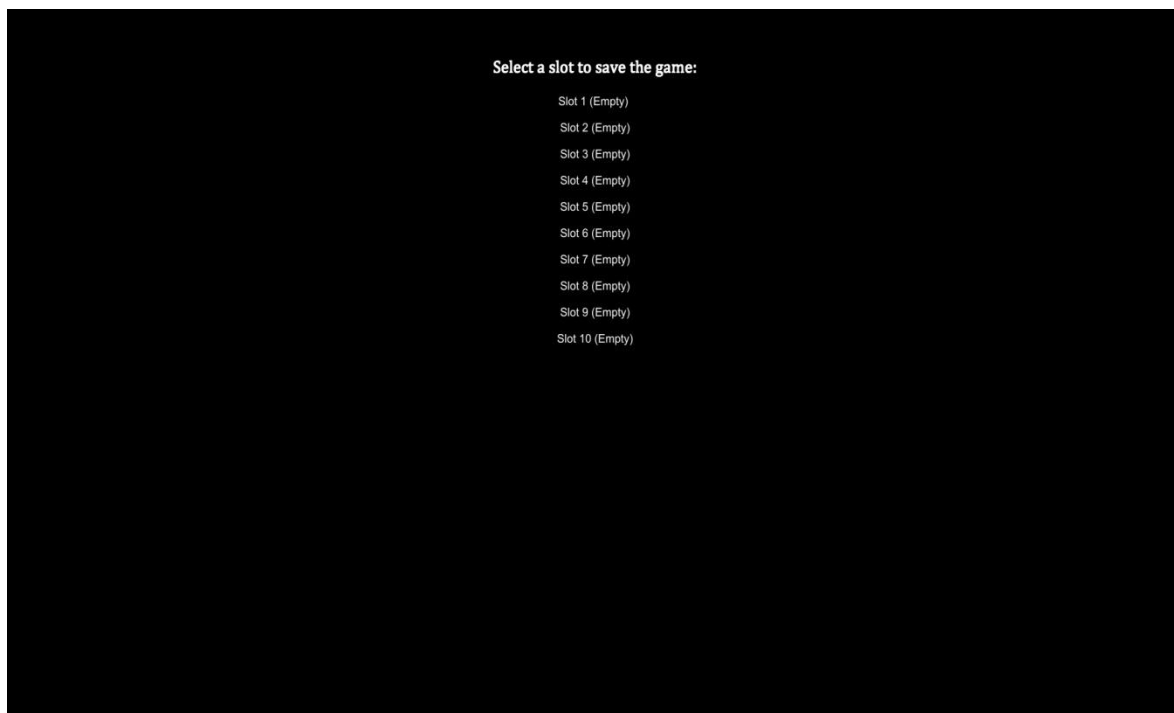
Εικόνα 53: Έξοδος από το πεδίο μικροσκοπικής παρατήρησης και επιστροφή στον χώρο εργασίας

Σε αυτό το βήμα, ο χρήστης επιλέγει έναν από τους αντικειμενικούς φακούς για να τερματίσει την οπτική παρατήρηση και να εξέλθει από το περιβάλλον της μικροσκόπησης. Η ενέργεια αυτή σηματοδοτεί την ολοκλήρωση της φάσης εξέτασης του δείγματος και την επιστροφή στην κύρια διεπαφή του εργαστηρίου. Στο πλαίσιο του Conformance Checking, η δραστηριότητα αυτή καταγράφεται ως το σημείο λήξης της παρατήρησης, κλείνοντας τον κύκλο της οπτικής ανάλυσης. Η σωστή έξοδος από το "microscoping mode" είναι απαραίτητη για τη μετάβαση σε επόμενα στάδια του πειράματος, όπως ο καθαρισμός ή η αποθήκευση του εξοπλισμού (**Εικόνα 53**).



Εικόνα 54: Ολοκλήρωση πειραματικής διαδικασίας και απενεργοποίηση εργαστηριακού εξοπλισμού

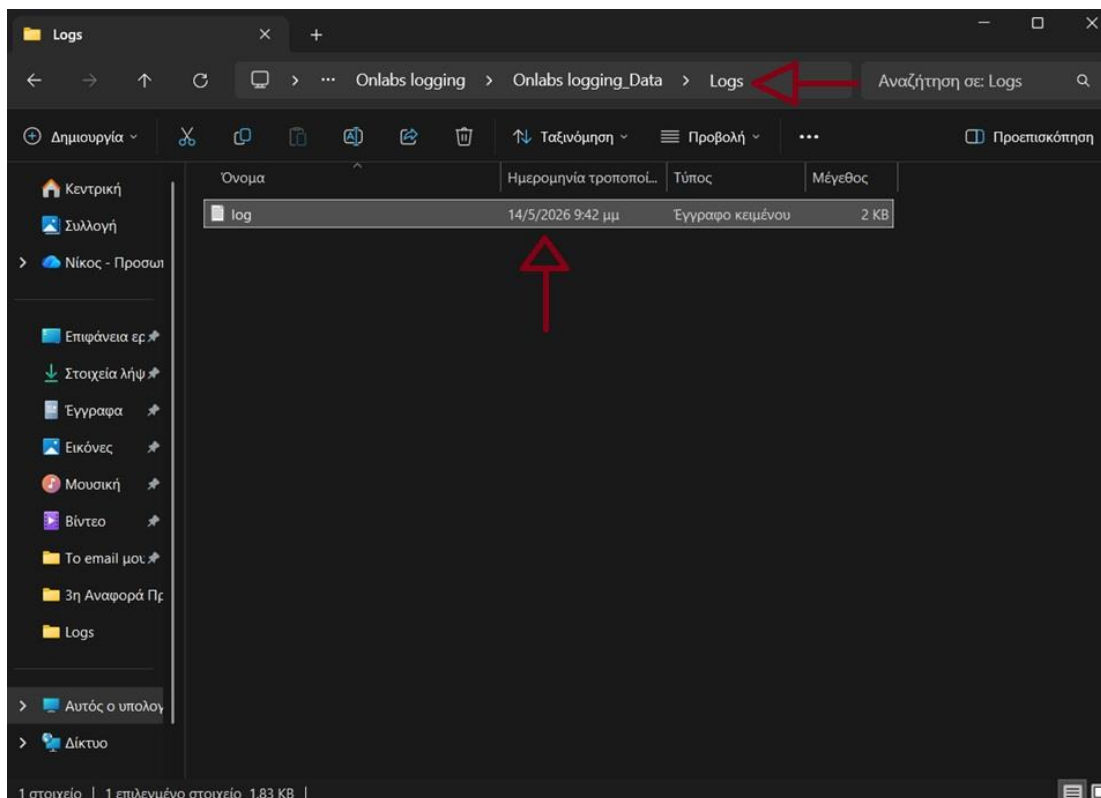
Σε αυτό το καταληκτικό στάδιο, ο χρήστης προβαίνει σε μια σειρά ενεργειών για την ορθή τακτοποίηση του εργαστηρίου, ξεκινώντας με την απομάκρυνση του παρασκευάσματος από την τράπεζα και την αποθήκευσή του στο απόθεμα (inventory). Στη συνέχεια, επαναφέρει το διάφραγμα στην αρχική του θέση, απενεργοποιεί την πηγή φωτισμού και αποσυνδέει το καλώδιο τροφοδοσίας, διασφαλίζοντας την ασφαλή κατάσταση της συσκευής. Από την πλευρά του Conformance Checking, οι ενέργειες αυτές είναι εξαιρετικά σημαντικές, καθώς επιβεβαιώνουν ότι ο χρήστης δεν εγκαταλείπει το πείραμα ημιτελές, αλλά ακολουθεί το πρωτόκολλο μέχρι και το στάδιο της απενεργοποίησης και προστασίας του εξοπλισμού (Εικόνα 54).



Εικόνα 55: Εξαγωγή της δραστηριότητας χρήστη και αποθήκευση του αρχείου καταγραφής

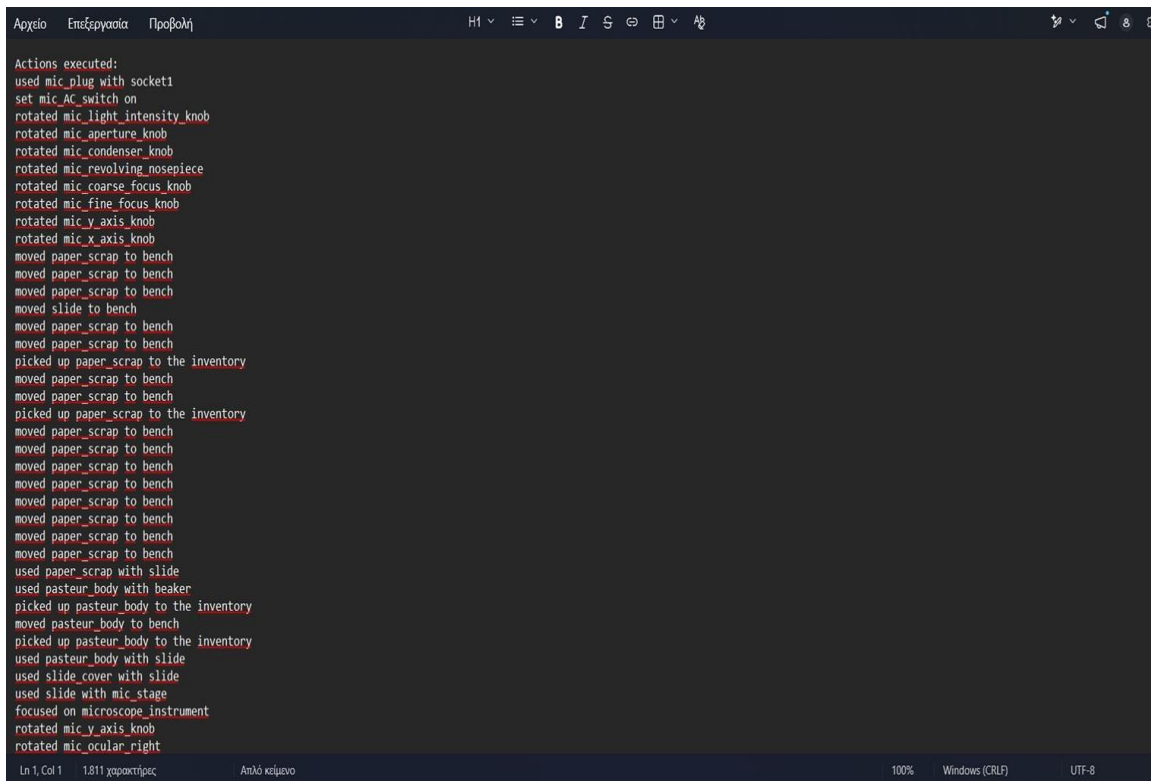
Στο τελευταίο αυτό βήμα, ο χρήστης επιλέγει το εικονίδιο της δισκέτας στο μενού, προκειμένου να εξάγει το ιστορικό των ενεργειών του σε ένα αρχείο καταγραφής γεγονότων (event log). Μέσω της διεπαφής επιλογής θέσης αποθήκευσης (save slot), εμφανίζεται μια λίστα διαθέσιμων θέσεων, όπου η ένδειξη «**Empty**» δηλώνει ότι η συγκεκριμένη θέση είναι κενή και δεν περιλαμβάνει προγενέστερη αποθήκευση, διασφαλίζοντας έτσι ότι δεν θα αντικατασταθούν κατά λάθος άλλα δεδομένα. Επιλέγοντας ένα διαθέσιμο slot, η διαδικασία οριστικοποιείται και το αρχείο αρχειοθετείται στον ειδικό φάκελο της εφαρμογής.

Αυτή η ενέργεια αποτελεί τη γέφυρα μεταξύ του εικονικού πειράματος και της ανάλυσης δεδομένων, καθώς το παραγόμενο αρχείο είναι η πρωτογενής πηγή για την εφαρμογή αλγορίθμων Process Mining. Με την αποθήκευση αυτή, διασφαλίζεται ότι κάθε βήμα που υλοποιήθηκε είναι πλέον διαθέσιμο για τον έλεγχο συμμόρφωσης (Conformance Checking) με το θεωρητικό μοντέλο (**Εικόνα 55**).



Εικόνα 56: Απεικόνιση του παραγόμενου αρχείου καταγραφής (Event Log) στον φάκελο αποθήκευσης

Στην παραπάνω εικόνα (**Εικόνα 56**), απεικονίζεται ο ειδικός φάκελος του συστήματος όπου έχει αποθηκευτεί το παραγόμενο αρχείο καταγραφής σε ψηφιακή μορφή. Το αρχείο αυτό περιλαμβάνει το σύνολο των δεδομένων της δραστηριότητας του χρήστη, αποτελώντας την απαραίτητη είσοδο (input) για την επεξεργασία και την ανάλυση της διαδικασίας μέσω των εργαλείων Process Mining.



```

Αρχείο  Επεξεργασία  Προβολή
H1  ≡  B  I  S  ☰  ↻

Actions executed:
used mic Plug with socket1
set mic AC switch on
rotated mic_light_intensity_knob
rotated mic_aperture_knob
rotated mic_condenser_knob
rotated mic_revolving_nosepiece
rotated mic_coarse_focus_knob
rotated mic_fine_focus_knob
rotated mic_y_axis_knob
rotated mic_x_axis_knob
moved paper_scrap to bench
moved paper_scrap to bench
moved paper_scrap to bench
moved slide to bench
moved paper_scrap to bench
moved paper_scrap to bench
picked up paper_scrap to the inventory
moved paper_scrap to bench
moved paper_scrap to bench
picked up paper_scrap to the inventory
moved paper_scrap to bench
moved paper_scrap to bench
moved paper_scrap to bench
moved paper_scrap to bench
moved paper_scrap to bench
moved paper_scrap to bench
moved paper_scrap to bench
used paper_scrap with slide
used pasteur_body with beaker
picked up pasteur_body to the inventory
moved pasteur_body to bench
picked up pasteur_body to the inventory
used pasteur_body with slide
used slide_cover with slide
used slide with mic_stage
focused on microscope_instrument
rotated mic_y_axis_knob
rotated mic_ocular_right

Ln 1, Col 1    1811 χαρακτήρες    Απλό κείμενο    100%    Windows (CRLF)    UTF-8

```

Εικόνα 57: Στιγμιότυπο του αρχείου καταγραφής συμβάντων

Στο παραπάνω στιγμιότυπο (**Εικόνα 57**), παρουσιάζεται η εσωτερική οργάνωση του αρχείου, όπου κάθε ενέργεια του χρήστη εντός του περιβάλλοντος OnLabs έχει καταγραφεί ως ένα διακριτό γεγονός. Το απόσπασμα αυτό αναδεικνύει την αλληλουχία των συμβάντων που αποτελούν τη βάση δεδομένων για τη μετέπειτα ανάλυση, επιτρέποντας τη σύγκριση της πραγματικής ροής του πειράματος με το προβλεπόμενο μοντέλο διαδικασίας.

ΚΕΦΑΛΑΙΟ 5

Μεθοδολογικό πλαίσιο, υλοποίηση και έλεγχος συμμόρφωσης της διεργασίας

5.1 Εισαγωγή

Στο κεφάλαιο αυτό παρουσιάζονται και αναλύονται τα τελικά αποτελέσματα που προέκυψαν από την εφαρμογή της μεθοδολογικής ροής (pipeline) που σχεδιάστηκε και υλοποιήθηκε για την παρούσα διπλωματική εργασία. Αφού ολοκληρώθηκε η προετοιμασία των δεδομένων, η ενότητα αυτή αποτελεί το πιο ουσιαστικό κομμάτι της έρευνας, καθώς δείχνει στην πράξη πώς οι ψηφιακές καταγραφές που συλλέχθηκαν από τον προσομοιωτή εικονικής πραγματικότητας OnLabs μετατρέπονται σε χρήσιμα συμπεράσματα. Μέσα από τη συγκεκριμένη ανάλυση, αποτιμάται η εκπαιδευτική συμπεριφορά με τη χρήση εξειδικευμένων εργαλείων, επιτρέποντας τη λεπτομερή μελέτη της ροής των εργαστηριακών ενεργειών.

Η παρουσίαση των αποτελεσμάτων ξεκινά με τον άξονα της ανακάλυψης και της ανάλυσης των διαδικασιών (Process Mining), όπου εστιάζω στα ευρήματα που εξήγαγα από το λογισμικό ProM. Μέσα από τη διερευνητική ανάλυση των δεδομένων του OnLabs, αποτυπώνεται η πραγματική πορεία που ακολούθησε η διαδικασία της μικροσκοπικής παρατήρησης. Στο σημείο αυτό παραθέτω όλα τα απαραίτητα στιγμιότυπα οθόνης (screenshots) από το περιβάλλον του ProM, τα οποία προσφέρουν μια καθαρή οπτική απεικόνιση των διαδρομών, αναδεικνύοντας τα βήματα που παραλείφθηκαν, τις άσκοπες επαναλήψεις, καθώς και τα σημεία συμφόρησης κατά την εστίαση των φακών.

Στη συνέχεια, το ενδιαφέρον μεταφέρεται στον Έλεγχο Συμμόρφωσης (Conformance Checking), ο οποίος αποτελεί την ποσοτική αξιολόγηση της συμπεριφοράς που καταγράφηκε. Η πραγματική ροή των ενεργειών συγκρίνεται απευθείας με το ιδανικό Μοντέλο Αναφοράς, χρησιμοποιώντας τον εξειδικευμένο κώδικα που ανέπτυξα στην Python μέσω της βιβλιοθήκης PM4Py. Στο στάδιο αυτό παρουσιάζονται και ερμηνεύονται οι τέσσερις βασικές μετρικές ποιότητας, δηλαδή η Καταλληλότητα, η Ακρίβεια, η Γενίκευση και η Απλότητα, προσφέροντας μια ξεκάθαρη εικόνα για το πόσο πιστά τηρήθηκε το επίσημο εργαστηριακό πρωτόκολλο, συνοδευόμενες από τα αντίστοιχα

screenshots των αποτελεσμάτων εκτέλεσης του κώδικα.

Σκοπός αυτής της εισαγωγής είναι να προετοιμάσει το έδαφος για την ανάγνωση των δεδομένων, των πινάκων και των οπτικών διαγραμμάτων που ακολουθούν στις επόμενες σελίδες. Τα αποτελέσματα αυτά αποδεικνύουν ότι η αξιολόγηση μιας εργαστηριακής άσκησης σε περιβάλλον εικονικής πραγματικότητας μπορεί να ξεφύγει από επιφανειακά στοιχεία, όπως ο συνολικός χρόνος παραμονής στην πλατφόρμα. Αντίθετα, ο συνδυασμός των δεδομένων του OnLabs, της ανάλυσης στο ProM και του ελέγχου συμμόρφωσης με τη βιβλιοθήκη PM4Py, προσφέρει μια αυτόματη, αντικειμενική και βαθιά αποτίμηση που εστιάζει αποκλειστικά στη λογική σειρά και στη σημασία των πράξεων.

5.2 «Μικροσκοπική Παρατήρηση Δείγματος» στο ψηφιακό περιβάλλον OnLabs

Το πείραμα «Μικροσκοπική Παρατήρηση Δείγματος» αποτελεί μια πρακτική διαδικασία που έχει ως αποκλειστικό σκοπό την εκμάθηση του σωστού χειρισμού των εξαρτημάτων του σύνθετου φωτονικού μικροσκοπίου και την προοδευτική εστίαση πάνω σε ένα έτοιμο βιολογικό παρασκεύασμα. Κατά τη διάρκεια του πειράματος, ο χρήστης καλείται να αλληλεπιδράσει με μια σειρά από μηχανικά και οπτικά μέρη της συσκευής, ρυθμίζοντας το ύψος της τράπεζας, την ένταση και τη συγκέντρωση του φωτός, καθώς και την εναλλαγή των αντικειμενικών φακών. Η όλη διαδικασία βασίζεται στην αυστηρή τήρηση μιας συγκεκριμένης αλληλουχίας κινήσεων, καθώς οποιοδήποτε λάθος στη σειρά των βημάτων — όπως η λανθασμένη χρήση των κοχλίων εστίασης — μπορεί να οδηγήσει σε καταστροφή του δείγματος ή των φακών. Στο ψηφιακό περιβάλλον του OnLabs, το πείραμα προσομοιώνει με ακρίβεια αυτές τις χειροκίνητες ενέργειες, απαιτώντας από τον χρήστη να οργανώσει τον πάγκο του, να σταθεροποιήσει τη γυάλινη αντικειμενοφόρο πλάκα, να εντοπίσει το δείγμα στη μικρότερη μεγέθυνση και στη συνέχεια να αυξήσει προοδευτικά τη μεγέθυνση για να μελετήσει τις λεπτομέρειες του παρασκευάσματος, ολοκληρώνοντας με την ασφαλή απενεργοποίηση και επαναφορά του εξοπλισμού.

Για την επιτυχή ολοκλήρωση της διαδικασίας στο OnLabs, οι ενέργειες χωρίζονται σε συγκεκριμένες κατηγορίες βημάτων, οι οποίες αναλύονται λεπτομερώς παρακάτω:

Κατηγορία 1: Προετοιμασία του Χώρου και Ενεργοποίηση του Μικροσκοπίου

Η διαδικασία ξεκινά με την οργάνωση του εργαστηριακού πάγκου και την ηλεκτρική έναρξη της συσκευής, ώστε να εξασφαλιστούν οι κατάλληλες συνθήκες φωτισμού πριν την τοποθέτηση του δείγματος.

- **Βήμα 1 (Καθαρισμός πάγκου):** Ο χρήστης ελέγχει τον εικονικό πάγκο εργασίας και απομακρύνει τυχόν αντικείμενα που εμποδίζουν, εξασφαλίζοντας ελεύθερο χώρο γύρω από το μικροσκόπιο.
- **Βήμα 2 (Σύνδεση τροφοδοσίας):** Επιλέγεται το καλώδιο ρεύματος του μικροσκοπίου και τοποθετείται στην ηλεκτρική πρίζα του πάγκου.
- **Βήμα 3 (Έναρξη φωτισμού):** Ο χρήστης πατά τον διακόπτη λειτουργίας (ON) που βρίσκεται στη βάση του μικροσκοπίου, ενεργοποιώντας την εσωτερική λυχνία.
- **Βήμα 4 (Ρύθμιση ροοστάτη):** Ο χρήστης περιστρέφει τον τροχό του ροοστάτη για να καθορίσει μια μέση ένταση φωτός, αποφεύγοντας την απότομη και υπερβολική φωτεινότητα που θα μπορούσε να κουράσει τα μάτια κατά την παρατήρηση.

Κατηγορία 2: Ασφαλής Τοποθέτηση και Σταθεροποίηση του Παρασκευάσματος

Σε αυτή την κατηγορία, ο χρήστης διαχειρίζεται τον μηχανισμό της τράπεζας και του οδηγού για να ασφαλίσει τη γυάλινη πλάκα χωρίς να προκληθεί φθορά.

- **Βήμα 5 (Επιλογή δείγματος):** Ο χρήστης ανοίγει το κουτί των δειγμάτων και επιλέγει την έτοιμη αντικειμενοφόρο πλάκα που περιέχει το προς εξέταση βιολογικό υλικό.
- **Βήμα 6 (Χαμήλωμα τράπεζας):** Πριν τοποθετηθεί η πλάκα, ο χρήστης περιστρέφει τον μεγάλο μακρομετρικό κοχλία για να κατεβάσει την μαύρη τράπεζα (stage) στο χαμηλότερο σημείο της, δημιουργώντας ασφαλή απόσταση από τους αντικειμενικούς φακούς.
- **Βήμα 7 (Χειρισμός ελάσματος):** Ο χρήστης πιέζει τον ειδικό μοχλό του μηχανικού οδηγού για να ανοίξει το μεταλλικό έλασμα (κλιπ) συγκράτησης.

- **Βήμα 8 (Τοποθέτηση πλάκας):** Η γυάλινη αντικειμενοφόρος πλάκα τοποθετείται προσεκτικά πάνω στην επίπεδη επιφάνεια της τράπεζας, φροντίζοντας να εφαρμόσει ακριβώς στις γωνίες του οδηγού.
- **Βήμα 9 (Ασφάλιση δείγματος):** Ο χρήστης απελευθερώνει αργά τον μοχλό, επιτρέποντας στο μεταλλικό κλιπ να κλείσει και να σταθεροποιήσει το παρασκεύασμα, εμποδίζοντας οποιαδήποτε ανεπιθύμητη μετακίνηση κατά τη διάρκεια των ρυθμίσεων.

Κατηγορία 3: Διαδικασία Αρχικής Εστίασης με τον Φακό Μικρής Μεγέθυνσης (4x)

Το στάδιο αυτό περιλαμβάνει τον εντοπισμό του δείγματος στο οπτικό πεδίο και την επίτευξη της πρώτης καθαρής εικόνας με τη χρήση του μικρότερου φακού.

- **Βήμα 10 (Επιλογή φακού 4x):** Ο χρήστης περιστρέφει την κυκλική κεφαλή (revolver) μέχρι να ακούσει τον ήχο ασφάλισης (κλικ), φέρνοντας σε ευθυγράμμιση τον αντικειμενικό φακό 4x (αυτόν με την κόκκινη λωρίδα).
- **Βήμα 11 (Κεντράρισμα δείγματος):** Με τη χρήση των δύο ομοαξονικών κοχλίων μετακίνησης της τράπεζας, ο χρήστης μετατοπίζει την πλάκα στον άξονα X και Y, ώστε η κηλίδα του δείγματος να βρεθεί ακριβώς στο κέντρο της φωτεινής δέσμης.
- **Βήμα 12 (Ανύψωση τράπεζας):** Κοιτάζοντας το μικροσκόπιο από το πλάι για λόγους ασφαλείας, ο χρήστης περιστρέφει τον μακρομετρικό κοχλία για να ανεβάσει την τράπεζα στο ανώτερο δυνατό σημείο της, φέρνοντας το δείγμα κοντά στον φακό.
- **Βήμα 13 (Μακρομετρική ρύθμιση):** Κοιτάζοντας πλέον μέσα από τους προσοφθάλμιους φακούς, ο χρήστης περιστρέφει αργά τον μακρομετρικό κοχλία προς την αντίθετη κατεύθυνση, κατεβάζοντας την τράπεζα μέχρι να εμφανιστεί το πρώτο θολό είδωλο του δείγματος.
- **Βήμα 14 (Μικρομετρική ρύθμιση):** Μόλις εμφανιστούν οι πρώτες γραμμές του δείγματος, ο χρήστης αφήνει τον μεγάλο κοχλία και περιστρέφει τον μικρό μακρομετρικό κοχλία, ρυθμίζοντας την εικόνα μέχρι να αποκτήσει απόλυτη ευκρίνεια.

Κατηγορία 4: Προοδευτική Αλλαγή Φακών (10x & 40x) και Βελτιστοποίηση Εικόνας

Σε αυτή την κατηγορία, ο χρήστης αυξάνει τη μεγέθυνση για να παρατηρήσει τις λεπτομέρειες, προσαρμόζοντας παράλληλα τα εξαρτήματα ελέγχου του φωτός.

- **Βήμα 15 (Αλλαγή σε φακό 10x):** Ο χρήστης περιστρέφει το revolver για να φέρει στη θέση εργασίας τον αντικειμενικό φακό 10x (κίτρινη λωρίδα).
- **Βήμα 16 (Αποκλειστική χρήση μικρομετρικού):** Από αυτό το βήμα και μετά, ο χρήστης διορθώνει την εστίαση **μόνο με τον μικρό μικρομετρικό κοχλία**. Ο μεγάλος μακρομετρικός κοχλίας δεν αγγίζεται ξανά, καθώς η απόσταση του φακού από το γυαλί είναι πολύ μικρή και υπάρχει κίνδυνος πρόσκρουσης.
- **Βήμα 17 (Ρύθμιση συμπακνωτή και διαφράγματος):** Ο χρήστης μετακινεί τον συμπακνωτή και ανοιγοκλείνει το διάφραγμα ίριδας κάτω από την τράπεζα, συγκεντρώνοντας τη φωτεινή δέσμη για να αυξήσει την αντίθεση (contrast) και να γίνουν ορατά τα όρια των δομών.
- **Βήμα 18 (Αλλαγή σε φακό 40x):** Το revolver περιστρέφεται στον αντικειμενικό φακό 40x (μπλε λωρίδα), ο οποίος πλησιάζει σε απόσταση αναπνοής από την καλυπτρίδα του δείγματος.
- **Βήμα 19 (Τελική εστίαση υψηλής μεγέθυνσης):** Κοιτάζοντας μέσα από τα προσοφθάλμια, ο χρήστης εκτελεί πολύ μικρές και προσεκτικές κινήσεις με τον μικρομετρικό κοχλία, ξεχωρίζοντας πλέον καθαρά τις λεπτομέρειες στο εσωτερικό του παρασκευάσματος, όπως τα κυτταρικά τοιχώματα ή τους πυρήνες.

Κατηγορία 5: Διαδικασία Τερματισμού και Επαναφοράς του Εργαστηρίου

Μετά το τέλος της παρατήρησης, ο χρήστης πρέπει να ακολουθήσει τα βήματα αποσύνδεσης και ασφαλούς φύλαξης του εξοπλισμού.

- **Βήμα 20 (Επαναφορά φακού):** Ο χρήστης περιστρέφει την κεφαλή και επαναφέρει τον μικρό φακό 4x στη θέση εργασίας.
- **Βήμα 21 (Απομάκρυνση τράπεζας):** Με τον μακρομετρικό κοχλία, η τράπεζα χαμηλώνει τελείως για να απελευθερωθεί χώρος.

- **Βήμα 22 (Αφαίρεση δείγματος):** Ο χρήστης πιέζει τον μοχλό του οδηγού, απελευθερώνει τη γυάλινη αντικειμενοφόρο πλάκα, την αφαιρεί από την τράπεζα και την τοποθετεί πίσω στο κουτί αποθήκευσης.
- **Βήμα 23 (Απενεργοποίηση φωτός):** Ο ροοστάτης του φωτός γυρίζει στο μηδέν και ο διακόπτης λειτουργίας τίθεται στη θέση OFF.
- **Βήμα 24 (Αποσύνδεση καλωδίου):** Το καλώδιο τροφοδοσίας αφαιρείται από την πρίζα του τοίχου, και ο χρήστης τακτοποιεί τον πάγκο εργασίας στην αρχική του κατάσταση, ολοκληρώνοντας το πείραμα στο OnLabs.

5.3: Αρχιτεκτονική και μεθοδολογική ροή υλοποίησης της πειραματικής διαδικασίας

Η αποτίμηση της εκπαιδευτικής συμπεριφοράς κατά την εκτέλεση του εργαστηριακού πειράματος «Μικροσκοπική παρατήρηση δείγματος» βασίστηκε σε μια ολοκληρωμένη μεθοδολογική ροή τεσσάρων διαδοχικών σταδίων, η οποία σχεδιάστηκε και υλοποιήθηκε εξολοκλήρου για τις ανάγκες της παρούσας μεταπτυχιακής διπλωματικής εργασίας. Η ανάπτυξη αυτής της αρχιτεκτονικής (pipeline) κρίθηκε αναγκαία, καθώς οι ενέργειες των χρηστών μέσα στο περιβάλλον εικονικής πραγματικότητας OnLabs παρουσιάζουν συχνά παρεκκλίσεις, επαναλήψεις και παραλείψεις βημάτων. Οι παραδοσιακές στατιστικές μέθοδοι αδυνατούν να αποτυπώσουν αυτή τη μη γραμμική συμπεριφορά. Μέσα από τη ροή που ανέπτυξα, τα ακατέργαστα δεδομένα καταγραφής μετασχηματίζονται σε δομημένη πληροφορία, καθιστώντας εφικτή την άμεση και αυστηρή σύγκριση της πραγματικής δραστηριότητας με τους κανόνες του επίσημου εργαστηριακού πρωτοκόλλου.

Το πρώτο στάδιο αφορά τη συγκέντρωση και τη θεμελιώδη διαμόρφωση των πρωτογενών δεδομένων (OnLabs logs). Κατά τη διάρκεια της εικονικής μικροσκόπησης, η πλατφόρμα OnLabs καταγράφει αυτόματα στο παρασκήνιο κάθε αλληλεπίδραση, από την αρχική τροφοδοσία ρεύματος και την τοποθέτηση της αντικειμενοφόρου πλάκας, μέχρι τη ρύθμιση του φωτισμού και τη διαδοχική αλλαγή των αντικειμενικών φακών. Στο σημείο αυτό, πραγματοποιήθηκε η μετατροπή αυτών των διάσπαρτων καταγραφών σε ένα πρότυπο ιστορικό συμβάντων (Event Log), το οποίο αποτελεί την απαραίτητη εισαγωγή για τις τεχνικές Process Mining. Για να επιτευχθεί, τα δεδομένα οργανώθηκαν γύρω από τρεις

βασικούς άξονες: τη συγκεκριμένη εκπαιδευτική συνεδρία ως Αναγνωριστικό Υπόθεσης (Case ID), την ονομασία της πράξης ως Όνομα Δραστηριότητας (Activity Name) και την ακριβή στιγμή της εκτέλεσης ως Χρονική Σήμανση (Timestamp), εξασφαλίζοντας έτσι τη χρονολογική αλληλουχία και τη σημασιολογική εγκυρότητα των δεδομένων.

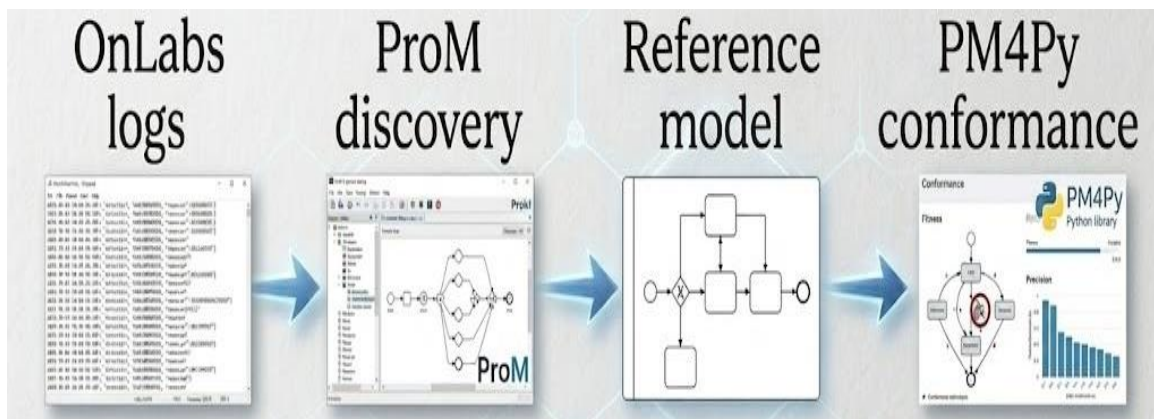
Το δεύτερο στάδιο εστιάζει στη διερευνητική ανάλυση και την οπτικοποίηση των δεδομένων (Process Discovery), διαδικασία που πραγματοποιήσα εισάγοντας το Event Log στο εξειδικευμένο λογισμικό ProM. Για την ανακάλυψη της πραγματικής ροής, εφάρμοσα τον αλγόριθμο Heuristic Miner, τον οποίο επέλεξα μετά από δική μου συγκριτική αξιολόγηση μεταξύ των διαθέσιμων αλγορίθμων της βιβλιογραφίας. Η αξιολόγηση αυτή έδειξε ότι ο συγκεκριμένος αλγόριθμος είναι ο πλέον κατάλληλος, καθώς βασίζεται στη συχνότητα των γεγονότων και απομονώνει τον θόρυβο των τυχαίων αποκλίσεων, παράγοντας ένα εύληπτο και αξιόπιστο μοντέλο. Η χρήση του Heuristic Miner μου επέτρεψε να χαρτογραφήσω την πραγματική πορεία των ενεργειών, φέρνοντας στο φως τις ακριβείς διαδρομές που ακολουθήθηκαν, τα βήματα που παραλείπονταν συστηματικά, τους βρόχους που δημιουργούνταν από επαναλαμβανόμενα λάθη, καθώς και τα λειτουργικά σημεία συμφόρησης (bottlenecks) κατά τη διαδικασία εστίασης των φακών.

Το τρίτο στάδιο αφορά τη σχεδίαση και τη δημιουργία του Μοντέλου Αναφοράς (Reference Model), το οποίο λειτουργεί ως ο αντικειμενικός οδηγός και το πρότυπο για τη σύγκριση. Αρχικά, σχεδίασα το μοντέλο στο περιβάλλον του Visual Paradigm σε μορφή επιχειρησιακής διαδικασίας (.bpmm), αποτυπώνοντας την ιδανική και επιστημονικά ορθή σειρά ενεργειών για την ολοκλήρωση της εργαστηριακής άσκησης. Στη συνέχεια, εισήγαγα το αρχείο αυτό στο λογισμικό ProM, όπου και το μετέτρεψα σε Δίκτυο Petri. Το τελικό δίκτυο αποτελείται από μια αυστηρή αλληλουχία 37 συγκεκριμένων βημάτων, η οποία ξεκινά από την προετοιμασία του οργάνου και την αφαίρεση του προστατευτικού καλύμματος, περνά από τη ρύθμιση των κοχλίων και τη στερέωση του δείγματος, και επεκτείνεται στη διαδοχική εστίαση με όλους τους φακούς (4x, 10x, 40x, 100x), πριν καταλήξει στην ασφαλή απενεργοποίηση. Για να εξασφαλίσω τον σωστό εσωτερικό συγχρονισμό, ενσωμάτωσα στο μοντέλο και τις απαραίτητες σιωπηλές μεταβάσεις (silent transitions), οι οποίες επιτρέπουν τη σωστή διαχείριση της λογικής ροής.

Το τέταρτο και τελικό στάδιο περιλαμβάνει την εκτέλεση του Ελέγχου Συμμόρφωσης (Conformance Checking), για την οποία ανέπτυξα έναν εξειδικευμένο κώδικα σε

περιβάλλον Python, αξιοποιώντας τις δυνατότητες της βιβλιοθήκης PM4Py. Στο στάδιο αυτό εφάρμοσα τον αλγόριθμο Token-Based Replay (Αναπαραγωγή Τεκμηρίων), τον οποίο επέλεξα για την εργασία μου κατόπιν δικής μου συστηματικής αξιολόγησης, καθώς τεκμηρίωσα ότι αποτελεί την πιο ακριβή μέθοδο για τον εντοπισμό αποκλίσεων σε επίπεδο επιμέρους ενεργειών. Ο αλγόριθμος αναπαράγει τις πραγματικές καταγραφές πάνω στο Δίκτυο Petri που σχεδίασα, παρακολουθώντας την κίνηση των δυναμικών στοιχείων του δικτύου, δηλαδή των τεκμηρίων (tokens). Με αυτόν τον τρόπο, το πρόγραμμα καταγράφει ποια τεκμήρια καταναλώθηκαν επιτυχώς, ποια έλειπαν λόγω παραλείψεων και ποια περίσσεψαν, επιτρέποντας τον υπολογισμό των τεσσάρων βασικών μετρικών ποιότητας: της Καταλληλότητας (Fitness), της Ακρίβειας (Precision), της Γενίκευσης (Generalization) και της Απλότητας (Simplicity) του μοντέλου.

Η ολοκλήρωση αυτής της μεθοδολογικής ροής εισάγει μια αντικειμενική και πλήρως αυτοματοποιημένη προσέγγιση στην εκπαιδευτική αξιολόγηση. Αντί η αποτίμηση της επίδοσης να στηρίζεται σε απλοϊκά ή έμμεσα κριτήρια, όπως ο συνολικός χρόνος παραμονής στο εικονικό περιβάλλον του OnLabs ή ο αριθμός των κλικ στην οθόνη, η υλοποίησή μου ελέγχει τη λογική και σημασιολογική ορθότητα των εργαστηριακών πράξεων. Προσφέρει στους εκπαιδευτές τη δυνατότητα να εντοπίζουν με απόλυτη ακρίβεια ποια βήματα παραλείφθηκαν, πού εμφανίστηκαν δυσκολίες εστίασης και πόσες διορθωτικές κινήσεις απαιτήθηκαν. Συνδυάζοντας τα δεδομένα του OnLabs, τη διερευνητική ανάλυση του Heuristic Miner στο ProM, το Μοντέλο Αναφοράς και τον έλεγχο συμμόρφωσης χρησιμοποιώντας τη βιβλιοθήκη PM4Py, η αρχιτεκτονική της εργασίας (**Εικόνα 58**) αποτελεί ένα έτοιμο και ισχυρό εργαλείο ανάλυσης μαθησιακών δεδομένων (Learning Analytics).



Εικόνα 58: Η ροή εργασιών της διπλωματικής εργασίας

5.4 Συλλογή δεδομένων και διαμόρφωση του ιστορικού συμβάντων (Event Log)

Η επιτυχής υλοποίηση των τεχνικών της εξόρυξης διεργασιών βασίζεται θεμελιωδώς στην ποιότητα και τη δομή των πρωτογενών δεδομένων. Μετά την ολοκλήρωση του εργαστηριακού πειράματος «Μικροσκοπική Παρατήρηση Δείγματος», τα βήματα του οποίου περιγράφονται αναλυτικά στην ενότητα 4.5.2 της παρούσας διπλωματικής εργασίας, εξήχθη το αρχικό αρχείο καταγραφής συμβάντων (event log) από τη βάση δεδομένων του τρισδιάστατου περιβάλλοντος OnLabs. Το ψηφιακό αυτό αποτύπωμα περιλάμβανε το σύνολο των αλληλεπιδράσεων και των ενεργειών που πραγματοποιήθηκαν κατά τη διάρκεια της εργαστηριακής άσκησης, καταγεγραμμένα σε πραγματικό χρόνο. Ωστόσο, η ακατέργαστη αυτή μορφή του ιστορικού δεν ήταν άμεσα αξιοποιήσιμη, καθώς οι ονομασίες των συμβάντων ήταν χαμηλού επιπέδου (low-level events) και δεν αντιστοιχούσαν άμεσα στις τυποποιημένες έννοιες της θεωρητικής διεργασίας, ενώ παράλληλα απαιτούνταν η κατάλληλη διαχείριση των χρονικών σημάνσεων (timestamps) για τη διαμόρφωση ενός πλήρους και συμβατού αρχείου καταγραφής συμβάντων.

Για να ξεπεραστεί αυτός ο περιορισμός και να αναβαθμιστεί ποιοτικά η πληροφορία, αναπτύχθηκε και εφαρμόστηκε κατάλληλος κώδικας σε γλώσσα προγραμματισμού Python. Μέσω της επεξεργασίας αυτής, πραγματοποιήθηκε η συστηματική αντιστοίχιση (mapping) της κάθε μεμονωμένης ενέργειας που περιλαμβανόταν στο αρχικό αρχείο με το αντίστοιχο έργο του μοντέλου (model task), τα οποία ήταν προκαθορισμένα και αποθηκευμένα σε ένα ξεχωριστό αρχείο αναφοράς. Αυτή η αυτόματη διαδικασία διόρθωσης και μετατροπής

οδήγησε στην παραγωγή της τελικής μορφής του αρχείου καταγραφής συμβάντων. Το τελικό ιστορικό, έχοντας πλέον ομοιόμορφη ονοματολογία και συνεκτική δομή, αποτελεί ένα έγκυρο σύνολο δεδομένων, απόλυτα κατάλληλο για την εκτέλεση της διαδικασίας ελέγχου συμμόρφωσης (conformance checking) που ακολουθεί στην ενότητα 5.7.

Στην παρακάτω εικόνα (**Εικόνα 59**), παρουσιάζεται η αρχική μορφή των δεδομένων, όπως αυτά εξήχθησαν αυτόματα σε μορφή κειμένου μετά την ολοκλήρωση του πειράματος. Το αρχείο περιλαμβάνει μια συνεχή χρονολογική λίστα με τις λεπτομερείς και ακατέργαστες ψηφιακές καταγραφές του συστήματος για κάθε εκτελεσθείσα κίνηση στο τρισδιάστατο περιβάλλον. Η συγκεκριμένη μορφή αποτελεί τη βάση των δεδομένων της παρούσας εργασίας, η οποία αναδεικνύει την ανάγκη για την επακόλουθη επεξεργασία και τον μετασχηματισμό του ιστορικού σε μια πιο δομημένη μορφή.

```

Actions executed:
take_off_cover
used mic_plug with socket1
set mic_AC_switch on
rotated mic_light_intensity_knob
rotated mic_aperture_knob
rotated mic_condenser_knob
lens_4x_check
rotated mic_revolving_nosepiece
lens_4x_confirm
test_coarse_knob
test_fine_knob
test_stage_knob
test_specimen_knob
moved paper_scrap to bench
moved paper_scrap to bench
moved paper_scrap to bench
moved slide to bench
moved paper_scrap to bench
moved paper_scrap to bench
picked up paper_scrap to the inventory
moved paper_scrap to bench
moved paper_scrap to bench
picked up paper_scrap to the inventory
moved paper_scrap to bench
moved paper_scrap to bench
moved paper_scrap to bench
moved paper_scrap to bench
moved paper_scrap to bench
moved paper_scrap to bench
moved paper_scrap to bench
moved paper_scrap to bench
used paper_scrap with slide
used pasteur_body with beaker
picked up pasteur_body to the inventory
moved pasteur_body to bench
picked up pasteur_body to the inventory
used pasteur_body with slide
used slide_cover with slide
used slide with mic stage
fix_specimen_clips
start_microscoping
center_specimen
focused on microscope_instrument
rotated mic_y_axis_knob
rotated mic_ocular_right
rotated mic_ocular_left
rotated mic_coarse_focus_knob
rotated mic_fine_focus_knob
rotated mic_revolving_nosepiece
rotated mic_coarse_focus_knob
rotated mic_fine_focus_knob
rotated mic_coarse_focus_knob
rotated mic_revolving_nosepiece
rotated mic_coarse_focus_knob
rotated mic_fine_focus_knob
rotated mic_coarse_focus_knob
rotated mic_coarse_focus_knob
rotated mic_fine_focus_knob
rotated mic_fine_focus_knob
rotated mic_fine_focus_knob
rotated mic_fine_focus_knob
rotated mic_fine_focus_knob
unfocused from microscope_instrument
picked up slide to the inventory
remove_100x
set_4x_end
close_iris
rotated mic_aperture_knob
set mic_AC_switch off
pulled mic_plug
put_cover_on

```

Ln 36, Col 28

1.917 χαρακτήρες

Απλό κείμενο

Εικόνα 59: Πρωτογενές αρχείο καταγραφής από το περιβάλλον OnLabs

Στη συνέχεια αυτής της ενότητας περιγράφεται αναλυτικά η διαδικασία που ακολουθήθηκε για την επεξεργασία και την τελική διαμόρφωση του αρχικού ιστορικού συμβάντων. Συγκεκριμένα, παρουσιάζεται ο κώδικας που γράφτηκε σε γλώσσα Python, ο οποίος αναλαμβάνει να καθαρίσει τα δεδομένα και να διορθώσει αυτόματα ορισμένα τεχνικά σφάλματα που προκύπτουν κατά την εξαγωγή από την πλατφόρμα OnLabs. Βασικό εργαλείο σε αυτή τη ροή εργασίας είναι το βοηθητικό αρχείο `mapping.csv` το οποίο δημιουργήθηκε και ελέγχθηκε από την ομάδα συνεργατών του OnLabs. Το αρχείο αυτό λειτουργεί ως ένας πίνακας αναφοράς, ο οποίος επιτρέπει στον κώδικα να αντιστοιχίσει μία προς μία τις ακατέργαστες καταγραφές του συστήματος με τις επίσημες δραστηριότητες του μοντέλου. Με τον συνδυασμό του κώδικα Python και του αρχείου αντιστοίχισης, παράγεται το τελικό, δομημένο αρχείο καταγραφής (`final event log`). Το περιεχόμενο αυτού του τελικού αρχείου παρουσιάζεται στη συνέχεια, επιβεβαιώνοντας τη δημιουργία ενός καθαρού και χρονολογικά σωστού συνόλου δεδομένων, έτοιμου για τη διαδικασία του ελέγχου συμμόρφωσης.

Αρχικά, παρουσιάζεται και αναλύεται διεξοδικά η προγραμματιστική μεθοδολογία και η δομή του κώδικα Python που αναπτύχθηκε για τον μετασχηματισμό των δεδομένων (**Εικόνα 60**). Σκοπός του κειμένου είναι η πλήρης επεξήγηση των βημάτων επεξεργασίας που εκτελούνται αυτόματα από το πρόγραμμα για την παραγωγή του τελικού αρχείου καταγραφής συμβάντων (`event log`).

Η αρχιτεκτονική του κώδικα βασίζεται σε μια γραμμική και αυστηρά δομημένη ροή εργασιών, η οποία χωρίζεται σε διακριτές φάσεις επεξεργασίας. Στο πρώτο στάδιο, το πρόγραμμα πραγματοποιεί την εισαγωγή των απαραίτητων βιβλιοθηκών και τον καθορισμό των απόλυτων μονοπατιών των αρχείων στο σύστημα, εξασφαλίζοντας την πρόσβαση στα πρωτογενή δεδομένα. Αμέσως μετά, ο κώδικας προχωρά στη δυναμική φόρτωση του εξωτερικού αρχείου `mapping.csv`. Χρησιμοποιώντας τη βιβλιοθήκη Pandas, διαβάζει τον πίνακα αντιστοίχισης και μετατρέπει τις στήλες των ακατέργαστων και των επίσημων ενεργειών σε ένα κλασικό λεξικό της Python (`dictionary`). Η δομή αυτή επιλέχθηκε σκόπιμα, καθώς επιτρέπει στο σύστημα να αναζητά και να ανακτά τις μεταφράσεις των δραστηριοτήτων άμεσα, βελτιστοποιώντας την υπολογιστική απόδοση.

Το επόμενο κρίσιμο βήμα αφορά την ανάγνωση και τον αρχικό καθαρισμό του αρχικού αρχείου καταγραφής συμβάντων (`event_log.txt`). Το πρόγραμμα ανοίγει το αρχείο κειμένου

με κωδικοποίηση UTF-8 και εφαρμόζει ένα φίλτρο γραμμή προς γραμμή. Κατά τη διαδικασία αυτή, απορρίπτονται αυτόματα όλες οι κενές γραμμές, καθώς και οι καταγραφές συστήματος που αποτελούν επικεφαλίδες ή σχόλια και ξεκινούν με ειδικούς χαρακτήρες, κρατώντας μόνο τις πραγματικές ενέργειες των χρηστών. Αμέσως μετά, εκτελείται ένας εξειδικευμένος αλγόριθμος διόρθωσης σφαλμάτων (bug fix) για την αντιμετώπιση των τεχνικών ασυνεχιών της πλατφόρμας OnLabs. Ο αλγόριθμος εντοπίζει τις ασύνδετες καταγραφές με την ένδειξη `mic_stage` και τις συγχωνεύει αυτόματα στο τέλος της προηγούμενης εγγραφής τους, αποκαθιστώντας πλήρως την ακεραιότητα του ιστορικού.

Στη συνέχεια, ο κώδικας εφαρμόζει έναν μηχανισμό αυστηρού ελέγχου εγκυρότητας (validation). Διατρέχει όλες τις διορθωμένες γραμμές του ιστορικού και επαληθεύει ότι κάθε μία από αυτές υπάρχει μέσα στο λεξικό αντιστοίχισης. Εάν εντοπιστεί έστω και μία ενέργεια που παραλείφθηκε από το `mapping.csv`, το πρόγραμμα διακόπτει αμέσως τη λειτουργία του και εμφανίζει προειδοποιητικό μήνυμα σφάλματος. Αυτή η αυστηρή λογική εγγυάται ότι δεν θα περάσουν αμετάφραστα δεδομένα στο τελικό αρχείο, προστατεύοντας την εγκυρότητα της διαδικασίας ελέγχου συμμόρφωσης (Conformance Checking). Εφόσον ο έλεγχος ολοκληρωθεί με επιτυχία, ο κώδικας μεταφράζει τις ενέργειες και φιλτράρει το αποτέλεσμα, αφαιρώντας οριστικά όλες τις βοηθητικές κινήσεις που έχουν χαρακτηριστεί ως αόρατα βήματα (`invisible_step`).

Στο τελικό στάδιο, ο κώδικας αναλαμβάνει να δώσει στα δεδομένα τη δομή ενός επίσημου αρχείου καταγραφής συμβάντων σύμφωνα με τα διεθνή πρότυπα του Process Mining. Δημιουργεί τη στήλη του Case ID αποδίδοντας τη σταθερή τιμή `Case_1` για όλη τη ροή του πειράματος, και παράγει αυτόματα συνθετικές χρονικές σημάνσεις (timestamps). Με αφετηρία μια συγκεκριμένη ημερομηνία και ώρα, ο κώδικας προσθέτει διαδοχικά 10 δευτερόλεπτα για κάθε επόμενη ενέργεια, εξασφαλίζοντας ένα σύνολο δεδομένων με συνεπή τεχνητή χρονική διάταξη. Επισημαίνεται ότι η σταθερή αυτή διαφορά των 10 δευτερολέπτων εξυπηρετεί αποκλειστικά τη διατήρηση της σωστής σειράς των ενεργειών και δεν επιτρέπει τη διεξαγωγή ανάλυσης απόδοσης (performance analysis) ή τη μελέτη χρονικών καθυστερήσεων. Η όλη διαδικασία ολοκληρώνεται με την ταξινόμηση των στηλών και την εξαγωγή του τελικού, καθαρού πίνακα δεδομένων σε ένα νέο αρχείο μορφής CSV, το οποίο είναι πλέον έτοιμο για την πραγματοποίηση του ελέγχου συμμόρφωσης.

```

import pandas as pd
import sys
from datetime import datetime, timedelta
import tkinter as tk
from tkinter import filedialog, messagebox

# 1. ΔΥΝΑΜΙΚΟΣ ΚΑΘΟΡΙΣΜΟΣ ΜΟΝΟΠΑΤΙΩΝ ΜΕΣΩ ΓΡΑΦΙΚΟΥ ΠΕΡΙΒΑΛΛΟΝΤΟΣ (GUI)
# Αρχικοποίηση της βιβλιοθήκης Tkinter και άμεση απόκρυψη του κυρίως παραθύρου (root window).
# Η διαδικασία αυτή επιτρέπει την κλήση των αυτόνομων παραθύρων διαλόγου των Windows,
# χωρίς την ανάγκη σχεδίασης μιας πλήρους γραφικής διεπαφής (user interface).
root = tk.Tk()
root.withdraw()

# Εμφάνιση αρχικού επεξηγηματικού μηνύματος για την καθοδήγηση του χρήστη στο πρώτο στάδιο.
messagebox.showinfo(
    "Ενημέρωση",
    "Επιλέξτε το αρχείο καταγραφής συμβάντων και το αρχείο αντιστοίχισης των ενεργειών."
)

# α) Δυναμική επιλογή του αρχικού αρχείου καταγραφής (Event Log) από τον χρήστη
log_path = filedialog.askopenfilename(
    title="1. Επιλέξτε το αρχικό αρχείο καταγραφής (event_log.txt)",
    filetypes=[("Text Files", "*.txt"), ("All Files", "*.*)")]
)

# Έλεγχος εγκυρότητας: Εάν ο χρήστης κλείσει το παράθυρο χωρίς επιλογή, η εκτέλεση διακόπτεται με ασφάλεια.
if not log_path:
    sys.exit("Η διαδικασία ακυρώθηκε: Δεν επιλέχθηκε αρχικό αρχείο Event Log.")

# β) Δυναμική επιλογή του αρχείου αντιστοίχισης (Mapping CSV) από τον χρήστη
mapping_csv_path = filedialog.askopenfilename(
    title="2. Επιλέξτε το αρχείο αντιστοίχισης (mapping.csv)",
    filetypes=[("CSV Files", "*.csv"), ("All Files", "*.*)")]
)

# Έλεγχος εγκυρότητας: Εάν δεν επιλεγεί αρχείο mapping, η διαδικασία τερματίζεται για την αποφυγή σφαλμάτων.
if not mapping_csv_path:
    sys.exit("Η διαδικασία ακυρώθηκε: Δεν επιλέχθηκε αρχείο Mapping.")

# 2. ΦΟΡΤΩΣΗ ΤΟΥ ΑΡΧΕΙΟΥ MAPPING ΚΑΙ ΜΕΤΑΤΡΟΠΗ ΣΕ ΛΕΞΙΚΟ (DICTIONARY)
# Η βιβλιοθήκη Pandas εισάγει τον πίνακα αντιστοίχισης από το αρχείο CSV σε μορφή DataFrame.
# Στη συνέχεια, οι δύο στήλες ("Log_Activity" και "Model_Activity") συνδυάζονται μέσω της συνάρτησης zip.
# Η δομή αυτή επιτρέπει την αναζήτηση και ανάκτηση της επίσημης ονομασίας κάθε ενέργειας,
# χρησιμοποιώντας ως μοναδικό κλειδί αναγνώρισης (key) την ακατέργαστη καταγραφή του onLabs.
mapping_df = pd.read_csv(mapping_csv_path)
mapping_dict = dict(zip(mapping_df["Log_Activity"], mapping_df["Model_Activity"]))

# 3. ΑΝΑΓΝΩΣΗ ΤΟΥ EVENT_LOG.TXT ΚΑΙ ΦΙΛΤΡΑΡΙΣΜΑ ΓΡΑΜΜΩΝ
# Το αρχείο κείμενου ανοίγει με ασφάλεια μέσω της εντολής "with", χρησιμοποιώντας κωδικοποίηση UTF-8.
# Προσδιορίζεται μια διαδικασία εκκαθάρισης (list comprehension) που εκτελεί τρεις ελέγχους ταυτόχρονα:
# α) Απομακρύνει τους περιττούς χαρακτήρες αλλαγής γραμμής και τα κενά διαστήματα μέσω της μεθόδου .strip().
# β) Απομακρύνει άμεσα τις νέες γραμμές που ενδέχεται να υπάρχουν στο πρωτόγενές αρχείο.
# γ) Φιλτράρει και αφαιρεί γραμμές συστήματος που ξεκινούν με "[" ή με τη λέξη "Actions", καθώς δεν αποτελούν ενέργειες.
with open(log_path, "r", encoding="utf-8") as file:
    lines = [line.strip() for line in file if line.strip() and not line.strip().startswith("[", "Actions")]

# 4. ΕΝΤΟΠΙΣΜΟΣ ΚΑΙ ΔΙΟΡΘΩΣΗ ΤΩΝ ΣΦΑΛΜΑΤΩΣ ΤΩΝ ONLABS (BUG FIX)
# Λόγω τεχνικού σφάλματος στην εξαγωγή των δεδομένων από το onLabs, η καταγραφή "mic_stage"
# αποσπάται από την ενότητα "line" και εμφανίζεται λανθασμένα ως ξεχωριστή, "ερρανή" γραμμή.
# Ο βρόχος επανάληψης "for" εντοπίζει αυτές τις "ερρανές" γραμμές και τις συγχωνεύει στο τέλος της προηγούμενης εγγραφής.
# Αμέσως μετά, δημιουργείται μια νέα φιλτραρισμένη λίστα που πετάει τις πλέον περιττές γραμμές "mic_stage".
# Με αυτόν τον τρόπο αποκτάται η ακεραιότητα της καταγραφής χωρίς να χαθεί καμία πληροφορία.
for i in range(len(lines)):
    if lines[i] == "mic_stage" and i > 0:
        lines[i-1] += lines[i]
        lines = [l for l in lines if l != "mic_stage"]

# 5. ΑΥΤΙΣΤΡΟΣ ΕΛΕΓΧΟΣ ΑΝΤΙΣΤΟΙΧΙΣΗΣ (MAPPING VALIDATION)
# Αυτό το στάδιο εγγυάται τη μεθυστική εγκυρότητα της διαδικασίας Conformance Checking.
# Ο κώδικας εξετάζει μία-μία όλες τις καταγραφές του log και ελέγχει αν είναι καταχωρημένες στο mapping.csv.
# Εάν εντοπίζεται ενώ και μία ενέργεια που δεν έχει αντιστοιχιστεί, η εκτέλεση διακόπτεται αμέσως (sys.exit).
# Αυτό σημαίνει το δημιουργία ενός μητέλους ή λανθασμένου Event Log, το οποίο θα οδηγούσε
# σε εσφαλμένα ή παραπλανητικά συμπεράσματα κατά την ανάλυση της συμμόρφωσης με χρήση της βιβλιοθήκης PM4Py.
for l in lines:
    if l not in mapping_dict:
        sys.exit(f"ΣΦΑΛΜΑ: Η ενέργεια '{l}' λείπει από το αρχείο mapping.csv! Η διαδικασία διακόπηκε.")

# 6. ΕΦΑΡΜΟΓΗ ΤΗΣ ΑΝΤΙΣΤΟΙΧΙΣΗΣ ΓΙΑ ΤΗ ΔΗΜΙΟΥΡΓΙΑ ΤΟΥ DataFrame
# Στο βήμα αυτό πραγματοποιείται η ουσιαστική μετάφραση των δεδομένων υψηλής ανάλυσης.
# Μέσω μιας επαναληπτικής διαδικασίας, κάθε ακατέργαστη ενέργεια του log αντικαθίσταται
# από την επίσημη, τοποθετημένη ονομασία της που ανακτήθηκε από το mapping_dict.
# Το αποτέλεσμα αναμένεται σε έναν πίνακα της Pandas (DataFrame) κάτω από την υποχρεωτική,
# βάσει των διεθνών προτύπων Process Mining (XES standard), στήλη με όνομα "concept:name".
df = pd.DataFrame({"concept:name": [mapping_dict[l] for l in lines]})

# 7. ΑΦΑΙΡΕΣΗ ΤΩΝ ΒΟΗΘΗΤΙΚΩΝ ΚΑΤΑΓΡΑΦΩΝ (INVISIBLE STEPS)
# Οι βοηθητικές ενέργειες (π.χ. απλή επίθεση ή μετακίνησης αντικείμενων στον πάγκο εργασίας)
# έχουν χαρακτηριστεί από το mapping με την ειδική ετικέτα "invisible_step".
# Στο σημείο αυτό, ο κώδικας φιλτράρει το dataframe αποκλώνοντας οριστικά αυτές τις γραμμές.
# Η μέθοδος .reset_index(drop=True) αναδομεί την αρίθμηση των γραμμών ώστε να ξεκινά πάλι ομαλά από το 0,
# διασφαλίζοντας ομοιομορφία των δεδομένων και χρησιμική σειρά του πίνακα.
df = df[df[["concept:name"]] != "invisible_step"].reset_index(drop=True)

# 8. ΠΡΟΣΘΗΚΗ ΤΗΣ ΥΠΟΧΡΕΩΤΙΚΗΣ ΣΤΗΛΗΣ CASE ID (ΥΠΟΘΕΣΗ)
# Κάθε έγκυρο αρχείο Event Log αποτελεί τον αση διαχωρισμό των διαφορετικών εκτελέσεων (cases).
# Επειδή η συγκεκριμένη καταγραφή αφορά τη μελέτη μιας μεμονωμένης, ολοκληρωμένης ροής πεπραμάτων,
# δημιουργείται μια νέα στήλη με την επίσημη ονομασία "case:concept:name".
# Σε όλες τις γραμμές του πίνακα, το σκελετό της στήλης "Case ID" δηλώνοντας ότι όλες οι
# καταγεγραμμένες δραστηριότητες ανήκουν στην ίδια ακριβώς διαδικασία εκτέλεσης.
df["case:concept:name"] = "Case-1"

# 9. ΔΗΜΙΟΥΡΓΙΑ ΣΤΗΜΑΤΙΚΩΝ ΧΡΟΝΙΚΩΝ ΣΗΜΑΝΣΕΩΝ (TIMESTAMPS)
# Το Process Mining απαιτεί την ύπαρξη χρονολογικής σειράς για την εύρεση της ροής της διεργασίας.
# Καθώς το onLabs δεν εδάζει πραγματικούς χρόνους, ορίζεται μια σταθερή έναρξη (26 Μαΐου 2026, 10:00:00).
# Μέσω της κλάσης timedelta, ο κώδικας υπολογίζει και προσέχει αντίκριμα 10 δευτερόλεπτα για κάθε γραμμή (1).
# Έτσι, η πρώτη ενέργεια λαμβάνει χρόνο 10:00:00, η δεύτερη 10:00:10, η τρίτη 10:00:20, εξασφαλίζοντας
# μια τεχνητή αλλά απόλυτα λογική και σωστή χρονολογική αλληλουχία στη στήλη "time:timestamp".
df["time:timestamp"] = [datetime(2026, 5, 26, 10, 0, 0) + timedelta(seconds=l*10) for l in range(len(df))]

# 10. ΔΗΜΙΟΥΡΓΙΑ ΑΠΟΘΗΚΕΥΣΗ ΚΑΙ ΕΞΑΓΩΓΗ ΤΟΥ ΤΕΛΙΚΟΥ CSV
# Πριν πραγματοποιηθεί η εγγραφή στον δίσκο, εκτελείται ένα τελικό ενημερωτικό παράθυρο
# που παρέχει τον χρήστη να προσδιορίσει την τοποθεσία εξαγωγής του τελικού δομημένου αρχείου.
messagebox.showinfo(
    "Ενημέρωση",
    "Επιλέξτε τον φάκελο και το όνομα αρχείου για την αποθήκευση του τελικού Event Log."
)

# Άνοιγμα του ειδικού παραθύρου διαλόγου των Windows για τον δυναμικό ορισμό φακέλου, ονόματος και κατάληξης αρχείου.
final_csv_path = filedialog.asksaveasfilename(
    title="3. Ορίστε όνομα και τοποθεσία αποθήκευσης του τελικού Event Log",
    defaultextension=".csv",
    filetypes=[("CSV Files", "*.csv"), ("All Files", "*.*)")]
)

# Έλεγχος εγκυρότητας: Εάν ο χρήστης ακυρώσει την αποθήκευση, ο κώδικας τερματίζει χωρίς να προκαλέσει runtime error.
if not final_csv_path:
    sys.exit("Η διαδικασία ακυρώθηκε: Δεν ορίστηκε αρχείο προορισμού.")

# Επιλογή των τριών υποχρεωτικών στήλων (Case ID, Activity Name, Timestamp) βάσει των διεθνών προτύπων.
# Τα δεδομένα εξάγονται στο δυναμικό μονοπάτι (final_csv_path) με κωδικοποίηση UTF-8 και απενεργοποιημένους τους εσωτερικούς
# δείκτες.
df[["case:concept:name", "concept:name", "time:timestamp"]].to_csv(final_csv_path, index=False, encoding="utf-8")

print("\n=====")
print("Το τελικό Event Log δημιουργήθηκε με απόλυτη επιτυχία!!!")
print(f"Αρχείο εισόδου: {log_path}")
print(f"Αρχείο αντιστοίχισης: {mapping_csv_path}")
print(f"Αρχείο εξόδου: {final_csv_path}")
print("=====")

```

Εικόνα 60: Προγραμματιστική επεξεργασία του αρχικού ιστορικού συμβάντων

Στην εικόνα που ακολουθεί (**Εικόνα 61**), παρουσιάζονται αναλυτικά τα περιεχόμενα του βοηθητικού αρχείου `mapping.csv`, το οποίο αποτέλεσε τη βάση για τη μετάφραση των δεδομένων. Μέσα από τον συγκεκριμένο πίνακα αναφοράς γίνεται εύκολα κατανοητός ο τρόπος με τον οποίο κάθε ακατέργαστη καταγραφή του συστήματος OnLabs συνδέεται άμεσα και μονοσήμαντα με μια επίσημη δραστηριότητα του μοντέλου διεργασίας. Η οπτικοποίηση αυτή αναδεικνύει τη δομή του αρχείου αντιστοίχισης, αποτυπώνοντας καθαρά τη μετάβαση από τις τεχνικές καταγραφές χαμηλού επιπέδου σε τυποποιημένα βήματα με σαφές επιστημονικό νόημα.



```
Log_Activity,Model_Activity
take_off_cover,take off cover from microscope
used mic_plug with socket1,connect microscope to socket
set mic_AC_switch on,turn on microscope light
rotated mic_light_intensity_knob,adjust light intensity control
rotated mic_aperture_knob,set iris fully open
rotated mic_condenser_knob,lift condenser knob to top
lens_4x_check,lens 4x active?
rotated mic_revolving_nosepiece,set lens 4x active
lens_4x_confirm,lens 4x active
test_coarse_knob,test the coarse focus adjustment knob
test_fine_knob,test the fine focus adjustment knob
test_stage_knob,test the stage knob
test_specimen_knob,test the specimen knob
used paper_scrap with slide,fold paper scrap to point
used pasteur_body with beaker,draw liquid with pasteur pipette
used pasteur_body with slide,add liquid specimen drop
used slide_cover with slide,cover specimen with slide cover
used slide with mic_stage,put test specimen on stage
fix_specimen_clips,fix test specimen on stage using stage clips
start_microscoping,start microscoping
center_specimen,use stage knob and specimen holder knob to place specimen at the center
rotated mic_coarse_focus_knob,adjust coarse focus knob to zoom
rotated mic_fine_focus_knob,adjust fine focus knob to focus
rotated mic_y_axis_knob,move stage y axis
rotated mic_ocular_right,adjust ocular lens
rotated mic_ocular_left,adjust ocular lens
rotated mic_x_axis_knob,move stage x axis
picked up slide to the inventory,remove specimen
remove_100x,remove 100x lens from microscope field of view
set_4x_end,set 4x lens active
close_iris,close iris
set mic_AC_switch off,turn off microscope light
pulled mic_plug,turn off microscope light
put_cover_on,put microscope cover on
focused on microscope_instrument,invisible_step
unfocused from microscope_instrument,invisible_step
moved paper_scrap to bench,invisible_step
moved slide to bench,invisible_step
moved pasteur_body to bench,invisible_step
moved slide_cover to bench,invisible_step
picked up paper_scrap to the inventory,invisible_step
picked up pasteur_body to the inventory,invisible_step
picked up slide_cover to the inventory,invisible_step
picked up beaker to the inventory,invisible_step
```

Εικόνα 61: Δομή και περιεχόμενα του αρχείου αντιστοίχισης

Στην εικόνα που ακολουθεί (**Εικόνα 62**), αποτυπώνεται το πλήρες περιεχόμενο από το τελικό αρχείο καταγραφής συμβάντων, όπως αυτό εξήχθη μετά την ολοκλήρωση της εκτέλεσης του αλγορίθμου. Στη συγκεκριμένη μορφή, τα δεδομένα έχουν πλέον διαμορφωθεί κατάλληλα, οι αποκομμένες γραμμές έχουν αποκατασταθεί και κάθε ενέργεια έχει αντιστοιχιστεί στην επίσημη ονοματολογία της. Η δομή αυτή, οργανωμένη σε καθαρές στήλες, αποτελεί το τελικό αποτέλεσμα της προεπεξεργασίας και θα χρησιμοποιηθεί για την πραγματοποίηση του ελέγχου συμμόρφωσης σε επόμενη ενότητα.

```

case:concept:name,concept:name,time:timestamp
Case_1,take off cover from microscope,2026-05-26 10:00:00
Case_1,connect microscope to socket,2026-05-26 10:00:10
Case_1,turn on microscope light,2026-05-26 10:00:20
Case_1,adjust light intensity control,2026-05-26 10:00:30
Case_1,set iris fully open,2026-05-26 10:00:40
Case_1,lift condenser knob to top,2026-05-26 10:00:50
Case_1,lens 4x active?,2026-05-26 10:01:00
Case_1,set lens 4x active,2026-05-26 10:01:10
Case_1,lens 4x active,2026-05-26 10:01:20
Case_1,test the coarse focus adjustment knob,2026-05-26 10:01:30
Case_1,test the fine focus adjustment knob,2026-05-26 10:01:40
Case_1,test the stage knob,2026-05-26 10:01:50
Case_1,test the specimen knob,2026-05-26 10:02:00
Case_1,fold paper scrap to point,2026-05-26 10:02:10
Case_1,draw liquid with pasteur pipette,2026-05-26 10:02:20
Case_1,add liquid specimen drop,2026-05-26 10:02:30
Case_1,cover specimen with slide cover,2026-05-26 10:02:40
Case_1,put test specimen on stage,2026-05-26 10:02:50
Case_1,fix test specimen on stage using stage clips,2026-05-26 10:03:00
Case_1,start microscoping,2026-05-26 10:03:10
Case_1,use stage knob and specimen holder knob to place specimen at the center,2026-05-26
10:03:20
Case_1,move stage y axis,2026-05-26 10:03:30
Case_1,adjust ocular lens,2026-05-26 10:03:40
Case_1,adjust ocular lens,2026-05-26 10:03:50
Case_1,adjust coarse focus knob to zoom,2026-05-26 10:04:00
Case_1,adjust fine focus knob to focus,2026-05-26 10:04:10
Case_1,set lens 4x active,2026-05-26 10:04:20
Case_1,adjust coarse focus knob to zoom,2026-05-26 10:04:30
Case_1,adjust fine focus knob to focus,2026-05-26 10:04:40
Case_1,adjust coarse focus knob to zoom,2026-05-26 10:04:50
Case_1,set lens 4x active,2026-05-26 10:05:00
Case_1,adjust coarse focus knob to zoom,2026-05-26 10:05:10
Case_1,adjust fine focus knob to focus,2026-05-26 10:05:20
Case_1,adjust coarse focus knob to zoom,2026-05-26 10:05:30
Case_1,adjust coarse focus knob to zoom,2026-05-26 10:05:40
Case_1,adjust fine focus knob to focus,2026-05-26 10:05:50
Case_1,set lens 4x active,2026-05-26 10:06:00
Case_1,adjust fine focus knob to focus,2026-05-26 10:06:10
Case_1,adjust fine focus knob to focus,2026-05-26 10:06:20
Case_1,adjust fine focus knob to focus,2026-05-26 10:06:30
Case_1,adjust fine focus knob to focus,2026-05-26 10:06:40
Case_1,adjust fine focus knob to focus,2026-05-26 10:06:50
Case_1,remove specimen,2026-05-26 10:07:00
Case_1,remove 100x lens from microscope field of view,2026-05-26 10:07:10
Case_1,set 4x lens active,2026-05-26 10:07:20
Case_1,close iris,2026-05-26 10:07:30
Case_1,set iris fully open,2026-05-26 10:07:40
Case_1,turn off microscope light,2026-05-26 10:07:50
Case_1,turn off microscope light,2026-05-26 10:08:00
Case_1,put microscope cover on,2026-05-26 10:08:10

```

Εικόνα 62: Δομή και περιεχόμενα του τελικού αρχείου καταγραφής συμβάντων

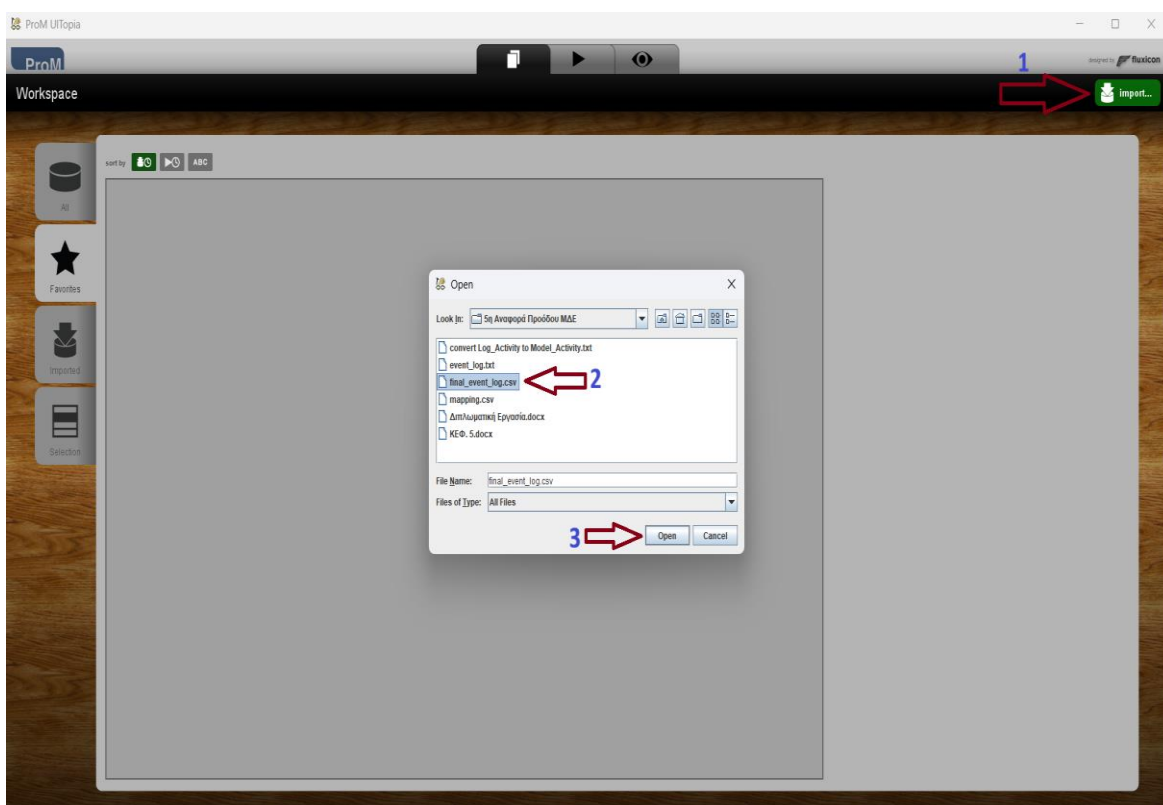
5.5 Ανακάλυψη διαδικασιών (Process Discovery) στο περιβάλλον ProM

Στην παρούσα ενότητα αναλύεται διεξοδικά η διαδικασία της ανακάλυψης διεργασιών (Process Discovery) για το εργαστηριακό πείραμα «Μικροσκοπική Παρατήρηση Δείγματος», με σκοπό την αυτόματη παραγωγή ενός αντιπροσωπευτικού μοντέλου συμπεριφοράς. Η υλοποίηση της φάσης αυτής βασίστηκε αποκλειστικά στο πραγματικό αρχείο καταγραφής συμβάντων (event log), το οποίο εξήχθη από το τρισδιάστατο εκπαιδευτικό περιβάλλον OnLabs κατά την προσωπική εκτέλεση του πειράματος. Μέσα από την επεξεργασία αυτών των ψηφιακών αποτυπωμάτων, η ανακάλυψη διεργασιών επιτρέπει τη μετατροπή των ασύνδετων καταγραφών σε μια ενιαία, δυναμική και μαθηματικά τεκμηριωμένη ροή εργασίας. Με τον τρόπο αυτό, καθίσταται δυνατή η οπτικοποίηση της πραγματικής αλληλουχίας των ενεργειών που πραγματοποιήθηκαν στο εργαστήριο, αποκαλύπτοντας την ακριβή συμπεριφορά του χρήστη, τις τυχόν παρακάμψεις, καθώς και τη συνολική δομή της εκτελεσθείσας διαδικασίας σε σχέση με τις προδιαγραφές του πειράματος.

Για την ανάδειξη της βέλτιστης δυνατής αναπαράστασης, η μεθοδολογία μου στηρίχθηκε στα συμπεράσματα της προηγούμενης έρευνας. Συγκεκριμένα, με βάση τη συγκριτική αξιολόγηση αλγορίθμων εξόρυξης διεργασιών που πραγματοποιήθηκε στην ενότητα 1.1.5.1 της παρούσας διπλωματικής εργασίας, για την πραγματοποίηση της συγκεκριμένης διαδικασίας επιλέχθηκε ο αλγόριθμος Heuristic Miner. Ο συγκεκριμένος αλγόριθμος κρίθηκε ως ο πλέον κατάλληλος, καθώς επιδεικνύει εξαιρετική ανθεκτικότητα στην παρουσία θορύβου (noise) στα δεδομένα και επικεντρώνεται στις κύριες, συχνές διαδρομές της διεργασίας, αποφεύγοντας την πολυπλοκότητα των σπάνιων εξαιρέσεων. Παρακάτω, παρουσιάζονται αναλυτικά και βήμα-βήμα τα στάδια εκτέλεσης της ανακάλυψης του μοντέλου διεργασίας στο λογισμικό ProM, η παραμετροποίηση των κατάλληλων δεικτών, καθώς και το τελικό παραγόμενο μοντέλο, όπως αυτά αποτυπώνονται μέσα από τα αντίστοιχα στιγμιότυπα οθόνης.

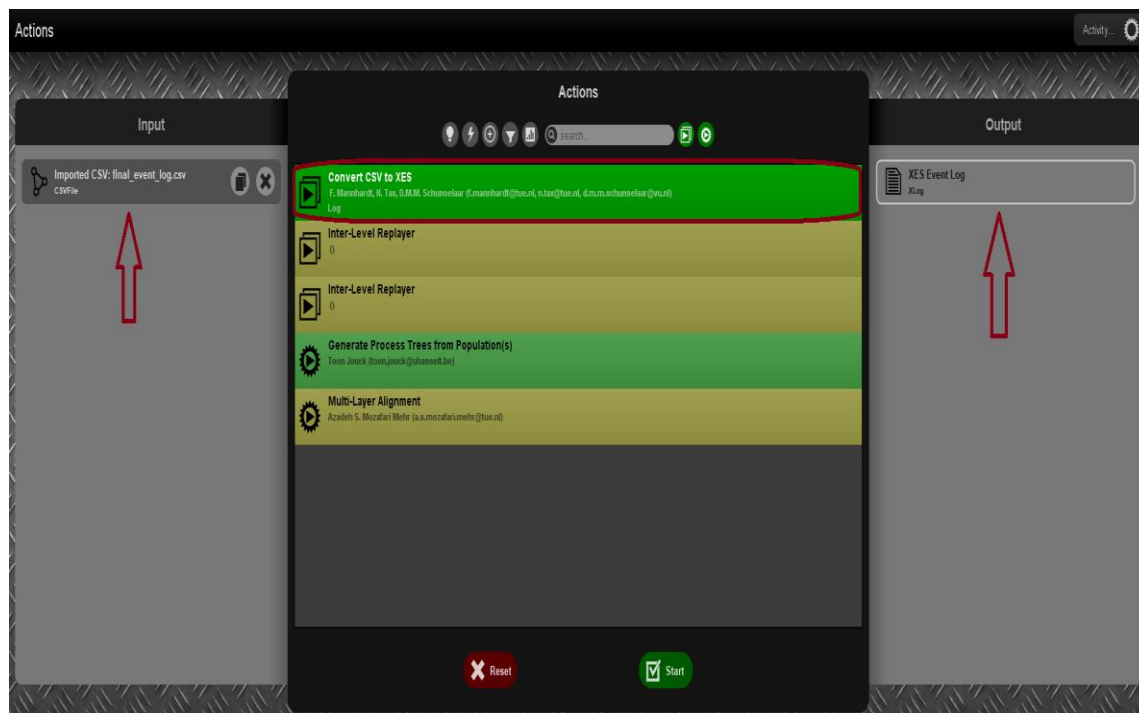
Στην παρακάτω εικόνα (**Εικόνα 63**) αποτυπώνεται το αρχικό στάδιο της διαδικασίας, το οποίο αφορά την εισαγωγή των δεδομένων στο λογισμικό ProM. Μέσω της επιλογής της εντολής εισαγωγής (Import) στην επάνω δεξιά γωνία της οθόνης, ανοίγει το παράθυρο πλοήγησης στα αρχεία του συστήματος για τον εντοπισμό του κατάλληλου αρχείου. Στη συγκεκριμένη περίπτωση, επιλέγεται το εξαχθέν αρχείο καταγραφής συμβάντων σε μορφή

CSV, το οποίο περιλαμβάνει τα ψηφιακά αποτυπώματα των ενεργειών που πραγματοποιήθηκαν κατά την εκτέλεση του πειράματος, αποτελώντας τη βάση δεδομένων για την επικείμενη ανακάλυψη της διεργασίας.



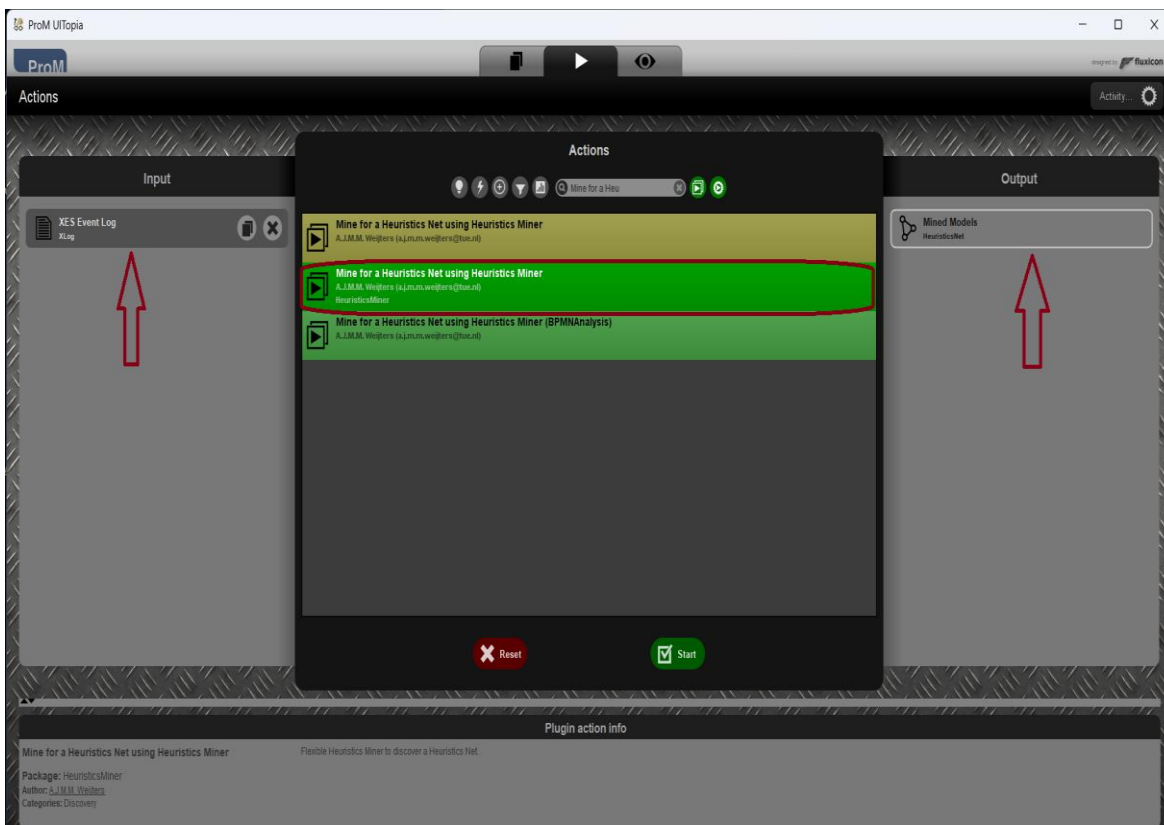
Εικόνα 63: Εισαγωγή του αρχείου καταγραφής συμβάντων στο περιβάλλον ProM

Στην παρακάτω εικόνα (**Εικόνα 64**) παρουσιάζεται το παράθυρο ενεργειών του ProM, όπου καθορίζεται η μέθοδος ψηφιακής μετάβασης του εισαχθέντος αρχείου. Για την αξιοποίηση των δεδομένων, επιλέχθηκε το πρόσθετο «Convert CSV to XES», το οποίο αναλαμβάνει τη δομική μεταβολή των καταγραφών από απλό κείμενο διαχωρισμένο με κόμματα σε αντικειμενοστραφή μορφή γεγονότων. Στο αριστερό τμήμα της οθόνης αποτυπώνεται το αρχικό αρχείο ως πηγή εισόδου (Input), ενώ στη δεξιά πλευρά προβάλλεται το αναμενόμενο αποτέλεσμα της επεξεργασίας (Output), προετοιμάζοντας το περιβάλλον για την έναρξη της διαδικασίας.



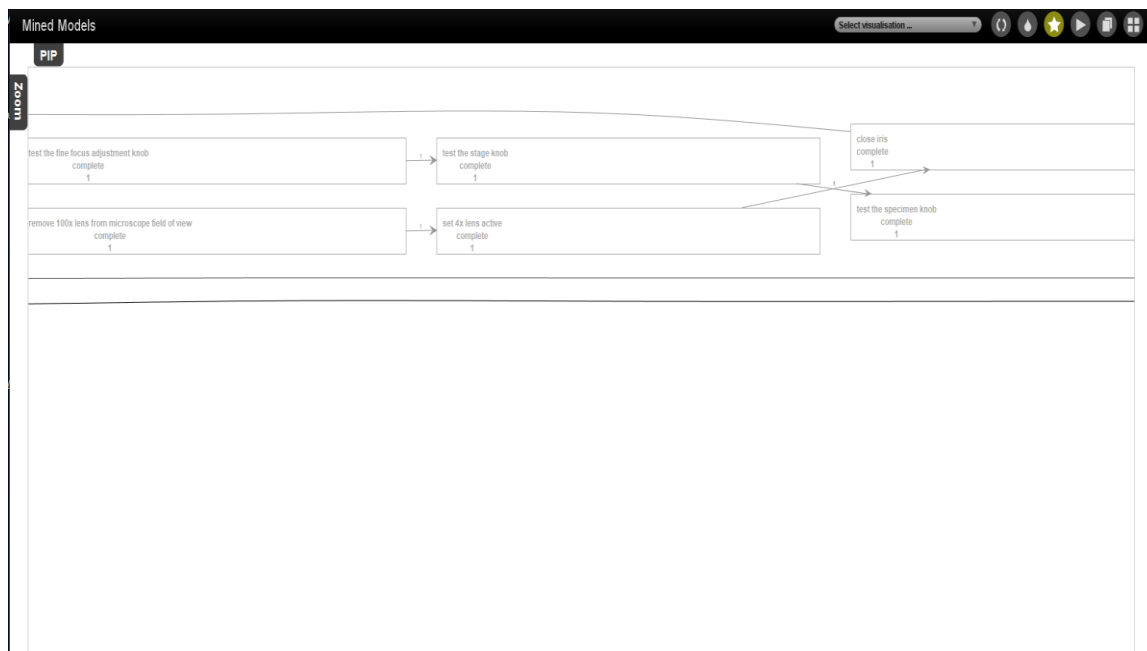
Εικόνα 64: Επιλογή του προσθέτου μετατροπής σε μορφή XES

Στην παρακάτω εικόνα (**Εικόνα 65**) αποτυπώνεται η διαδικασία ανάθεσης της κύριας υπολογιστικής εργασίας ανακάλυψης στο περιβάλλον Actions του λογισμικού ProM. Έχοντας μετατρέψει επιτυχώς τα δεδομένα στη δομή συμβάντων, επιλέχθηκε το εξειδικευμένο πρόσθετο «Mine for a Heuristics Net using Heuristics Miner». Στο αριστερό τμήμα ορίζεται ως είσοδος (Input) το διαμορφωμένο αρχείο συμβάντων, ενώ στη δεξιά πλευρά προσδιορίζεται το αναμενόμενο αντικείμενο εξόδου (Output), το οποίο θα είναι ένα ευρετικό δίκτυο (Heuristics Net). Η επιλογή αυτή, όπως προέκυψε από τη συγκριτική αξιολόγηση, καθορίζει την έναρξη της δομικής ανάλυσης της διεργασίας με βάση τις συχνότητες των καταγεγραμμένων ενεργειών.



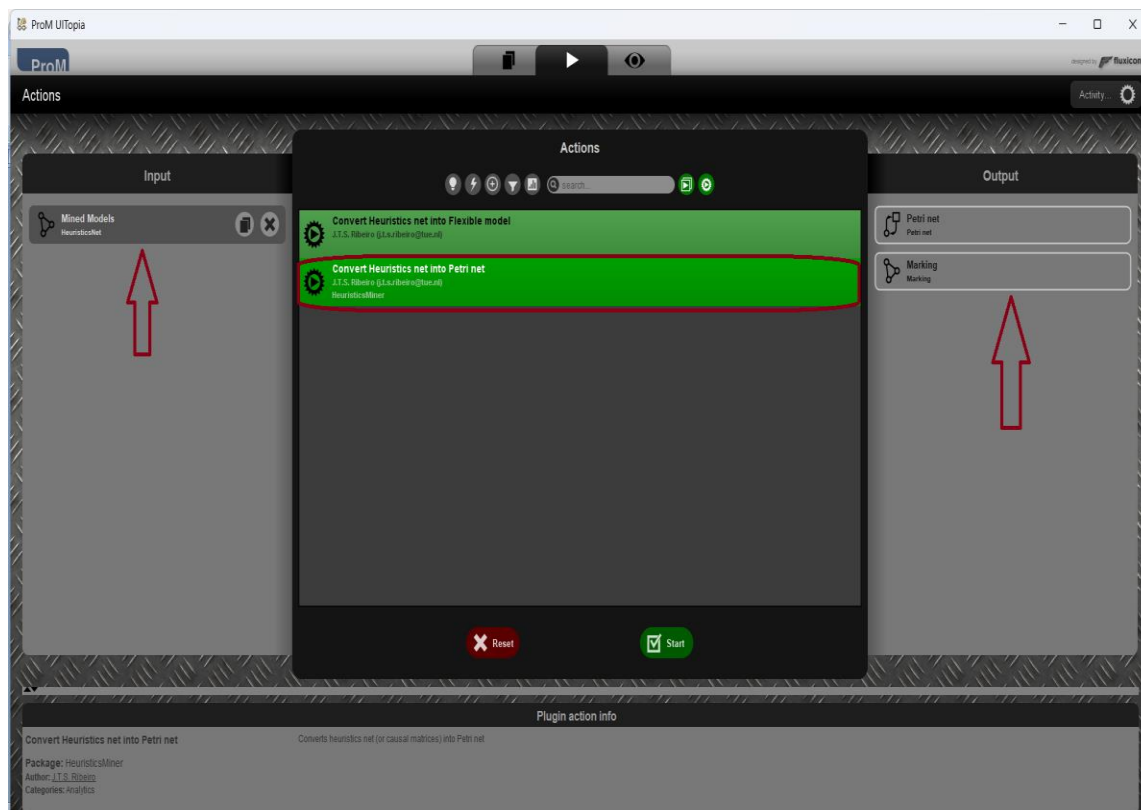
Εικόνα 65: Επιλογή του αλγορίθμου Heuristic Miner για την ανακάλυψη μοντέλου

Στην παρακάτω εικόνα (Εικόνα 66) παρουσιάζεται ένα αντιπροσωπευτικό τμήμα του τελικού αποτελέσματος της επεξεργασίας, το οποίο αφορά το ανακαλυφθέν ευρετικό δίκτυο (Heuristics Net) της εργαστηριακής άσκησης. Λόγω του εύρους της συνολικής διεργασίας, αποτυπώνεται εστιασμένα ένα μέρος των επιμέρους δραστηριοτήτων της μικροσκοπικής παρατήρησης, όπως οι δοκιμές των κομβίων εστίασης και της τράπεζας, καθώς και οι αλλαγές των αντικειμενικών φακών. Οι κατευθυνόμενες ακμές μεταξύ των ορθογώνιων πλαισίων αναδεικνύουν την πραγματική αλληλουχία και τις αιτιατές εξαρτήσεις των συγκεκριμένων βημάτων, προσφέροντας μια καθαρή και εποπτική εικόνα για τη δομή της ροής που εκτελέστηκε στο ψηφιακό περιβάλλον.



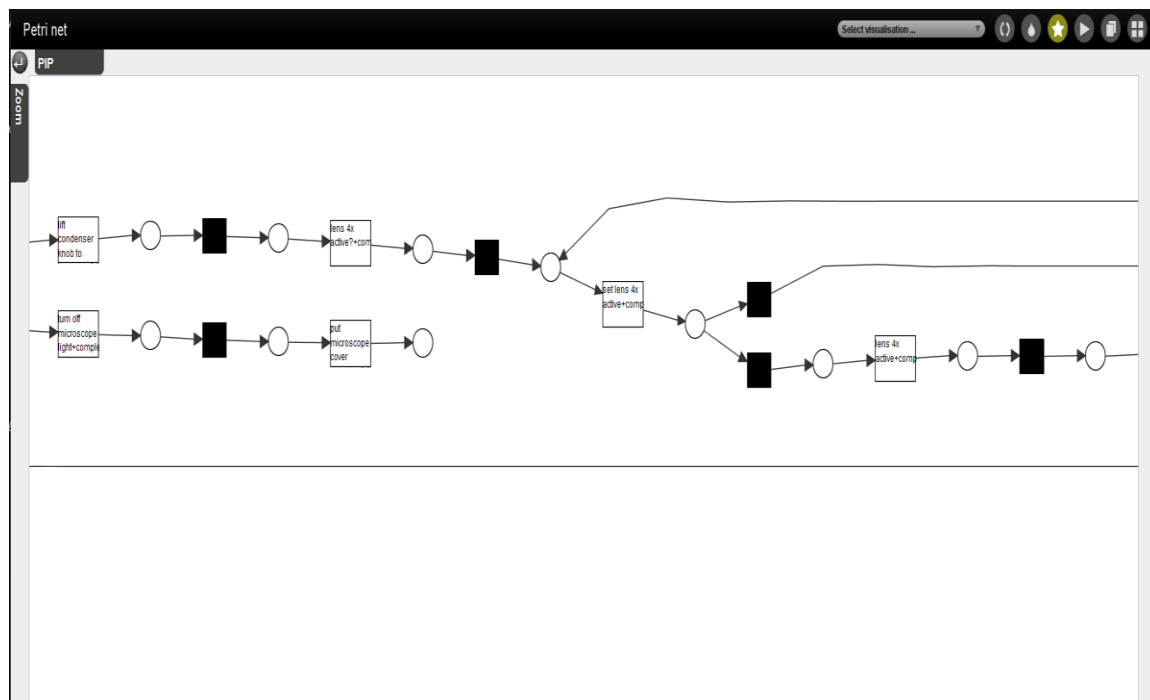
Εικόνα 66: Μερική απεικόνιση του παραγόμενου ευρετικού δικτύου

Στην παρακάτω εικόνα (**Εικόνα 67**) παρουσιάζεται η διαδικασία μετασχηματισμού του ανακαλυφθέντος μοντέλου σε μια τυποποιημένη μαθηματική δομή, κατάλληλη για τον έλεγχο συμμόρφωσης. Εφόσον παρήχθη το αρχικό ευρετικό δίκτυο, επιλέχθηκε η ενέργεια «Convert Heuristics net into Petri net» μέσα από το περιβάλλον του ProM. Στο αριστερό τμήμα ορίζεται ως είσοδος (Input) το Heuristics Net, ενώ στη δεξιά πλευρά προσδιορίζονται τα αναμενόμενα αντικείμενα εξόδου (Output), τα οποία περιλαμβάνουν το δομικό δίκτυο Petri (Petri net) και την αρχική του αρχιτεκτονική κατάσταση (Marking). Η επιλογή αυτή είναι απαραίτητη, καθώς επιτρέπει τη μετάβαση από ένα μοντέλο συχνοτήτων σε ένα αυστηρό μοντέλο καταστάσεων και μεταβάσεων.



Εικόνα 67: Μετατροπή του ευρετικού δικτύου σε δίκτυο Petri

Στην παρακάτω εικόνα (**Εικόνα 68**) παρουσιάζεται ένα εστιασμένο τμήμα του τελικού δικτύου Petri που προέκυψε από τη διαδικασία της ανακάλυψης. Λόγω του μεγάλου μεγέθους και της πολυπλοκότητας της συνολικής εργαστηριακής διεργασίας, αποτυπώνεται ένα μέρος της δομής του, αναδεικνύοντας τις θέσεις (places) ως κύκλους και τις μεταβάσεις (transitions) ως ορθογώνια πλαίσια. Στο συγκεκριμένο στιγμιότυπο διακρίνονται οι διασυνδέσεις ενεργειών, όπως η ρύθμιση του πυκνωτή και η ενεργοποίηση του αντικειμενικού φακού, προσφέροντας μια καθαρή εικόνα της μαθηματικής λογικής και των συνθηκών συγχρονισμού που διέπουν την πραγματική ροή.



Εικόνα 68: Μερική απεικόνιση του παραγόμενου δικτύου Petri

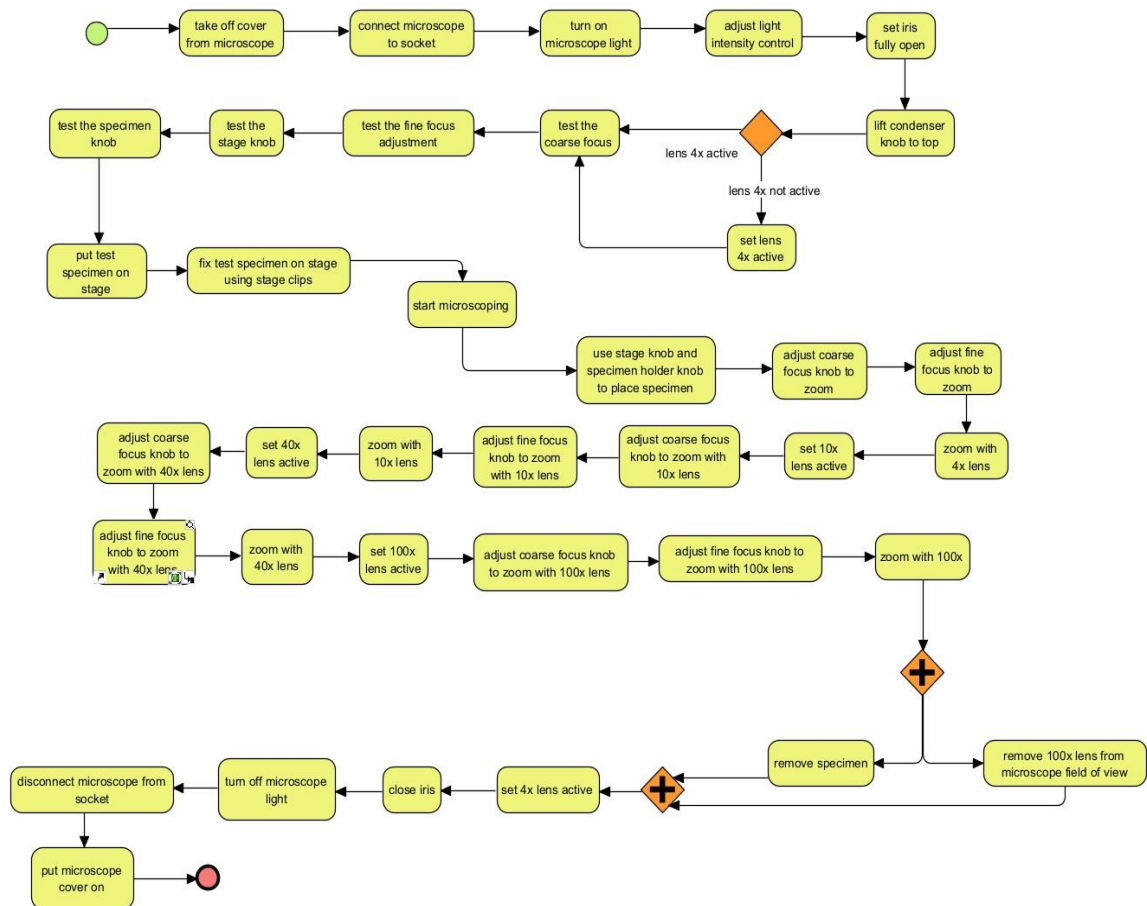
5.6 Σχεδιασμός του μοντέλου αναφοράς (Reference Model)

Η συγκεκριμένη ενότητα παρουσιάζει αναλυτικά τη διαδικασία σχεδιασμού του μοντέλου αναφοράς (Reference Model) για το πείραμα «Μικροσκοπική Παρατήρηση Δείγματος». Το μοντέλο αναφοράς αποτελεί τη θεωρητική βάση και την ιδανική αποτύπωση των βημάτων που πρέπει να ακολουθηθούν, λειτουργώντας ως το πρότυπο σύγκρισης για τη μετέπειτα φάση της ανάλυσης. Για την ολοκληρωμένη ανάπτυξή του, η μεθοδολογία μου βασίστηκε σε δύο διακριτά στάδια ψηφιακής μοντελοποίησης, με σκοπό τη μετατροπή των θεωρητικών κανόνων του εργαστηρίου σε μια αυστηρά δομημένη γραφική και μαθηματική αναπαράσταση. Αρχικά, οι κανόνες και η ροή της εργαστηριακής άσκησης σχεδιάστηκαν στο εξειδικευμένο λογισμικό Visual Paradigm, όπου αποτυπώθηκαν σε μορφή BPMN (Business Process Model and Notation). Αυτή η γραφική αναπαράσταση προσέφερε την απαραίτητη οπτική σαφήνεια, ορίζοντας με ακρίβεια τη σειρά των ενεργειών, τις πύλες αποφάσεων και τις επιτρεπτές επαναλήψεις (loops) κατά τη χρήση του μικροσκοπίου.

Στη συνέχεια, η ροή εργασίας πέρασε στο δεύτερο στάδιο επεξεργασίας, προκειμένου το μοντέλο να καταστεί αξιοποιήσιμο από τα εργαλεία εξόρυξης διεργασιών. Με τη χρήση του

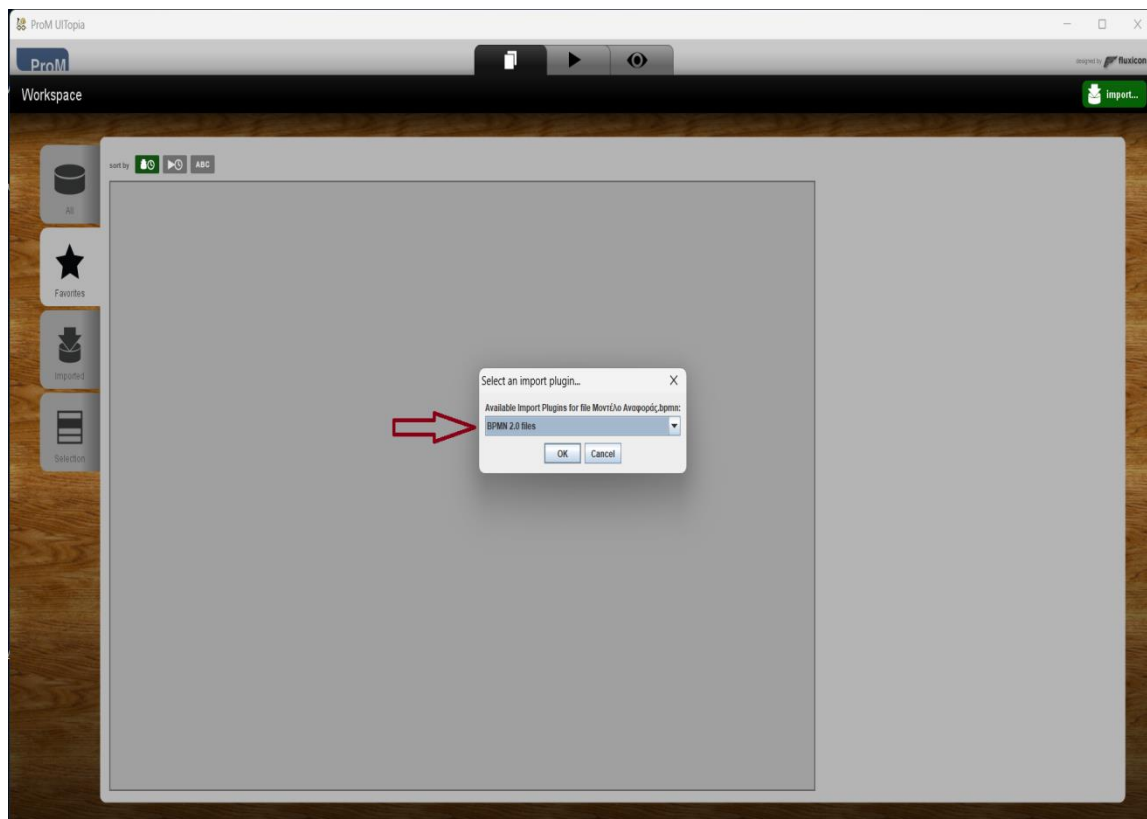
ακαδημαϊκού λογισμικού ProM, το αρχικό μοντέλο BPMN μετασχηματίστηκε και μετατράπηκε σε δίκτυο Petri (Petri Net), παίρνοντας την ψηφιακή μορφή αρχείου με κατάληξη .pnml (Petri Net Markup Language). Η μορφή αυτή είναι η πλέον κατάλληλη και διεθνώς αναγνωρισμένη για την εκτέλεση του ελέγχου συμμόρφωσης (conformance checking), καθώς επιτρέπει την αυτοματοποιημένη σύγκριση της ιδανικής διεργασίας με τις πραγματικές καταγραφές του χρήστη. Παρακάτω, αναλύεται βήμα-βήμα αυτή η σχεδιαστική διαδρομή, αποτυπώνοντας την εξέλιξη του μοντέλου μέχρι την τελική του μορφή, η οποία παρουσιάζεται λεπτομερώς μέσα από τα κατάλληλα στιγμιότυπα οθόνης (screenshots) των λογισμικών που χρησιμοποιήθηκαν.

Στην παρακάτω εικόνα (**Εικόνα 69**) απεικονίζεται η πλήρης ροή της εργαστηριακής άσκησης, όπως αυτή σχεδιάστηκε αρχικά στο λογισμικό Visual Paradigm. Η διαδικασία ξεκινά με την προετοιμασία του μικροσκοπίου, προχωρά στη σταδιακή μεγέθυνση του δείγματος με τη χρήση των φακών και ολοκληρώνεται με την απενεργοποίηση της συσκευής. Οι ενσωματωμένες πύλες ελέγχου καθορίζουν με ακρίβεια τη λογική σειρά των βημάτων και τις επιτρεπτές επαναλήψεις, προσφέροντας έτσι μια ολοκληρωμένη και έγκυρη βάση για τη μετέπειτα διαδικασία του ελέγχου συμμόρφωσης.



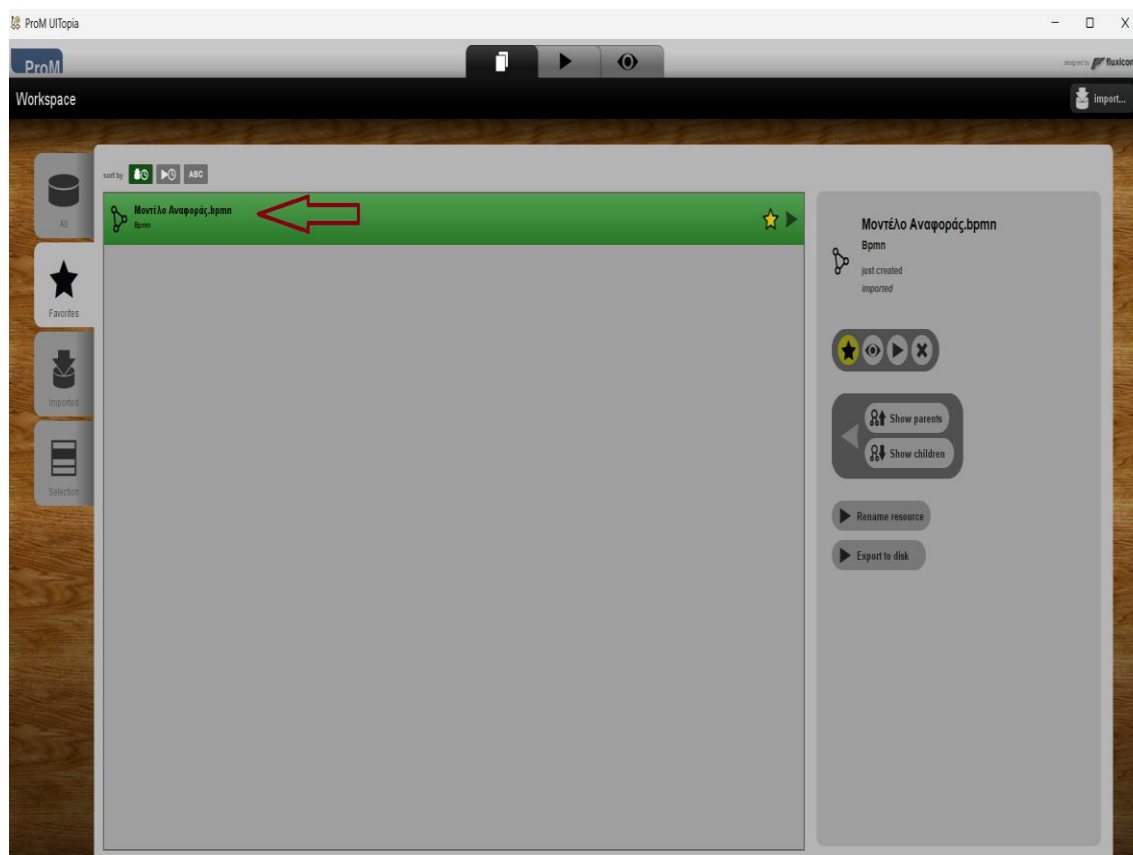
Εικόνα 69: Ψηφιακή μοντελοποίηση της διεργασίας με χρήση του προτύπου BPMN

Στην παρακάτω εικόνα (**Εικόνα 70**) αποτυπώνεται το πρώτο βήμα της διαδικασίας μετασχηματισμού, το οποίο αφορά την εισαγωγή του αρχικού διαγράμματος στο περιβάλλον εργασίας του ProM. Κατά τη φάση αυτή, το σύστημα αναγνωρίζει αυτόματα τη δομή των δεδομένων και ζητά την επιλογή του κατάλληλου προσθέτου (plugin) για την ανάγνωσή τους. Η επιλογή της ρύθμισης για αρχεία τύπου BPMN 2.0 εξασφαλίζει την ορθή φόρτωση της ροής και προετοιμάζει το λογισμικό για την επακόλουθη μετατροπή της διεργασίας σε δίκτυο Petri.



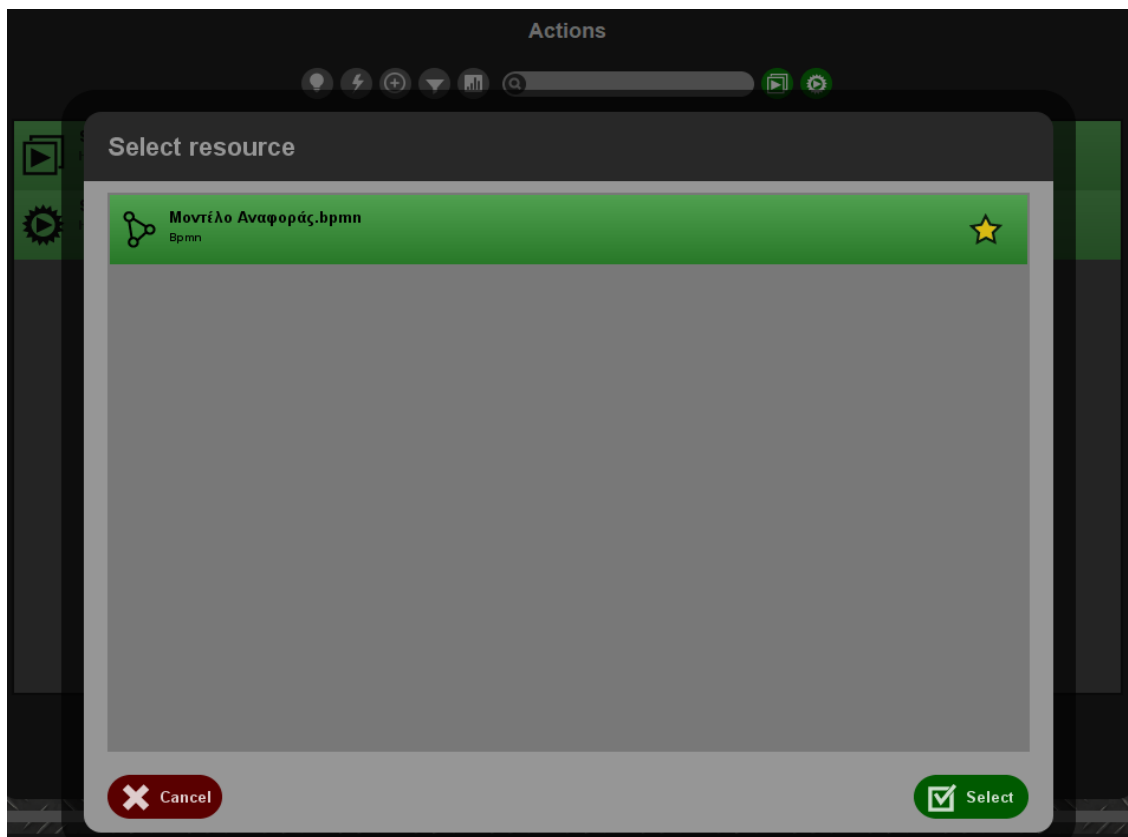
Εικόνα 70: Εισαγωγή του αρχείου BPMN στο λογισμικό ProM

Στην παρακάτω εικόνα (**Εικόνα 71**) παρουσιάζεται η ολοκλήρωση της διαδικασίας εισαγωγής, όπου το ψηφιακό μοντέλο εμφανίζεται πλέον ως ενεργός πόρος (resource) μέσα στο χώρο εργασίας (workspace) του λογισμικού ProM. Στη δεξιά πλευρά της οθόνης παρέχονται οι βασικές μεταπληροφορίες και οι διαθέσιμες επιλογές διαχείρισης του αρχείου, επιβεβαιώνοντας ότι η δομή αναγνωρίστηκε σωστά από το σύστημα. Το στάδιο αυτό είναι απαραίτητο, καθώς ορίζει το σημείο εκκίνησης για την εφαρμογή των αλγορίθμων μετασχηματισμού της διεργασίας.



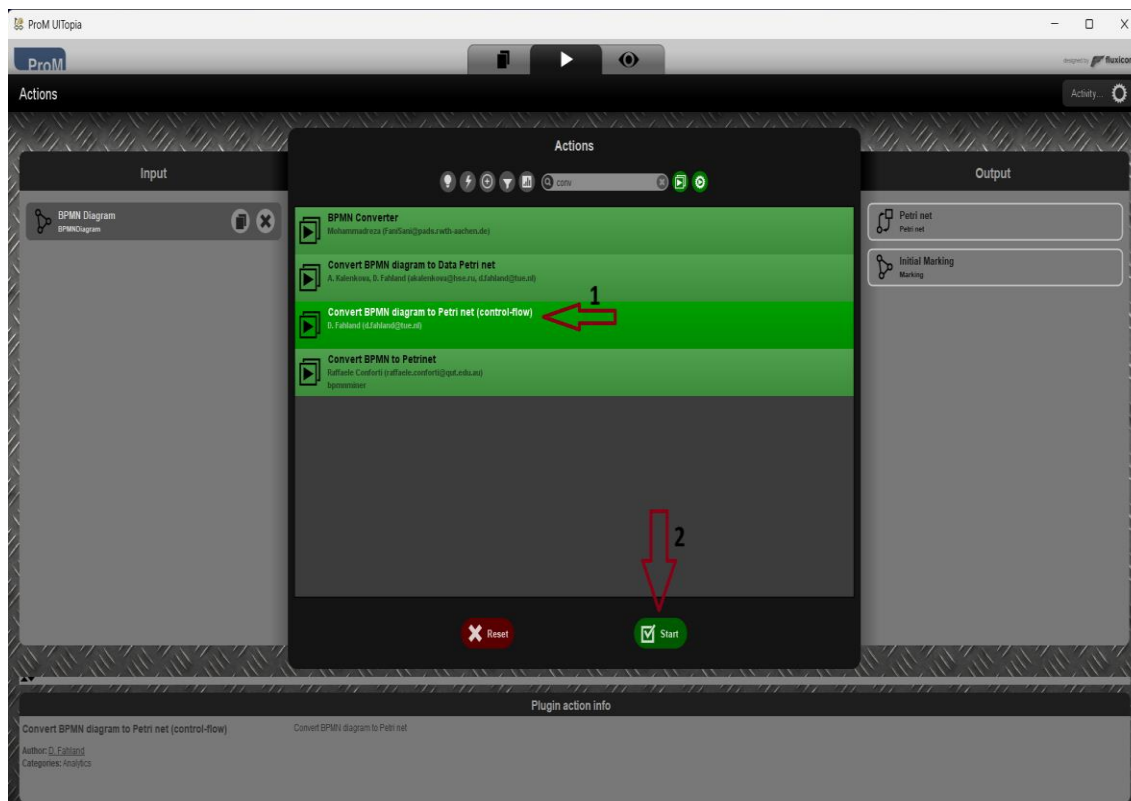
Εικόνα 71: Επιτυχής επιβεβαίωση εισαγωγής του μοντέλου στο περιβάλλον εργασίας

Στην παρακάτω εικόνα (**Εικόνα 72**) απεικονίζεται το παράθυρο επιλογής πόρου (Select resource) κατά την εκκίνηση μιας νέας ενέργειας στο λογισμικό ProM. Ο χρήστης καλείται να προσδιορίσει το συγκεκριμένο αρχείο εισόδου πάνω στο οποίο θα εφαρμοστεί η διαδικασία του μετασχηματισμού. Η επιλογή του εισαχθέντος μοντέλου αναφοράς αποτελεί το απαραίτητο βήμα για τη σύνδεση των δεδομένων με τον αλγόριθμο μετατροπής της διεργασίας σε δίκτυο Petri.



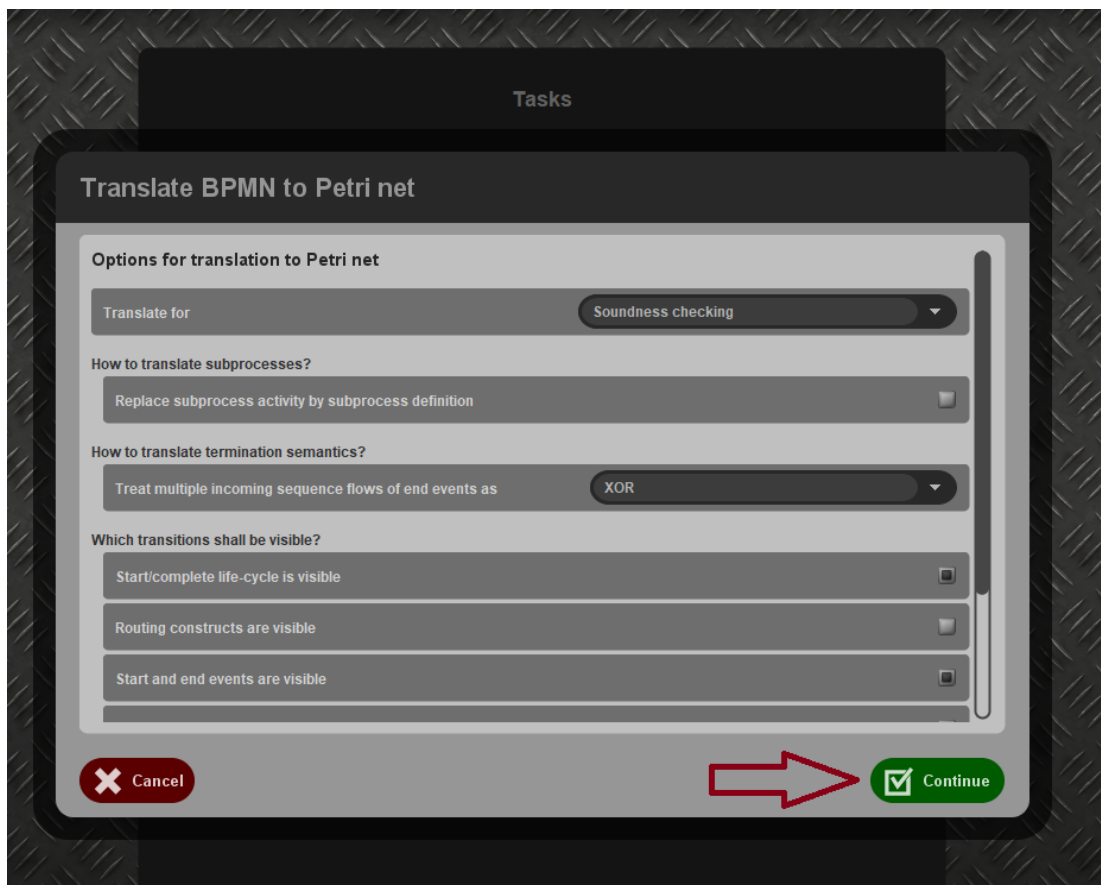
Εικόνα 72: Επιλογή του μοντέλου αναφοράς για την εκτέλεση ενέργειας

Στην παρακάτω εικόνα (**Εικόνα 73**) παρουσιάζεται η διαδικασία επιλογής του κατάλληλου αλγορίθμου για τον μετασχηματισμό της διεργασίας. Για την υλοποίηση αυτού του βήματος, επιλέχθηκε το εξειδικευμένο πρόσθετο «Convert BPMN diagram to Petri net (control-flow)», το οποίο αναλαμβάνει να μεταφράσει τη δομή του διαγράμματος ροής σε μαθηματικό δίκτυο Petri. Με την ενεργοποίηση της συγκεκριμένης επιλογής, το λογισμικό ProM προετοιμάζει την εξαγωγή των δύο βασικών συνιστωσών του μοντέλου, δηλαδή του ίδιου του δικτύου (Petri net) και της αρχικής του κατάστασης (Initial Marking).



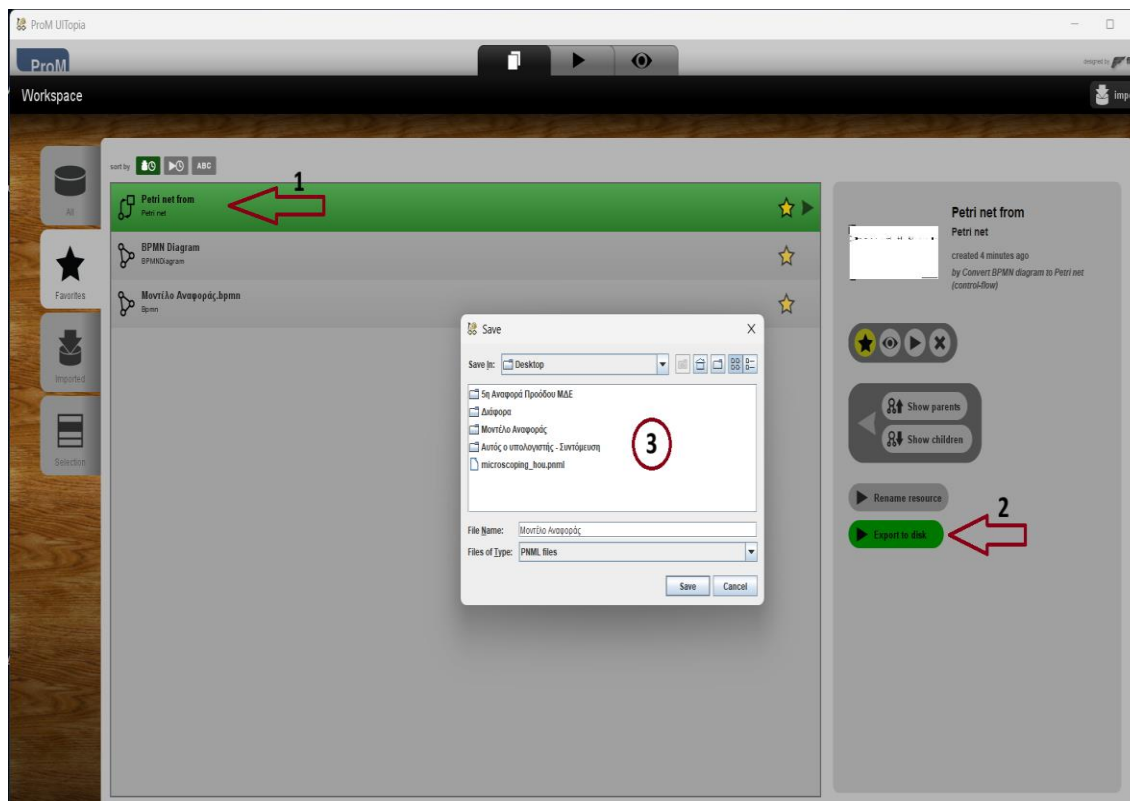
Εικόνα 73: Επιλογή του αλγορίθμου μετατροπής σε δίκτυο Petri

Στην παρακάτω εικόνα (**Εικόνα 74**), φαίνεται το παράθυρο ρυθμίσεων για τη μετατροπή του μοντέλου, όπου καθορίζονται οι τεχνικές παράμετροι της μετάφρασης από BPMN σε δίκτυο Petri. Στη συγκεκριμένη φάση, διατηρήθηκε η προεπιλεγμένη ρύθμιση για τον έλεγχο αρτιότητας (Soundness checking), η οποία διασφαλίζει ότι το παραγόμενο δίκτυο δεν θα παρουσιάζει δομικά σφάλματα, όπως αδιέξοδα (dead locks). Με την επιλογή του πλήκτρου συνέχειας (Continue), ορίζεται η οριστική έναρξη του μαθηματικού μετασχηματισμού της ροής με βάση τα επιλεγμένα κριτήρια.



Εικόνα 74: Παραμετροποίηση των ιδιοτήτων μετάφρασης της διεργασίας

Στην παρακάτω εικόνα (**Εικόνα 75**) αποτυπώνεται η ολοκλήρωση του μετασχηματισμού και η διαδικασία τοπικής αποθήκευσης του νέου μοντέλου. Μετά τη δημιουργία του δικτύου Petri στο ProM, χρησιμοποιήθηκε η επιλογή «Export to disk» με σκοπό την εξαγωγή της διεργασίας σε αρχείο κατάλληλου τύπου. Όπως φαίνεται στο αναδυόμενο παράθυρο, επιλέχθηκε η μορφή αρχείου PNML (Petri Net Markup Language), εξασφαλίζοντας έτσι ότι το μοντέλο αναφοράς είναι πλέον αποθηκευμένο σε μια τυποποιημένη ψηφιακή δομή, έτοιμη να χρησιμοποιηθεί στη φάση του ελέγχου συμμόρφωσης.



Εικόνα 75: Εξαγωγή και αποθήκευση του παραγόμενου δικτύου Petri

5.7 Έλεγχος συμμόρφωσης (Conformance Checking) με χρήση της βιβλιοθήκης PM4Py

Μετά την επιτυχή ολοκλήρωση της δημιουργίας του μοντέλου αναφοράς (reference model) και την τελική διαμόρφωση του δομημένου αρχείου καταγραφής συμβάντων (event log), η παρούσα ενότητα εστιάζει στην εκτέλεση της διαδικασίας του ελέγχου συμμόρφωσης (conformance checking). Η συγκεκριμένη φάση αποτελεί έναν από τους κεντρικότερους πυλώνες της εξόρυξης διεργασιών (Process Mining), καθώς επιτρέπει την απευθείας σύγκριση και τον συσχετισμό μεταξύ της προδιαγεγραμμένης ροής της διεργασίας και της πραγματικής συμπεριφοράς, όπως αυτή καταγράφηκε αυτόματα από το σύστημα OnLabs κατά την πειραματική διαδικασία. Για την υλοποίηση αυτού του ελέγχου αξιοποιήθηκε η εξειδικευμένη βιβλιοθήκη PM4Py της γλώσσας προγραμματισμού Python, η οποία παρέχει προηγμένους και διεθνώς αναγνωρισμένους αλγορίθμους για την αποτίμηση της συμβατότητας μεταξύ δεδομένων και μοντέλων.

Η διαδικασία βασίστηκε στην ταυτόχρονη ανάλυση του τελικού ιστορικού συμβάντων και του τυποποιημένου μοντέλου αναφοράς, προκειμένου να εντοπιστούν με ακρίβεια τυχόν αποκλίσεις, παραλείψεις ενεργειών ή εσφαλμένες αλληλουχίες βημάτων κατά την εκτέλεση των εργαστηριακών ασκήσεων. Με βάση τη συγκριτική αξιολόγηση αλγορίθμων που πραγματοποιήθηκε στην ενότητα 1.1.5.1 της παρούσας διπλωματικής εργασίας, επιλέχθηκε ο αλγόριθμος Token-Based Replay για την πραγματοποίηση του ελέγχου συμμόρφωσης. Η εφαρμογή της συγκεκριμένης τεχνικής μέσω της βιβλιοθήκης PM4Py επέτρεψε την ποσοτική και ποιοτική αξιολόγηση της ροής εργασιών μέσω της αναπαραγωγής των ιχνών. Αυτή η διαδικασία σύγκρισης κατέστησε εφικτό τον προσδιορισμό του βαθμού στον οποίο οι πραγματοποιηθείσες εκτελέσεις ακολούθησαν τις οδηγίες, αναδεικνύοντας τα σημεία εκείνα όπου η πραγματική ροή διαφοροποιήθηκε από τον προβλεπόμενο σχεδιασμό.

Από την ολοκλήρωση του ελέγχου συμμόρφωσης εξήχθησαν ιδιαίτερα χρήσιμα και κρίσιμα συμπεράσματα σχετικά με τη δυναμική των εκτελέσεων και τη λειτουργικότητα της εκπαιδευτικής πλατφόρμας. Τα αποτελέσματα προσέφεραν μια ξεκάθαρη αποτύπωση των συχνότερων λαθών, των παρακάμψεων στη ροή της διαδικασίας, αλλά και των κοινών μοτίβων συμπεριφοράς που αναπτύχθηκαν κατά τη διάρκεια των δοκιμών. Τα συμπεράσματα αυτά δεν περιορίζονται μόνο στον εντοπισμό των αποκλίσεων, αλλά αποτελούν τη βάση για την κατανόηση των αιτιών που τις προκάλεσαν, προσφέροντας πολύτιμη ανατροφοδότηση για τη μελλοντική βελτίωση του σχεδιασμού των ασκήσεων και την ενίσχυση της εκπαιδευτικής διαδικασίας.

Στο σημείο αυτό παρουσιάζεται το βασικό απόσπασμα του κώδικα προγραμματισμού σε γλώσσα Python (**Απόσπασμα κώδικα 1**), ο οποίος αναπτύχθηκε και εκτελέστηκε με τη χρήση της εξειδικευμένης βιβλιοθήκης PM4Py, προκειμένου να πραγματοποιηθεί ο έλεγχος συμμόρφωσης των δεδομένων. Το συγκεκριμένο υπολογιστικό σενάριο (script) αποτελεί τον κεντρικό πυρήνα του πρακτικού μέρους, καθώς αναλαμβάνει να συνδέσει τις θεωρητικές προδιαγραφές του μοντέλου με τις πραγματικές εγγραφές του ιστορικού συμβάντων, φορτώνοντας τα δεδομένα, εκτελώντας τον αλγόριθμο Token-Based Replay και υπολογίζοντας τις σχετικές μετρικές. Ο πλήρης εκτελέσιμος κώδικας παρατίθεται αναλυτικά στο **Παράρτημα Α**. Στην συνέχεια, ακολουθεί μια αναλυτική περιγραφή της δομής και των επιμέρους τμημάτων του κώδικα, επεξηγώντας βήμα προς βήμα τις διαδικασίες φόρτωσης, την αρχιτεκτονική του αλγορίθμου Token-Based Replay, καθώς και τον τρόπο με τον οποίο

εξάγονται και μορφοποιούνται οι τελικές αριθμητικές μετρικές αξιολόγησης.

1. Φόρτωση Event Log και Petri Net

```
df = pd.read_csv(final_csv_path)
df["time:timestamp"] = pd.to_datetime(df["time:timestamp"])
log = log_converter.apply(df)
net, initial_marking, final_marking = pnml_importer.apply(pnml_path)
```

2. Εκτέλεση Token-Based Replay & Υπολογισμός Fitness

```
replay_results = fitness_evaluator.apply(
    log, net, initial_marking, final_marking,
    variant=fitness_evaluator.Variants.TOKEN_BASED
)
```

3. Υπολογισμός Tokens για Precision & Generalization

```
pure_replay = token_replay.apply(log, net, initial_marking, final_marking)
```

4. Υπολογισμός Δομικής Απλότητας (Simplicity)

```
simplicity_val = simplicity_evaluator.apply(net)
```

Απόσπασμα κώδικα 1

Ο κώδικας που αναπτύχθηκε αποτελεί το λειτουργικό υπόβαθρο για τη διεξαγωγή του ελέγχου συμμόρφωσης και βασίζεται εξολοκλήρου στις δυνατότητες της βιβλιοθήκης PM4Py. Σκοπός του προγράμματος είναι η αυτοματοποιημένη διαχείριση των δεδομένων, η συστηματική αντιπαραβολή τους με το μοντέλο αναφοράς και ο ακριβής υπολογισμός των τεσσάρων βασικών μετρικών ποιότητας που ορίζει η επιστήμη της εξόρυξης διεργασιών, δηλαδή της καταλληλότητας, της ακρίβειας, της γενίκευσης και της απλότητας. Η αρχιτεκτονική του σεναρίου κώδικα ακολουθεί μια αυστηρά σειριακή δομή εννέα φάσεων, μετατρέποντας σταδιακά τα πρωτογενή δεδομένα σε μια ολοκληρωμένη ποσοτική αναφορά, έτοιμη προς επιστημονική αξιολόγηση.

Η ροή του κώδικα εκκινεί με την εισαγωγή των απαραίτητων προγραμματιστικών

εργαλείων και τον καθορισμό των διαδρομών συστήματος για την άντληση του ιστορικού συμβάντων και του δικτύου Petri. Στη συνέχεια, το προεπεξεργασμένο αρχείο καταγραφής εισάγεται στη μνήμη και η στήλη των χρονικών σημάνσεων τυποποιείται κατάλληλα, ώστε να εξαλειφθούν τυχόν ασυμβατότητες κατά την ανάγνωση των ωρών και των ημερομηνιών. Αυτό το βήμα επιτρέπει στον ειδικό μετατροπέα της βιβλιοθήκης να μετασχηματίσει τη στατική δομή του αρχικού πίνακα σε ένα δυναμικό, αντικειμενοστρεφές ιστορικό συμβάντων, το οποίο είναι πλέον πλήρως αναγνωρίσιμο από τους αλγορίθμους προσομοίωσης. Παράλληλα, το μοντέλο αναφοράς ανακτάται και αναλύεται συντακτικά, διαχωρίζοντας τη μαθηματική οντότητα του δικτύου στη γραφική του δομή, στην αρχική κατάσταση έναρξης της διεργασίας και στην επιθυμητή τελική κατάσταση τερματισμού.

Πριν από την έναρξη του κυρίως ελέγχου, ο κώδικας ενεργοποιεί μια εξωτερική μηχανή σχεδίασης, η οποία αναλαμβάνει να οπτικοποιήσει τη δομή του δικτύου Petri, προβάλλοντάς το άμεσα στην οθόνη και αποθηκεύοντάς το ταυτόχρονα ως αυτόνομο αρχείο εικόνας, γεγονός που διασφαλίζει την οπτική επαλήθευση του μοντέλου. Αμέσως μετά, ξεκινά η κεντρική υπολογιστική διαδικασία, όπου εφαρμόζεται ο αλγόριθμος αναπαραγωγής tokens. Ο αλγόριθμος αυτός προσομοιώνει την πραγματική ροή των καταγεγραμμένων ενεργειών πάνω στο μοντέλο αναφοράς, προκειμένου να αποτιμηθεί άμεσα ο βαθμός της καταλληλότητας. Για να επιτευχθεί όμως μια βαθύτερη και πιο αναλυτική προσέγγιση, ο κώδικας δεν περιορίζεται στις έτοιμες συναρτήσεις του συστήματος, αλλά εκτελεί έναν επαναληπτικό έλεγχο σε κάθε καταγεγραμμένο ίχνος ξεχωριστά, εισχωρώντας στο χαμηλότερο επίπεδο πληροφορίας των tokens.

Κατά τη διάρκεια αυτής της αναλυτικής προσπάθειας, ο κώδικας καταμετρά αθροιστικά τέσσερα θεμελιώδη μεγέθη που αφορούν τα tokens του δικτύου, δηλαδή εκείνα που καταναλώθηκαν για την ενεργοποίηση των βημάτων, όσα παρήχθησαν κατά τη ροή, εκείνα που έλειπαν λόγω παρακάμψεων και όσα τελικά περίσρεψαν στο σύστημα μετά τον τερματισμό. Με βάση αυτά τα αθροίσματα, το πρόγραμμα εφαρμόζει τους επίσημους μαθηματικούς τύπους της θεωρίας του Process Mining για να υπολογίσει δυναμικά την ακρίβεια και την ικανότητα γενίκευσης του μοντέλου, εισάγοντας παράλληλα ειδικές δομές ελέγχου για την αποφυγή υπολογιστικών σφαλμάτων. Στο τελικό στάδιο, η δομική απλότητα του δικτύου αποτιμάται αυτόματα και όλα τα τελικά αποτελέσματα, μαζί με τα αναλυτικά στατιστικά των tokens, μορφοποιούνται με ακρίβεια τεσσάρων δεκαδικών

ψηφίων και τυπώνονται συγκεντρωτικά στην κονσόλα εξόδου.

Η ολοκλήρωση του ελέγχου συμμόρφωσης με τον αλγόριθμο Token-Based Replay παρέχει μια αντικειμενική εικόνα, τόσο σε ποσοτικό όσο και σε ποιοτικό επίπεδο, για το πόσο κοντά βρίσκεται το θεωρητικό Μοντέλο Αναφοράς στην πραγματική ροή του πειράματος. Από τη μία πλευρά, το Μοντέλο Αναφοράς σχεδιάστηκε ως δίκτυο Petri και ενσωματώνει όλες τις αυστηρές προδιαγραφές, τους περιορισμούς και τα σωστά βήματα για τις εργαστηριακές ασκήσεις μικροσκοπίας στην πλατφόρμα OnLabs. Από την άλλη πλευρά, το αρχείο καταγραφής (log file) αποτυπώνει τις πραγματικές ενέργειες που εκτελέστηκαν κατά τη διάρκεια του πειράματος. Η σύγκριση αυτών των δύο στοιχείων έγινε μέσω της βιβλιοθήκης PM4Py και απέδωσε μια σειρά από μετρικές και στατιστικά δεδομένα. Τα στοιχεία αυτά αναλύονται παρακάτω, με σκοπό να αξιολογηθεί η ροή της διαδικασίας και να προσδιοριστεί η ακριβής φύση των αποκλίσεων.

Πριν από την παρουσίαση και την ερμηνεία των πειραματικών αποτελεσμάτων, είναι απαραίτητο να οριστούν οι τέσσερις βασικές μετρικές ποιότητας που διέπουν τον έλεγχο συμμόρφωσης (Conformance Checking) στην Εξόρυξη Διεργασιών (Process Mining), καθώς και ο μαθηματικός τρόπος υπολογισμού τους μέσω του αλγορίθμου Token-Based Replay.

Η πρώτη και θεμελιώδης μετρική είναι η **Καταλληλότητα (Fitness)**, η οποία μετρά το ποσοστό της συμπεριφοράς που είναι καταγεγραμμένο στο αρχείο συμβάντων (event log) και μπορεί να αναπαραχθεί πιστά από το Μοντέλο Αναφοράς. Στο πλαίσιο του Token-Based Replay, η καταλληλότητα αξιολογεί την ικανότητα του δικτύου Petri να «καταναλώσει» τα συμβάντα του ίχνους χωρίς να δημιουργούνται ελλείποντα ή ανεπιθύμητα εναπομείναντα tokens. Ο μαθηματικός υπολογισμός της βασίζεται στον τύπο:

$$\text{Fitness} = \frac{1}{2} \cdot \left(1 - \frac{m}{c}\right) + \frac{1}{2} \cdot \left(1 - \frac{r}{p}\right)$$

όπου **m** συμβολίζει τα ελλείποντα tokens (missing) που χρειάστηκε να δημιουργηθούν τεχνητά για την εκτέλεση μιας μετάβασης, **c** τα συνολικά καταναλωθέντα tokens (consumed), **r** τα tokens που παρέμειναν εγκλωβισμένα στο δίκτυο μετά τον τερματισμό του ίχνους (remaining) και **p** τα συνολικά παραχθέντα tokens (produced).

Η δεύτερη μετρική, η **Ακρίβεια (Precision)**, αξιολογεί αν το Μοντέλο Αναφοράς είναι

υπερβολικά «χαλαρό», δηλαδή αν επιτρέπει θεωρητικές συμπεριφορές και ροές που απέχουν σημαντικά από τις πραγματικές καταγραφές του αρχείου συμβάντων. Για τον υπολογισμό της αξιοποιείται η επίσημη φόρμουλα De Medeiros, η οποία εκφράζει τον λόγο των tokens που καταναλώθηκαν ορθώς προς το άθροισμα των καταναλωθέντων και εναπομεινάντων tokens:

$$\text{Precision} = \frac{c}{c + r}$$

όπου c είναι τα καταναλωθέντα tokens και r τα εναπομεινάντα tokens στο δίκτυο Petri.

Η τρίτη μετρική είναι η **Γενίκευση (Generalization)**, η οποία εκτιμά την ικανότητα του Μοντέλου Αναφοράς να καλύπτει μελλοντικές, μη καταγεγραμμένες εκτελέσεις της διεργασίας χωρίς να οδηγεί σε υπερπροσαρμογή (overfitting) στα ήδη υπάρχοντα δεδομένα. Υπολογίζεται από τη σχέση των ελλειπόντων tokens ως προς τα καταναλωθέντα tokens:

$$\text{Generalization} = 1 - \frac{m}{c}$$

όπου m είναι τα ελλείποντα tokens και c τα καταναλωθέντα tokens.

Τέλος, η **Απλότητα (Simplicity)** αποτιμά τη δομική πολυπλοκότητα του ίδιου του Μοντέλου Αναφοράς (Petri net), εξετάζοντας αν το μοντέλο είναι ευανάγνωστο, συμπαγές και απαλλαγμένο από περιττές μεταβάσεις ή σύνθετες δομές. Η μετρική υπολογίζεται ως ο λόγος του συνολικού πλήθους των κόμβων προς τις ακμές του δικτύου:

$$\text{Simplicity} = \frac{|P| + |T|}{|A|}$$

όπου $|P|$ παριστάνει το πλήθος των θέσεων (places), $|T|$ το πλήθος των μεταβάσεων (transitions) και $|A|$ το συνολικό πλήθος των κατευθυνόμενων συνδέσεων/ακμών (arcs) μεταξύ θέσεων και μεταβάσεων. Όλες οι παραπάνω μετρικές λαμβάνουν τιμές στο κλειστό διάστημα $[0, 1]$, με την τιμή 1.0 να αντιπροσωπεύει την βέλτιστη επίδοση.

Η πρώτη μετρική, η **καταλληλότητα (Fitness)**, διαμορφώθηκε στην τιμή **0.5076**. Το αποτέλεσμα αυτό υποδεικνύει ότι περίπου οι μισές από τις ενέργειες και τις αλληλουχίες που καταγράφηκαν στο πείραμα εμφάνισαν δομικές αποκλίσεις από την προδιαγεγραμμένη ροή του Μοντέλου Αναφοράς. Το εύρημα αυτό φανερώνει σημαντικές διαφορές, οι οποίες οφείλονται είτε στο ότι παραλείφθηκαν υποχρεωτικά βήματα, είτε στο ότι κάποιες ενέργειες

πραγματοποιήθηκαν με διαφορετική σειρά, είτε στις επαναλαμβανόμενες προσπάθειες σταθεροποίησης και εστίασης του φακού στο μικροσκόπιο.

Η μετρική της **ακρίβειας (Precision)** διαμορφώθηκε στο **0.6800**. Η τιμή αυτή υποδηλώνει ότι το Μοντέλο Αναφοράς είναι σωστά δομημένο και περιορίζει τις αυθαίρετες ροές, αν και αφήνει ορισμένα περιθώρια ελευθερίας ή εναλλακτικών διαδρομών που δεν ενεργοποιήθηκαν στη συγκεκριμένη πειραματική διαδικασία. Το αποτέλεσμα αυτό επιβεβαιώνει ότι το δίκτυο Petri λειτουργεί ικανοποιητικά ως ρυθμιστικό πλαίσιο, αποφεύγοντας την υιοθέτηση υπερβολικά χαλαρών κανόνων.

Η μετρική της **γενίκευσης (Generalization)** σημείωσε την τιμή **0.5000**, η οποία όμως παρουσιάζεται μόνο ενδεικτικά και με σαφή επιφύλαξη. Επειδή το αρχείο καταγραφής περιλαμβάνει μία μόνο πειραματική προσπάθεια (Case_1), η μετρική αυτή δεν μπορεί να ερμηνευτεί με ασφάλεια. Η τιμή 0.5000 αποτελεί απλώς ένα ουδέτερο υπολογιστικό αποτέλεσμα του αλγορίθμου, καθώς απουσιάζουν επιπλέον δεδομένα από άλλες εκτελέσεις ή χρήστες που θα επέτρεπαν να αξιολογηθεί αν το Μοντέλο Αναφοράς μπορεί να καλύψει μελλοντικές συμπεριφορές. Συνεπώς, από το συγκεκριμένο εύρημα δεν μπορούν να εξαχθούν βέβαια συμπεράσματα για τη γενίκευση του μοντέλου.

Τέλος, η μετρική της **απλότητας (Simplicity)** κατέγραψε την εξαιρετικά υψηλή τιμή **0.9789**. Η επίδοση αυτή αποδεικνύει ότι το Μοντέλο Αναφοράς σχεδιάστηκε με υποδειγματική σαφήνεια και δομική οικονομία. Παρά τις αποκλίσεις που εντοπίστηκαν στα δεδομένα εκτέλεσης, το ίδιο το μοντέλο παραμένει μαθηματικά απλό, άρτιο, εύκολα αναγνώσιμο και απαλλαγμένο από περιττές δομές που θα δυσκόλευαν την ανάλυση.

Για να κατανοηθούν βαθύτερα αυτοί οι δείκτες, εξετάστηκαν τα πρωτογενή στατιστικά στοιχεία από την κίνηση των tokens. Καταγράφηκαν συνολικά 34 καταναλωθέντα tokens (Consumed Tokens) και 33 παραχθέντα tokens (Produced Tokens). Οι αριθμοί αυτοί δείχνουν τον συνολικό όγκο της υπολογιστικής κίνησης στο δίκτυο Petri κατά την προσομοίωση. Η σχετική ισορροπία μεταξύ παραγωγής και κατανάλωσης φανερώνει ότι το μοντέλο κατάφερε να καθοδηγήσει τη ροή από την αρχή μέχρι το τέλος, ενεργοποιώντας τις αντίστοιχες μεταβάσεις για κάθε γεγονός που ήταν συμβατό με τους κανόνες του.

Ωστόσο, η κύρια αιτία της χαμηλής καταλληλότητας (Fitness) εντοπίζεται στα 17 ελλείποντα tokens (Missing Tokens). Στα δίκτυα Petri, ένα token χαρακτηρίζεται ως

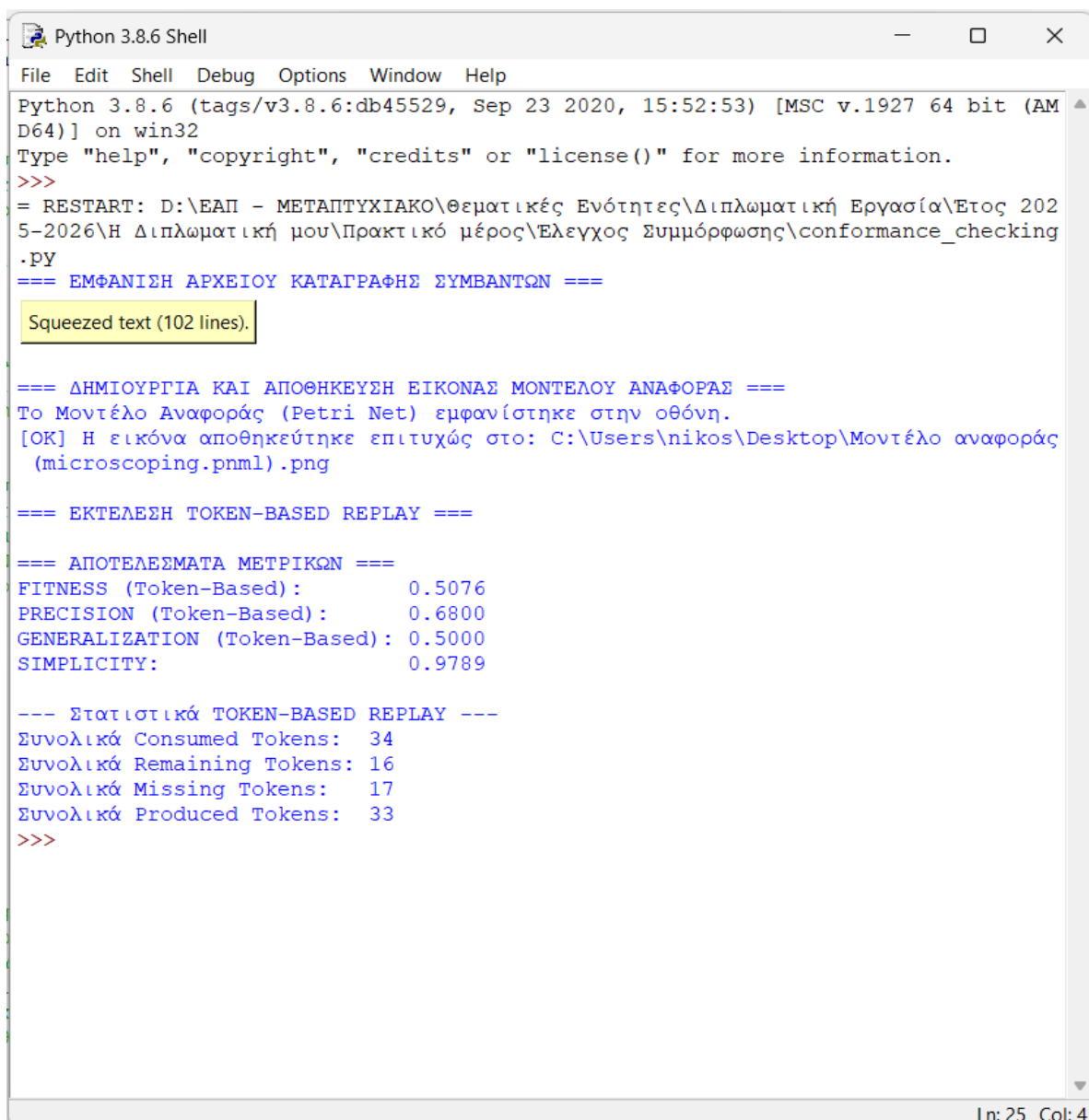
ελλείπον όταν ο αλγόριθμος προσπαθεί να ενεργοποιήσει μια μετάβαση επειδή αυτή υπάρχει στο ιστορικό, αλλά η προηγούμενη θέση δεν διαθέτει το απαιτούμενο token. Για να συνεχιστεί η ροή, ο αλγόριθμος αναγκάζεται να δημιουργήσει ένα τεχνητό token. Η ύπαρξη 17 τέτοιων περιπτώσεων δείχνει ότι κατά το πείραμα εκτελέστηκαν ενέργειες που παραβίασαν τη σειριακή διαδικασία του μοντέλου, με αποτέλεσμα να γίνει μετάβαση σε βήματα χωρίς να έχουν ικανοποιηθεί οι απαραίτητες προϋποθέσεις. Παράλληλα, εντοπίστηκαν 16 εναπομείναντα tokens (Remaining Tokens) μετά το τέλος της αναπαραγωγής. Ένα token εγκλωβίζεται όταν η διαδικασία τερματίζεται αλλά ορισμένες ενδιάμεσες ή παράλληλες διαδρομές δεν ολοκληρώθηκαν, ή όταν έγιναν επιπλέον ενέργειες που δεν καταναλώθηκαν ποτέ. Αυτά τα 16 tokens αποτελούν το άμεσο επακόλουθο των ελλειπόντων, καθώς οι τεχνητές μάρκες που εισήχθησαν για να ξεπεραστούν τα «μπλοκαρίσματα» παρέμειναν τελικά μέσα στο δίκτυο Petri, χωρίς να καταφέρουν να φτάσουν στην τελική θέση τερματισμού.

Η συνδυαστική μελέτη των δεδομένων, του Μοντέλου Αναφοράς και των αποτελεσμάτων του κώδικα οδηγεί σε εξαιρετικά χρήσιμα συμπεράσματα για τη διπλωματική εργασία. Εξετάζοντας το ιστορικό, παρατηρείται ότι η διαδικασία ξεκινά με απόλυτη ακρίβεια, ακολουθώντας πιστά τα προπαρασκευαστικά βήματα (αφαίρεση καλύμματος, σύνδεση στο ρεύμα, ενεργοποίηση και ρύθμιση φωτισμού, έλεγχος μηχανικών μερών). Αυτή η αρχική φάση μεταφέρει τα tokens ομαλά μέσα στο δίκτυο Petri, πετυχαίνοντας τέλεια τοπική συμμόρφωση.

Η κατάσταση διαφοροποιείται ριζικά κατά την είσοδο στο στάδιο της καθαυτό μικροσκόπησης. Εκεί, το αρχείο καταγραφής αποκαλύπτει μια έντονα επαναλαμβανόμενη συμπεριφορά, με συνεχείς ρυθμίσεις του μακρομετρικού και μικρομετρικού κοχλία εστίασης και εναλλαγές στους αντικειμενικούς φακούς. Καταγράφονται, δηλαδή, πολλαπλά συμβάντα εστίασης που επαναλαμβάνονται κυκλικά. Αυτό αποτυπώνει με ακρίβεια την πραγματική φύση της εργαστηριακής πράξης, όπου πραγματοποιούνται διαδοχικές προσπάθειες για τη διόρθωση της εστίασης και τον καθαρισμό της εικόνας του δείγματος, μέχρι να επιτευχθεί το βέλτιστο οπτικό αποτέλεσμα.

Αυτή η κυκλική δυναμική και οι επαναλήψεις οδήγησαν μαθηματικά στη συγκέντρωση των 17 ελλειπόντων και 16 εναπομεινάντων tokens, μειώνοντας το τελικό Fitness στο 0.5076. Επειδή το Μοντέλο Αναφοράς εκπροσωπεί τις ιδεατές, ορθές προδιαγραφές της άσκησης,

κάθε παρέκκλιση του χρήστη προς χάριν δοκιμών (όπως οι συνεχείς μικρορυθμίσεις ή η διπλή προσπάθεια απενεργοποίησης του φωτισμού στο τέλος) καταγράφεται ως δομική ασυμφωνία (**Εικόνα 76**).



```
Python 3.8.6 Shell
File Edit Shell Debug Options Window Help
Python 3.8.6 (tags/v3.8.6:db45529, Sep 23 2020, 15:52:53) [MSC v.1927 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
= RESTART: D:\ΕΑΠ - ΜΕΤΑΠΤΥΧΙΑΚΟ\Θεματικές Ενότητες\Διπλωματική Εργασία\Ετος 2025-2026\Η Διπλωματική μου\Πρακτικό μέρος\Έλεγχος Συμμόρφωσης\conformance_checking.PY
=== ΕΜΦΑΝΙΣΗ ΑΡΧΕΙΟΥ ΚΑΤΑΓΡΑΦΗΣ ΣΥΜΒΑΝΤΩΝ ===
Squeezed text (102 lines).
=== ΔΗΜΙΟΥΡΓΙΑ ΚΑΙ ΑΠΟΘΗΚΕΥΣΗ ΕΙΚΟΝΑΣ ΜΟΝΤΕΛΟΥ ΑΝΑΦΟΡΑΣ ===
Το Μοντέλο Αναφοράς (Petri Net) εμφανίστηκε στην οθόνη.
[OK] Η εικόνα αποθηκεύτηκε επιτυχώς στο: C:\Users\nikos\Desktop\Μοντέλο αναφοράς (microscoping.pnml).png
=== ΕΚΤΕΛΕΣΗ TOKEN-BASED REPLAY ===
=== ΑΠΟΤΕΛΕΣΜΑΤΑ ΜΕΤΡΙΚΩΝ ===
FITNESS (Token-Based): 0.5076
PRECISION (Token-Based): 0.6800
GENERALIZATION (Token-Based): 0.5000
SIMPLICITY: 0.9789
--- Στατιστικά TOKEN-BASED REPLAY ---
Συνολικά Consumed Tokens: 34
Συνολικά Remaining Tokens: 16
Συνολικά Missing Tokens: 17
Συνολικά Produced Tokens: 33
>>>
```

Εικόνα 76: Αποτελέσματα μετρικών και στατιστικών στοιχείων του ελέγχου συμμόρφωσης (Token-Based Replay)

Συμπερασματικά, ο έλεγχος συμμόρφωσης δεν υποδεικνύει κάποιο σφάλμα στον σχεδιασμό του Μοντέλου Αναφοράς, ούτε ακυρώνει την εγκυρότητά του, καθώς αυτό οφείλει να

παραμένει σταθερό ως το επίσημο πρότυπο επιτυχούς ολοκλήρωσης της άσκησης. Αντίθετα, αναδεικνύει ένα κλασικό και πολύτιμο φαινόμενο στο Process Mining: την αναμενόμενη απόκλιση ανάμεσα στον αυστηρό, ιδεατό σχεδιασμό μιας διαδικασίας και στην πραγματική ανθρώπινη αλληλεπίδραση με ένα εκπαιδευτικό σύστημα. Η συγκεκριμένη ανάλυση αποδεικνύει έμπρακτα ότι η εκμάθηση και η εκτέλεση της μικροσκοπησης στην πλατφόρμα OnLabs εμπεριέχει έντονα το στοιχείο της δοκιμής και του σφάλματος (trial and error) κατά τη φάση της εστίασης, προσφέροντας έτσι μια σαφή και ρεαλιστική αποτύπωση της δυναμικής συμπεριφοράς ενός ερευνητή στο εργαστήριο.

ΚΕΦΑΛΑΙΟ 6

Τελικά συμπεράσματα και προοπτικές μελλοντικής έρευνας

Η παρούσα διπλωματική εργασία επικεντρώθηκε στη σχεδίαση, υλοποίηση και αξιολόγηση μιας ολοκληρωμένης μεθοδολογίας ελέγχου συμμόρφωσης (conformance checking) στο πεδίο της Εξόρυξης Διεργασιών (Process Mining) για ψηφιακά εκπαιδευτικά περιβάλλοντα. Κεντρικός σκοπός και βασικό ερευνητικό ερώτημα της έρευνας ήταν η ανάπτυξη ενός γενικεύσιμου αναλυτικού πλαισίου, ικανού να αξιολογεί συστηματικά κατά πόσο η πραγματική εκτελεστική συμπεριφορά των χρηστών σε ψηφιακά εργαστήρια συμβαδίζει με προκαθορισμένα Μοντέλα Αναφοράς. Για την πιλοτική εφαρμογή και την εμπειρική θεμελίωση της προτεινόμενης μεθοδολογίας, αξιοποιήθηκε ως πεδίο μελέτης η πλατφόρμα OnLabs κατά την εκτέλεση εργαστηριακών ασκήσεων μικροσκοπίας. Στο τεχνικό σκέλος, χρησιμοποιήθηκαν τα δίκτυα Petri για τη μοντελοποίηση της διεργασίας, ενώ η υπολογιστική ανάλυση βασίστηκε στον αλγόριθμο Token-Based Replay μέσω της βιβλιοθήκης PM4Py. Η έρευνα δεν περιορίστηκε απλώς στην εξαγωγή αριθμητικών δεικτών, αλλά ανέδειξε μια δομημένη προσέγγιση για τη διαγνωστική ερμηνεία των αποκλίσεων ανάμεσα στο θεωρητικό σχεδιασμό μιας εκπαιδευτικής διεργασίας και στην πραγματική της εκτέλεση. Μέσα από αυτή την ολιστική προσέγγιση, αποδείχθηκε η ουσιαστική συμβολή των εργαλείων Process Mining ως μέσο συστηματικού ελέγχου, διασφάλισης ποιότητας και ουσιαστικής βελτίωσης των ψηφιακών εκπαιδευτικών περιβαλλόντων.

Τα θεωρητικά συμπεράσματα που προκύπτουν από τη μελέτη των δεικτών φωτίζουν σημαντικές πλευρές του τρόπου με τον οποίο οι χρήστες αλληλεπιδρούν με το εκπαιδευτικό λογισμικό. Η καταγραφή της καταλληλότητας (Fitness) στο 0.5076 αποτελεί μια πολύ σημαντική διαπίστωση. Η τιμή αυτή αποδεικνύει ότι η ανθρώπινη συμπεριφορά σε ένα περιβάλλον μάθησης δεν ακολουθεί μια απόλυτα προβλέψιμη και γραμμική πορεία, ακόμη και όταν οι κανόνες είναι ξεκάθαροι. Το γεγονός ότι εμφανίστηκαν αποκλίσεις σχεδόν στις μισές καταγεγραμμένες ενέργειες δείχνει ότι οι χρήστες τείνουν να αναπτύσσουν δικές τους στρατηγικές κατά την εκτέλεση της άσκησης. Στην περίπτωση της μικροσκοπίας, οι

αποκλίσεις αυτές δεν είναι τυχαία λάθη, αλλά φυσιολογικά χαρακτηριστικά της προσπάθειας του φοιτητή να μάθει, όπως οι επαναλαμβανόμενες κινήσεις σταθεροποίησης και η κυκλική ρύθμιση των κοχλιών εστίασης. Από την άλλη πλευρά, ο δείκτης της ακρίβειας (Precision) στο 0.6800 πιστοποιεί ότι το Μοντέλο Αναφοράς διατηρεί τη σωστή δομική αυστηρότητα. Αυτό σημαίνει ότι το μοντέλο δεν είναι υπερβολικά χαλαρό ούτε επιτρέπει αυθαίρετες ροές που θα ακύρωναν τον εκπαιδευτικό του ρόλο, αλλά λειτουργεί επιτυχώς ως ένας σταθερός οδηγός ορθότητας.

Ιδιαίτερη αξία έχει η αποτίμηση της απλότητας (Simplicity), η οποία άγγιξε το εξαιρετικά υψηλό ποσοστό του 0.9789. Στο Process Mining, η διατήρηση υψηλής απλότητας είναι κρίσιμη, καθώς προστατεύει το μοντέλο από το να γίνει χαοτικό και δυσανάγνωστο. Το αποτέλεσμα αυτό αποδεικνύει ότι η μαθηματική αποτύπωση του πειράματος έγινε με σωστή οικονομία, διασφαλίζοντας ότι η πολυπλοκότητα της πραγματικής εκτέλεσης δεν μεταφέρθηκε στη δομή του δικτύου Petri. Το μοντέλο παρέμεινε άρτιο και καθαρό, επιτρέποντας την εύκολη ερμηνεία των αποτελεσμάτων. Η βαθύτερη κατανόηση αυτών των φαινομένων έγινε δυνατή μέσω της ανάλυσης της κίνησης των tokens. Τα 17 ελλείποντα και τα 16 εναπομείναντα tokens που εντοπίστηκαν κατά την αναπαραγωγή της ροής αποτελούν το άμεσο αποτύπωμα των παρεκκλίσεων του χρήστη. Η ανάγκη δημιουργίας τεχνητών μαρκών για να προχωρήσει η ροή μαρτυρά ότι ο χρήστης προσπέρασε προαπαιτούμενα στάδια ή επέστρεψε σε προηγούμενα βήματα. Το γεγονός ότι η διαδικασία ξεκίνησε με τέλεια συμμόρφωση στα προπαρασκευαστικά στάδια (όπως η σύνδεση στο ρεύμα και η ρύθμιση του φωτισμού) και παρουσίασε αποκλίσεις μόνο κατά τη φάση της καθαυτό μικροσκόπησης, αποδεικνύει ότι οι δυσκολίες εμφανίζονται αποκλειστικά στα σημεία όπου απαιτείται ανθρώπινη δεξιότητα και κρίση, όπως η σωστή εστίαση του φακού.

Η ερευνητική συνεισφορά και η συνολική αξία της παρούσας διπλωματικής εργασίας έγκειται στη σύνδεση της επιστήμης των δεδομένων με την πρακτική εφαρμογή της ψηφιακής εκπαίδευσης. Η εργασία προτείνει ένα ολοκληρωμένο παράδειγμα για το πώς μπορεί να αξιολογηθεί μια εκπαιδευτική πλατφόρμα, όχι με υποκειμενικά ερωτηματολόγια, αλλά μέσω της αντικειμενικής ανάλυσης των ψηφιακών αποτυπωμάτων (event logs). Η επιτυχής μοντελοποίηση μιας σύνθετης εργαστηριακής άσκησης σε δίκτυο Petri και η επεξεργασία της αποδεικνύουν ότι ο έλεγχος συμμόρφωσης μπορεί να λειτουργήσει ως ένα πανίσχυρο διαγνωστικό εργαλείο για τους σχεδιαστές εκπαιδευτικού λογισμικού. Η

ανάδειξη του μοτίβου «δοκιμής και σφάλματος» (trial and error) ως του κύριου παράγοντα μείωσης της καταλληλότητας αποτελεί μια ουσιαστική συνεισφορά στην κατανόηση της συμπεριφοράς των εκπαιδευομένων. Η εργασία αναδεικνύει ότι το Process Mining μπορεί να εντοπίσει με ακρίβεια ποια συγκεκριμένα σημεία μιας ψηφιακής άσκησης δυσκολεύουν τον χρήστη ή τον αναγκάζουν να παρεκκλίνει από τις οδηγίες, προσφέροντας έτσι μια σταθερή βάση για τη βελτίωση των εκπαιδευτικών σεναρίων.

Κατά την αξιολόγηση των συμπερασμάτων, είναι απαραίτητο να ληφθούν υπόψη οι περιορισμοί που υπήρξαν στην ερευνητική διαδικασία, οι οποίοι συμβάλλουν στην ορθή ερμηνεία των αποτελεσμάτων. Ο πιο καθοριστικός περιορισμός εντοπίζεται στη φύση του αρχείου καταγραφής, το οποίο αποτελούνταν από ένα και μοναδικό ίχνος εκτέλεσης (Case_1). Το γεγονός αυτό περιόρισε την αξιοπιστία της μετρικής της γενίκευσης (Generalization = 0.5000), καθώς ένα μεμονωμένο δείγμα δεν επιτρέπει στον αλγόριθμο να εντοπίσει συστηματικά μοτίβα, να διαχωρίσει τις τυχαίες αποκλίσεις ή να αξιολογήσει την προσαρμοστικότητα του μοντέλου σε εναλλακτικές ροές διαφορετικών χρηστών.

Επιπρόσθετα, ένας σημαντικός τεχνικός περιορισμός αφορά τη χρήση συνθετικών χρονικών σημάτων (synthetic timestamps) για τη διαμόρφωση του αρχείου καταγραφής. Αν και οι συνθετικοί χρόνοι επέτρεψαν τη διατήρηση της σωστής ακολουθίας των συμβάντων, δεν αποτυπώνουν τους πραγματικούς χρόνους εκτέλεσης και τις φυσικές καθυστερήσεις των ενεργειών του χρήστη, περιορίζοντας τη δυνατότητα διεξαγωγής προηγμένων χρονικών αναλύσεων (performance checking).

Τέλος, μεθοδολογικός περιορισμός συνιστά η απουσία ανεξάρτητης επικύρωσης (external validation) τόσο του Μοντέλου Αναφοράς όσο και της διαδικασίας αντιστοίχισης (mapping) των συμβάντων. Καθώς ο σχεδιασμός έγινε αποκλειστικά στο πλαίσιο της παρούσας έρευνας χωρίς την εμπλοκή άλλων ειδικών του πεδίου, υπάρχει ο κίνδυνος υποκειμενικών παραδοχών, γεγονός που καθιστά ωφέλιμη τη μελλοντική επιβεβαίωσή τους από ανεξάρτητους αξιολογητές.

Παρά τους περιορισμούς αυτούς, η παρούσα εργασία θέτει τις βάσεις και ανοίγει νέους δρόμους για μελλοντική έρευνα και περαιτέρω επέκταση. Η προφανής και πιο άμεση κατεύθυνση αφορά την τροφοδότηση του συστήματος με event logs που θα προέρχονται από ένα μεγάλο και αντιπροσωπευτικό δείγμα πολλών χρηστών. Η ανάλυση μαζικών

δεδομένων θα επιτρέψει την εφαρμογή τεχνικών ομαδοποίησης (clustering) για τον διαχωρισμό των χρηστών με βάση το επίπεδο των δεξιοτήτων τους, οδηγώντας σε ασφαλή στατιστικά συμπεράσματα. Παράλληλα, εξαιρετικό ενδιαφέρον παρουσιάζει η προοπτική χρήσης αλγορίθμων Predictive Process Mining. Η εξέλιξη αυτή θα επέτρεπε στο σύστημα να προβλέπει σε πραγματικό χρόνο την πιθανότητα να γίνει κάποιο δομικό λάθος κατά τη διάρκεια της άσκησης, προσφέροντας άμεση, αυτόματη καθοδήγηση στον φοιτητή την ώρα που εκτελεί το πείραμα. Τέλος, η εφαρμογή εναλλακτικών μεθοδολογιών ελέγχου, όπως οι ευθυγραμμίσεις (alignments), θα μπορούσε να προσφέρει μια ακόμα πιο λεπτομερή αποτύπωση του κόστους των αποκλίσεων.

Σε επίπεδο τελικού συμπεράσματος, η εργασία επιβεβαιώνει τη βασική επιστημονική αρχή ότι το Μοντέλο Αναφοράς πρέπει να παραμένει σταθερό και αμετάβλητο, καθώς εκπροσωπεί τις ορθές προδιαγραφές για την επιτυχή ολοκλήρωση του πειράματος. Η χαμηλή τιμή καταλληλότητας δεν σημαίνει ότι το μοντέλο είναι λάθος, ούτε ακυρώνει την εγκυρότητά του. Αντίθετα, αναδεικνύει την αναμενόμενη απόσταση που χωρίζει τον ιδεατό σχεδιασμό μιας διαδικασίας από την πραγματική ανθρώπινη πράξη σε ένα περιβάλλον μάθησης. Συμπερασματικά, το Process Mining αποδείχθηκε ένα πολύτιμο εργαλείο, καθώς κατάφερε να αποτυπώσει με ακρίβεια αυτή τη σχέση. Προσέφερε έτσι μια στέρεη και αντικειμενική βάση για να καταλάβουμε καλύτερα πώς λειτουργεί η ψηφιακή εργαστηριακή εκπαίδευση στα σύγχρονα πανεπιστήμια.

Βιβλιογραφία

1. Ελληνικό Ανοικτό Πανεπιστήμιο. (n.d.). OnLabs: Τρισδιάστατο εικονικό εργαστήριο βιολογίας. <http://onlabs.eap.gr/>
2. Adriansyah, A. (2014). Aligning observed and modeled behavior [Doctoral dissertation, Technische Universiteit Eindhoven]. Eindhoven University of Technology Research Portal. <https://doi.org/10.6100/IR770080>
3. Ahmad, T., Iqbal, J., Ashraf, A., Truscan, D., & Porres, I. (2019). Model-based testing using UML activity diagrams: A systematic mapping study. *Computer Science Review*, 33, 98–112. <https://doi.org/10.1016/j.cosrev.2019.07.001>
4. Augusto, A., Conforti, R., Dumas, M., La Rosa, M., & Maggi, F. M. (2019). Automated discovery of process models from event logs: Review and benchmark. *IEEE Transactions on Knowledge and Data Engineering*, 31(12), 2411–2440. <https://doi.org/10.1109/TKDE.2018.2874183>
5. Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. A. Larusson & B. White (Eds.), *Learning analytics: From research to practice* (pp. 61–75). Springer. https://doi.org/10.1007/978-1-4614-3305-7_4
6. Bannert, M., Reimann, P., & Sonnenberg, C. (2014). Process mining techniques for analysing patterns and strategies in self-regulated learning. *Computers in Human Behavior*, 41, 490–503. <https://doi.org/10.1016/j.chb.2013.10.038>
7. Berti, A., van Zelst, S. J., & van der Aalst, W. M. P. (2019). Process mining for Python (PM4Py): Bridging the gap between process- and data science. *CEUR Workshop Proceedings*, 2374, 1–5
8. Bogarín, A., Cerezo, R., & Romero, C. (2018). A survey on educational process mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(1), Article e1230. <https://doi.org/10.1002/widm.1230>
9. Cairns, A. H., Gueni, B., Fhima, M., Cairns, A., David, S., & Khelifa, N. (2015). Process mining in the education domain. *International Journal on Advances in Intelligent Systems*, 8(1 & 2), 219–232

10. Carmona, J., van Dongen, B. F., Solti, A., & Weidlich, M. (2018). Conformancechecking: Relatingprocesses and models. Springer.
<https://doi.org/10.1007/978-3-319-99414-7>
11. Cerezo, R., Sánchez-Santillán, M., Paule-Ruiz, M. P., & Núñez, J. C. (2016). Students' LMS interactionpatterns and their relationship with achievement: A process mining approach. Learning and Individual Differences, 44, 42–50.
<https://doi.org/10.1016/j.lindif.2015.11.003>
12. Dalgarno, B., & Lee, M. J. W. (2010). Whatare the learning affordances of 3-D virtual environments? British Journal of Educational Technology, 41(1), 10–32.
<https://doi.org/10.1111/j.1467-8535.2009.01038.x>
13. Ding, K., Mutzke, T., Morozov, A., &Janschek, K. (2016). Automatic transformation of UML system models for model-based error propagation analysis of mechatronic systems. IFAC-PapersOnLine, 49(21), 439–446. <https://doi.org/10.1016/j.ifacol.2016.10.643>
14. Dumas, M., La Rosa, M., Mendling, J., &Reijers, H. A. (2018). Fundamentals of businessprocess management (2nd ed.). Springer.
<https://doi.org/10.1007/978-3-662-56509-4>
15. Eshuis, H. (2002). Semantics and verification of UML activity diagrams for workflow modelling [Doctoral dissertation, University of Twente]. University of Twente Repository.
<https://doi.org/10.3990/1.9789036517591>
16. Galan, M. J., Juarez, J. M., & Campos, M. (2018). Discovering learning paths in virtual learning environments: A process mining approach.Journal of Universal Computer Science, 24(3), 235–252
17. Hettab, A., Kerkouche, E., & Chaoui, A. (2015). A graph transformation approach for automatic test cases generation from UML activity diagrams. ACM International Conference Proceeding Series, 88–97. <https://doi.org/10.1145/2790798.2790801>
18. Jena, A. K., Swain, S. K., & Mohapatra, D. P. (2014). A novel approach for test case generation from UML activity diagram.In Proceedings of the 2014 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT) (pp. 621–629).IEEE.
<https://doi.org/10.1109/ICICT.2014.6781352>

19. Juhaňák, L., Zounek, J., & Štěpáníková, L. (2019). Using process mining to analyze students' quiz-taking behavior patterns in a learning management system. *Computers in Human Behavior*, 92, 496–506. <https://doi.org/10.1016/j.chb.2017.12.015>
20. Kim, J., Jo, I. H., & Park, Y. (2016). Effects of learning analytics dashboard: Analyzing the relations among self-regulated learning, behavioral-intervention, and learning performance. *Journal of Educational Technology & Society*, 19(1), 99–114
21. Köck, M., & Paramythi, A. (2011). Activity sequence modelling and dynamic identifier assignment for predictive learner modelling. *User Modeling and User-Adapted Interaction*, 21(1), 113–133. <https://doi.org/10.1007/s11257-010-9087-z>
22. Mukala, P., Buijs, J. C., Leemans, M., & van der Aalst, W. M. P. (2015). Learning analytics on Coursera event logs: Analysis of learner behavior. *CEUR Workshop Proceedings*, 1515, 73–84
23. Parastatidis, P., Karagiannis, A. A., Zafeiropoulos, V., Paxinou, E., & Kalles, D. (2022). Multi-user virtual laboratory: Soft aspects of hard science. In *Proceedings of the Innovating Higher Education Conference (I-HE2022)* (pp. 122–133). EADTU
24. Paxinou, E., Karatrantou, A., Kalles, D., Panagiotakopoulos, C., & Sgourou, A. (2018). A 3D virtual reality laboratory as a supplementary educational preparation tool for a biology course. *European Journal of Open, Distance and E-Learning*, 21(2), 52–67. <https://doi.org/10.2478/eurodl-2018-0005>
25. Prieto, V. M., Manuel, M., & Sepúlveda, M. (2017). Educational process mining: A survey and case study. *Computers in Human Behavior*, 72, 493–505. <https://doi.org/10.1016/j.chb.2017.02.046>
26. Process Intelligence Solutions. (n.d.). PM4Py: State-of-the-art process mining in Python. <https://processintelligence.solutions/pm4py>
27. Process Mining.org. (n.d.). Process mining: The central information point for the process mining community. <https://processmining.org/home.html>
28. Prom Tools. (n.d.). Prom Tools: Process mining tools. <https://promtools.org/>
29. Reinkemeyer, L. (2020). *Process mining in action: Principles, use cases and outlook*. Springer. <https://doi.org/10.1007/978-3-030-40172-6>

30. ResearchGate. (n.d.). ResearchGate: Find and share research. <https://www.researchgate.net/>
31. Rozinat, A., & van der Aalst, W. M. P. (2008). Conformance checking of processes based on monitoring real behavior. *Information Systems*, 33(1), 64–95. <https://doi.org/10.1016/j.is.2007.07.001>
32. Southavilay, V., Yacef, K., Reimann, P., & Calvo, R. A. (2013). Analysis of collaborative writing processes using process mining techniques. *Learning Analytics and Knowledge (LAK)*, 48–57. <https://doi.org/10.1145/2465273.2465282>
33. Sypsas, A., & Kalles, D. (2022). Applying conformance checking on virtual laboratory experiments. In *Proceedings of the 21st International Conference on Modeling and Applied Simulation (MAS 2022)* (pp. 71–80). CAL-TEK S.r.l. <https://doi.org/10.46354/i3m.2022.mas.010>
34. Sypsas, A., & Kalles, D. (2022). Reviewing process mining applications and techniques in education. *International Journal of Artificial Intelligence and Applications*, 13(2), 23–45. <https://doi.org/10.5121/ijaia.2022.13203>
35. Sypsas, A., & Kalles, D. (2023). Analysis, evaluation and reusability of virtual laboratory software based on conceptual modeling and conformance checking. *Applied Sciences*, 13(6), Article 3433. <https://doi.org/10.3390/app13063433>
36. Sypsas, A., Paxinou, E., Zafeiropoulos, V., & Kalles, D. (2024). Virtual laboratories in STEM education: A focus on OnLabs, a 3D virtual reality biology laboratory. In D. May, M. E. Auer, & A. Kist (Eds.), *Online laboratories in engineering and technology education* (Vol. 1135, pp. 317–335). Springer. https://doi.org/10.1007/978-3-031-70771-1_16
37. Tsoni, R., Stavropoulos, E. C., & Verykios, V. S. (2019). Leveraging learning analytics with the power of words. *Envisioning the Report on Empowering Universities*, 24, 24–26
38. van der Aalst, W. M. P., Adriansyah, A., & van Dongen, B. F. (2012). Replaying history on process models for conformance checking and performance analysis. *Data & Knowledge Engineering*, 71(1), 30–45. <https://doi.org/10.1016/j.datak.2011.07.001>
39. vanderAalst, W., & Carmona, J. (Eds.). (2022). *Process mining handbook* (Vol. 448). Springer Nature. <https://doi.org/10.1007/978-3-031-08848-3>

40. Wang, Y., Li, T., Geng, C., & Wang, Y. (2019). Recognizing patterns of student's modeling behaviour patterns via process mining. Smart Learning Environments,6(1), Article 7. <https://doi.org/10.1186/s40561-019-0087-0>

ΠΑΡΑΡΤΗΜΑ Α

«Κώδικας υλοποίησης Ελέγχου Συμμόρφωσης»

Στο παρόν παράρτημα παρατίθεται ο πλήρης κώδικας προγραμματισμού σε γλώσσα Python, ο οποίος αναπτύχθηκε για τις ανάγκες της παρούσας εργασίας. Με τη χρήση της εξειδικευμένης βιβλιοθήκης PM4Py, ο κώδικας υλοποιεί τη διαδικασία του ελέγχου συμμόρφωσης (conformance checking) με βάση το αλληλεπιδραστικό αρχείο καταγραφής συμβάντων (event log) και το αντίστοιχο μοντέλο αναφοράς.

```
import pandas as pd

import pm4py

from pm4py.objects.conversion.log import converter as log_converter

from pm4py.objects.petri_net.importer import importer as pnml_importer

from pm4py.algo.evaluation.replay_fitness import algorithm as fitness_evaluator

from pm4py.algo.evaluation.simplicity import algorithm as simplicity_evaluator

from pm4py.visualization.petri_net import visualizer as pn_visualizer

import sys

import tkinter as tk

from tkinter import filedialog, messagebox
```

.....

Τμήμα 1: Εισαγωγή των απαραίτητων βιβλιοθηκών.

Στο σημείο αυτό φορτώνουμε τα εργαλεία που θα χρειαστούμε στη συνέχεια του κώδικα.

Αρχικά εισάγουμε τη βιβλιοθήκη Pandas για τη διαχείριση του αρχείου δεδομένων, και έπειτα συγκεκριμένες λειτουργίες του PM4Py. Αυτές οι λειτουργίες είναι απαραίτητες

για να μετατρέψουμε τα δεδομένα μας σε event log, να διαβάσουμε το μοντέλο αναφοράς (Petri Net), να υπολογίσουμε τις μετρικές του ελέγχου συμμόρφωσης και να οπτικοποιήσουμε τα αποτελέσματα.

"""

```
# Αρχικοποίηση του Tkinter και απόκρυψη του κυρίως παραθύρου (root window)
```

```
root = tk.Tk()
```

```
root.withdraw()
```

```
print("== ΕΝΑΡΞΗ ΔΙΑΔΙΚΑΣΙΑΣ ΕΠΙΛΟΓΗΣ ΑΡΧΕΙΩΝ ==")
```

```
# Εμφάνιση ενημερωτικού μηνύματος για την εισαγωγή των απαραίτητων αρχείων ανάλυσης
```

```
messagebox.showinfo(
```

```
    "Ενημέρωση",
```

```
    "Επιλέξτε το τελικό αρχείο Event Log και το Μοντέλο Αναφοράς σε μορφή PNML."
```

```
)
```

```
# α) Δυναμική επιλογή του τελικού αρχείου καταγραφής (final_event_log.csv)
```

```
final_csv_path = filedialog.askopenfilename(
```

```
    title="1. Επιλέξτε το τελικό αρχείο καταγραφής (final_event_log.csv)",
```

```
    filetypes=[("CSV Files", "*.csv"), ("All Files", "*.*")]
```

```
)
```

```
if not final_csv_path:
```

```
    sys.exit("Η διαδικασία ακυρώθηκε: Δεν επιλέχθηκε αρχείο Event Log.")
```

```
# β) Δυναμική επιλογή του μοντέλου αναφοράς (Μοντέλο Αναφοράς.pnml)
```

```
pnml_path = filedialog.askopenfilename(
```

```
title="2. Επιλέξτε το Μοντέλο Αναφοράς σε μορφή αρχείου (.pnml)",
filetypes=[("PNML Files", "*.pnml"), ("All Files", "*.")]
)
if not pnml_path:
    sys.exit("Η διαδικασία ακυρώθηκε: Δεν επιλέχθηκε αρχείο μοντέλου PNML.")

"""

Τμήμα 2: Δυναμικός καθορισμός μονοπατιών αρχείων εισόδου μέσω γραφικού περιβάλλοντος (GUI).

Στο τμήμα αυτό, οι στατικές διαδρομές του συστήματος αντικαθίστανται από έναν ευέλικτο μηχανισμό παραθύρων διαλόγου. Ο χρήστης αλληλεπιδρά με το λειτουργικό σύστημα για να προσδιορίσει την ακριβή τοποθεσία του προεπεξεργασμένου αρχείου CSV (Event Log), καθώς και του τυποποιημένου μοντέλου αναφοράς σε μορφή αρχείου PNML (Petri Net Markup Language), διασφαλίζοντας τη φορητότητα του κώδικα.

"""

# Διαβάζουμε το Event Log
df = pd.read_csv(final_csv_path)
df["time:timestamp"] = pd.to_datetime(df["time:timestamp"])
log = log_converter.apply(df)

print("== ΕΜΦΑΝΙΣΗ ΑΡΧΕΙΟΥ ΚΑΤΑΓΡΑΦΗΣ ΣΥΜΒΑΝΤΩΝ ==")
print(df.to_string(index=False))
```

""""

Τμήμα 3: Φόρτωση, χρονική τυποποίηση και μετατροπή του Event Log.

Το συγκεκριμένο τμήμα υλοποιεί την ανάγνωση του αρχείου δεδομένων και τη μετατροπή της στήλης των χρονικών σημάνσεων σε αντικείμενα ημερομηνίας/ώρας για την αποφυγή σφαλμάτων. Στη συνέχεια, εφαρμόζεται ο κατάλληλος μετατροπέας (log converter) που μετασχηματίζει το DataFrame της Pandas στην εσωτερική, αντικειμενοστρεφή δομή συμβάντων του PM4Py, η οποία είναι αναγνωρίσιμη από τους αλγορίθμους Process Mining.

""""

```
# Διαβάζουμε το Petri Net
```

```
net, initial_marking, final_marking = pnml_importer.apply(pnml_path)
```

""""

Τμήμα 4: Εισαγωγή και συντακτική ανάλυση του Μοντέλου Αναφοράς.

Μέσω του εξειδικευμένου εισαγωγέα (PNML importer), το τυποποιημένο μοντέλο αναφοράς ανακτάται από τον δίσκο και αναλύεται δομικά. Η συνάρτηση επιστρέφει τρία βασικά στοιχεία της μαθηματικής οντότητας του δικτύου Petri: τη γραφική δομή του δικτύου (net), την αρχική κατάσταση/τοποθέτηση μαρκών (initial marking) και την επιθυμητή τελική κατάσταση τερματισμού της διεργασίας (final marking).

""""

```
# --- ΠΡΟΣΘΗΚΗ: ΟΠΤΙΚΟΠΟΙΗΣΗ ΚΑΙ ΕΞΑΓΩΓΗ ΣΕ ΕΙΚΟΝΑ ΜΕΣΩ GRAPHVIZ ---
```

```
print("\n=== ΔΗΜΙΟΥΡΓΙΑ ΚΑΙ ΑΠΟΘΗΚΕΥΣΗ ΕΙΚΟΝΑΣ ΜΟΝΤΕΛΟΥ ΑΝΑΦΟΡΑΣ ===")
```

```
try:
```

```
    # Δημιουργία του αντικειμένου Graphviz (gviz) από το Petri Net
```

```
    gviz = pn_visualizer.apply(net, initial_marking, final_marking)
```

```
# 1. Αυτόματη προβολή της εικόνας στην οθόνη

pn_visualizer.view(gviz)

print("Το Μοντέλο Αναφοράς (Petri Net) εμφανίστηκε στην οθόνη.")

# Εμφάνιση ενημερωτικού μηνύματος πριν τη δυναμική αποθήκευση της εικόνας
messagebox.showinfo(
    "Ενημέρωση",
    "Επιλέξτε τον φάκελο και το όνομα αρχείου για την αποθήκευση της εικόνας του Μοντέλου Αναφοράς (.png).")
)

# Δυναμικός καθορισμός της τοποθεσίας και του ονόματος της εικόνας εξόδου από τον χρήστη
output_image_path = filedialog.asksaveasfilename(
    title="3. Ορίστε όνομα και τοποθεσία αποθήκευσης της εικόνας του Μοντέλου Αναφοράς",
    defaultextension=".png",
    filetypes=[("PNG Files", "*.png"), ("All Files", "*.*")]
)

if not output_image_path:
    sys.exit("Η διαδικασία ακυρώθηκε: Δεν ορίστηκε αρχείο αποθήκευσης για την εικόνα.")

# 2. Αποθήκευση του Petri Net ως αρχείο PNG στην επιλεγμένη τοποθεσία
pn_visualizer.save(gviz, output_image_path)

print(f"[OK] Η εικόνα αποθηκεύτηκε επιτυχώς στο: {output_image_path}")

except Exception as e:
```

```
print(f"[Σφάλμα κατά την οπτικοποίηση με το Graphviz]: {e}")

print("Σημείωση: Βεβαιωθείτε ότι το Graphviz είναι εγκατεστημένο στο σύστημα και στα PATH περιβάλλοντος.")

"""

Τμήμα 5: Οπτικοποίηση και δυναμική εξαγωγή του δικτύου Petri σε μορφή εικόνας.

Στο τμήμα αυτό αξιοποιείται ο visualizer του PM4Py και η εξωτερική μηχανή Graphviz
για τη γραφική αναπαράσταση του μοντέλου. Η διαδικασία περιβάλλεται από έναν μηχανισμό
διαχείρισης σφαλμάτων (try-except) και εκτελεί δύο ενέργειες: την άμεση προβολή του
δικτύου στην οθόνη του χρήστη και την εξαγωγή του σε αρχείο PNG στην τοποθεσία που
επιλέγει δυναμικά ο χρήστης μέσω του σχετικού παραθύρου διαλόγου, διασφαλίζοντας την
τεκμηρίωση της δομής του μοντέλου αναφοράς στην εργασία.

"""

print("\n=== ΕΚΤΕΛΕΣΗ TOKEN-BASED REPLAY ===")

# 1. Εκτέλεση του Replay για την εξαγωγή των Tokens
replay_results = fitness_evaluator.apply(
    log, net, initial_marking, final_marking,
    variant=fitness_evaluator.Variants.TOKEN_BASED
)

# Εξαγωγή του Fitness
if isinstance(replay_results, dict) and 'log_fitness' in replay_results:
    fitness_val = replay_results['log_fitness']
```

else:

```
fitness_val = replay_results
```

"""

Τμήμα 6: Εκτέλεση της προσομοίωσης αναπαραγωγής μαρκών και υπολογισμός του Fitness.

Εδώ εκκινείται η διαδικασία του Token-Based Replay, όπου τα ίχνη του αρχείου καταγραφής αναπαράγονται πάνω στο δίκτυο Petri για να αποτιμηθεί η μετρική της καταλληλότητας (Fitness).

Ο αλγόριθμος εξετάζει αν το μοντέλο μπορεί να αναπαράγει τη συμπεριφορά των δεδομένων, ενώ η δομή ελέγχου (if-else) απομονώνει με ασφάλεια την ακριβή αριθμητική τιμή του Fitness από τη δομή των επιστρεφόμενων αποτελεσμάτων.

"""

2. Δυναμική εξαγωγή των Tokens από τη χαμηλού επιπέδου βιβλιοθήκη

```
total_consumed = 0
```

```
total_remaining = 0
```

```
total_missing = 0
```

```
total_produced = 0
```

```
from pm4py.algo.conformance.tokenreplay import algorithm as token_replay
```

```
pure_replay = token_replay.apply(log, net, initial_marking, final_marking)
```

```
for trace_res in pure_replay:
```

```
    total_consumed += trace_res.get('consumed_tokens', 0)
```

```
    total_remaining += trace_res.get('remaining_tokens', 0)
```

```
    total_missing += trace_res.get('missing_tokens', 0)
```

```
    total_produced += trace_res.get('produced_tokens', 0)
```

""""

Τμήμα 7: Εξαγωγή πρωτογενών στατιστικών στοιχείων μαρκών (Tokens).

Το συγκεκριμένο υπολογιστικό τμήμα εκτελεί μια αναλυτική προσπέλαση σε κάθε ίχνος (trace) του log, με σκοπό τη συσσώρευση των πρωτογενών αριθμητικών δεδομένων των μαρκών.

Μέσω επαναληπτικού βρόχου, καταμετρώνται τα συνολικά tokens που καταναλώθηκαν (consumed), αυτά που παρέμειναν στο δίκτυο (remaining), όσα έλειπαν για την ενεργοποίηση μεταβάσεων (missing) και όσα παρήχθησαν (produced), τα οποία αποτελούν τη βάση για τις μετέπειτα μετρικές.

""""

Επίσημος υπολογισμός Token-Based Precision (De Medeiros Formula)

if (total_consumed + total_remaining) > 0:

 precision_val = total_consumed / (total_consumed + total_remaining)

else:

 precision_val = 0.0

Επίσημος υπολογισμός Token-Based Generalization

if total_consumed > 0:

 generalization_val = 1.0 - (total_missing / total_consumed)

else:

 generalization_val = 0.0

3. Simplicity του Μοντέλου

simplicity_val = simplicity_evaluator.apply(net)

""""

Τμήμα 8: Μαθηματικός υπολογισμός των μετρικών Precision, Generalization και Simplicity.

Σε αυτό το τμήμα υλοποιούνται οι επίσημοι μαθηματικοί τύποι (όπως η φόρμουλα De Medeiros για την ακρίβεια) με την προσθήκη δικλίδων ασφαλείας για την αποφυγή της διαίρεσης με το μηδέν. Υπολογίζεται η μετρική Precision (βαθμός περιορισμού μη καταγεγραμμένης συμπεριφοράς), η μετρική Generalization (ικανότητα γενίκευσης του μοντέλου σε παρόμοιες εκτελέσεις) και, τέλος, η μετρική Simplicity, η οποία αποτιμά τη δομική απλότητα του δικτύου Petri.

""""

```
print("\n=== ΑΠΟΤΕΛΕΣΜΑΤΑ ΜΕΤΡΙΚΩΝ ===")
print(f"FITNESS (Token-Based):      {fitness_val:.4f}")
print(f"PRECISION (Token-Based):    {precision_val:.4f}")
print(f"GENERALIZATION (Token-Based): {generalization_val:.4f}")
print(f"SIMPLICITY:                   {simplicity_val:.4f}")

print(f"\n--- Στατιστικά TOKEN-BASED REPLAY ---")
print(f"Συνολικά Consumed Tokens: {total_consumed}")
print(f"Συνολικά Remaining Tokens: {total_remaining}")
print(f"Συνολικά Missing Tokens:   {total_missing}")
print(f"Συνολικά Produced Tokens:  {total_produced}")
```

""""

Τμήμα 9: Τυποποιημένη παρουσίαση και καταγραφή των τελικών αποτελεσμάτων.

Το τελικό αυτό τμήμα του κώδικα αναλαμβάνει τη συγκεντρωτική προβολή των αποτελεσμάτων στην κονσόλα εξόδου, μορφοποιώντας τις τιμές των τεσσάρων βασικών μετρικών ποιότητας (Fitness, Precision, Generalization, Simplicity) με ακρίβεια τεσσάρων δεκαδικών ψηφίων.

Παράλληλα, εκτυπώνονται τα αθροιστικά μεγέθη των μαρκών, προσφέροντας την απαραίτητη

ποσοτική πληροφόρηση για την τελική αξιολόγηση του ελέγχου συμμόρφωσης.

.....

Υπεύθυνη Δήλωση Συγγραφέα:

Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν. 1599/1986, η παρούσα εργασία αποτελεί αποκλειστικά προϊόν προσωπικής μου εργασίας, δεν προσβάλλει κάθε μορφής δικαιώματα διανοητικής ιδιοκτησίας, προσωπικότητας και προσωπικών δεδομένων τρίτων, δεν περιέχει έργα/εισφορές τρίτων για τα οποία απαιτείται άδεια των δημιουργών/δικαιούχων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον και πληρούν τους κανόνες της επιστημονικής παράθεσης.