

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ



ΣΧΟΛΗ ΕΠΙΣΤΗΜΗΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ
ΠΛΗΡΟΦΟΡΙΚΗΣ

ΤΜΗΜΑ ΑΡΧΕΙΟΝΟΜΙΑΣ, ΒΙΒΛΙΟΘΗΚΟΜΙΑΣ ΚΑΙ
ΜΟΥΣΕΙΟΛΟΓΙΑΣ

Πρόγραμμα Μεταπτυχιακών Σπουδών

«Διαχείριση Ψηφιακής Πληροφορίας – Υπηρεσίες
Πληροφόρησης»

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

«Χρήση μεθόδων και εργαλείων Τεχνητής Νοημοσύνης στην
Πολιτιστική Πληροφορία»

Αποστόλου Ευφροσύνη

Κέρκυρα Μάιος 2026

Η σελίδα αυτή είναι σκόπιμα λευκή

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

ΣΧΟΛΗ ΕΠΙΣΤΗΜΗΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗΣ

ΤΜΗΜΑ ΑΡΧΕΙΟΝΟΜΙΑΣ, ΒΙΒΛΙΟΘΗΚΟΜΙΑΣ ΚΑΙ ΜΟΥΣΕΙΟΛΟΓΙΑΣ

Πρόγραμμα Μεταπτυχιακών Σπουδών

«Διαχείριση Ψηφιακής Πληροφορίας – Υπηρεσίες Πληροφόρησης»

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

«Χρήση μεθόδων και εργαλείων Τεχνητής Νοημοσύνης στην Πολιτιστική Πληροφορία»

Αποστόλου Ευφροσύνη

Αριθμός Μητρώου: alm.dimis2308

Επιβλέπων Καθηγητής

Ματθαίος Δαμίγος, Τμήμα Αρχειονομίας, Βιβλιοθηκονομίας και Μουσειολογίας, Ιόνιο
Πανεπιστήμιο, Επίκουρος Καθηγητής

Εξεταστική Επιτροπή

Ματθαίος Δαμίγος, Τμήμα Αρχειονομίας, Βιβλιοθηκονομίας και Μουσειολογίας, Ιόνιο
Πανεπιστήμιο, Επίκουρος Καθηγητής

Μιχάλης Σφακάκης, Τμήμα Αρχειονομίας, Βιβλιοθηκονομίας και Μουσειολογίας, Ιόνιο
Πανεπιστήμιο, Καθηγητής

Χρήστος Παπαθεοδώρου, Τμήμα Ιστορίας και Φιλοσοφίας της Επιστήμης του Εθνικού
και Καποδιστριακού Πανεπιστημίου Αθηνών, Καθηγητής

Κέρκυρα, Μάιος 2026

Υπεύθυνη Δήλωση Συγγραφέα:

Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν. 1599/1986 και τα άρθρα 2,4,6 παρ. 3 του Ν. 1256/1982, η παρούσα εργασία αποτελεί αποκλειστικά προϊόν προσωπικής εργασίας και δεν προσβάλλει κάθε μορφής πνευματικά δικαιώματα τρίτων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον.

ΥΠΟΓΡΑΦΗ: Αποστόλου Ευφροσύνη



Η παρούσα Διπλωματική Εργασία καθώς και τα αποτελέσματά της, αποτελούν συνιδιοκτησία του Ιονίου Πανεπιστημίου και της φοιτήτριας, ο καθένας από τους οποίους έχει το δικαίωμα ανεξάρτητης χρήσης, αναπαραγωγής και αναδιανομής τους (στο σύνολο ή τμηματικά) για διδακτικούς και ερευνητικούς σκοπούς, σε κάθε περίπτωση αναφέροντας τον τίτλο και το συγγραφέα της Εργασίας καθώς και το όνομα του Ιονίου Πανεπιστημίου όπου εκπονήθηκε.

Ευχαριστίες / Αφιέρωση

Φτάνοντας στο τέλος αυτής της διαδρομής, αισθάνομαι την ανάγκη να σταθώ για λίγο και να ευχαριστήσω όσους τη μοιράστηκαν μαζί μου.

Ευχαριστώ θερμά τον επιβλέποντα καθηγητή μου, κ. Ματθαίο Δαμίγο, για την αμέριστη συμπαράστασή του, τις εύστοχες υποδείξεις του και για την κατανόηση που επέδειξε σε όλη τη διάρκεια της εκπόνησης. Η καθοδήγησή του ξεπέρασε τα όρια της απλής επιστημονικής εποπτείας.

Ευχαριστώ επίσης τον φίλο και συνάδελφο Λευτέρη Καλόγερο, μέλος ΕΔΙΠ του Τμήματος Αρχειονομίας, Βιβλιοθηκονομίας και Μουσειολογίας, για την πολύτιμη συνεργασία μας και για την ιδέα που μοιράστηκε μαζί μου, μια ιδέα γύρω από την οποία χτίστηκε το πειραματικό μέρος αυτής της εργασίας. Οι συζητήσεις μαζί του υπήρξαν πηγή έμπνευσης και ώθησης.

Το πιο μεγάλο ευχαριστώ ανήκει στην οικογένειά μου. Στα παιδιά μου, για την υπομονή και την αγάπη τους όλους αυτούς τους μήνες. Και πάνω απ' όλα στον σύζυγό μου, που γνωρίζει καλύτερα από οποιονδήποτε τον χρόνο που του «έκλεψα» για να φτάσω ως εδώ και που, παρ' όλα αυτά, στάθηκε δίπλα μου με μεγαλοψυχία και πίστη σε κάθε βήμα. Σ' αυτόν χρωστώ περισσότερα απ' όσα μπορούν να χωρέσουν σε μια σελίδα.

Η σελίδα αυτή είναι σκόπιμα κενή

Περίληψη

Η αρχειακή περιγραφή αποτελεί μία χρονοβόρα και εξειδικευμένη εργασία που, σε συνθήκες τεράστιου όγκου συλλογών και περιορισμένων πόρων, δημιουργεί συχνά μεγάλες αναμονές ακατέργαστου υλικού. Η παρούσα εργασία εξετάζει κατά πόσο τα Μεγάλα Γλωσσικά Μοντέλα (LLMs) μπορούν να υποστηρίξουν αυτή τη διαδικασία, παράγοντας δομημένες εγγραφές συμβατές με το Διεθνές Πρότυπο Αρχειακής Περιγραφής ISAD(G)/ΔΙΠΑΠ(Γ).

Για τον σκοπό αυτό σχεδιάστηκε και εκτελέστηκε ένα πιλοτικό πείραμα σύγκρισης έντεκα (11) μοντέλων, ανοιχτού κώδικα και εμπορικών, διαφορετικού μεγέθους και αρχιτεκτονικής. Ως δεδομένα εισόδου χρησιμοποιήθηκε η αρχειακή περιγραφή του Αρχείου της Επτανήσου Πολιτείας (1800–1807), που φυλάσσεται στα Γενικά Αρχεία του Κράτους, Αρχεία Νομού Κέρκυρας. Κάθε μοντέλο έλαβε ελεύθερο κείμενο ως είσοδο και κλήθηκε να παραγάγει δομημένες εγγραφές για πέντε ιεραρχικά επίπεδα, με βάση ένα prompt έξι κανόνων που στόχευε στην αποτροπή hallucinations και στη διασφάλιση συμμόρφωσης με το πρότυπο. Η αξιολόγηση πραγματοποιήθηκε από ειδικό αρχειονόμο, σε σύγκριση με πέντε εγγραφές αναφοράς.

Τα αποτελέσματα δείχνουν σημαντική διακύμανση ανάμεσα στα μοντέλα. Τα μεγαλύτερα μοντέλα, κυρίως εκείνα με ισχυρό instruction tuning, παρήγαγαν γενικά πλουσιότερες περιγραφές και ισχυρότερη πολυεπίπεδη συλλογιστική, αλλά ήταν επίσης πιο πιθανό να εισάγουν μη τεκμηριωμένες μονάδες σε επίπεδο σειράς, κιβωτίου ή μητρώου. Τα μικρότερα μοντέλα απέφυγαν ορισμένες παραισθήσεις, αλλά συχνά απέτυχαν να ανακατασκευάσουν την αναμενόμενη ιεραρχία αρχείου, φακέλου και τεκμηρίου. Ωστόσο, κανένα μοντέλο δεν κατάφερε να διαφοροποιήσει τον παραγωγό ανά τεκμήριο ή να σκεφτεί ιεραρχικά, αδυναμίες που αντανακλούν μια βαθύτερη ασυμφωνία ανάμεσα στη λογική των LLMs και στις αρχές της αρχειακής επιστήμης.

Τα ευρήματα λοιπόν υποδηλώνουν ότι τα LLMs μπορούν να συνδράμουν στη δόμηση αρχειακών μεταδεδομένων, αλλά παραμένει αναγκαία η εποπτεία από ειδικό για τη διασφάλιση της συμμόρφωσης με το ISAD(G). Η ανθρώπινη επικύρωση παραμένει αναντικατάστατη, τόσο για δεοντολογικούς λόγους όσο και για λόγους ευθύνης,

τεκμηρίωσης και ερμηνευσιμότητας. Κλείνοντας, η εργασία προτείνει υβριδικά συστήματα που συνδυάζουν αυτοματοποίηση και ανθρώπινη κρίση, με στόχο την ευρύτερη προσβασιμότητα του αρχειακού υλικού χωρίς να τίθεται σε κίνδυνο η αρχειακή ακεραιότητα.

Λέξεις - Κλειδιά

Μεγάλα Γλωσσικά Μοντέλα (LLMs), αρχειακή περιγραφή, ISAD(G)/ΔΙΠΑΠ(Γ), εξαγωγή μεταδεδομένων, αρχειακή τεκμηρίωση, Επεξεργασία Φυσικής Γλώσσας (NLP), hallucinations, υβριδικά συστήματα AI, Πολιτιστική Πληροφορία

Using AI methods & tools in Cultural Information

Abstract

Archival description is a time-consuming and specialized task which, in the context of expanding collections and limited resources, often leads to large backlogs of unprocessed material. This study examines whether Large Language Models (LLMs) can support this process by generating structured records compatible with the International Standard for Archival Description ISAD(G).

To this end, a pilot comparative experiment was designed and conducted, involving eleven (11) models, both open-source and commercial, of different sizes and architectures. The archival description of the Archive of the Septinsular Republic (1800–1807), held at the General State Archives, Archives of the Prefecture of Corfu, was used as input data. Each model received free text as input and was asked to generate structured records for five hierarchical levels, based on a six-rule prompt designed to prevent hallucinations and ensure compliance with the standard. The evaluation was carried out by a professional archivist, in comparison with five reference records.

The results show that some models, especially those with strong instruction tuning, perform satisfactorily at the fonds level, covering the main areas of the standard. However, no model managed to differentiate the creator at the item level, or reason hierarchically, weaknesses that reflect a deeper mismatch between the logic of LLMs and the principles of archival science.

The central conclusion is that LLMs do not replace the archivist; they support the archivist. Human validation remains indispensable, both for ethical reasons and for reasons of responsibility, documentation, and interpretability. Finally, the study proposes hybrid systems that combine automation with human judgment, aiming to broaden access to archival material without compromising archival integrity.

Keywords

Large Language Models (LLMs), archival description, ISAD(G), metadata extraction, archival documentation, Natural Language Processing (NLP), hallucination, hybrid AI systems, cultural heritage

Περιεχόμενα

Περίληψη	i
Using AI methods & tools in Cultural Information.....	iii
Abstract.....	iii
Περιεχόμενα	v
Κατάλογος Σχημάτων	viii
Κατάλογος Πινάκων	viii
Κατάλογος Συντομογραφιών.....	ix
Κεφάλαιο 1 Εισαγωγή.....	1
Κεφάλαιο 2 Βιβλιογραφική Επισκόπηση.....	7
2.1 Τεχνικές εξαγωγής μεταδεδομένων	7
2.1.1 Κλασικές rule-based και ML pipelines	7
2.1.2 Pipelines με LLMs και γενετική (generative) μοντελοποίηση.....	11
2.1.3 Μέθοδοι NLP και ΑΙ για εξαγωγή μεταδεδομένων	14
2.1.4 Υβριδικά σενάρια και ποιότητα σε αρχεία	19
2.2 Πλαίσια και Πρότυπα μεταδεδομένων.....	21
2.3 Εφαρμογές TN σε πολιτιστική πληροφορία	27
Κεφάλαιο 3 Θεωρητικό Πλαίσιο	33
3.1 Μεταδεδομένα και Εξαγωγή Μεταδεδομένων	33
3.1.1 Structured vs unstructured extraction	35

3.1.2	Rule-based vs ML vs LLM.....	36
3.2	Αρχειακή Περιγραφή και ISAD(G).....	40
3.2.1	Βασικές αρχές.....	40
3.2.2	Ιεραρχικότητα.....	40
3.2.3	Οι επτά περιοχές της περιγραφής ISAD(G) - Υποχρεωτικά και προαιρετικά πεδία	41
3.2.4	Η σημασία της πιστότητας στο πρωτογενές κείμενο	44
3.3	Μεγάλα Γλωσσικά Μοντέλα (LLMs).....	44
3.3.1	Αρχιτεκτονική Transformer	45
3.3.2	Το πρόβλημα των hallucinations.....	46
3.3.3	Instruction following και prompt engineering.....	46
3.3.4	Αξιολόγηση δεδομένων και αποτελεσμάτων (Data quality metrics)	47
Κεφάλαιο 4 Δημιουργία αρχειακής περιγραφής με βάση το ΔΙΠΑΠ(Γ)/ISAD(G) με τη χρήση LLMs		53
4.1	Θεωρητικό Πλαίσιο και Μεθοδολογία	53
4.1.1	Δεδομένα εισόδου	53
4.1.2	Μεθοδολογική ροή.....	54
4.1.3	Μοντέλα που αξιολογήθηκαν	58
4.2	Σχεδιασμός Prompt (Prompt Design).....	58
4.3	Μελέτη Περίπτωσης: Αρχείο Επανήσου Πολιτείας.....	60
4.3.1	Περιγραφή της μελέτης περίπτωσης.....	60
4.3.2	Αποτελέσματα ανά μοντέλο	61
4.3.3	Συγκριτική επισκόπηση αποτελεσμάτων.....	65

Κεφάλαιο 5	Συμπεράσματα – Μελλοντική Έρευνα	69
5.1	Χρήση LLM ως βοηθητικό εργαλείο.....	69
5.2	Ανάγκη human validation.....	71
5.3	Προοπτικές για hybrid archival AI systems	72
	Βιβλιογραφία.....	75
	Παράρτημα Α Αρχειακή Περιγραφή βάσει δεδομένων και μεταδεδομένων Αρχείου Επιανήσου Πολιτείας.....	81
	Παράρτημα Β Prompt used for generating fonds – and sub-unit-level ISAD(G).....	85
	Παράρτημα Γ Ευρετήριο Όρων	87

Κατάλογος Σχημάτων

Σχήμα 1. Flow chart εξαγωγής αλγοριθμικών μεταδεδομένων από επιστημονικά άρθρα [NR23].....	9
Σχήμα 2. Αρχιτεκτονική συστήματος εξαγωγής αλγοριθμικών μεταδεδομένων από επιστημονικά άρθρα [NR23]	10
Σχήμα 3. Παράδειγμα εξαγωγής μεταδεδομένων από ακαδημαϊκό έγγραφο: εντοπισμός δομής, κεφαλαίων και μεταδεδομένων εξωφύλλου [RB24].	11
Σχήμα 4. Κύκλος ζωής μεταδεδομένων και βασικές δραστηριότητες σε AI-assisted πλαίσιο διαχείρισης [YFA+25].....	12
Σχήμα 5. Τα 15 βασικά στοιχεία του Dublin Core οργανωμένα ως απλό σύνολο πεδίων για βασική περιγραφική μεταδεδομένων και διαλειτουργικότητα μεταξύ ετερογενών συστημάτων.	21
Σχήμα 6. Ιεραρχική δομή και περιοχές περιγραφής του ISAD(G) [ICA00, Gil08].....	22
Σχήμα 7. Δομή EAD ως ιεραρχικό XML finding aid - αντιστοίχιση με τα επίπεδα ISAD(G) [Gil08]	23
Σχήμα 8. Βασικές κλάσεις και ιδιότητες του CIDOC-CRM - οντολογικό μοντέλο για τη διασύνδεση δεδομένων πολιτιστικής κληρονομιάς [Gil08, ATO20]	25

Κατάλογος Πινάκων

Πίνακας 1. Συγκριτική επισκόπηση προτύπων μεταδεδομένων ως προς τον χώρο εφαρμογής, τη δομή και τη συμβατότητα με LLMs [Gil08, ICA00, ATO20, LCD+20]	27
Πίνακας 2 Οι επτά περιοχές περιγραφής του ISAD(G) και τα βασικά τους στοιχεία [ICA00].....	43

Πίνακας 3 Μεγάλα Γλωσσικά Μοντέλα που χρησιμοποιήθηκαν στην Αξιολόγηση.....	56
Πίνακας 4. Αξιολόγηση της Ιεραρχικής Κατανόησης και της Δομικής Συμμόρφωσης μεταξύ Ανοικτών LLMs.	65
Πίνακας 5 Καλύτερα αποδίδοντα LLMs ανά περιοχή ISAD(G)	67

Κατάλογος Συντομογραφιών

ΣΥΝΤΟΜΟΓΡΑΦΙΑ	ΕΞΗΓΗΣΗ
AI	Artificial Intelligence
BERT	Bidirectional Encoder Representations from Transformers
CIDOC CRM	International Committee for Documentation
CKAN	Comprehensive Knowledge Archive Network
CNN	Convolutional Neural Network / Συνελικτικό Νευρωνικό Δίκτυο
CRF	Conditional Random Fields
CSV	Comma-Separated Values
DAMS	Digital Asset Management System

DC	Dublin Core
DCAT	Data Catalog Vocabulary
DCMI	Dublin Core Metadata Initiative
EAD	Encoded Archival Description
GDPR	General Data Protection Regulation
GPT	Generative Pre-trained Transformer
GROBID	GeneRation Of Bibliographic Data
GRU	Gated Recurrent Unit
ICA	International Council on Archives / Διεθνές Συμβούλιο Αρχείων
ICOM	International Council of Museums
ISAD(G)	International Standard Archival Description (General)
ISBN	International Standard Book Number
ISO	International Organization for Standardization
JSON	JavaScript Object Notation

KNN	K-Nearest Neighbors
LLM	Large Language Model / Μεγάλο Γλωσσικό Μοντέλο
LSTM	Long Short-Term Memory
MARC	Machine-Readable Cataloging
MARXML	Machine-Readable Cataloging in XML format
METS	Metadata Encoding and Transmission Standard
ML	Machine Learning / Μηχανική Μάθηση
MODS	Metadata Object Description Schema
NER	Named Entity Recognition / Αναγνώριση Ονοματικών Οντοτήτων
NLP	Natural Language Processing / Επεξεργασία Φυσικής Γλώσσας
OCR	Optical Character Recognition / Οπτική Αναγνώριση Χαρακτήρων
OWL	Web Ontology Language
PREMIS	Preservation Metadata: Implementation Strategies
RDF	Resource Description Framework

RLHF	Reinforcement Learning from Human Feedback
SKOS	Simple Knowledge Organization System
SVM	Support Vector Machine
XML	Extensible Markup Language
ΓΑΚ-ΑΝΚ	Γενικά Αρχεία του Κράτους - Αρχεία Νομού Κέρκυρας
ΔΙΠΑΠ(Γ)	Διεθνές Πρότυπο Αρχειακής Περιγραφής (Γενικό)

Κεφάλαιο 1 Εισαγωγή

Τα αρχεία συχνά αποτελούν τη συλλογική μνήμη μιας κοινωνίας, επειδή διατηρούν τεκμήρια που καταγράφουν πράξεις, αποφάσεις και διαδικασίες: έγγραφα, αλληλογραφία, φωτογραφίες, χάρτες, δημόσια αρχεία κ.ά. Για να μιλήσουμε όμως με ακρίβεια, χρειάζεται να ξεχωρίσουμε δύο έννοιες που συχνά συγχέονται, το τεκμήριο και το αρχείο.

Με τον όρο **τεκμήριο (record)** αναφερόμαστε σε ένα επιμέρους πληροφοριακό αποτύπωμα που παράγεται ή λαμβάνεται στο πλαίσιο μιας δραστηριότητας και διατηρείται ως μαρτυρία αυτής της δραστηριότητας. Η οπτική αυτή είναι κεντρική στη διαχείριση τεκμηρίων (records management), στην οποία η έμφαση δίνεται στην παραγωγή, καταγραφή και διαχείριση των τεκμηρίων ανεξαρτήτως μορφής (έντυπης ή ψηφιακής) [ISO16]. Παράλληλα, σε επίπεδο βασικού λεξιλογίου τεκμηρίωσης, το πρότυπο ISO 5127 ορίζει το εννοιολογικό πλαίσιο όρων που χρησιμοποιούνται σε βιβλιοθήκες, αρχεία και διαχείριση τεκμηρίων, στηρίζοντας τη διάκριση μεταξύ αντικειμένων πληροφορίας και των συστημάτων που τα διαχειρίζονται [ISO17].

Με τον όρο **αρχείο (archive)** δεν εννοούμε απλώς «μια συλλογή τεκμηρίων», αλλά ένα οργανωμένο σύνολο τεκμηρίων που αποκτά νόημα, κυρίως μέσα από το πλαίσιο παραγωγής και τις σχέσεις των τεκμηρίων μεταξύ τους. Με απλά λόγια, το τεκμήριο είναι η μονάδα, ενώ το αρχείο είναι το σύνολο που αποδίδει πλαίσιο και δομή. Αυτή η λογική «πλαισίου» αποτυπώνεται και στο διεθνές πρότυπο αρχειακής περιγραφής ISAD(G), το οποίο οργανώνει την περιγραφή ιεραρχικά και επιδιώκει να διατηρεί την ενότητα και το συγκείμενο του αρχειακού συνόλου [ICA00].

Ακολούθως, τα τεκμήρια αυτά αποκτούν νόημα μόνο όταν περιγράφονται σωστά, δηλαδή όταν ο αρχειονόμος δημιουργεί μεταδεδομένα που εξηγούν τι είναι το κάθε τεκμήριο, πότε δημιουργήθηκε, από ποιόν και σε ποιο πλαίσιο ανήκει [Gil08]. Τα μεταδεδομένα δεν είναι απλά «ετικέτες»· εξυπηρετούν πολλαπλές λειτουργίες, όπως το να περιγράφουν τι είναι (περιγραφικά), πώς οργανώνονται (δομικά), ποιος τα διαχειρίζεται (διοικητικά), πώς θα επιβιώσουν μακροπρόθεσμα (διατήρησης) και πώς αξιοποιούνται (χρήσης) [Gil08]. Χωρίς αυτές τις λειτουργίες, ένα αρχείο δεν είναι παρά μία αποθήκη γεμάτη τεκμήρια που κανείς δεν μπορεί να βρει, να κατανοήσει ή να εμπιστευτεί.

Η αρχειακή περιγραφή είναι μια εργασία που απαιτεί πολύ χρόνο και υψηλή εξειδίκευση. Κάθε τεκμήριο πρέπει να εξεταστεί προσεκτικά, να κατανοηθεί το περιεχόμενό του και το δομικό του πλαίσιο, και στη συνέχεια να συμπληρωθούν τα κατάλληλα πεδία σύμφωνα με αυστηρά διεθνή πρότυπα. Το ISAD(G), που εκδόθηκε από το Διεθνές Συμβούλιο Αρχείων (ICA), ορίζει 26 στοιχεία περιγραφής οργανωμένα σε 7 περιοχές, με ιεραρχική λογική από το αρχείο (fonds) μέχρι το τεκμήριο (item) [Gil08]. Στα ελληνικά αρχεία, η αντίστοιχη εφαρμογή γίνεται μέσω του ISAD(G). Ο αρχειονόμος καλείται να μελετήσει το υλικό, να εντοπίσει κρίσιμες πληροφορίες (παραγωγό, χρονολογία, θεματική, μορφή, έκταση), να αποφασίσει το κατάλληλο επίπεδο περιγραφής και να τις αποτυπώσει με συνέπεια στα πεδία του προτύπου. Σε συνθήκες αυξανόμενου όγκου τεκμηρίων και περιορισμένων ανθρώπινων και οικονομικών πόρων, αυτή η διαδικασία δημιουργεί εύκολα μια «ουρά» ακατέργαστου υλικού, τεκμήρια που περιμένουν περιγραφή για χρόνια ή και δεκαετίες, με αποτέλεσμα μεγάλα τμήματα των συλλογών να παραμένουν πρακτικά «αόρατα» [GWF24].

Το φαινόμενο αυτό δεν είναι καινούργιο. Ήδη από το 2005, οι Greene και Meissner διαπίστωσαν ότι τα αρχεία αντιμετωπίζουν «ανεπεξέργαστες συσσωρεύσεις». Η παραδοσιακή, λεπτομερής επεξεργασία δεν προλαβαίνει να καλύψει την εισροή νέου υλικού, με αποτέλεσμα σημαντικά σώματα τεκμηρίων να παραμένουν ουσιαστικά απρόσιτα [GM05]. Προτείνοντας τη φιλοσοφία «More Product, Less Process», υποστήριξαν ότι απαιτείται αναθεώρηση της αρχειακής εργασίας προς πιο «ελαφρές» μορφές περιγραφής, ώστε να αυξηθεί η προσβασιμότητα ακόμη και με κάποιο τίμημα στη λεπτομέρεια [GM05]. Η ένταση ανάμεσα στην ανάγκη για ολοκληρωμένη, υψηλής ποιότητας περιγραφή και στην πίεση για γρήγορη διάθεση περισσότερου υλικού παραμένει μέχρι σήμερα άλυτη· αντί να αμβλύνεται, εντείνεται όσο οι συλλογές, ιδίως οι «born-digital», συνεχίζουν να διογκώνονται.

Τα ψηφιακά παραγόμενα (born-digital) αρχεία (emails, ψηφιακά έγγραφα, social media) αυξάνονται εκθετικά, ενώ οι πόροι ψηφιοποίησης και περιγραφής παραμένουν σταθερά περιορισμένοι [JC22]. Σε πολλούς πολιτιστικούς οργανισμούς, η ψηφιοποίηση γίνεται αποσπασματικά, ως απάντηση σε μία άμεση ανάγκη (π.χ. ένας ερευνητής ζητά συγκεκριμένο υλικό και τότε μόνο ψηφιοποιείται), αντί να ακολουθείται μια στρατηγική προτεραιοτήτων και προγραμματισμού [GWF24]. Παράλληλα, τα μεταδεδομένα στον χώρο της πολιτιστικής κληρονομιάς πάσχουν από γνωστά προβλήματα όπως να είναι ελλιπή ή ασυνεπή, χρησιμοποιούν διαφορετικά πρότυπα από φορέα σε φορέα και συχνά δεν περιγράφουν το περιεχόμενο με αρκετό «νόημα» ώστε να υποστηρίξουν την αποτελεσματική αναζήτηση και ανάκτηση [ATO20]. Σε εκτενή ανασκόπηση 78 εργασιών, εντοπίστηκε ότι αυτά τα ζητήματα δεν είναι απλώς τεχνικά, αλλά αφορούν τη

δυνατότητα των χρηστών (ερευνητών, πολιτών, εκπαιδευτικών) να βρουν και να κατανοήσουν τα πολιτιστικά τεκμήρια [ATO20]. Η ανάγκη λοιπόν για εργαλεία που μπορούν να επιταχύνουν τη δημιουργία μεταδεδομένων χωρίς να θυσιάζουν την ακρίβεια δεν ήταν ποτέ πιο επείγουσα.

Σε αυτό το σημείο εισέρχεται η Τεχνητή Νοημοσύνη (AI). Τα τελευταία χρόνια, η εξέλιξη στην Επεξεργασία Φυσικής Γλώσσας (NLP) και στη Μηχανική Μάθηση (ML) έχει ανοίξει νέες δυνατότητες για την αυτοματοποίηση εργασιών που σχετίζονται με μεταδεδομένα [YFA+25]. Τεχνικές όπως η αναγνώριση οντοτήτων (NER), η αυτόματη ταξινόμηση κειμένων και η εξαγωγή μεταδεδομένων μέσω κανόνων ή εποπτευόμενης μάθησης χρησιμοποιούνται πλέον σε πολλά πεδία: στη βιβλιογραφική εξαγωγή από ακαδημαϊκά άρθρα [NR23, GHK+12], στην αυτόματη σήμανση ερευνητικών δεδομένων [WBM+23] και στην αξιολόγηση της ποιότητας των περιγραφών σε πολιτιστικά ιδρύματα [LRT21]. Ένα παράδειγμα από τον αρχειακό χώρο αποτελεί η εργασία των Mardiaty κ.ά. (2023), οι οποίοι εξερευνούν τη δυνατότητα χρήσης AI για τη συλλογή μεταδεδομένων σε δραστηριότητες περιγραφής φωτογραφικών αρχείων, καταδεικνύοντας ότι η αυτοματοποίηση μπορεί να εφαρμοστεί και σε εξειδικευμένες αρχειακές συλλογές [MAG+23]. Τα αποτελέσματα αυτών των προσεγγίσεων είναι ενθαρρυντικά, αλλά έχουν σαφή όρια. Κάθε σύστημα NLP/ML χρειάζεται εκπαίδευση σε εξειδικευμένα annotated datasets, σύνολα δεδομένων με χειροκίνητη επισήμειωση από ανθρώπους που έχουν «μαρκάρει» σε κάθε παράδειγμα ποιο είναι το σωστό πεδίο, κάτι που στον αρχειακό χώρο σπανίζει τόσο σε ποσότητα όσο και σε ομοιογένεια [WBM+23]. Επίσης, τα περισσότερα συστήματα σχεδιάζονται για σχετικά επίπεδα σχήματα μεταδεδομένων (π.χ. Dublin Core) και όχι για σύνθετα, ιεραρχικά πρότυπα όπως το ISAD(G).

Η πραγματική αλλαγή παραδείγματος ήρθε με τα Μεγάλα Γλωσσικά Μοντέλα (Large Language Model – LLMs). Μοντέλα όπως τα GPT-4, Claude, LLaMA ή Mistral βασίζονται στην αρχιτεκτονική Transformer που εισήγαγαν οι Vaswani κ.ά. (2017) [VSP+17] και εκπαιδεύονται σε τεράστια σώματα κειμένων, αποκτώντας γενικές αναπαραστάσεις της γλώσσας. Έτσι, αντί να απαιτούν ειδική εκπαίδευση για κάθε νέο dataset, μπορούν μέσω κατάλληλων prompts, να κατανοούν εντολές, να αναλύουν κείμενα και να παράγουν δομημένα αποτελέσματα [Bag24]. Για τα αρχεία, αυτό σημαίνει ότι ένα LLM θα μπορούσε θεωρητικά να «διαβάσει» ένα πρωτογενές έγγραφο ή μια ήδη υπάρχουσα αδόμητη περιγραφή και να παράγει μια αρχειακή εγγραφή σύμφωνα με τα πεδία του ISAD(G), με ελάχιστη ή χωρίς ανθρώπινη παρέμβαση. Πρόσφατα, οι Groppe κ.ά. [GMW+25], πρότειναν μια προσέγγιση ομοσπονδιακής νοημοσύνης (federated intelligence) μεταξύ πολλαπλών LLMs για την αυτόματη παραγωγή αρχειακών περιγραφών, δείχνοντας ότι ο συνδυασμός μοντέλων μπορεί να αντισταθμίσει τις

αδυναμίες κάθε μεμονωμένου συστήματος. Παράλληλα, οι Yang κ.ά. [YFA+25], καταγράφουν ότι τα LLMs και γενικότερα τα GenAI συστήματα ήδη χρησιμοποιούνται σε ροές εργασίας μεταδεδομένων για αυτόματη παραγωγή, εμπλουτισμό και επικύρωση, σε κλίμακες που ήταν πρακτικά αδύνατες με τις κλασικές τεχνικές. Για μικρούς πολιτιστικούς φορείς χωρίς πόρους για εκπαίδευση ή fine-tuning εξειδικευμένων μοντέλων, η αξιοποίηση έτοιμων LLMs μέσω zero-shot ή few-shot prompting είναι ιδιαίτερα ελκυστική [GWF24].

Η δυνατότητα, όμως, δεν σημαίνει αυτομάτως αξιοπιστία, και εδώ έγκειται το βασικό ερευνητικό κενό που επιχειρεί να φωτίσει η παρούσα εργασία. Το πιο γνωστό και πολυσυζητημένο πρόβλημα των LLMs είναι οι «ψευδαισθήσεις» (hallucinations) οι οποίες είναι η παραγωγή πληροφοριών που φαίνονται πειστικές αλλά δεν αντιστοιχούν στην πραγματικότητα [RSD23]. Στο πλαίσιο μιας αρχειακής περιγραφής, αυτό μπορεί να σημαίνει ότι το μοντέλο «εφευρίσκει» παραγωγούς, χρονολογίες ή γεγονότα με βάση κάποια στερεοτυπικά μοτίβα, χωρίς αυτά να τεκμηριώνονται όμως στο πρωτογενές υλικό.

Εξίσου σημαντικό είναι το πρόβλημα της δομικής ασυνέπειας, το οποίο συμβαίνει, αν η δομή εξόδου από το LLM θα είναι σταθερή μετά από κάθε εκτέλεση. Αυτό λοιπόν δεν είναι απόλυτα σίγουρο ότι θα οδηγήσει σε σωστό αποτέλεσμα και θα τηρηθεί πιστά η δομή του ή θα συμπληρωθούν όλα τα πεδία με συνεπή τρόπο, ακόμα και αν το μοντέλο λάβει πολύ συγκεκριμένες οδηγίες [WBM+23]. Επίσης τα LLMs αναπαράγουν και μεροληψίες (biases) που υπάρχουν στα δεδομένα εκπαίδευσής τους και σύμφωνα με τους Griffin κ.ά. [GWF24], δείχνουν χαρακτηριστικά πως αυτό μπορεί να οδηγήσει σε λανθασμένες αποφάσεις ακόμα και σε απλές εργασίες περιγραφής που στην πραγματικότητα ο αρχειονόμος δεν θα έκανε ποτέ. Ακόμα υπάρχει η δυσκολία κατανόησης του πλαισίου: τα αρχεία δεν είναι τυποποιημένα, κάθε τεκμήριο μπορεί να απαιτεί διαφορετική ανάγνωση, και η «μηχανή» δεν αντιλαμβάνεται πάντα τις λεπτές αποχρώσεις που ο «άνθρωπος», ο αρχειονόμος αντλεί μέσα από το πλαίσιο δημιουργίας [JC22]. Τέλος, όπως επισημαίνουν οι Toth κ.ά. [TAP25], σε μελέτη για την εφαρμογή explainable AI σε αρχειακή πολιτιστική κληρονομιά, η ερμηνευσιμότητα των αποτελεσμάτων είναι εξίσου κρίσιμη με την ακρίβεια, δηλαδή δεν αρκεί ένα μοντέλο να δίνει «σωστές» απαντήσεις, ο αρχειονόμος πρέπει να μπορεί να καταλάβει γιατί το μοντέλο κατέληξε σε αυτό το αποτέλεσμα. Αν δεν το καταλαβαίνει, δεν μπορεί ούτε να το ελέγξει ούτε να το εγκρίνει με εμπιστοσύνη [TAP25]. Σε έναν χώρο όπου η ακρίβεια, η πιστότητα στο πρωτογενές κείμενο και η συμμόρφωση με πρότυπα είναι θεμελιώδεις αρχές, αυτά τα ζητήματα δεν είναι δευτερεύοντα. Αν τα μεταδεδομένα που παράγει ένα

μοντέλο οδηγούν σε λάθος ερμηνεία ή δυσχεραίνουν τον εντοπισμό τεκμηρίων, η ζημιά μπορεί να είναι μεγαλύτερη από την πλήρη απουσία περιγραφής.

Παράλληλα, τα θεσμικά και ρυθμιστικά πλαίσια δεν έχουν προλάβει αυτήν την τεχνολογική εξέλιξη. Οι Svärd κ.ά. [SGB+24], σε βιβλιογραφική ανασκόπηση τοπικών κανονισμών για τη χρήση ΑΙ στη διαχείριση δημοσίων αρχείων σε Σουηδία, Φινλανδία και Νότια Αφρική, διαπιστώνουν ότι σε πολλές χώρες δεν υπάρχει ακόμη σαφές νομοθετικό πλαίσιο· τα αρχεία πειραματίζονται με ΑΙ σε ένα κανονιστικό κενό [SGB+24]. Οι κατευθυντήριες γραμμές της Ευρωπαϊκής Επιτροπής για «αξιόπιστη ΑΙ» ελάχιστα αναφέρονται ρητά στη διαχείριση δημοσίων αρχείων, παρά τις κρίσιμες επιπτώσεις που μπορεί να έχει μια αυτοματοποιημένη λανθασμένη περιγραφή σε ζητήματα πρόσβασης, διαφάνειας και προστασίας προσωπικών δεδομένων [SGB+24]. Αυτό όμως εγείρει σοβαρά ερωτήματα λογοδοσίας. Αν ένα LLM παράγει εσφαλμένα ή παραπλανητικά μεταδεδομένα, ποιος έχει την ευθύνη; Ο φορέας που το χρησιμοποίησε, ο πάροχος του μοντέλου, ή ο εκάστοτε αρχειονόμος που ενέκρινε την έξοδο;

Οι Alma'aitah κ.ά. [ATO20], υπογραμμίζουν ότι οι αδυναμίες μεταδεδομένων δεν είναι μόνο τεχνικό ζήτημα. Επηρεάζουν άμεσα τη δυνατότητα των χρηστών να εντοπίζουν και να ερμηνεύουν τεκμήρια, αλλά και τη διαλειτουργικότητα μεταξύ των συστημάτων [ATO20]. Η παρούσα εργασία λοιπόν τοποθετείται μέσα σε αυτό το πλαίσιο, ανάμεσα στην ανάγκη αυτοματοποίησης της εξαγωγής μεταδεδομένων σε αρχειακές πηγές, στις εντυπωσιακές δυνατότητες των LLMs αλλά και στα αναπάντητα ερωτήματα αξιοπιστίας τους. Μέσα σε αυτό το περιβάλλον διαμορφώνονται τα τρία βασικά ερευνητικά ερωτήματα που ακολουθούν:

- i. Μπορούν τα LLMs να μετατρέψουν περιγραφές αρχειακού ελεύθερου κειμένου σε περιγραφικά στοιχεία ευθυγραμμισμένα με το ISAD(G); Δηλαδή, αν τους δοθεί ένα κείμενο ή στοιχεία από ένα αρχειακό τεκμήριο, μπορούν να συμπληρώσουν τα πεδία που ζητάει το πρότυπο με ακρίβεια και πληρότητα;
- ii. Συμμορφώνονται με αυστηρούς κανόνες χωρίς παρεκκλίσεις; Ακολουθούν πιστά τις οδηγίες που τους δίνονται μέσω prompts ή «αυτοσχεδιάζουν», προσθέτουν πληροφορίες που δεν υπάρχουν στο πρωτογενές κείμενο, αλλάζουν τη δομή, ή παραλείπουν πεδία; Με πόση ακρίβεια, λοιπόν, διαφορετικά LLMs διατηρούν την αρχειακή ιεραρχία κατά τον μετασχηματισμό αυτό;
- iii. Πώς διαφοροποιείται η απόδοση ανά μέγεθος μοντέλου; Τα μεγαλύτερα μοντέλα (π.χ. 70B παραμέτρων) δίνουν καλύτερα αποτελέσματα, ή μπορεί ένα μικρότερο, ανοιχτού κώδικα μοντέλο (π.χ. 7B ή 13B) να πλησιάσει ικανοποιητικά κάτι που θα ήταν πρακτικά εφαρμόσιμο σε φορείς με περιορισμένους πόρους;

Τα ερευνητικά ερωτήματα πηγάζουν από μια κεντρική παρατήρηση. Ενώ η βιβλιογραφία περιλαμβάνει ολοένα και περισσότερες εργασίες για τη χρήση τεχνητής νοημοσύνης στα μεταδεδομένα, από rule-based εξαγωγή [AA18] μέχρι deep learning [NR23] και generative AI [Bag24], ελάχιστες εστιάζουν ειδικά στα αρχεία και ακόμα λιγότερες αξιολογούν συστηματικά τη συμμόρφωση με αρχειακά πρότυπα. Στόχος της παρούσας έρευνας δεν είναι η ανάπτυξη ενός νέου εργαλείου, αλλά η συστηματική μελέτη της υπάρχουσας βιβλιογραφίας σχετικά με τη χρήση τεχνητής νοημοσύνης και ιδιαίτερα των τεχνικών εξαγωγής μεταδεδομένων μέσω NLP, ML και LLMs στον χώρο της πολιτιστικής πληροφορίας, με εφαρμογή στα αρχεία [ATO20, Pav22]. Μέσα από αυτήν τη βιβλιογραφική ανασκόπηση, η εργασία επιδιώκει να χαρτογραφήσει τι λειτουργεί, τι δεν λειτουργεί, και κυρίως τι λείπει, ώστε να θέσει τις βάσεις για ένα πειραματικό κεφάλαιο (Κεφάλαιο 4) όπου ανοιχτά LLMs δοκιμάζονται πρακτικά στην παραγωγή αρχειακών περιγραφών κατά ISAD(G) και τα αποτελέσματά τους αξιολογούνται από ομάδα ειδικών αρχειονόμων σε σύγκριση με ανθρώπινες εγγραφές.

Η εργασία οργανώνεται ως εξής. Στο Κεφάλαιο 2, παρουσιάζεται η βιβλιογραφική επισκόπηση, η οποία καλύπτει τις μεθόδους NLP/AI που έχουν χρησιμοποιηθεί ή χρησιμοποιούνται στην εξαγωγή μεταδεδομένων, στις τεχνικές εξαγωγής (rule-based, ML, deep learning, LLMs), στα πλαίσια και πρότυπα μεταδεδομένων (Dublin Core, ISAD(G), EAD, CIDOC-CRM), και τις εφαρμογές τεχνητής νοημοσύνης στον χώρο της πολιτιστικής πληροφορίας. Στο Κεφάλαιο 3 αναπτύσσεται το θεωρητικό πλαίσιο. Ορισμοί μεταδεδομένων και εξαγωγής μεταδεδομένων, η αρχειακή περιγραφή κατά ISAD(G), τα Μεγάλα Γλωσσικά Μοντέλα και η σχέση τους με την πολιτιστική κληρονομιά. Στο Κεφάλαιο 4 παρουσιάζεται το πρακτικό σκέλος, δηλαδή, η μεθοδολογία, ο σχεδιασμός του prompt και τα αποτελέσματα της μελέτης περίπτωσης, εκτιμώντας την ποιότητα και τη συνέπεια του παραγόμενου κειμένου πεδίου, ευθυγραμμισμένου με το ISAD(G) και συζητώντας τις πρακτικές επιπτώσεις για τη δημιουργία εργαλείων ανεύρεσης βασισμένων σε πρότυπα και για τη μείωση των καθυστερήσεων. Τέλος, στο Κεφάλαιο 5 κλείνει η εργασία με συμπεράσματα και προτάσεις για μελλοντική έρευνα.

Κεφάλαιο 2 Βιβλιογραφική Επισκόπηση

2.1 Τεχνικές εξαγωγής μεταδεδομένων

Στη βιβλιογραφία, η εξαγωγή μεταδεδομένων δεν υλοποιείται ως ένα «μαύρο κουτί», αλλά ως αλυσίδα από σχετικά απλά βήματα που συνεργάζονται. Τυπικά, ξεκινά με την απόκτηση και προετοιμασία των δεδομένων (π.χ. σαρώσεις και OCR για χαρτώα αρχειακά τεκμήρια, εξαγωγή κειμένου από PDF, εξαγωγή headers από email), συνεχίζει με προ-επεξεργασία (καθαρισμός κειμένου, tokenization, κανονικοποίηση, απομάκρυνση θορύβου) και περνά σε κάποιο στάδιο «νοημοσύνης» (κανόνες, μοντέλα μηχανικής μάθησης ή deep learning). Στο τέλος υπάρχει σχεδόν πάντα ένα στάδιο μετα-επεξεργασίας, κατά το οποίο τα αποτελέσματα διορθώνονται, κανονικοποιούνται, ελέγχονται ως προς ένα συγκεκριμένο σχήμα (π.χ. EAD, Dublin Core, εθνικά αρχειακά πρότυπα) και ενίοτε υποβάλλονται σε ανθρώπινο έλεγχο [AA18].

2.1.1 Κλασικές rule-based και ML pipelines

Ήδη από το 2008, ο Underwood [Und08] εξέτασε τη δυνατότητα αυτόματης εξαγωγής μεταδεδομένων ειδικά για αρχειακές περιγραφές, αναδεικνύοντας τόσο τις δυνατότητες όσο και τους περιορισμούς των τεχνικών της εποχής. Τα πρώτα συστήματα ήταν καθαρά βασισμένα πάνω σε κανόνες, υπό την έννοια ότι η αναγνώριση των πεδίων των μεταδεδομένων γινόταν αποκλειστικά μέσω ρητών κανόνων τύπου «αν...τότε...» που είχαν οριστεί από τον σχεδιαστή του συστήματος. Για παράδειγμα, μπορεί να ορίζεται ότι «αν το κείμενο με τη μεγαλύτερη γραμματοσειρά τοποθετημένο πριν τη λέξη Abstract είναι σε bold, τότε αυτό το τμήμα κειμένου είναι ο τίτλος», ή ότι «αν μετά τη φράση "Keywords:" ακολουθεί μια λίστα όρων χωρισμένων με κόμματα, τότε αυτοί οι όροι αντιστοιχούν στις λέξεις-κλειδιά». Τέτοιου τύπου κανόνες, βασισμένοι σε απλά μοτίβα διάταξης και κειμένου, συναποτελούν το «μοντέλο» του συστήματος και εκτελούνται αυτούσια χωρίς διαδικασία εκπαίδευσης ή στατιστικής γενίκευσης [AA18]. Οι Azimjonov και Alikhanov [AA18] περιγράφουν χαρακτηριστικά μια τέτοια pipeline για ακαδημαϊκά άρθρα, όπου συλλέγονται μαζικά PDFs, εντοπίζονται οι πρώτες/τελευταίες σελίδες, εφαρμόζονται κανόνες για τίτλο, abstract, keywords, κύριο κείμενο, συμπεράσματα και βιβλιογραφία και, στο τέλος, τα εξαγόμενα πεδία αποθηκεύονται σε βάση και XML. Το σημαντικό τους εύρημα είναι ότι με προσεκτικά σχεδιασμένους

κανόνες και ισχυρό στάδιο μετα-επεξεργασίας (π.χ. χειρισμός εξαιρέσεων, fallback σε χειροκίνητο review) μπορούν να φτάσουν σε ακρίβειες άνω του 90% για βασικά πεδία, ξεπερνώντας εργαλεία όπως GROBID και ParsCit σε συγκεκριμένα corpora [AA18].

Αν μεταφέρουμε αυτή τη λογική σε ένα αρχείο, ένα αντίστοιχο rule-based workflow θα μπορούσε να είναι ομαδική επεξεργασία σαρωμένων φακέλων, OCR, εφαρμογή κανόνων τύπου «αν στην πρώτη γραμμή εμφανίζεται "ΥΠΟΥΡΓΕΙΟ..." και ακριβώς από κάτω ημερομηνία, τότε αυτά είναι ο παραγωγός και η χρονολογία», «αν εμφανίζεται το string "Αρ. πρωτ.", τότε τα επόμενα strings μέχρι το τέλος της γραμμής είναι ο αριθμός πρωτοκόλλου» κ.λπ. · στη συνέχεια, τα πεδία αυτά χαρτογραφούνται σε μια EAD ή ISAD(G) δομή σε επίπεδο τεκμηρίου. Σε καλά τυποποιημένες σειρές (π.χ. διοικητικά έγγραφα ενός φορέα με σταθερή μακέτα), τέτοιες pipelines μπορούν να δώσουν «βιομηχανική» απόδοση, αλλά είναι ευαίσθητες σε αλλαγές φόρμας και απαιτούν συνεχή συντήρηση των κανόνων [Gil08].

Η επόμενη γενιά pipelines χρησιμοποίησε συστηματικά μοντέλα μηχανικής μάθησης. Οι Granitzer κ.ά. [GHK+12] συγκρίνουν τεχνικές όπως Conditional Random Fields (CRF) και Support Vector Machines (SVM) για την εξαγωγή βιβλιογραφικών μεταδεδομένων σε ρεαλιστικά, «βρώμικα - ακαθάριστα» σύνολα από κοινωνικά δίκτυα έρευνας. Η pipeline τους περιλαμβάνει πρώτα αναλυτική εξαγωγή κειμένου και layout (διάταξη) από PDF, ώστε κάθε λέξη ή γραμμή να συνοδεύεται όχι μόνο από το περιεχόμενό αλλά και από χαρακτηριστικά όπως μέγεθος, οικογένεια γραμματοσειράς, πλάτος γραμματοσειράς, μέση απόσταση χαρακτήρων, x και y συντεταγμένες στη σελίδα, καθώς και τις «κλίσεις» αυτών των συντεταγμένων, δηλαδή αν το κείμενο κινείται προς τα δεξιά ή προς τα κάτω. Στη συνέχεια το κείμενο του άρθρου μετατρέπεται σε ακολουθίες tokens, όπου κάθε token είναι ένα d -tuple «κείμενο και 34 layout/ορθογραφικά χαρακτηριστικά καθώς και ετικέτα πεδίου». Πάνω σε αυτές τις ακολουθίες εκπαιδεύονται μοντέλα σε CRFs (sequence labeling) και γραμμικά SVMs, τα οποία αποφασίζουν για κάθε token αν ανήκει σε τίτλο, σε όνομα συγγραφέα, σε πεδίο περιοδικού ή σε «απλό» κείμενο. Αφού όλα τα tokens έχουν πάρει μια αρχική ετικέτα, ακολουθεί ένα στάδιο μετα-επεξεργασίας όπου εφαρμόζονται domain-specific heuristics για να λυθούν οι ασάφειες: για παράδειγμα, κανόνες που απαιτούν ο τίτλος να είναι ενιαίο συνεχές μπλοκ, οι συγγραφείς να εμφανίζονται μετά τον τίτλο και πριν από την περίληψη, ή που συγχωνεύουν συνεχόμενες γραμμές με παρόμοιο layout σε ένα πεδίο, ώστε να αποφευχθούν διπλές/αντικρουόμενες ετικέτες. Οι συγγραφείς επισημαίνουν ότι αυτά τα post-processing heuristics, οι απλοί εμπειρικοί-ευρετικοί κανόνες που ορίζονται από τον σχεδιαστή για να λυθούν ασάφειες, είναι απαραίτητοι. Χωρίς αυτά, η απόδοση των

καθαρά στατιστικών μοντέλων πέφτει αισθητά. Συνοψίζοντας, το πρόβλημα της βιβλιογραφικής εξαγωγής μεταδεδομένων σε crowdsourced περιβάλλον (Mendeley, CiteULike κ.λπ.) μπορεί να αντιμετωπιστεί αποτελεσματικά με έναν συνδυασμό κατάλληλων χαρακτηριστικών διάταξης (layout features), ενός ισχυρού ταξινομητή (two-stage SVM) και καλά οριοθετημένων μετα-επεξεργαστικών κανόνων. Δηλαδή δεν αρκεί ένα καλό μοντέλο, αλλά χρειάζεται και το σωστό encoding της διάταξης του εγγράφου και ένα επίπεδο κανόνων πάνω από τα στατιστικά αποτελέσματα [GHK+12].

Σε αρχειακό πλαίσιο, ανάλογη ML pipeline μπορεί να εφαρμοστεί σε μαζικές ψηφιοποιήσεις φακέλων ή σε ψηφιακά παραγόμενο υλικό (π.χ. email). Ενδεικτικά, σε ένα αρχείο email θα μπορούσε να εφαρμοστεί workflow στο οποίο γίνεται parsing των headers (From, To, Date, Subject), των συνημμένων και του σώματος του μηνύματος, εξάγονται χαρακτηριστικά (features) όπως θέση, μοτίβα τύπου «Re:», «Fwd:» και υπογραφές, και εκπαιδεύεται classifier για να ξεχωρίζει «επίσημα» από «μη επίσημα» μηνύματα ή για να αποδίδει θεματικές επικεφαλίδες. Στο τελικό στάδιο, εφαρμόζονται κανόνες για τη δημιουργία αρχειακών εγγραφών σε επίπεδο φακέλου ή σειράς κατά ISAD(G) (π.χ. ομαδοποίηση κατά project ή αποστολέα), ακολουθώντας τη λογική των layout-based pipelines με classifiers που περιγράφουν οι Granitzer κ.ά. [GHK+12], προσαρμοσμένη στις ιδιαιτερότητες των born-digital αρχειακών συνόλων [CBJ+21].

Οι Nayaka και Ranjan [NR23] πηγαίνουν ένα βήμα παραπέρα, προτείνοντας pipeline με deep νευρωνικά δίκτυα για την εξαγωγή «αλγοριθμικών» μεταδεδομένων

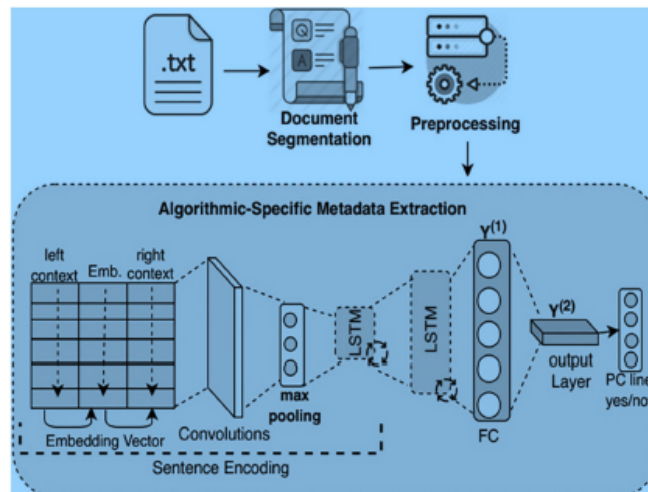


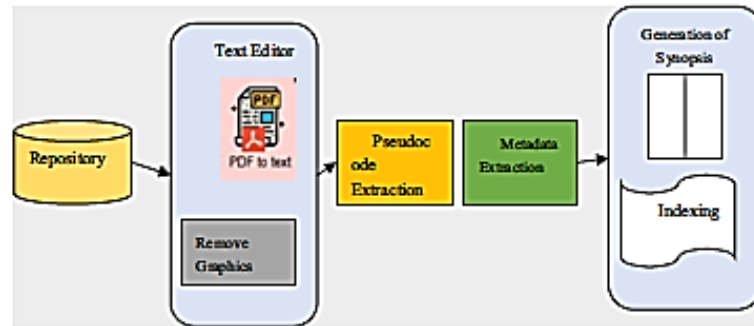
Fig. 3 Proposed system architecture

SN Computer Science
A SpringerOpen Journal

Σχήμα 1. Flow chart εξαγωγής αλγοριθμικών μεταδεδομένων από επιστημονικά άρθρα [NR23]

(ψευδοκώδικες, datasets, μετρικές, πολυπλοκότητα) από επιστημονικά άρθρα. Εδώ η ροή είναι: parsing PDF → δημιουργία γραμμών με περίπου 58 χαρακτηριστικά (περιεχόμενο, στυλ, δομή) → ταξινόμηση γραμμών με Random Forest/deep δίκτυα → δεύτερο στάδιο για εξαγωγή προτάσεων με αλγοριθμικά μεταδεδομένα (Σχ. 1,2). Σε αρχειακό πλαίσιο,

Fig. 1 Proposed flow chart

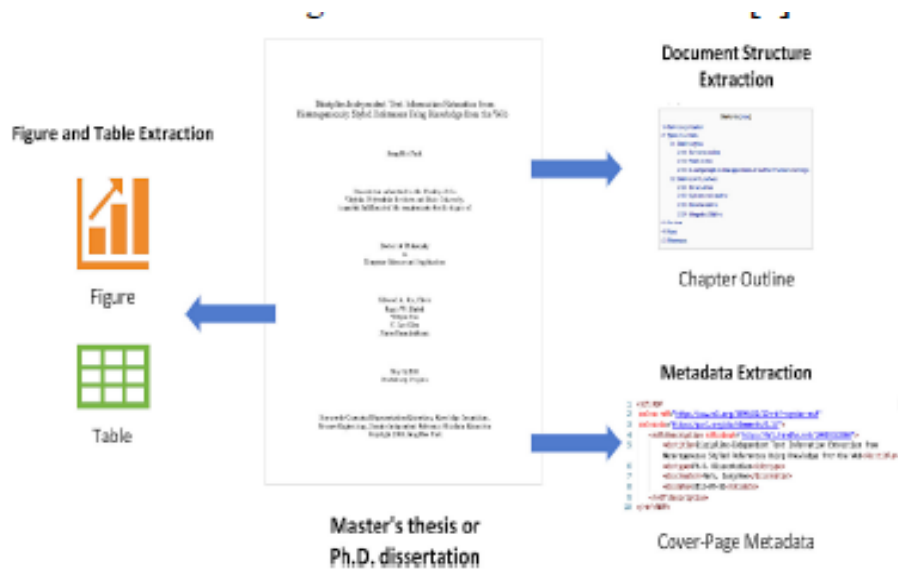


SN Computer Science
A SCIENCE DIRECT JOURNAL

Σχήμα 2. Αρχιτεκτονική συστήματος εξαγωγής αλγοριθμικών μεταδεδομένων από επιστημονικά άρθρα [NR23]

ανάλογη προσέγγιση θα μπορούσε να χρησιμοποιηθεί για να εντοπίζονται αυτόματα πληροφορίες όπως τα εμπλεκόμενα όργανα, οι αρμοδιότητες ή οι ενημερώσεις, εμπλουτίζοντας έτσι τα διαχειριστικά μεταδεδομένα μιας σειράς χωρίς ο αρχειονόμος να χρειάζεται να διαβάσει χειροκίνητα εκατοντάδες σελίδες, ένα πρόβλημα που οι Greene και Meissner [GM05] είχαν επισημάνει ως κεντρική πρόκληση για τα αρχεία.

Οι Raval και Bhaidasna [RB24] συνοψίζουν ότι αυτές οι pipelines έχουν πλέον ωριμάσει σε πολλούς τομείς, αλλά τα ανοιχτά προβλήματα παραμένουν σημαντικά. Δεν υπάρχουν κοινά benchmarks, δηλαδή τυποποιημένα σύνολα δεδομένων και μετρικές αξιολόγησης που να επιτρέπουν αντικειμενική σύγκριση μεταξύ συστημάτων. Κάθε ερευνητής δοκιμάζει το σύστημά του σε διαφορετικά δεδομένα και με διαφορετικές μετρικές, με αποτέλεσμα να είναι δύσκολο να κριθεί ποια προσέγγιση αποδίδει πραγματικά καλύτερα. Ακόμη, υπάρχουν δυσκολίες με θορυβώδη ή ελλιπή PDFs, ελλιπής υποστήριξη για πολυτροπικά δεδομένα (κείμενο, εικόνες, πίνακες) και υψηλό υπολογιστικό κόστος για τα πιο προηγμένα μοντέλα. Παρόμοια, οι Wu κ.ά. [WBM+23] επισημαίνουν ότι πιο «ερμηνευτικά» στοιχεία, όπως υπαινιγμοί ή σχέσεις μεταξύ εγγράφων, εξακολουθούν να διαφεύγουν από τα τυπικά ML μοντέλα, κάτι που σημαίνει ότι η ανθρώπινη κρίση δεν μπορεί να απουσιάζει. Οι Lorenzini κ.ά. [LRT21] το επιβεβαιώνουν αυτό από διαφορετική οπτική: ακόμη και μετά την εξαγωγή, χρειάζονται ξεχωριστά εργαλεία για να αξιολογείται η ποιότητα των περιγραφών διακρίνοντας, για παράδειγμα, «υψηλής» από «χαμηλής» ποιότητας εγγραφές σε συλλογές πολιτιστικής κληρονομιάς. (Σχ.3)



Σχήμα 3. Παράδειγμα εξαγωγής μεταδεδομένων από ακαδημαϊκό έγγραφο: εντοπισμός δομής, κεφαλαίων και μεταδεδομένων εξωφύλλου [RB24].

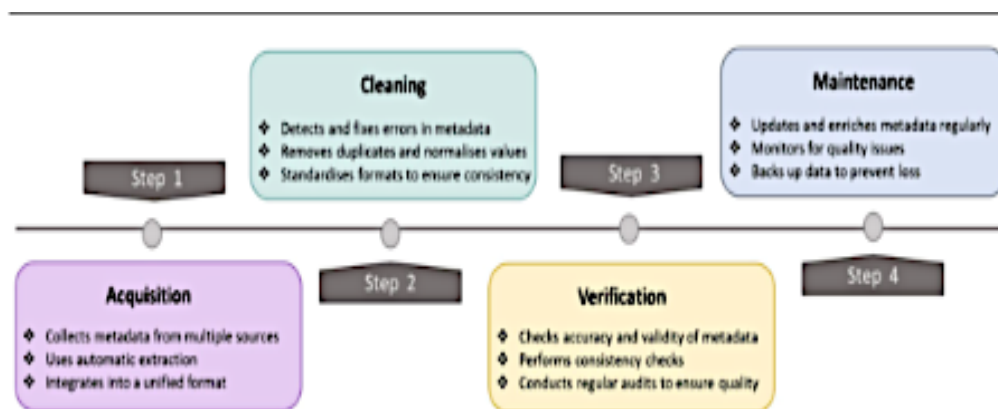
2.1.2 Pipelines με LLMs και γενετική (generative) μοντελοποίηση

Με την είσοδο των LLMs, ένα μεγάλο μέρος της «νοημοσύνης» μεταφέρεται στο ίδιο το μοντέλο. Αντί το σύστημα να εντοπίζει τμήματα κειμένου και να τα ταξινομεί σε πεδία μέσα από εκπαιδευμένους ταξινομητές, το LLM λαμβάνει το πλήρες κείμενο και ίσως κάποια ήδη υπάρχοντα πεδία, και καλείται, μέσω ειδικών prompts, να παράξει απευθείας μια δομημένη εγγραφή μεταδεδομένων σύμφωνα με ένα περιγραφικό μοντέλο. Αυτό σημαίνει ότι στη βιβλιογραφία διακρίνονται τρεις βασικές ερευνητικές κατευθύνσεις για την αξιοποίηση των LLMs στα μεταδεδομένα, που αξίζει να παρουσιαστούν.

Στην πρώτη και στην πιο άμεση κατεύθυνση που αφορά στην αυτόματη εξαγωγή και συμπλήρωση μεταδεδομένων, δίνεις στο μοντέλο ένα κείμενο και αυτό παράγει δομημένα πεδία, έτοιμα για καταχώριση σε ένα σχήμα. Ο Bagchi [Bag24] αντιμετωπίζει τα μεταδεδομένα ως αποτέλεσμα συνολικής κατανόησης του κειμένου από το μοντέλο και όχι ως απλά εξαγόμενες ετικέτες μεταδεδομένων. Σε ένα αρχειακό περιβάλλον η αντίστοιχη pipeline περιλαμβάνει ελαφρά προ-επεξεργασία (καθαρισμός, segmentation σε επίπεδο τεκμηρίου ή φακέλου), κλήση LLM με οδηγίες τύπου «παράγαγε μεταδεδομένα σύμφωνα με το τάδε schema» ή «εντόπισε δημιουργό, ημερομηνία, γεωγραφικές αναφορές», ισχυρό validation/normalization (έλεγχος συμμόρφωσης με το

schema, χρήση ελεγχόμενων λεξιλογίων, συνέπεια τιμών μεταξύ πεδίων, εντοπισμός προφανών hallucinations) και επλεκτική ανθρώπινη επιθεώρηση για «δύσκολες» ή ευαίσθητες περιπτώσεις [RB24]. Σε αρχειακό πλαίσιο, αυτό μπορεί να σημαίνει ότι, αντί να περνά κάθε έγγραφο από κανόνες και classifiers, δημιουργείται ένα self-contained prompt που περιέχει το κείμενο ενός φακέλου (ή αντιπροσωπευτικά τεκμήρια μιας σειράς) και το LLM καλείται να προτείνει τίτλο, χρονολογικά όρια, σύντομη περιγραφή και θεματικές επικεφαλίδες. Τα αποτελέσματα ελέγχονται αυτόματα (π.χ. αν οι ημερομηνίες είναι εντός των ορίων της σειράς) και, όπου χρειάζεται, χειρωνακτικά [LRT21]. Πρόσφατα συστήματα που αξιοποιούν LLMs, όπως το MOLE για επιστημονικά άρθρα [AAG25] ή σχήματα ομοσπονδιακής νοημοσύνης για αρχειακές περιγραφές [GMW+25], δείχνουν πώς τα μεγάλα γλωσσικά μοντέλα μπορούν να προτείνουν και να ελέγχουν μεταδεδομένα σε κλίμακα, με δυνατότητες validation και προσαρμογής σε τοπικά σχήματα.

Οι Yang κ.ά. [YFA+25] περιγράφουν γενικότερα ένα AI-assisted πλαίσιο διαχείρισης μεταδεδομένων, με modules όπως «Metadata Extraction and Generation», «Metadata Quality Management» και «Metadata Lineage and Governance» (βλ. Σχήμα 4). Σε αυτήν την αρχιτεκτονική, τα LLMs εντάσσονται ως υπηρεσίες που μπορούν να κληθούν από υπάρχοντα αρχειακά συστήματα (MMS, DAMS) για προτάσεις περιγραφών, θεματικών όρων, γεωγραφικών οντοτήτων κ.λπ., ενώ η συμμόρφωση με τα πρότυπα και οι πολιτικές πρόσβασης παραμένουν στην ευθύνη των εργαλείων διακυβέρνησης και των επαγγελματιών τεκμηρίωσης.



Σχήμα 4. Κύκλος ζωής μεταδεδομένων και βασικές δραστηριότητες σε AI-assisted πλαίσιο διαχείρισης [YFA+25]

Στον χώρο των επιστημονικών άρθρων, οι Alyafeai κ.ά. [AAG25] προτείνουν το MOLE, μια pipeline όπου το LLM καλείται, μέσω προσεκτικά σχεδιασμένων prompts, να εξαγάγει και να επικυρώνει μεταδεδομένα (τίτλος, συγγραφείς, affiliations, θεματικά πεδία,

πειραματικές ρυθμίσεις). Τα αποτελέσματα περνούν από ένα στάδιο κανόνων και consistency checks (π.χ. σύγκριση με CrossRef ή υπάρχουσες εγγραφές), ακριβώς για να αντιμετωπιστεί το φαινόμενο των hallucinations. Σε συνδυασμό με τα παραπάνω, οι Osti & Roke [OR24] εξετάζουν τη χρήση LLMs για την αξιολόγηση και τη βελτίωση της συμπεριληπτικότητας των μεταδεδομένων (metadata inclusivity) σε ψηφιακές συλλογές. Δείχνουν ότι τα μοντέλα μπορούν να εντοπίσουν ελλείψεις ή προκαταλήψεις στις υπάρχουσες εγγραφές και να προτείνουν τρόπους εμπλουτισμού τους, στοιχείο ιδιαίτερα σημαντικό για αρχεία που επιδιώκουν να γίνουν πιο προσβάσιμα σε ευρύτερες κοινότητες. Παράλληλα οι Wu κ.ά. [WBM+23] επισημαίνουν ότι τέτοιες προσεγγίσεις που βασίζονται σε LLMs μπορούν να παράγουν πλούσια μεταδεδομένα, με την προϋπόθεση ότι συνοδεύονται από σαφείς μηχανισμούς ελέγχου και συνδυάζονται με παραδοσιακά εργαλεία εξαγωγής, ώστε να διασφαλίζεται η αξιοπιστία τους.

Η δεύτερη κατεύθυνση αφορά τον μετα-σημασιολογικό εμπλουτισμό. Εργασίες που χαρτογραφούν υπάρχοντα σχήματα (Dublin Core, τοπικά σχήματα μουσείων) όπως CIDOC-CRM και άλλα οντολογικά μοντέλα δίνουν το θεωρητικό υπόβαθρο για «έξυπνη» διαλειτουργικότητα. Δεν πρόκειται απλώς για περισσότερα πεδία, αλλά για πιο πλούσιες σχέσεις, οντότητα, γεγονός, πλαίσιο [SKD+25, ATO20]. Συνδυασμοί με εξωτερικές πηγές γνώσης, όπως το Wikidata, επιτρέπουν αυτόματο γέμισμα ιδιοτήτων (π.χ. χρονολογίες, τόπους, ρόλους) και προτάσεις προς τον αρχειονόμο, με ημι-αυτόματο workflow, το σύστημα προτείνει, ο άνθρωπος εγκρίνει.

Η τρίτη κατεύθυνση αφορά τα ερμηνευτικά και επεξηγηματικά (explainability) μεταδεδομένα. Πέρα από τα «τεχνικά» πεδία (τίτλος, δημιουργός, ημερομηνία) ένα αυξανόμενο ερευνητικό ρεύμα βλέπει τα μεταδεδομένα ως εργαλείο επεξήγησης και ελέγχου των ίδιων των AI μοντέλων, να καταγράφονται δηλαδή οι υποθέσεις κατά τη δημιουργία της εγγραφής, τα σύνολα εκπαίδευσης που χρησιμοποιήθηκαν, οι κίνδυνοι προκατάληψης, ώστε τα μοντέλα και οι προβλέψεις τους να είναι ελέγξιμα από επιμελητές και ερευνητές. Οι Toth κ.ά. [TAP25] υπογραμμίζουν αυτή τη διάσταση ειδικά για αρχειακά συστήματα: ένα αποτέλεσμα που δεν μπορεί να εξηγηθεί στον αρχειονόμο δεν μπορεί ούτε να αξιολογηθεί ούτε να εγκριθεί με εμπιστοσύνη. Οι Wu κ.ά. [WBM+23] επισημαίνουν ότι τέτοιες LLM-based pipelines μπορούν να παράγουν πλούσια μεταδεδομένα, αλλά χρειάζονται σαφείς μηχανισμούς ελέγχου και συνδυασμό με παραδοσιακά εργαλεία extraction για να παραμένουν αξιόπιστες. Στο υπόβαθρο και των τριών γραμμών βρίσκεται η ίδια παρατήρηση, ότι δηλαδή, χωρίς ποιοτικά και πλούσια μεταδεδομένα, ούτε τα πιο προηγμένα μοντέλα TN μπορούν να λειτουργήσουν με τρόπο αξιόπιστο και κοινωνικά χρήσιμο.

Για αρχειακά σενάρια, ανάλογες pipelines αρχίζουν να εμφανίζονται σε πιλοτικά έργα. Χαρακτηριστικά παραδείγματα είναι η επεξεργασία ιστορικών εφημερίδων, όπου LLMs χρησιμοποιούνται μετά το OCR για αυτόματη διόρθωση κειμένου, δημιουργία τίτλων άρθρων και εντοπισμό οντοτήτων (πρόσωπα, τοπωνύμια), αλλά και η διαχείριση born-digital email, όπου NER/LLM μοντέλα βοηθούν στην αναγνώριση σχέσεων μεταξύ προσώπων, projects και θεματικών όρων, ώστε να προταθούν δομές φακέλων και σειρών κατά ISAD(G) [CBJ+21, JC22]. Σε ό,τι αφορά την αξιοπιστία αυτών των pipelines, αξίζει να σημειωθεί και η εργασία του Bodenhamer [Bod23], ο οποίος εξετάζει τη χρήση ChatGPT για τη δημιουργία βιβλιογραφικών μεταδεδομένων, επισημαίνοντας ότι ενώ τα LLMs μπορούν να παράγουν ευέλικτα και φαινομενικά πλούσια μεταδεδομένα, παρουσιάζουν σοβαρά προβλήματα συνέπειας όταν δεν βασίζονται αποκλειστικά στο πρωτογενές κείμενο της πηγής, μια παρατήρηση που έχει άμεση συνάφεια με το σχεδιασμό prompts για αρχειακές περιγραφές [Bod23].

2.1.3 Μέθοδοι NLP και AI για εξαγωγή μεταδεδομένων

Η αυτόματη εξαγωγή μεταδεδομένων από κείμενα και άλλα ψηφιακά τεκμήρια αποτελεί αντικείμενο έρευνας εδώ και τουλάχιστον δύο δεκαετίες. Ο γενικός στόχος παραμένει σταθερός: από τα δεδομένα ενός εγγράφου, έντυπου ή ψηφιακού, να εντοπιστούν αυτόματα κρίσιμα στοιχεία όπως τίτλος, δημιουργός, ημερομηνία, λέξεις-κλειδιά, βιβλιογραφία κ.λπ. και να αποτυπωθούν αυτά τα στοιχεία σε δομημένη μορφή, σύμφωνα με κάποιο σχήμα μεταδεδομένων. Η ακρίβεια και η αποδοτικότητα αυτής της διαδικασίας έχουν άμεση επίπτωση στη διαχείριση συλλογών, στην αναζήτηση και στην αξιοποίηση του περιεχομένου, είτε πρόκειται για ακαδημαϊκή βιβλιογραφία, είτε για νομικά έγγραφα, είτε για αρχειακά τεκμήρια.

Οι πρώτες συστηματικές προσεγγίσεις βασίστηκαν σε κανόνες. Οι Azimjonov και Alikhanov [AA18] ανέπτυξαν ένα rule-based framework που επεξεργάζεται PDFs ακαδημαϊκών άρθρων, αξιοποιώντας πληροφορία διάταξης (layout), γραμματοσειράς και γλωσσικά μοτίβα για να εξάγει βασικά πεδία μεταδεδομένων. Τα αποτελέσματά τους δείχνουν πολύ υψηλή ακρίβεια και ανάκληση για τίτλο, περίληψη, λέξεις-κλειδιά, κύριο κείμενο, συμπεράσματα και βιβλιογραφία, ξεπερνώντας καθιερωμένα εργαλεία όπως το GROBID και το ParsCit [AA18], ενώ το σύστημα μπορεί να επεξεργαστεί 15–20 άρθρα το λεπτό. Το πλεονέκτημα τέτοιων rule-based συστημάτων έγκειται στη διαφάνεια και τη σταθερότητα. Όταν η δομή των εγγράφων είναι ομοιογενής (π.χ. άρθρα ενός συγκεκριμένου περιοδικού), οι κανόνες μπορούν να επιτύχουν σταθερή απόδοση σε μεγάλο όγκο εγγράφων, δηλαδή να υπάρχει μία αξιόπιστη μαζική τους επεξεργασία.

Όμως αυτά τα συστήματα είναι ευαίσθητα σε αλλαγές μορφότυπου, απαιτούν χειρωνακτική δημιουργία και συντήρηση των κανόνων και δεν μεταφέρονται εύκολα σε νέους τύπους εγγράφων.

Για να αντιμετωπιστεί αυτή η «ακαμψία», η έρευνα στράφηκε σε μεθόδους μηχανικής μάθησης. Οι Granitzer και συνεργάτες [GHK+12] πραγματοποίησαν μια ιδιαίτερα συστηματική μελέτη, συγκρίνοντας διαφορετικές τεχνικές εξαγωγής βιβλιογραφικών μεταδεδομένων βασισμένες στη διάταξη της σελίδας (layout-based metadata extraction). Στο πλαίσιο αυτό δοκίμασαν διάφορα μοντέλα μηχανικής μάθησης, μεταξύ των οποίων Conditional Random Fields (CRF) και Support Vector Machines (SVM), σε συνδυασμό με χαρακτηριστικά που περιέγραφαν τη μορφή και τη θέση στοιχείων στη σελίδα (μέγεθος και στυλ γραμματοσειράς, συντεταγμένες κειμένου, απόσταση από τα περιθώρια, σχέση με άλλα μπλοκ). Η βασική τους διαπίστωση ήταν ότι η σωστή επιλογή και σχεδίαση αυτών των χαρακτηριστικών διάταξης είναι καθοριστική. Όταν τα μοντέλα «βλέπουν» όχι μόνο το κείμενο αλλά και το πού βρίσκεται και πώς μορφοποιείται, μπορούν να αναγνωρίσουν με αρκετή αξιοπιστία πεδία όπως τίτλος, ονόματα συγγραφέων, στοιχεία περιοδικού και βιβλιογραφικές αναφορές, ακόμη και σε έγγραφα διαφορετικών κλάδων και εκδοτικών μορφότυπων. Παρά την καλή συνολική απόδοση, κανένα μοντέλο δεν αποδείχθηκε «καλύτερο σε όλα»· άλλα απέδιδαν καλύτερα στον εντοπισμό τίτλων, άλλα στις αναφορές ή στα στοιχεία περιοδικού, δείχνοντας ότι η επιλογή μεθόδου πρέπει να προσαρμόζεται στο είδος μεταδεδομένων που ενδιαφέρει κάθε εφαρμογή.

Σε επόμενο στάδιο, η βιβλιογραφία στράφηκε σε βαθιά νευρωνικά δίκτυα με εξειδικευμένη αρχιτεκτονική. Οι Nayaka και Ranjan [NR23] συνδυάζουν δύο διαφορετικές αρχιτεκτονικές που λειτουργούν συμπληρωματικά ή μία στην άλλη. Από τη μία έχουμε ένα συνελκτικό νευρωνικό δίκτυο (Convolutional Neural Network - CNN) [ZZL15] που εφαρμόζεται σε επίπεδο χαρακτήρων, δηλαδή διαβάζει το κείμενο γράμμα γράμμα και μαθαίνει να αναγνωρίζει τοπικά μοτίβα, προθέματα, καταλήξεις ή ακολουθίες χαρακτήρων που είναι ενδεικτικές συγκεκριμένων κατηγοριών περιεχομένου όπως μεταβλητές κώδικα ή μαθηματικά σύμβολα. Από την άλλη, γίνεται χρήση ενός αμφίδρομου LSTM (bidirectional Long Short-Term Memory, bi-LSTM) [HS97]. Το LSTM είναι ένας τύπος αναδρομικού νευρωνικού δικτύου σχεδιασμένο να επεξεργάζεται ακολουθίες δεδομένων, διατηρώντας «μνήμη» από προηγούμενα σημεία της ακολουθίας. Μπορεί δηλαδή να συνδέσει μια λέξη στο τέλος μιας πρότασης με κάποια άλλη που εμφανίστηκε πολύ νωρίτερα, κάτι που τα κλασικά αναδρομικά δίκτυα δυσκολεύονται να κάνουν. Το χαρακτηριστικό «αμφίδρομο» σημαίνει ότι το δίκτυο διαβάζει την πρόταση τόσο από αριστερά προς δεξιά όσο και αντίστροφα, ώστε κάθε λέξη να ερμηνεύεται με

βάση όσα προηγούνται αλλά και όσα ακολουθούν. Ο συνδυασμός των δύο αρχιτεκτονικών επιτρέπει στο σύστημα να αντιλαμβάνεται ταυτόχρονα τη μορφή των λέξεων και τη συντακτική τους σχέση εντός της πρότασης, και έτσι να εξάγει πιο σύνθετα μεταδεδομένα από ακαδημαϊκά άρθρα, όπως ψευδοκώδικες και προτάσεις που περιγράφουν datasets, μετρικές απόδοσης και υπολογιστική πολυπλοκότητα. Χρησιμοποιώντας λοιπόν οι Nayaka και Ranjan [NR23], 58 χαρακτηριστικά που καλύπτουν περιεχόμενο, στυλ γραμματοσειράς και δομή, πετυχαίνουν ακρίβεια 94,23% στην ταξινόμηση γραμμών ψευδοκώδικα και 82% στην εξαγωγή προτάσεων με «αλγοριθμικά» μεταδεδομένα, βελτιώνοντας σημαντικά τόσο rule-based όσο και SVM-based προσεγγίσεις. Η μετάβαση αυτή από τα κλασικά ML σε deep learning επιτρέπει την αποτύπωση πιο σύνθετων, μη γραμμικών μοτίβων, αλλά απαιτεί επίσης μεγαλύτερα και καλύτερα επισημειωμένα datasets.

Παράλληλα, έχουν αναπτυχθεί ειδικές τεχνικές για την εξαγωγή μη κειμενικών (όπως, εικόνες, σχήματα, διαγράμματα, πίνακες, ήχο, βίντεο) ή «πολυτροπικών» (πολλαπλών τύπων περιεχομένου) μεταδεδομένων, όπου ένα τεκμήριο συνδυάζει περισσότερους από έναν τύπους περιεχομένου, και η εξαγωγή πληροφορίας απαιτεί χειρισμό καθενός με διαφορετικό τρόπο. Ο Choudhury κ.ά. [CMK+13] μελετούν την αυτόματη εξαγωγή μεταδεδομένων για σχήματα (figures) σε επιστημονικά άρθρα PDF, αντιμετωπίζοντας ξεχωριστά τρία βήματα: τον εντοπισμό και εξαγωγή της εικόνας, την εξαγωγή της λεζάντας και τον εντοπισμό των σημείων στο κείμενο όπου το σχήμα αναφέρεται. Για τη λεζάντα συνδυάζουν χαρακτηριστικά διάταξης (θέση, γραμματοσειρά, απόσταση από την εικόνα) με λεξιλογικά χαρακτηριστικά, δοκιμάζοντας τόσο rule-based όσο και ML προσεγγίσεις, και στη συνέχεια αντιστοιχίζουν λεζάντες και εικόνες με γεωμετρικές και δομικές ενδείξεις. Πάνω στα εξαγόμενα figure metadata υλοποιούν μια μηχανή αναζήτησης (Solr/Lucene) για σχήματα από περίπου 150.000 άρθρα, δείχνοντας ότι η στοχευμένη εξαγωγή μεταδεδομένων μπορεί να υποστηρίξει εξειδικευμένη αναζήτηση σε ψηφιακές βιβλιοθήκες.

Σε άλλους εφαρμόσιμους τομείς, όπως τα κατασκευαστικά έργα, οι Ozogul και Ergen [OE24] αντιμετωπίζουν το πρόβλημα της ελλιπούς ποιότητας μεταδεδομένων σε έγγραφα Request for Information (RFIs). Διαπιστώνουν ότι ο κωδικός ειδικότητας (discipline code) καταχωρείται χειροκίνητα στα συστήματα, με αποτέλεσμα καθυστερήσεις και λανθασμένες ή ελλιπείς εγγραφές, και προτείνουν για αυτό το λόγο, μια pipeline NLP και ML όπου τα κείμενα των RFIs προ-επεξεργάζονται, μετατρέπονται σε διανυσματικά χαρακτηριστικά και ταξινομούνται αυτόματα με Naïve Bayes (NB) και K-Nearest Neighbors (KNN). Τα πειραματικά αποτελέσματα δείχνουν ότι τα μοντέλα επιτυγχάνουν

ικανοποιητική ακρίβεια στην αυτόματη ανάθεση του *discipline code*, με διαφοροποιήσεις ανά ετικέτα, ανοίγοντας τον δρόμο για συστήματα διαχείρισης RFI's που ενσωματώνουν αυτόματη εξαγωγή μεταδεδομένων και βελτιωμένη ανάλυση ρίσκου. Η εφαρμογή παρόμοιων τεχνικών στον αρχειακό χώρο έχει επίσης μελετηθεί για εξειδικευμένες συλλογές. Οι Mardiati κ.ά. [MAG+23] εξετάζουν τις δυνατότητες χρήσης τεχνητής νοημοσύνης για τη συλλογή μεταδεδομένων σε δραστηριότητες περιγραφής φωτογραφικών αρχείων, καταδεικνύοντας ότι η αυτοματοποίηση μπορεί να επεκταθεί πέρα από τα έντυπα κείμενα και σε οπτικό αρχειακό υλικό. Στο ίδιο πεδίο, οι Zoutsou κ.ά. [ZSP+24] εφαρμόζουν θεματική μοντελοποίηση με τον αλγόριθμο LDA σε σώμα 40.881 περιλήψεων ερευνητικής βιβλιογραφίας για την πολιτιστική κληρονομιά, διερευνώντας πώς οι αυτόματες μέθοδοι NLP μπορούν να εντοπίσουν τα βασικά θέματα που διαπερνούν μια μεγάλη συλλογή κειμένων, χωρίς όμως την ανθρώπινη ανάγνωση. Τα αποτελέσματά τους υπογραμμίζουν τόσο τις δυνατότητές του - η βέλτιστη ρύθμιση του αλγορίθμου παρήγαγε 19 ερμηνεύσιμα θέματα με υψηλή συνοχή - όσο και τους περιορισμούς. Η ανθρώπινη ερμηνεία και επικύρωση των αποτελεσμάτων παραμένει αναντικατάστατη, καθώς ο αλγόριθμος αναγνωρίζει μεν μοτίβα κοινής εμφάνισης λέξεων, αλλά δεν κατανοεί το νόημά τους. Αυτή η διαπίστωση δεν είναι αποκλειστικό χαρακτηριστικό της θεματικής μοντελοποίησης, εμφανίζεται σε ολόκληρο το φάσμα των αυτόματων μεθόδων, από τα *rule-based* συστήματα μέχρι τα LLMs, και αποτελεί ένα από τα κεντρικά ζητήματα που εξετάζει η παρούσα εργασία.

Σε επίπεδο βιβλιογραφικής σύνοψης, οι Raval και Bhaidasna [RB24] παρέχουν μια συστηματική ανασκόπηση τεχνικών εξαγωγής μεταδεδομένων από επιστημονικά άρθρα, καταγράφοντας τόσο παραδοσιακές μεθόδους NLP/ML όσο και πιο πρόσφατες προσεγγίσεις με BERT-based μοντέλα και LLMs. Τα συμπεράσματά τους είναι ότι, παρά τη σαφή πρόοδο, πολλά προβλήματα παραμένουν ανοικτά: έλλειψη καθολικών προτύπων αξιολόγησης, δυσκολίες διαχείρισης θορυβωδών ή ελλιπών PDFs, ελλιπής υποστήριξη για πολυτροπικά δεδομένα δηλαδή περιεχόμενο που περιέχει κείμενο, εικόνες, πίνακες, αλλά και υψηλό υπολογιστικό κόστος για τις πιο προηγμένες μεθόδους. Συναφείς εργασίες, όπως αυτή των Wu κ.ά. [WBM+23], τονίζουν τα όρια της «καθαρής» μηχανικής μάθησης στην αυτόματη επισημείωση μεταδεδομένων και υπογραμμίζουν ότι ορισμένα στοιχεία (π.χ. ερμηνευτικές σχέσεις, λεπτές σημασιολογικές αποχρώσεις) παραμένουν δύσκολα. Οι Lorenzini κ.ά. [LRT21] δείχνουν ότι ακόμη και όταν η εξαγωγή έχει γίνει, η αυτόματη αξιολόγηση της ποιότητας των περιγραφών σε αρχεία πολιτιστικής κληρονομιάς απαιτεί προσεκτικό σχεδιασμό δεικτών, δεδομένου ότι η «καλή» περιγραφή δεν είναι μόνο ζήτημα ορθότητας αλλά και πληρότητας, σαφήνειας και συνέπειας.

Πιο πρόσφατα, η συζήτηση μετατοπίζεται στα Μεγάλα Γλωσσικά Μοντέλα (LLMs), τα οποία αντιμετωπίζουν την εξαγωγή μεταδεδομένων περισσότερο ως πρόβλημα κατανόησης και γένεσης γλώσσας παρά απλής αναγνώρισης μοτίβων. Όπως αναφέρθηκε ήδη, η προσέγγιση του Bagchi [Bag24] εντάσσεται ακριβώς σε αυτήν την κατεύθυνση. Προτείνει λοιπόν μια «Generative AI-driven Metadata Modelling» προσέγγιση, όπου αντιμετωπίζει τα μεταδεδομένα ως αποτέλεσμα συνολικής κατανόησης του κειμένου από το μοντέλο. Οι Alyafeai κ.ά. [AAG25] παρουσιάζουν το MOLE, ένα σύστημα που χρησιμοποιεί LLMs για εξαγωγή και επικύρωση μεταδεδομένων από επιστημονικά άρθρα, συνδυάζοντας prompts για δομημένη έξοδο με κανόνες ελέγχου συνέπειας, ώστε να μειώνεται ο κίνδυνος «παραισθήσεων» (hallucinations). Αξίζει επίσης να αναφερθεί η εργασία του Bodenhamer [Bod23], ο οποίος εξετάζει συγκεκριμένα τη χρήση ChatGPT για τη δημιουργία βιβλιογραφικών μεταδεδομένων σε βιβλιοθηκονομικό πλαίσιο, επισημαίνοντας ότι ενώ τα LLMs μπορούν να παράγουν εύελικτα και φαινομενικά πλούσια μεταδεδομένα, παρουσιάζουν σοβαρά προβλήματα αξιοπιστίας και συνέπειας όταν δεν βασίζονται αποκλειστικά στο πρωτογενές κείμενο της πηγής [Bod23].

Σε πιο εξειδικευμένα περιβάλλοντα, οι Jacyszyn κ.ά. [JJG+25] αναπτύσσουν το AI4DiTraRe, ένα LLM-based πλαίσιο για εξαγωγή πληροφοριών από κλιματικά ερευνητικά αποθετήρια (repositories), το οποίο εστιάζει στη στοιχίση της ελεύθερης περιγραφής με ελεγχόμενα λεξιλόγια και πρότυπα μεταδεδομένων. Χαρακτηριστικό παράδειγμα εξειδικευμένου LLM για τον αρχειακό χώρο είναι το ArcGPT των Zhang κ.ά. [ZHP+23], το οποίο εκπαιδεύεται εκ νέου πάνω σε εκτεταμένα αρχειακά δεδομένα και συνοδεύεται από το benchmark AMBLE για την αξιολόγηση τυπικών αρχειακών εργασιών (περιγραφή, ευρετηρίαση, αναζήτηση). Οι συγγραφείς δείχνουν ότι το ArcGPT υπερέχει σε αυτές τις εργασίες έναντι γενικών LLMs, αναγνωρίζοντας ταυτόχρονα ζητήματα ποιότητας δεδομένων, πιθανών παραληρημάτων και συμμόρφωσης με αρχειακά πρότυπα, τα οποία αφήνουν σημαντικό χώρο για περαιτέρω έρευνα.

Οι εργασίες αυτές [Bag24, AAG25, Bod23, JJG+25, ZHP+23] δείχνουν τόσο το δυναμικό όσο και τα όρια των LLMs: μπορούν, με κατάλληλο prompt engineering, να παράγουν πλούσια και κατά κανόνα εύστοχα μεταδεδομένα, αλλά η συνέπεια στη δομή, η συμμόρφωση με συγκεκριμένα πρότυπα και η αποφυγή επινοημένων πληροφοριών παραμένουν ζητούμενα [RSD23, WBM+23].

Συνολικά, η εικόνα που προκύπτει είναι ότι οι μέθοδοι NLP και AI για εξαγωγή μεταδεδομένων έχουν φτάσει σε επίπεδο ωριμότητας επαρκές για βιομηχανική χρήση σε αρκετούς τομείς (βιβλιογραφία, κατασκευές, ιατρική τεκμηρίωση, επιστημονικά

repositories), ιδίως όταν τα έγγραφα είναι σχετικά τυποποιημένα και υπάρχουν μεγάλα datasets για εκπαίδευση. Η μεταφορά τους όμως στον αρχειακό χώρο δεν είναι αυτόματη. Τα αρχειακά τεκμήρια είναι συχνά λιγότερο τυποποιημένα, τα διαθέσιμα datasets μικρότερα και περισσότερο «θορυβώδη», ενώ τα πρότυπα περιγραφής (όπως το ISAD(G)) εισάγουν απαιτήσεις ιεραρχικότητας, πλαισίου και πιστότητας στο πρωτογενές κείμενο που σπάνια λαμβάνονται υπόψη στα γενικά μοντέλα εξαγωγής μεταδεδομένων. Αυτή η απόσταση ανάμεσα στην τεχνολογική ωριμότητα και στις συγκεκριμένες ανάγκες των αρχείων αποτελεί ένα από τα κεντρικά κίνητρα της παρούσας εργασίας.

Παρά την ποικιλία και την τεχνική ωριμότητα των μεθόδων που περιεγράφηκαν παραπάνω, η συντριπτική πλειονότητα των εφαρμογών στοχεύει σε σχετικά επίπεδα, «μονοεπίπεδα» σχήματα μεταδεδομένων, όπου η μονάδα περιγραφής είναι συνήθως το μεμονωμένο άρθρο ή έγγραφο. Αντίθετα, στα αρχεία η περιγραφή οργανώνεται ιεραρχικά, ακολουθεί αρχές προέλευσης και τάξης του αρχικού σχηματισμού και υλοποιείται μέσω προτύπων όπως το ISAD(G) και τα συναφή εθνικά σχήματα. Η μετάβαση από τεχνικές εξαγωγής μεταδεδομένων που λειτουργούν σε επίπεδο «ενός εγγράφου» σε τεχνικές που μπορούν να παραγάγουν συνεπείς, ιεραρχικές και πλαίσιο-συνειδητές περιγραφές αποτελεί άρα, ένα κρίσιμο και ακόμη σε μεγάλο βαθμό ανοιχτό ζητούμενο, το οποίο εξετάζεται τόσο στο θεωρητικό πλαίσιο (Κεφάλαιο 3) όσο και στο πειραματικό μέρος της παρούσας εργασίας (Κεφάλαιο 4).

2.1.4 Υβριδικά σενάρια και ποιότητα σε αρχεία

Ένα χαρακτηριστικό παράδειγμα υβριδικής (hybrid) και LLM-ενισχυμένης pipeline είναι το AI4DiTraRe [JJG+25] για αποθετήρια κλιματικής έρευνας, το οποίο λειτουργεί με πολλαπλούς «πράκτορες» (agents) πάνω σε μια κοινή υποδομή. Ο όρος «υβριδική» αναφέρεται στον συνδυασμό διαφορετικών μεθόδων μέσα στην ίδια ροή εργασίας όπως αλγοριθμικές τεχνικές για μηχανικές διεργασίες, LLMs για εργασίες που απαιτούν κατανόηση φυσικής γλώσσας, κανόνες συμμόρφωσης για επικύρωση δομής, και επιλεκτική ανθρώπινη παρέμβαση για τις περιπτώσεις όπου η αυτόματη απόφαση δεν είναι αξιόπιστη. Σε απλοποιημένη μορφή, η ροή του ξεκινά με τη συλλογή και τον αρχικό εμπλουτισμό μεταδεδομένων (μηχανικά πεδία, identifiers, βασική βιβλιογραφική πληροφορία), συνεχίζει με κλήση LLM-agents για περίληψη περιεχομένου, εξαγωγή εννοιών και χαρτογράφηση σε ελεγχόμενα λεξιλόγια, και ολοκληρώνεται με εφαρμογή schema validators και κανόνων συνέπειας, για παράδειγμα, έλεγχος αν οι θεματικοί όροι προέρχονται από συγκεκριμένο θησαυρό ή αν οι ημερομηνίες ταιριάζουν με τα

πρωτογενή δεδομένα. Εγγραφές που σημαίνονται ως αμφίβολες παραπέμπονται για ανθρώπινη ανασκόπηση.

Το ArcGPT [ZHP+23] δείχνει πώς μια τέτοια λογική μπορεί να εφαρμοστεί ειδικά σε αρχειακά συμφραζόμενα: πρόκειται για ένα LLM προσαρμοσμένο σε αρχειακά δεδομένα και αξιολογημένο σε benchmark (AMBLE) για εργασίες όπως δημιουργία περιγραφών, θεματική ευρετηρίαση και ερωτοαπαντήσεις πάνω σε finding aids. Η προτεινόμενη αρχιτεκτονική υπονοεί μια pipeline όπου το LLM παράγει προτάσεις περιγραφών, οι οποίες περνούν από εργαλεία ελέγχου ποιότητας πριν ενσωματωθούν σε αρχειακά συστήματα. Στην ίδια λογική οι Groppe κ.ά. [GMW+25] προτείνουν μια προσέγγιση ομοσπονδιακής νοημοσύνης (federated intelligence) μεταξύ πολλαπλών LLMs για την αυτόματη παραγωγή αρχειακών περιγραφών, δείχνοντας ότι ο συνδυασμός μοντέλων μπορεί να αντισταθμίσει τις αδυναμίες κάθε μεμονωμένου συστήματος. Παράλληλα, οι Toth κ.ά. [TAP25] υπογραμμίζουν τη σημασία της ερμηνευσιμότητας (explainability) στα LLM-based συστήματα για αρχειακή πολιτιστική κληρονομιά, επισημαίνοντας ότι ο έλεγχος ποιότητας δεν αρκεί να είναι μόνο τεχνικός, πρέπει να είναι κατανοητός από τους επαγγελματίες που καλούνται να αξιολογήσουν και να εγκρίνουν τα αποτελέσματα.

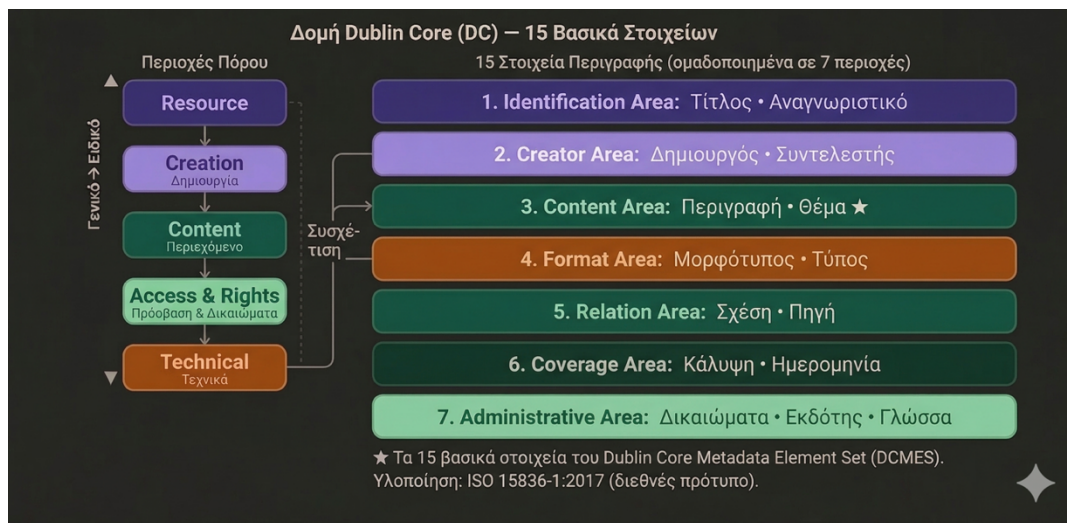
Συνολικά, η σημερινή τάση δεν είναι να αντικατασταθούν οι υπάρχουσες pipelines, αλλά να επεκταθούν. Οι κλασικές φάσεις (OCR, parsing, feature extraction) παραμένουν, ενώ το «μοντέλο» στη μέση μπορεί πλέον να είναι ένας συνδυασμός συμβατικών ταξινομητών και LLMs, και το στάδιο μετα-επεξεργασίας αποκτά κεντρικό ρόλο ως φίλτρο αξιοπιστίας, συμμόρφωσης με πρότυπα και ευθυγράμμισης με την αρχειακή πρακτική. Ειδικά σε χώρους όπως τα αρχεία, όπου οι περιγραφές είναι ιεραρχικές, ενταγμένες σε πλαίσιο προέλευσης και υπόκεινται σε ηθικές και νομικές δεσμεύσεις, αυτή η υβριδική προσέγγιση φαίνεται πιο ρεαλιστική από ένα πλήρως end-to-end σενάριο και αποτελεί το υπόβαθρο πάνω στο οποίο δοκιμάζεται η παρούσα εργασία.

Οι παραπάνω pipelines, είτε βασίζονται σε κανόνες, είτε σε ML/LLM και λειτουργούν κυρίως σε επίπεδο μεμονωμένου τεκμηρίου ή αρχείου PDF και στοχεύουν στην εξαγωγή πεδίων «τύπου βιβλιοθήκης» (τίτλος, δημιουργός, ημερομηνία, θεματικοί όροι) [GHK+12, RB24, WBM+23]. Στο αρχειακό πλαίσιο, όμως, τα μεταδεδομένα δεν αφορούν μόνο το ίδιο το τεκμήριο αλλά και τη θέση του μέσα σε μια ιεραρχική δομή (fonds-series-file-item) και στο πλαίσιο προέλευσης, όπως ορίζουν τα πρότυπα ISAD(G) και συναφή σχήματα [Gil08, ATO20].

2.2 Πλαίσια και Πρότυπα μεταδεδομένων

Η εξαγωγή μεταδεδομένων αποκτά πραγματική αξία μόνο όταν συνδέεται με συγκεκριμένα πρότυπα και σχήματα, τα οποία ορίζουν τι θεωρείται «σημαντική» πληροφορία, πώς δομείται και με ποιον τρόπο μπορεί να ανταλλαχθεί και να ερμηνευθεί από άλλα συστήματα.

Στον χώρο των ψηφιακών βιβλιοθηκών, το πιο διαδεδομένο βασικό σχήμα είναι το Dublin Core [DCMI98, ISO17], ένα σύνολο 15 στοιχείων (τίτλος, δημιουργός, θέμα, περιγραφή κ.λπ.) σχεδιασμένο για να είναι όσο το δυνατόν πιο ουδέτερο και επαναχρησιμοποιήσιμο. Η απλότητα και η ευρεία αποδοχή του το καθιστούν συχνά τον «κοινό παρονομαστή» σε συστήματα που χρειάζονται βασική διαλειτουργικότητα μεταξύ βιβλιοθηκών, αρχείων και μουσείων [Gil08]. Σε πιο σύνθετα περιβάλλοντα, ιδίως σε πλατφόρμες ανοικτών κυβερνητικών δεδομένων, χρησιμοποιείται το DCAT (Data Catalog Vocabulary) της W3C [W3C24], το οποίο επιτρέπει την αναπαράσταση καταλόγων δεδομένων με πληροφορίες για θεματική, γεωγραφική κάλυψη και νομικούς

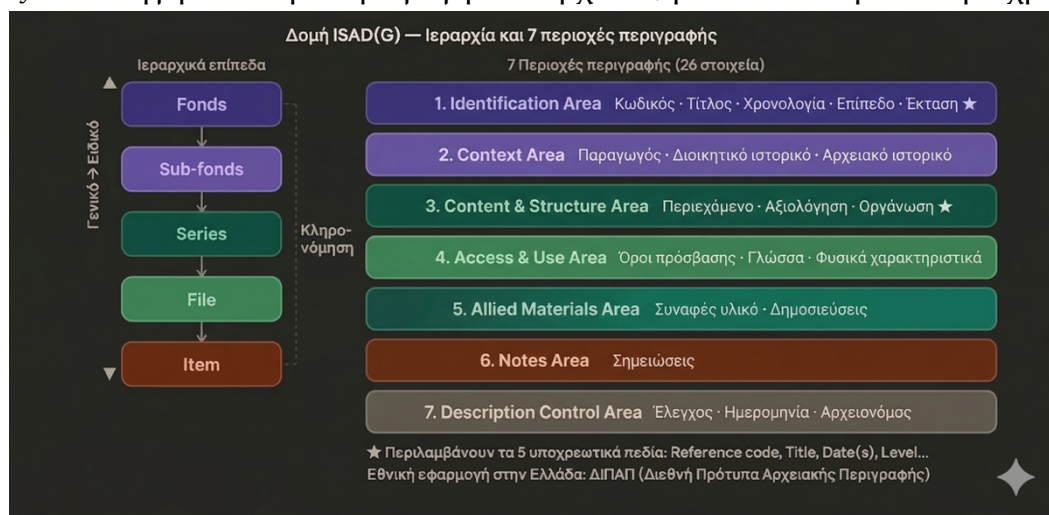


Σχήμα 5. Τα 15 βασικά στοιχεία του Dublin Core οργανωμένα ως απλό σύνολο πεδίων για βασική περιγραφική μεταδεδομένων και διαλειτουργικότητα μεταξύ ετερογενών συστημάτων.

όρους χρήσης, εξασφαλίζοντας ότι τα μεταδεδομένα είναι μηχανικά αναγνώσιμα και μπορούν να συγκομιστούν από τρίτα συστήματα (Σχ. 5). Στο ευρωπαϊκό πλαίσιο, η ειδικότερη εφαρμογή του DCAT είναι το DCAT-AP (DCAT Application Profile for data portals in Europe) [DCAT-AP24], το οποίο προσαρμόζει το γενικό λεξιλόγιο στις ανάγκες των ευρωπαϊκών πυλών ανοικτών δεδομένων, εξασφαλίζοντας διαλειτουργικότητα μεταξύ καταλόγων διαφορετικών κρατών μελών και υποστηρίζοντας τη συμμόρφωση με τις αρχές των FAIR data. Στον αρχειακό χώρο, κεντρική θέση κατέχει το ISAD(G)

(General International Standard Archival Description), που εκδόθηκε από το Διεθνές Συμβούλιο Αρχείων (ICA) σε δεύτερη έκδοση το 2000 [ICA00]. Το ISAD(G) ορίζει 26 στοιχεία περιγραφής οργανωμένα σε επτά περιοχές: Identification (ταυτότητα), Context (πλαίσιο), Content and Structure (περιεχόμενο και δομή), Conditions of Access and Use (όροι πρόσβασης και χρήσης), Allied Materials (συναφές υλικό), Notes (σημειώσεις) και Description Control (έλεγχος περιγραφής). Από αυτά, πέντε θεωρούνται ελάχιστα υποχρεωτικά για κάθε εγγραφή (ο κωδικός αναφοράς, ο τίτλος, η χρονολογία, το επίπεδο περιγραφής και η έκταση του υλικού). Το βασικό χαρακτηριστικό του προτύπου είναι η ιεραρχική του λογική, η περιγραφή δηλαδή, οργανώνεται από το γενικό προς το ειδικό, σε επίπεδα όπως *fonds*, *subfonds*, *series*, *file* και *item*, με την πληροφορία να «κληρονομείται» από τα ανώτερα στα κατώτερα επίπεδα. Αυτό σημαίνει ότι το παράγωγο ή το διοικητικό πλαίσιο δηλώνονται μία φορά στο επίπεδο του *fonds* και δεν επαναλαμβάνονται σε κάθε επιμέρους τεκμήριο.

Το ISAD(G) δεν λειτουργεί αυτόνομα, αλλά εντάσσεται σε μια ευρύτερη οικογένεια προτύπων του ICA που καλύπτουν συμπληρωματικές διαστάσεις της αρχειακής περιγραφής. Το ISAAR(CPF) (International Standard Archival Authority Record for Corporate Bodies, Persons and Families) [ICA04] εξειδικεύεται στην περιγραφή των ιδίων των παραγωγών των αρχείων, δηλαδή των φυσικών προσώπων, οικογενειών και νομικών προσώπων που δημιούργησαν ή συγκέντρωσαν το υλικό, επιτρέποντας τη συστηματική τεκμηρίωση της βιογραφικής και διοικητικής τους ιστορίας, ανεξάρτητα από τα ίδια τα αρχεία. Το ISDF (International Standard for Describing Functions) [ICA07] συμπληρώνει την εικόνα περιγράφοντας τις λειτουργίες, τις δραστηριότητες και τις διεργασίες που οδήγησαν στην παραγωγή των αρχείων, μια διάσταση ιδιαίτερα χρήσιμη

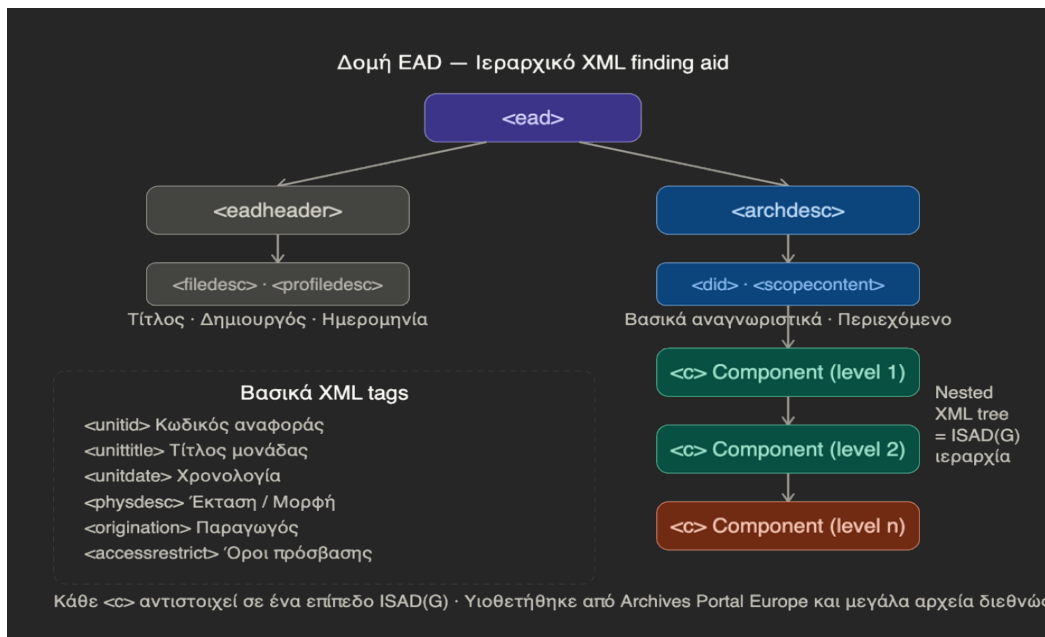


Σχήμα 6. Ιεραρχική δομή και περιοχές περιγραφής του ISAD(G) [ICA00, Gil08]

σε φορείς όπου το θεσμικό πλαίσιο αλλάζει συχνά και η ίδια λειτουργία μπορεί να ασκείται από διαφορετικούς οργανισμούς διαχρονικά. Μαζί, τα τρία πρότυπα συγκροτούν ένα ολοκληρωμένο πλαίσιο όπου το αρχειακό υλικό, ο παραγωγός και το πλαίσιο παραγωγής αντιμετωπίζονται ως αλληλένδετες οντότητες και όχι ως αυτόνομες εγγραφές.

Στην Ελλάδα, η εφαρμογή του ISAD(G) γίνεται μέσω του ΔΙΠΑΠ (Διεθνών Προτύπων Αρχειακής Περιγραφής), το οποίο αποτελεί την εθνική προσαρμογή του προτύπου στις ανάγκες των ελληνικών αρχείων, διατηρώντας τη βασική δομή αλλά εισάγοντας τοπικές εξειδικεύσεις στην ορολογία και τις οδηγίες εφαρμογής [ICA00, Gil08]. (Σχ.6)

Σημειώνεται επίσης, πως η συζήτηση για αυθεντικότητα στις ψηφιακές αρχειακές περιγραφές αποτελεί ένα αυξανόμενο ερευνητικό πεδίο. Οι Pacheco κ.ά. [PGF23] προτείνουν ένα μοντέλο μεταδεδομένων ειδικά σχεδιασμένο για την τεκμηρίωση της αυθεντικότητας σε ψηφιακές αρχειακές περιγραφές, επισημαίνοντας ότι η απλή αντιγραφή υπάρχοντων σχημάτων όπως το ISAD(G) δεν επαρκεί για να εξασφαλιστεί η αξιοπιστία των ψηφιακών εγγραφών, χρειάζονται περισσότερα στοιχεία που τεκμηριώνουν την προέλευση, την ακεραιότητα και τη διαδοχή της περιγραφής. Αυτό το



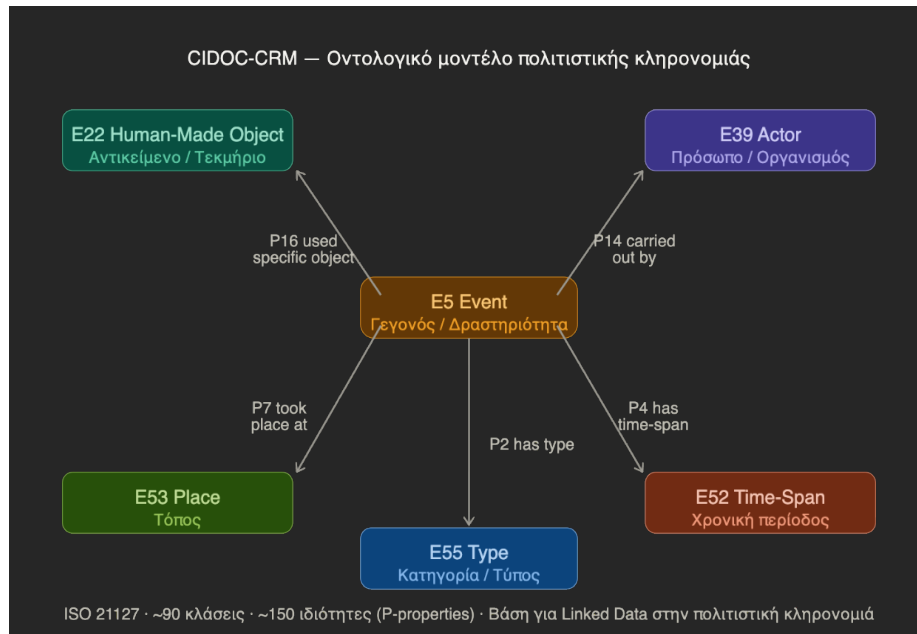
Σχήμα 7. Δομή EAD ως ιεραρχικό XML finding aid - αντιστοίχιση με τα επίπεδα ISAD(G) [Gil08]

ζήτημα αποκτά ακόμη μεγαλύτερη σημασία στην εποχή των LLMs, όπου η αυτόματη δημιουργία περιγραφών δημιουργεί νέες ανάγκες για μεταδεδομένα που να τεκμηριώνουν πώς και από ποιον/τι παράχθηκε μια εγγραφή [PGF23]. Για την

κωδικοποίηση των ISAD(G), συμβατών περιγραφών σε ανταλλάξιμη ψηφιακή μορφή, έχει αναπτυχθεί το EAD (Encoded Archival Description) [LoC23], ένα XML πρότυπο που επιτρέπει την αναπαράσταση πολυσύνθετων, ιεραρχικών αρχειακών ευρετηρίων. Το EAD ορίζει συγκεκριμένα XML tags για κάθε στοιχείο περιγραφής, για παράδειγμα το <c> για τα components κάθε ιεραρχικού επιπέδου, το <did> για τα βασικά αναγνωριστικά στοιχεία και το <scopecontent> για την περιγραφή περιεχομένου, επιτρέποντας έτσι τη δημιουργία ευρετηρίων (finding aids) που μπορούν να δημοσιευτούν διαδικτυακά, να αναζητηθούν πλήρως και να ανταλλαχθούν μεταξύ ιδρυμάτων. Το EAD έχει υιοθετηθεί ευρέως από μεγάλα αρχεία και κοινοπρακτικά ευρετήρια, χαρακτηριστικό παράδειγμα είναι το Archives Portal Europe, παρά την εγγενή του πολυπλοκότητα, η οποία συχνά αποτελεί εμπόδιο για μικρότερους φορείς [Gil08]. (Σχ. 7)

Πιο πρόσφατα, η αρχειακή περιγραφή έχει μετατοπιστεί από καταλόγους πεδίων προς εννοιολογική μοντελοποίηση. Από το 2012, το ICA έχει αναπτύξει το Records in Contexts – Conceptual Model (RiC-CM) [ICA23], ένα υψηλού επιπέδου μοντέλο επικεντρωμένο στην περιγραφή εγγραφών μαζί με τους φορείς (πρόσωπα/οργανισμούς) που τις δημιουργούν και τις χρησιμοποιούν, καθώς και των δραστηριοτήτων/λειτουργιών που τεκμηριώνουν. Το RiC-CM προορίζεται ως ένα πλαίσιο ενοποίησης που καλύπτει το ουσιαστικό περιεχόμενο των προηγούμενων προτύπων του ICA (ISAD(G), ISAAR(CPF), ISDF, ISDIAH) και επιτρέπει πλουσιότερη, σχεσιακά προσανατολισμένη περιγραφή – συχνά ευθυγραμμισμένη με αναπαραστάσεις βασισμένες σε οντολογίες για βελτιωμένη σημασιολογική πρόσβαση · αυτό λειτουργικοποιείται περαιτέρω στο Records in Contexts – Ontology (RiC-O) [ICA25], το οποίο εκφράζει τις οντότητες και τις σχέσεις του RiC-CM σε μορφή OWL/RDF για την υποστήριξη υλοποιήσεων συνδεδεμένων δεδομένων. Η εφαρμογή του RiC-CM σε πραγματικά αρχειακά σύνολα έχει ήδη ξεκινήσει και χαρακτηριστικό παράδειγμα είναι η περιγραφή του Αρχείου του Ποινικού Δικαστηρίου της Κέρκυρας την περίοδο της Ιονίου Πολιτείας (1815–1864) από τους Dimitriou κ.ά. [DDK+25], όπου το μοντέλο χρησιμοποιήθηκε για την αποτύπωση τόσο των εγγραφών όσο και των δραστηριοτήτων και προσώπων που συμμετέχουν σε κάθε υπόθεση.

Για τη μουσειακή τεκμηρίωση και ευρύτερα την πολιτιστική κληρονομιά, το κυρίαρχο εννοιολογικό μοντέλο είναι το CIDOC-CRM (Conceptual Reference Model) [CRM24], διεθνές πρότυπο ISO 21127, που αναπτύχθηκε από την ICOM-CIDOC. Σε αντίθεση με σχήματα τύπου «πεδίο-τιμή» όπως το Dublin Core, το CIDOC-CRM είναι ένα πλήρες οντολογικό μοντέλο με κλάσεις (π.χ. E22 Human-Made Object, E39 Actor, E5 Event) και ιδιότητες (P-properties) που επιτρέπουν τη λεπτομερή περιγραφή σύνθετων ιστορικών



Σχήμα 8. Βασικές κλάσεις και ιδιότητες του CIDOC-CRM - οντολογικό μοντέλο για τη διασύνδεση δεδομένων πολιτιστικής κληρονομιάς [Gil08, ATO20]

διαδικασιών, δημιουργία, μεταβίβαση κυριότητας, έκθεση, καταστροφή (Σχ. 8). Η δύναμη του έγκειται στο ότι επιτρέπει τη διασύνδεση δεδομένων από μουσεία, αρχεία και βιβλιοθήκες σε γραφήματα γνώσης (knowledge graphs), ανοίγοντας τον δρόμο για σημασιολογικές αναζητήσεις που ξεπερνούν τα όρια ενός μεμονωμένου ιδρύματος. Η υιοθέτησή του απαιτεί υψηλό επίπεδο εννοιολογικής κατανόησης και συνεκτικό mapping από υφιστάμενα σχήματα, κάτι που αποτελεί σημαντικό εμπόδιο για πολλούς φορείς [Gil08, ATO20]. Ενδιαφέρουσα σε αυτό το πλαίσιο είναι η εργασία των Stamou κ.ά. [SKD+25], οι οποίοι εξετάζουν τη δυνατότητα χρήσης LLMs για τη δημιουργία αφηγηματικών κειμένων από δεδομένα κωδικοποιημένα σε CIDOC CRM, δείχνοντας ότι η γέφυρα μεταξύ σύνθετων οντολογικών μοντέλων και φυσικής γλώσσας μπορεί να χτιστεί μέσω LLMs, μια προοπτική που ανοίγει νέες δυνατότητες για αυτόματη παραγωγή αρχειακών ή μουσειακών περιγραφών από δομημένα δεδομένα [SKD+25]. Έτσι η σχέση μεταξύ RiC-CM και CIDOC-CRM, ως δύο μοντέλων που εξυπηρετούν συμπληρωματικούς σκοπούς στη διαχείριση της πολιτιστικής πληροφορίας, αποτελεί ενεργό ερευνητικό πεδίο. Οι Bountouri κ.ά. [BDD+23] προτείνουν συστηματική σημασιολογική αντιστοίχιση των κύριων οντοτήτων, ιδιοτήτων και σχέσεων του RiC-CM στις αντίστοιχες έννοιες του CIDOC-CRM, διευκολύνοντας τη διαλειτουργικότητα μεταξύ αρχειακών και ευρύτερων πολιτιστικών δεδομένων.

Γύρω από αυτά τα μεγάλα σχήματα υπάρχει ένα ευρύτερο οικοσύστημα συμπληρωματικών προτύπων που καλύπτουν εξειδικευμένες ανάγκες. Το MARC 21 [LoC99] παραμένει το κυρίαρχο πρότυπο για την ανταλλαγή και τη διαλειτουργικότητα βιβλιογραφικών εγγραφών μεταξύ βιβλιοθηκών παγκοσμίως, με πλούσια κωδικοποίηση για ονόματα, θεματικές επικεφαλίδες και ταξινομικούς αριθμούς. Η δομή του όμως, σχεδιασμένη αποκλειστικά για βιβλιογραφικές εγγραφές, δεν είναι εύκολα προσαρμόσιμη στις ανάγκες αρχειακής περιγραφής. Το MODS (Metadata Object Description Schema) [LoC18], προσφέρει μια πιο ευέλικτη XML εναλλακτική, πιο κοντά στον κόσμο του MARC αλλά καλύτερα προσαρμοσμένη σε ψηφιακά αντικείμενα και μεικτές συλλογές. Σε επίπεδο ψηφιακής διατήρησης, τα METS (Metadata Encoding and Transmission Standard) [LoC10] και PREMIS (Preservation Metadata: Implementation Strategies) [PRM15], αποτελούν τα βασικά εργαλεία για τη δέσμευση και ανταλλαγή μεταδεδομένων διατήρησης, πληροφορίες για αρχειακές πράξεις, τεχνικές παραμέτρους, ιστορικό αλλαγής αρχείων και δικαιώματα πρόσβασης. Στην πράξη, οι φορείς συνδυάζουν συχνά περισσότερα από ένα πρότυπα: οι Silva κ.ά. [Sil21] δείχνουν πώς συνδυασμοί Dublin Core, PREMIS και METS μπορούν να καλύψουν ταυτόχρονα περιγραφικά, διοικητικά και μεταδεδομένα διατήρησης σε περιβάλλοντα ανοικτών δεδομένων, ενώ οι Lin κ.ά. [LCD+20] υπογραμμίζουν ότι η συμμόρφωση με αναγνωρισμένα πρότυπα αποτελεί βασική προϋπόθεση για την αξιοπιστία και την εμπιστοσύνη στα ψηφιακά αποθετήρια.

Η Gilliland [Gil08] επισημαίνει ότι η επιλογή προτύπου ή συνδυασμού προτύπων πρέπει πάντα να καθοδηγείται από τους στόχους του φορέα και τις ανάγκες των χρηστών. Ένα ελαφρύ σχήμα όπως το Dublin Core είναι κατάλληλο για βασική αναζήτηση και διαλειτουργικότητα, αλλά ανεπαρκές όταν απαιτούνται λεπτομερείς διοικητικές, νομικές ή τεχνικές πληροφορίες. Αντίστροφα, η πλήρης υιοθέτηση ISAD(G), EAD ή CIDOC-CRM μπορεί να είναι υπερβολικά απαιτητική για μικρούς φορείς. Για τα LLMs, αυτό σημαίνει κάτι πολύ συγκεκριμένο, ότι το σύνολο πεδίων, κανόνων και σχέσεων κάθε προτύπου, δεν μπορεί να αντιμετωπίζεται ως αφηρημένη «καλή πρακτική». Πρέπει να ενσωματώνεται ρητά στις οδηγίες, να ορίζεται ποια πεδία είναι υποχρεωτικά, ποια προαιρετικά, ποια λεξιλόγια προτιμώνται και πώς όλα αυτά χαρτογραφούνται σε συγκεκριμένο φορμάτ (EAD XML, RDF/DCAT, CRM-based γραφήματα κ.λπ.). Χωρίς αυτή τη ρητή σύνδεση, το αποτέλεσμα ενός LLM κινδυνεύει να είναι γενικά «λογικό» αλλά πρακτικά ασύμβατο με τα πραγματικά συστήματα που θα το χρησιμοποιήσουν [Gil08, LCD+20].

Ακολουθεί ένας συγκριτικός πίνακας με επισκόπηση προτύπων μεταδεδομένων. (Πιν.1)

Πίνακας 1. Συγκριτική επισκόπηση προτύπων μεταδεδομένων ως προς τον χώρο εφαρμογής, τη δομή και τη συμβατότητα με LLMs [Gil08, ICA00, ATO20, LCD+20]

ΠΡΟΤΥΠΟ	ΑΡ. ΠΕΔΙΩΝ	ΕΙΔΟΣ ΣΧΗΜΑΤΟΣ	ΧΩΡΟΣ ΕΦΑΡΜΟΓΗΣ	ΙΕΡΑΡΧΙΚΟΤΗΤΑ	ΣΥΜΒΑΤΟΤΗΤΑ ΜΕ LLMs
ISAD(G)/ ΔΙΠΑΠ(Γ) ICA, 2000	26	Εννοιολογικό πρότυπο πεδίων	Αρχεία	Ναι fonds → item	Μέτρια Απαιτεί ρητό prompt engineering για ιεραρχία
EAD SAA/LC, 2015	~150 tags	XML DTD/Schema	Αρχεία	Ναι Nested XML components	Μέτρια Πολύπλοκη δομή XML, χρειάζεται validation
CIDOC-CRM ISO 21127	~90 κλάσεις / ~150 ιδιότητες	Οντολογία (RDF/OWL)	Μουσεία Πολιτιστική κληρονομιά	Ναι Γραφήματα γνώσης	Δύσκολη Υψηλή πολυπλοκότητα, απαιτεί εξειδίκευση
Dublin Core DCMI, 1998	15	Απλό σχήμα πεδίων	Βιβλιοθήκες Αρχεία Μουσεία	Όχι	Υψηλή Απλή δομή, εύκολο για zero-shot prompting
MARC 21 LC	~200 πεδία	Διαδικό / MARCXML	Βιβλιοθήκες	Όχι	Χαμηλή Πολύπλοκη κωδικοποίηση, δύσκολο για LLMs
MODS LC	~20 κορυφαία στοιχεία	XML Schema	Βιβλιοθήκες Ψηφιακές συλλογές	Όχι	Μέτρια XML αλλά πιο ευανάγνωστο από MARC
METS/ PREMIS LC	METS: ~10 τμήματα / PREMIS: ~5 οντότητες	XML Schema	Ψηφιακή διατήρηση	Όχι	Χαμηλή Τεχνικά μεταδεδομένα, όχι περιγραφικά
DCAT W3C	~15 κορυφαία στοιχεία	RDF/OWL Vocabulary	Ανοιχτά δεδομένα Κυβερνητικά portals	Όχι	Μέτρια RDF αλλά σχετικά απλή δομή

2.3 Εφαρμογές ΤΝ σε πολιτιστική πληροφορία

Η εφαρμογή των LLMs και της Τεχνητής Νοημοσύνης γενικότερα στον χώρο της πολιτιστικής κληρονομιάς, πληροφορίας και αρχείων αποτελεί ένα σχετικά νέο αλλά ταχέως αναπτυσσόμενο ερευνητικό πεδίο. Η βιβλιογραφία κινείται σε δύο παράλληλες κατευθύνσεις που η παρούσα ενότητα επιχειρεί να αναδείξει. Η πρώτη εξετάζει τις δυνατότητες που προσφέρει η Τεχνητή Νοημοσύνη (AI) για την αυτοματοποίηση εργασιών τεκμηρίωσης, εμπλουτισμού μεταδεδομένων και πρόσβασης, ενώ η δεύτερη,

εστιάζει στους περιορισμούς, τα ηθικά ζητήματα και τις πρακτικές προϋποθέσεις για αξιόπιστη εφαρμογή της.

Τα τελευταία 20–25 χρόνια, τα αρχεία, τα μουσεία και οι βιβλιοθήκες πέρασαν από ένα πρώτο στάδιο απλής ψηφιοποίησης, όπως σκαναρισμένες εικόνες, βασικά μεταδεδομένα, online catalogs, σε ένα πολύ πιο απαιτητικό στάδιο όπου χρειάζεται «νοημοσύνη» πάνω στα δεδομένα. Την αυτόματη περιγραφή, αναζήτηση σε πολύ μεγάλα σύνολα, την εξατομικευμένη πρόσβαση και την προληπτική συντήρηση. Αυτή η μετάβαση δεν είναι ομαλή ούτε ενιαία για όλους τους φορείς. Οι Colavizza κ.ά. [CBJ+21] και οι Jaillant & Caruto [JC22] επισημαίνουν ότι στα ψηφιακά παραγόμενα αρχεία όπως email, web αρχεία, προσωπικά αρχεία δημιουργών, η βασική «κρίση» είναι ότι, για λόγους ιδιωτικότητας, πνευματικών δικαιωμάτων και τεχνικής πολυπλοκότητας, τεράστιοι όγκοι υλικού παραμένουν «σκοτεινοί» (dark archives): ούτε αναζητήσιμοι ούτε προσβάσιμοι στο ευρύ ή και στο ερευνητικό κοινό. Παράλληλα, οι Griffin κ.ά. [GWF24] δείχνουν ότι στα πολιτιστικά ιδρύματα η ΤΝ εμφανίζεται κατά κανόνα περισσότερο πειραματικά, με μικρές πιλοτικές εφαρμογές, παρά ως ώριμη, συστημική υποδομή, οι επαγγελματίες συχνά αισθάνονται ότι «τρέχουν να προλάβουν» την τεχνολογία με περιορισμένο προσωπικό και δεξιότητες. Η μεγάλη υπόσχεση της ΤΝ είναι ότι μπορεί να βοηθήσει να «ξεκλειδωθούν» αυτά τα δεδομένα, αρκεί αυτό να γίνει με τρόπο ηθικό, διάφανο και συνεργατικό.

Ο Pavlidis [Pav22] καταγράφει, σε ανασκόπηση των τάσεων της ΑΙ στις ψηφιακές ανθρωπιστικές, εφαρμογές που εκτείνονται από τη 3D ψηφιοποίηση και την αυτοματοποίηση τεκμηρίωσης μέχρι την ερμηνεία και την εξατομικευμένη πρόσβαση, συμπεριλαμβανομένων deep learning εφαρμογών για ανάγνωση επιγραφών, πρόβλεψη θέσεων αρχαιολογικών ταφών, αναγνώριση υλικών μέσω φασματοσκοπίας και δημιουργία έξυπνων εικονικών μουσείων. Στο επίπεδο της τεκμηρίωσης, ιδιαίτερο ενδιαφέρον παρουσιάζουν οι προσπάθειες για αυτόματη δημιουργία ή εμπλουτισμό μεταδεδομένων: αναγνώριση αντικειμένων σε εικόνες, σύνδεση με οντολογίες και ανοιχτά λεξιλόγια (π.χ. Wikidata), εξαγωγή οντοτήτων από ιστορικά κείμενα για δημιουργία σημασιολογικών δικτύων.

Οι Ardissono κ.ά. [ARM22] περιγράφουν σειρά εφαρμογών όπου τεχνικές ΑΙ και HCI συνδυάζονται για να βελτιώσουν την πρόσβαση σε πολιτιστικές συλλογές: συστήματα που προσαρμόζουν το περιεχόμενο σε πραγματικό χρόνο ανάλογα με το προφίλ του χρήστη, αυτοματοποιημένα εργαλεία ερμηνείας εκθεμάτων και έξυπνα περιβάλλοντα έκθεσης που αξιοποιούν δεδομένα χρήσης για να αναδιαμορφώσουν την παρουσίαση. Αντίστοιχα, οι Münster κ.ά. [MML+24] θέτουν ως στόχο την ανάπτυξη ερευνητικής ατζέντας για την ΑΙ στην ψηφιακή κληρονομιά, αναγνωρίζοντας ότι οι υπάρχουσες

υποδομές συλλογής και ψηφιοποίησης μπορούν να γίνουν πολύ περισσότερο «έξυπνες» με την έγκαιρη ενσωμάτωση εργαλείων TN, και προτείνουν προτεραιότητες όπως η κλιματική ανθεκτικότητα μνημείων, η βελτιωμένη τεκμηρίωση, η ένταξη κοινοτήτων και η προσβασιμότητα. Ένα ενδιαφέρον σημείο που αναδύεται από τη σύγκριση αυτών των εργασιών είναι ότι, ενώ σε άρθρα «επίδειξης» περιπτώσεων χρήσης η TN εμφανίζεται σχεδόν μαγική, στις έρευνες που εστιάζουν στην καθημερινότητα των φορέων εμφανίζεται ως κάτι εύθραστο, γεμάτο σφάλματα, που χρειάζεται πολλή «χειροκίνητη φροντίδα» για να βγάλει πρακτικό νόημα [GWF24, MML+24].

Στο επίπεδο των μεταδεδομένων, οι Almaaitah κ.ά. [ATO20] επισημαίνουν ότι τα μεγαλύτερα εμπόδια στην πρόσβαση σε συλλογές πολιτιστικής κληρονομιάς είναι πρακτικής φύσης: ανομοιογενή σχήματα, ελλιπής ή ασυνεπής περιγραφή, έλλειψη σημασιολογικού πλούτου και περιορισμένη διαλειτουργικότητα. Παρότι προτείνονται διάφορες τεχνικές λύσεις (semantic web, οντολογίες, linked data), η χρήση ΑΙ για αυτοματοποίηση παραμένει σε μεγάλο βαθμό πιλοτική. Οι Colavizza κ.ά. [CBJ+21] και οι Jaillant & Caruto [JC22] δείχνουν ότι, σε επίπεδο αρχειακών πρακτικών, η ΑΙ χρησιμοποιείται κυρίως για εργασίες όπως η αναγνώριση ονομάτων και η ανίχνευση ευαίσθητων δεδομένων σε born-digital συλλογές, αλλά σπανίως για πλήρη παραγωγή αρχειακών περιγραφών. Στο υπόβαθρο όλων αυτών βρίσκεται μια κεντρική παρατήρηση: χωρίς ποιοτικά και πλούσια μεταδεδομένα, ούτε τα πιο προηγμένα μοντέλα TN μπορούν να λειτουργήσουν με τρόπο αξιόπιστο και κοινωνικά χρήσιμο [ATO20, WBM+23].

Πέρα από τα ζητήματα διαλειτουργικότητας και ποιότητας δεδομένων, η βιβλιογραφία αναδεικνύει και μία σειρά πιο θεμελιωδών περιορισμών που σχετίζονται άμεσα με τη φύση των LLMs. Ένα βασικό πρόβλημα είναι το πρόβλημα των hallucinations. Στον αρχειακό χώρο μία λανθασμένη χρονολογία ή ένα επινοημένος παραγωγός δεν συνιστούν απλώς τεχνικό σφάλμα, αλλά μπορούν να αλλοιώσουν την ιστορική ερμηνεία ενός συνόλου τεκμηρίων [RSD23, ZHP+23]. Ακόμη και με αυστηρά prompts, τα μοντέλα παρουσιάζουν δομική ασυνέπεια, δεν εγγυώνται δηλαδή μία σταθερή συμπλήρωση όλων των πεδίων ενός προτύπου όπως το ISAD(G) σε κάθε εκτέλεση [WBM+23]. Σε αυτά προστίθεται και η εξάρτηση από την ποιότητα των δεδομένων εκπαίδευσης. Τα μοντέλα αναπαράγουν μεροληψίες των σωμάτων κειμένων στα οποία εκπαιδεύτηκαν, κάτι που στον αρχειακό χώρο μπορεί να σημαίνει ότι ορισμένες γλώσσες, ιστορικές περιοδοί ή κοινότητες, δεν εκπροσωπούνται όσο πρέπει, με αποτέλεσμα να υπάρχει χαμηλότερη αξιοπιστία στην επεξεργασία του αντίστοιχου υλικού [GWF24, OR24]. Τέλος οι Toth et al. [TAP25] επισημαίνουν μία λεπτή αλλά κρίσιμη πτυχή. Ακόμα και όταν ένα LLM παράγει τεχνικά σωστά αποτελέσματα, αν ο αρχειονόμος δεν μπορεί να κατανοήσει γιατί το μοντέλο κατέληξε σε μία συγκεκριμένη εγγραφή, δηλαδή αν λείπει η ερμηνευσιμότητα

(explainability), τότε δεν μπορεί να αξιολογήσει ούτε να εγκρίνει το αποτέλεσμα με εμπιστοσύνη. Σε αυτό το πλαίσιο, τα explainable AI συστήματα δεν αποτελούν απλώς τεχνική επιλογή, αλλά επαγγελματική και δεοντολογική απαίτηση για τον αρχειακό χώρο.

Σε συνέχεια, ένα από τα πιο ευαίσθητα ζητήματα που αναδύεται στη βιβλιογραφία αφορά την πρόσβαση στα born-digital αρχεία και τη σχέση της με την ιδιωτικότητα. Οι Colavizza κ.ά. [CBJ+21] τονίζουν ότι οι θεσμοί συχνά επιλέγουν τη νομικά ασφαλή λύση να κλείνουν ολόκληρες συλλογές επ' αόριστον, αντί να επενδύσουν σε διαδικασίες λεπτομερούς ελέγχου ευαίσθητου περιεχομένου. Τεχνικά, εργαλεία ML μπορούν να βοηθήσουν στη «sensitivity review»: να επισημαίνουν προσωπικά αναγνωρίσιμες πληροφορίες (PII), νομικά ή ηθικά ευαίσθητο υλικό, να προτείνουν ανωνυμοποίηση ή μερική απόκρυψη, πάντα όμως με τελικό ανθρώπινο έλεγχο. Σε αυτή την κατεύθυνση κινούνται οι Hathurusinghe κ.ά. [HNB21], οι οποίοι προτείνουν μια privacy-preserving προσέγγιση για την αυτόματη εξαγωγή προσωπικών πληροφοριών μέσω ανώνυμης επισημείωσης και federated learning, δείχνοντας ότι είναι τεχνικά εφικτό να εξαχθούν μεταδεδομένα από ευαίσθητο υλικό χωρίς να εκτεθούν τα ίδια τα δεδομένα. Παράλληλα, οι Shepherd κ.ά. [SCK+24] μελετούν τα «σημεία πόνου» στη διαχείριση προσωπικών αρχείων, εντοπίζοντας σημαντικές αδυναμίες στις υπάρχουσες πρακτικές και επισημαίνοντας την ανάγκη για πιο συστηματικές και ηθικά τεκμηριωμένες προσεγγίσεις που να λαμβάνουν υπόψη τόσο τις ανάγκες των δημιουργών των αρχείων όσο και των μελλοντικών ερευνητών. Σε θεσμικό επίπεδο, η βιβλιογραφία προτείνει «μέσες λύσεις» ανάμεσα στο πλήρως κλειστό και το πλήρως ανοικτό, ελεγχόμενη πρόσβαση μέσω ασφαλών πλατφορμών με ταυτοποιημένους χρήστες, αλλά και διεθνείς ή εθνικές κοινοπραξίες για born-digital υλικό, ώστε μικρότεροι φορείς να μοιράζονται υποδομές και τεχνογνωσία και να μη βρίσκονται σε άνιση θέση [CBJ+21, SCK+24].

Σημαντική διάσταση αποτελούν επίσης οι ρυθμιστικές και ηθικές πτυχές. Οι Svärd κ.ά. [SGB+24] αναλύουν πώς διάφορες δικαιοδοσίες επιχειρούν να ρυθμίσουν τη χρήση AI σε διαχείριση δημόσιων αρχείων, καταλήγοντας ότι βρισκόμαστε ακόμη σε πρώιμο στάδιο: κυριαρχούν γενικές αρχές (διαφάνεια, λογοδοσία, προστασία δεδομένων) αλλά σπάνια μεταφράζονται σε συγκεκριμένες οδηγίες για εργαλεία AI σε αρχειακό περιβάλλον. Οι Griffin κ.ά. [GWF24], στην περίπτωση σουηδικών πολιτιστικών οργανισμών, δείχνουν ότι τα πρακτικά εμπόδια όπως η έλλειψη προσωπικού, οι τεχνικές δεξιότητες, η στρατηγική ψηφιοποίησης, είναι εξίσου σημαντικά με τα τεχνικά: χωρίς ανθρώπινο και θεσμικό υπόβαθρο, ακόμη και τα πιο προηγμένα μοντέλα μένουν στο στάδιο «πλότου». Οι ίδιοι συγγραφείς αναφέρουν ότι οι επαγγελματίες εντοπίζουν ως βασικά εμπόδια την έλλειψη στρατηγικής, τον κατακερματισμό συστημάτων, τα μικρά και εξειδικευμένα σύνολα δεδομένων που δυσκολεύουν την εκπαίδευση μοντέλων, αλλά και τον φόβο για

το πώς η TN θα επηρεάσει τον ρόλο και την ευθύνη τους [GWF24]. Η δεοντολογική διάσταση είναι σαφής: η TN δεν είναι μόνο τεχνικό εργαλείο αλλά και μηχανισμός εξουσίας πάνω στη συλλογική μνήμη· άρα οι αποφάσεις για το τι είναι ορατό, σε ποιον και με ποιους όρους, πρέπει να είναι ρητές, διαφανείς και συζητήσιμες [SGB+24, TAP25].

Αν κανείς επιχειρήσει να συνθέσει τα νήματα που αναδύονται από τη βιβλιογραφία, διακρίνονται ορισμένες κοινές κατευθύνσεις που συνδέουν θεωρία με πράξη. Πρώτον, η ενδυνάμωση επαγγελματιών: χωρίς επαγγελματίες που κατανοούν τόσο το υλικό όσο και τα βασικά της TN, δεν υπάρχει ουσιαστική υιοθέτηση, η βιβλιογραφία επισημαίνει την ανάγκη για μακρόχρονα προγράμματα κατάρτισης και δημιουργία «υβριδικών» ρόλων (curator-data scientist) μέσα στους ίδιους τους φορείς [GWF24, MML+24]. Επιπροσθέτως για την συνεπιμέλεια ανθρώπου-μηχανής, τα πιο πειστικά σενάρια που εμφανίζονται στη βιβλιογραφία είναι εκεί όπου η TN μειώνει τον όγκο της εργασίας (π.χ. προτάσεις μεταδεδομένων, σηματοδότηση ευαίσθητων τμημάτων, clustering μεγάλων σωμάτων κειμένου), ενώ η τελική κρίση παραμένει σε αρχειονόμους και επιμελητές που γνωρίζουν το ιστορικό πλαίσιο [CBJ+21, JC22]. Επίσης, η διαλειτουργικότητα και οι κοινές υποδομές, η επένδυση σε κοινά μοντέλα (CIDOC CRM, SKOS, κοινά ελεγχόμενα λεξιλόγια) και σε συλλογικές υποδομές τύπου Europeana εμφανίζεται ως βασικός όρος για να μη μείνουν μικροί φορείς απομονωμένοι [ATO20, SKD+25]. Εξίσου σημαντική είναι η δεοντολογική διάσταση που αντισταθμίζει την τεχνολογική αισιοδοξία. Η συζήτηση για bias, αδιαφάνεια και δικαιώματα των υποκειμένων των αρχείων είναι έντονη, από το ποιος «χάνεται» όταν εκπαιδεύουμε μοντέλα πάνω σε δεδομένα που υπεραντιπροσωπεύουν ορισμένες ομάδες, μέχρι το πώς διαχειριζόμαστε τα σφάλματα ενός μοντέλου που κάνει λανθασμένη κρίση για το τι είναι ευαίσθητο [SGB+24, HNB21]. Τέλος, η στόχευση σε κοινωνικό αντίκτυπο, αρκετές εργασίες στις ψηφιακές ανθρωπιστικές υποστηρίζουν ότι η TN αποκτά νόημα όταν βοηθά να τεθούν νέα κοινωνικά και ιστορικά ερωτήματα, να φωτιστούν υποεκπροσωπούμενες ομάδες, να επαναδιαβαστούν αρχεία με οπτικές φύλου, φυλής και κοινωνικής τάξης, να γίνουν οι συλλογές πιο συμπεριληπτικές και προσβάσιμες, όχι απλώς πιο «γρήγορα επεξεργάσιμες» [OR24, Pav22].

Σε αυτό το πλαίσιο, αναδύεται και ένα πιο ειδικό πεδίο που αφορά τη βελτίωση της συμπεριληπτικότητας (inclusivity) των αρχειακών μεταδεδομένων. Οι Osti & Roke [OR24] δείχνουν ότι τα LLMs μπορούν να εντοπίσουν ελλείψεις ή προκαταλήψεις στις υπάρχουσες εγγραφές, όπως η απουσία πληροφοριών για περιθωριοποιημένες ομάδες, για ομάδες που δεν έχουν εκπροσωπηθεί ιδιαίτερα στην ιστορική καταγραφή (γυναίκες, μειονότητες, εργατικές τάξεις κ.ά.) και να προτείνουν εμπλουτισμό. Η κατεύθυνση αυτή συνδέεται με ευρύτερες συζητήσεις στις ψηφιακές ανθρωπιστικές για το πώς τα αρχεία μπορούν να γίνουν πιο αντιπροσωπευτικά και προσβάσιμα σε ευρύτερες κοινότητες,

πέρα από τους παραδοσιακούς χρήστες τους. Στην ίδια λογική, οι Stamou et. Al [SKD+25] αναδεικνύουν τη συμπληρωματική όψη του προβλήματος, πως δηλαδή τα δεδομένα που είναι ήδη κωδικοποιημένα σε σύνθετα μοντέλα όπως το CIDOC-CRM μπορούν να μετασχηματιστούν, μέσω LLMs, σε αναγνώσιμες αφηγηματικές περιγραφές που απευθύνονται σε ανθρώπινο κοινό, γεφυρώνοντας την απόσταση ανάμεσα στην οντολογική πληρότητα και την προσβασιμότητα.

Η συνολική εικόνα από τη βιβλιογραφία είναι διττή. Από τη μία πλευρά, οι μέθοδοι NLP/AI για εξαγωγή μεταδεδομένων έχουν ωριμάσει σε αρκετούς τομείς και προσφέρουν έτοιμα, εκμεταλλεύσιμα παραδείγματα. Από την άλλη, στον αρχαιακό χώρο, ιδίως όταν το ζητούμενο είναι η συμμόρφωση με πρότυπα όπως το ISAD(G), οι εφαρμογές παραμένουν περιορισμένες, αποσπασματικές και συχνά εστιασμένες σε επιμέρους βήματα και όχι σε ολοκληρωμένη παραγωγή περιγραφών. Αυτό το κενό ανάμεσα στο «τι μπορεί να κάνει» η AI γενικά και στο «τι κάνει» ήδη για τα αρχεία, είναι ακριβώς ο χώρος μέσα στον οποίο τοποθετείται η παρούσα εργασία.

Κεφάλαιο 3 Θεωρητικό Πλαίσιο

3.1 Μεταδεδομένα και Εξαγωγή Μεταδεδομένων

Ο όρος «μεταδεδομένα» χρησιμοποιείται συχνά με απλοϊκό τρόπο ως «δεδομένα για τα δεδομένα», όμως η βιβλιογραφία αναδεικνύει ότι πρόκειται για μία πολύ πιο σύνθετη έννοια. Η Gilliland ορίζει τα μεταδεδομένα ως δομημένη πληροφορία που περιγράφει, εξηγεί, εντοπίζει ή διευκολύνει τη διαχείριση, την πρόσβαση και τη χρήση ενός πόρου, υπογραμμίζοντας ότι δεν είναι μόνο περιγραφικά αλλά μπορεί να είναι και δομικά, διοικητικά, τεχνικά ή μεταδεδομένα διατήρησης [Gil08]. Οι Yang κ.ά. [YFA+25], επεκτείνουν αυτή την οπτική σε ένα περιβάλλον «data-driven», όπου τα μεταδεδομένα λειτουργούν ως κρίσιμη υποδομή για τη διακυβέρνηση δεδομένων, την ιχνηλασιμότητα, την ποιότητα και τη λογοδοσία σε πολύπλοκα οικοσυστήματα πληροφορίας.

Στη βιβλιογραφία διακρίνονται παραδοσιακά πέντε βασικές κατηγορίες μεταδεδομένων, καθεμία από τις οποίες εξυπηρετεί διαφορετικό σκοπό στον κύκλο ζωής ενός πόρου [Gil08, RB24]. Αυτές είναι, τα περιγραφικά μεταδεδομένα, τα δομικά, τα διοικητικά, τα τεχνικά μεταδεδομένα και τα μεταδεδομένα διατήρησης.

Τα περιγραφικά μεταδεδομένα (descriptive metadata) είναι ίσως ο πιο διαδεδομένος τύπος και αφορούν την αναγνώριση και ανάκτηση ενός πόρου: τίτλος, δημιουργός/παραγωγός, ημερομηνία, θέμα, λέξεις-κλειδιά, περίληψη. Παραδείγματα σχημάτων αυτής της κατηγορίας είναι το Dublin Core [DCMI98] και το MODS, ενώ στον αρχειακό χώρο το ISAD (G) [ICA00] και το EAD [LCD+20] εξυπηρετούν τον ίδιο σκοπό με ιεραρχική λογική.

Τα δομικά μεταδεδομένα (structural metadata) περιγράφουν τις εσωτερικές σχέσεις μεταξύ των μερών ενός πόρου ή μεταξύ πολλαπλών πόρων. Πως οργανώνονται τα κεφάλαια ενός βιβλίου, πως συνδέονται οι σελίδες ενός σαρωμένου τεκμηρίου, ή πως ένα τεκμήριο εντάσσεται σε μία ιεραρχία αρχειακών επιπέδων (fonds, series, file, item). Το METS αποτελεί χαρακτηριστικό παράδειγμα προτύπου που εξυπηρετεί τη δομική αναπαράσταση πολυσύνθετων ψηφιακών αντικειμένων [Sil21].

Τα διοικητικά μεταδεδομένα (administrative metadata) αφορούν τη διαχείριση του πόρου ως διοικητικού αγαθού. Ποιος είναι ο κάτοχος των δικαιωμάτων, ποιες είναι οι

άδειες χρήσης, ποιος έχει πρόσβαση και υπό ποιες προϋποθέσεις, ποιο είναι το ιστορικό τροποποιήσεων [Gil08]. Στον δημόσιο αρχειακό χώρο, αυτή η κατηγορία αποκτά ιδιαίτερη βαρύτητα λόγω νομοθετικών απαιτήσεων που σχετίζονται με το GDPR και την πρόσβαση σε ευαίσθητο υλικό [SGB+24].

Τα μεταδεδομένα διατήρησης (preservation metadata) εξυπηρετούν τη μακροπρόθεσμη διατήρηση και αυθεντικότητα του ψηφιακού πόρου. Καταγράφουν τη φυσική και λογική του δομή, το ιστορικό μετατροπών σε διαφορετικά formats, τις τεχνικές προδιαγραφές του, και την αλυσίδα φύλαξης (chain of custody). Το πρότυπο PREMIS είναι το πιο διαδεδομένο σε αυτήν την κατηγορία, ενώ στην πράξη συνδυάζεται συχνά με Dublin Core και METS για ολοκληρωμένη κάλυψη [Sil21].

Τα τεχνικά μεταδεδομένα (technical metadata) τέλος, περιγράφουν τα τεχνικά χαρακτηριστικά του ψηφιακού πόρου, δηλαδή, τη μορφή αρχείου, το μέγεθος, την ανάλυση, την κωδικοποίηση και το λογισμικό που το παρήγαγε. Συχνά παράγονται αυτόματα κατά την ψηφιοποίηση ή δημιουργία του τεκμηρίου και αποτελούν προαπαιτούμενο για κάθε μεταγενέστερη επεξεργασία.

Όμως πέρα από την κλασική τυπολογία, η σύγχρονη βιβλιογραφία αναδεικνύει και νεότερες κατηγορίες. Τα παραδεδομένα (paradata) καταγράφουν τη διαδικασία παραγωγής του πόρου, δηλαδή, ποιος, με ποια εργαλεία, με ποιες αποφάσεις και αποκτούν ιδιαίτερη σημασία στο πλαίσιο της αυτόματης παραγωγής μεταδεδομένων από ΑΙ συστήματα, όπου η τεκμηρίωση του «πώς» και «από ποιον/τι» είναι κρίσιμη για την αξιοπιστία [PGF23, TAP25]. Οι Yang κ.ά. [YFA+25] εισάγουν επιπλέον την έννοια των μεταδεδομένων διακυβέρνησης (governance metadata), τα οποία υποστηρίζουν τη λογοδοσία, την ιχνηλασιμότητα και τη συμμόρφωση με κανονιστικά πλαίσια σε σύνθετα οικοσυστήματα πληροφορίας.

Στην πράξη, οι φορείς συνδυάζουν συχνά περισσότερα από ένα πρότυπα: οι Silva κ.ά. [Sil21] δείχνουν πώς συνδυασμοί Dublin Core, PREMIS και METS μπορούν να καλύψουν ταυτόχρονα περιγραφικά, διοικητικά και μεταδεδομένα διατήρησης σε περιβάλλοντα ανοικτών δεδομένων, ενώ οι Lin κ.ά. [LCD+20] υπογραμμίζουν ότι η συμμόρφωση με αναγνωρισμένα πρότυπα αποτελεί βασική προϋπόθεση για την αξιοπιστία και την εμπιστοσύνη στα ψηφιακά αποθετήρια.

Στο πλαίσιο της παρούσας εργασίας τώρα, εστιάζονται πρωτίστως τα περιγραφικά μεταδεδομένα (descriptive metadata), δηλαδή εκείνα που βοηθούν στον εντοπισμό και την κατανόηση αρχειακών τεκμηρίων (τίτλος, δημιουργός/παραγωγός, ημερομηνία, θεματικοί όροι, περιλήψεις κ.λπ.), αλλά αναγνωρίζεται επίσης ότι η αρχειακή περιγραφή

εμπλέκει και δομικά (σχέσεις μεταξύ μονάδων), διοικητικά (συνθήκες πρόσβασης, δικαιώματα) και μεταδεδομένα διατήρησης. Η βιβλιογραφία της ψηφιακής διατήρησης, μέσα από σχήματα όπως τα PREMIS, και METS, επισημαίνει ότι αυτά τα επίπεδα μεταδεδομένων αλληλοεπικαλύπτονται και συχνά χρειάζονται συνδυαστικά σχήματα, όπως για παράδειγμα, Dublin Core σε συνδυασμό με PREMIS και METS, ώστε να καλυφθεί όλος ο κύκλος ζωής των ψηφιακών πόρων [ATO20, NR23]. Για τα αρχεία, αυτό σημαίνει ότι καμία προσέγγιση εξαγωγής μεταδεδομένων δεν μπορεί να αγνοήσει εντελώς το πλαίσιο χρήσης. Άλλο είδος μεταδεδομένων απαιτείται για βασική αναζήτηση, άλλο για μακροχρόνια διατήρηση, άλλο για διαλειτουργικότητα με άλλα ιδρύματα όπως βιβλιοθήκες και μουσεία [ATO20].

3.1.1 Structured vs unstructured extraction

Η διαδικασία εξαγωγής μεταδεδομένων μπορεί να χωριστεί, σε υψηλό επίπεδο, σε δύο κατηγορίες, structured και unstructured extraction. Στην πρώτη περίπτωση, εργαζόμαστε με ήδη δομημένα ή ημι-δομημένα δεδομένα, για παράδειγμα, πίνακες βάσεων δεδομένων, CSV, XML ή JSON αρχεία, όπου τα πεδία είναι σαφώς ορισμένα και η εξαγωγή συνίσταται κυρίως σε mapping μεταξύ διαφορετικών σχημάτων ή στη δημιουργία πιο πλούσιων, ενοποιημένων εγγραφών. Η Yang κ.ά. (2025) περιγράφουν πώς παραδοσιακά metadata management systems (π.χ. CKAN, Amundsen, DataHub) βασίζονται σε structured extraction για να συλλέξουν μεταδεδομένα από ποικίλες πηγές, συνήθως μέσω connectors που «διαβάζουν» υπάρχοντα σχήματα (π.χ. DCAT, relational schemas) και τα ευθυγραμμίζουν με το εσωτερικό μοντέλο του συστήματος.

Στην unstructured extraction, αντίθετα, το σημείο εκκίνησης είναι μη δομημένα ή χαλαρά δομημένα δεδομένα, όπως, PDF επιστημονικών άρθρων, σαρωμένα έγγραφα, ελεύθερο κείμενο, εικόνες, ηχητικά, ακόμη και βίντεο. Εδώ η πρώτη πρόκληση είναι να εντοπιστούν οι περιοχές ενδιαφέροντος (π.χ. τίτλοι, λεζάντες, σώμα κειμένου) και στη συνέχεια να μετασχηματιστούν σε μορφή από την οποία μπορούν να εξαχθούν μεταδεδομένα. Οι Azimjonov & Alikhanov (2018) δείχνουν πόσο κρίσιμο είναι το στάδιο της ανάλυσης του layout για PDF άρθρα. Η επιτυχία της εξαγωγής εξαρτάται από το να ξεχωρίσει κανείς οπτικά και λογικά τμήματα (τίτλος, συγγραφείς, περίληψη) πριν καν φτάσει στον εντοπισμό συγκεκριμένων πεδίων. Οι Granitzer κ.ά. (2012) καταλήγουν σε παρόμοιο συμπέρασμα στο οποίο, χωρίς αξιόπιστη structured αναπαράσταση της σελίδας, ακόμη και τα καλύτερα ταξινομητικά μοντέλα αποτυγχάνουν να αναγνωρίσουν σωστά, π.χ., βιβλιογραφικές αναφορές.

Στον αρχειακό χώρο, η διάκριση αυτή έχει ιδιαίτερη σημασία. Πολλά σύγχρονα αρχεία διαθέτουν «δομημένες» βάσεις δεδομένων καταλόγων (structured data), αλλά τα πρωτογενή τεκμήρια (επιστολές, εκθέσεις, πρακτικά) συνήθως υπάρχουν ως σαρωμένα PDF, εικόνες ή μη επισημειωμένα κείμενα (unstructured data). Η πρόκληση λοιπόν είναι διπλή. Αφενός, να μπορέσουμε να εξαγάγουμε μεταδεδομένα απευθείας από τα πρωτογενή τεκμήρια, αφετέρου να «γεφυρώσουμε» τα υπάρχοντα μεταδεδομένα (structured), τα οποία μπορεί να είναι ελλιπή ή ασυνεπή, με νέες πληροφορίες που προκύπτουν από unstructured extraction. Ειδικά για born-digital αρχεία (emails, αρχεία συστημάτων), όπως δείχνουν οι Jaillant & Caputo (2022) και οι Colavizza κ.ά. (2021), το μεγαλύτερο μέρος της χρήσιμης πληροφορίας παραμένει σε unstructured μορφή και η έλλειψη εργαλείων εξαγωγής αποτελεί έναν από τους βασικούς λόγους που οι συλλογές αυτές μένουν κλειστές ή ελάχιστα περιγεγραμμένες.

3.1.2 Rule-based vs ML vs LLM

Η βιβλιογραφία διακρίνει τρεις κύριες κατηγορίες προσεγγίσεων για την εξαγωγή μεταδεδομένων. Τα rule-based συστήματα, τα συστήματα μηχανικής μάθησης και τις προσεγγίσεις βασισμένες σε LLMs. Οι τρεις αυτές κατηγορίες δεν αντιπροσωπεύουν απλώς τεχνολογική εξέλιξη, αλλά αντιπροσωπεύουν διαφορετικές φιλοσοφίες για το πώς ένα σύστημα «μαθαίνει» να αναγνωρίζει και να οργανώνει την πληροφορία, με διαφορετικές συνέπειες για την αξιοπιστία, τη γενίκευση και τον ανθρώπινο έλεγχο. Η κατανόηση των δυνατοτήτων και των ορίων κάθε κατηγορίας είναι απαραίτητη για να αξιολογηθεί η καταλληλότητά τους στον αρχειακό χώρο.

Τα rule-based ή knowledge-based συστήματα είναι η παλαιότερη και η πιο διαφανής κατηγορία. Η αναγνώριση πεδίων γίνεται αποκλειστικά μέσω ρητών κανόνων τύπου «αν...τότε...» που έχουν οριστεί από τον σχεδιαστή, αν το κείμενο με τη μεγαλύτερη γραμματοσειρά εμφανίζεται πριν τη λέξη «Abstract» και είναι bold, τότε είναι ο τίτλος, αν μετά τη φράση «Keywords:» ακολουθεί λίστα όρων με κόμματα, τότε αυτές είναι οι λέξεις-κλειδιά. Οι κανόνες βασίζονται σε χαρακτηριστικά διάταξης σελίδας (θέση, μέγεθος και στυλ γραμματοσειράς), σε γλωσσικά μοτίβα και σε συγκεκριμένες μορφές (regex) για ημερομηνίες, DOI, ISBN κ.λπ. Οι Azimjonov και Alikhanov [AA18] περιγράφουν χαρακτηριστικά ένα τέτοιο framework για PDF ακαδημαϊκών άρθρων, μαζική συλλογή αρχείων, εντοπισμός πρώτης και τελευταίας σελίδας, εφαρμογή κανόνων για τίτλο, abstract, keywords, κύριο κείμενο, συμπεράσματα και βιβλιογραφία, και αποθήκευση σε βάση και XML. Με προσεκτικά σχεδιασμένους κανόνες και ισχυρό στάδιο μετα-επεξεργασίας, χειρισμός εξαιρέσεων, fallback σε χειρωνακτικό review, φτάνουν σε ακρίβειες άνω του 90% για βασικά πεδία, ξεπερνώντας εργαλεία όπως

GROBID και ParsCit [AA18]. Παρόμοια, οι Choudhury κ.ά. [CMK+13] για figure metadata συνδυάζουν heuristics πάνω σε binarized σελίδες με απλού τύπου ταξινόμηση, δείχνοντας ότι rule-based λογική μπορεί να αποδώσει ακόμα και σε μη κειμενικά δεδομένα.

Το πλεονέκτημα αυτών των συστημάτων είναι ακριβώς η διαφάνεια και η σταθερότητα, η κάθε απόφαση εξηγείται με βάση έναν κανόνα, δεν υπάρχει «μαύρο κουτί», και σε ομοιογενή περιβάλλοντα, π.χ. άρθρα ενός συγκεκριμένου περιοδικού μπορούν να επιτύχουν «βιομηχανική» απόδοση. Το μειονέκτημα είναι η ακαμψία, κάθε αλλαγή μορφότυπου απαιτεί χειρωνακτική αναθεώρηση κανόνων, η δημιουργία και συντήρησή τους είναι εργατοβόρα, και η μεταφορά τους σε νέους τομείς ή γλώσσες είναι ιδιαίτερα δύσκολη. Στον αρχαικό χώρο, ένα αντιστοιχο rule-based workflow θα μπορούσε να λειτουργεί ως εξής, ομαδική επεξεργασία σαρωμένων φακέλων, OCR, εφαρμογή κανόνων τύπου «αν στην πρώτη γραμμή εμφανίζεται "ΥΠΟΥΡΓΕΙΟ..." και ακριβώς από κάτω ημερομηνία, τότε αυτά είναι ο παραγωγός και η χρονολογία», και στη συνέχεια χαρτογράφηση σε EAD ή ISAD(G) δομή σε επίπεδο τεκμηρίου. Σε καλά τυποποιημένες σειρές η απόδοση μπορεί να είναι εντοπωσιακή, αλλά τα αρχεία σπάνια παρουσιάζουν αυτή την ομοιογένεια [Gil08].

Η επόμενη γενιά συστημάτων επιχείρησε να ξεπεράσει αυτή την ακαμψία μέσω μηχανικής μάθησης. Αντί ο σχεδιαστής να ορίζει χειρωνακτικά τους κανόνες, το μοντέλο τους μαθαίνει αυτόματα από annotated παραδείγματα. Οι Granitzer κ.ά. [GHK+12] πραγματοποίησαν μια ιδιαίτερα συστηματική μελέτη, συγκρίνοντας Conditional Random Fields (CRF) και Support Vector Machines (SVM) για την εξαγωγή βιβλιογραφικών μεταδεδομένων από μεγάλες συλλογές κειμένων. Η pipeline τους ξεκινά με αναλυτική εξαγωγή κειμένου και layout από PDF. Κάθε λέξη ή γραμμή συνοδεύεται όχι μόνο από το περιεχόμενό της αλλά και από χαρακτηριστικά όπως μέγεθος και οικογένεια γραμματοσειράς, συντεταγμένες x και y στη σελίδα και κλίσεις αυτών των συντεταγμένων. Το κείμενο μετατρέπεται σε ακολουθίες tokens με d-tuple χαρακτηριστικά, πάνω στις οποίες εκπαιδεύονται Conditional Random Fields (CRFs) και γραμμικά Support Vector Machines (SVMs) που αποφασίζουν για κάθε token αν ανήκει σε τίτλο, ονόματα συγγραφέων, πεδίο περιοδικού ή «απλό» κείμενο. Ένα κρίσιμο εύρημα της εργασίας τους είναι ότι τα post-processing heuristics, απλοί εμπειρικοί κανόνες για να λυθούν ασάφειες, π.χ. ότι ο τίτλος πρέπει να είναι ενιαίο συνεχές μπλοκ ή ότι οι συγγραφείς εμφανίζονται μετά τον τίτλο, είναι απαραίτητοι. Χωρίς αυτά, η απόδοση των καθαρά στατιστικών μοντέλων πέφτει αισθητά [GHK+12].

Σε επόμενο στάδιο, οι Nayaka και Ranjan [NR23] υιοθετούν συνδυασμό CNN σε επίπεδο χαρακτήρων και bi-LSTM για την εξαγωγή σύνθετων μεταδεδομένων από ακαδημαϊκά άρθρα, ψευδοκώδικες, datasets, μετρικές απόδοσης, υπολογιστική πολυπλοκότητα. Χρησιμοποιώντας περίπου 58 χαρακτηριστικά που καλύπτουν περιεχόμενο, στυλ γραμματοσειράς και δομή, επιτυγχάνουν ακρίβεια 94,23% στην ταξινόμηση γραμμών ψευδοκώδικα και 82% στην εξαγωγή προτάσεων με «αλγοριθμικά» μεταδεδομένα [NR23]. Η μετάβαση αυτή από κλασικά ML σε deep learning επιτρέπει την αποτύπωση πιο σύνθετων, μη γραμμικών μοτίβων, αλλά απαιτεί επίσης μεγαλύτερα και καλύτερα επισημειωμένα datasets, κάτι που παραμένει ο βασικός περιοριστικός παράγοντας. Όπως επισημαίνουν οι Raval και Bhaidasna [RB24] σε εκτενή ανασκόπηση τεχνικών εξαγωγής μεταδεδομένων, τα ανοιχτά προβλήματα παραμένουν σημαντικά, δεν υπάρχουν κοινά benchmarks, υπάρχουν δυσκολίες με θορυβώδη ή ελλιπή PDFs, ελλιπής υποστήριξη για πολυτροπικά δεδομένα (κείμενο, εικόνες, πίνακες), και υψηλό υπολογιστικό κόστος για τα πιο προηγμένα μοντέλα. Οι Wu κ.ά. [WBM+23] προσθέτουν ότι πιο «ερμηνευτικά» στοιχεία υπαινιγμοί, σχέσεις μεταξύ εγγράφων εξακολουθούν να διαφεύγουν από τα τυπικά ML μοντέλα, κάτι που σημαίνει ότι η ανθρώπινη κρίση δεν μπορεί να απουσιάζει. Στον αρχαικό χώρο, ανάλογες ML pipelines έχουν εφαρμοστεί σε μαζικές ψηφιοποιήσεις φακέλων ή σε born-digital υλικό όπως, parsing headers email (From, To, Date, Subject), εξαγωγή χαρακτηριστικών, εκπαίδευση classifiers για διάκριση «επίσημων» από «μη επίσημα» μηνύματα ή για απόδοση θεματικών επικεφαλίδων, και τελικά δημιουργία αρχαικών εγγραφών κατά ISAD(G) [CBJ+21]. Πάνω από όλα όμως, τα ML/DL συστήματα έχουν ένα κοινό αδύναμο σημείο με τα rule-based, δεν σχεδιάστηκαν για ιεραρχικές περιγραφές. Λειτουργούν κυρίως σε επίπεδο ενός εγγράφου, χωρίς να λαμβάνουν υπόψη τη σχέση μεταξύ ιεραρχικών επιπέδων ακριβώς η λογική που διέπει το ISAD(G).

Η τρίτη κατηγορία, που ενδιαφέρει ιδιαίτερα την παρούσα εργασία, είναι οι προσεγγίσεις βασισμένες σε LLMs μοντέλα όπως GPT-4, LLaMA, Mistral κ.ά. Το κρίσιμο διαφοροποιητικό χαρακτηριστικό τους είναι ότι εκπαιδεύονται γενικά σε πάρα πολύ μεγάλες συλλογές κειμένων (δισεκατομμύρια λέξεις από βιβλία, ιστοσελίδες, επιστημονικά άρθρα) και αποκτούν γενικές αναπαραστάσεις της γλώσσας που τους επιτρέπουν να λειτουργούν αποτελεσματικά με ελάχιστα ή καθόλου παραδείγματα. Η αρχιτεκτονική Transformer που εισήγαγαν οι Vaswani κ.ά. [VSP+17] και η τεχνική RLHF για instruction-following που ανέπτυξαν οι Ouyang κ.ά. [OWJ+22] δημιούργησαν μοντέλα που μπορούν να κατανοούν σύνθετες οδηγίες σε φυσική γλώσσα, να εκτελούν ποικίλες εργασίες όπως, σύνοψη, εξαγωγή πεδίων, ταξινόμηση και να παράγουν δομημένη έξοδο χωρίς ειδική εκπαίδευση για κάθε use case. Στο πεδίο των

μεταδεδομένων, ο Bagchi [Bag24] περιγράφει μια «Generative AI-driven Metadata Modelling» προσέγγιση όπου τα μεταδεδομένα δεν εξάγονται απλώς αλλά «συντίθενται» με βάση τη συνολική κατανόηση του κειμένου. Οι Alyafeai κ.ά. [AAG25] παρουσιάζουν το MOLE, σύστημα που συνδυάζει prompts για δομημένη έξοδο με κανόνες ελέγχου συνέπειας ώστε να μειώνεται ο κίνδυνος hallucination. Σε αρχειακό πλαίσιο, χαρακτηριστικό παράδειγμα είναι το ArcGPT [ZHP+23], ένα LLM προσαρμοσμένο σε αρχειακά δεδομένα και αξιολογημένο σε benchmark (AMBLE) για εργασίες όπως δημιουργία περιγραφών, θεματική ευρετηρίαση και ερωτοαπαντήσεις πάνω σε finding aids, που υπερέρχει σε αυτές τις εργασίες έναντι γενικών LLMs.

Το μεγάλο πλεονέκτημα των LLMs είναι η δυνατότητα λειτουργίας σε zero-shot ή few-shot σενάρια. Αντί να χρειάζεται ειδικό training set για κάθε αρχείο ή πρότυπο, μπορούν να «καθοδηγηθούν» με προσεκτικά σχεδιασμένα prompts και λίγα παραδείγματα, κάτι ιδιαίτερα ελκυστικό για αρχεία όπου η δημιουργία annotated datasets είναι δύσκολη και δαπανηρή [GWF24]. Από την άλλη, οι κίνδυνοι είναι εξίσου σημαντικοί. Οι Rawte κ.ά. [RSD23] αναλύουν συστηματικά το πρόβλημα της hallucination σε Large Foundation Models, διακρίνοντας intrinsic hallucinations, το μοντέλο παράγει πληροφορίες που έρχονται σε αντίθεση με το δοσμένο πλαίσιο από extrinsic, όπου παράγει πληροφορίες που δεν μπορούν να επαληθευτούν ούτε να διαμφευστούν. Σε συνδυασμό με την αστάθεια στη δομή εξόδου, ασυνεπής χρήση πεδίων, παράλειψη υποχρεωτικών στοιχείων και την εγγενή περιορισμένη επεξηγησιμότητα αυτό καθιστά αναγκαίο τον σχεδιασμό ισχυρών μηχανισμών ελέγχου και αξιολόγησης, ειδικά όταν εφαρμόζονται σε αρχειακή περιγραφή όπου η πιστότητα στο πρωτογενές κείμενο δεν είναι επιλογή αλλά δεοντολογική απαίτηση [WBM+23].

Συνοψίζοντας, η εικόνα που προκύπτει από τη βιβλιογραφία είναι ότι καμία κατηγορία δεν είναι ιδανική για τον αρχειακό χώρο από μόνη της. Τα rule-based προσφέρουν διαφάνεια και σταθερότητα αλλά σπάνε μπροστά στην ποικιλομορφία των αρχειακών τεκμηρίων. Τα ML/DL γενικεύουν καλύτερα αλλά απαιτούν annotated datasets που σπανίζουν. Και τα LLMs, που φαίνονται η πιο ελκυστική επιλογή για αρχεία χωρίς πόρους για εξειδικευμένη εκπαίδευση, φέρουν μαζί τους τον κίνδυνο της επινοημένης πληροφορίας. Η παρούσα εργασία τοποθετείται ακριβώς σε αυτό το τριγωνικό πρόβλημα, αξιοποιεί τη βιβλιογραφία των προηγούμενων προσεγγίσεων για να ορίσει τι σημαίνει «καλή» εξαγωγή μεταδεδομένων, και δοκιμάζει αν και υπό ποιες προϋποθέσεις τα LLMs μπορούν να πλησιάσουν αυτή την απόδοση στην ειδική περίπτωση της αρχειακής περιγραφής κατά ISAD(G).

3.2 Αρχειακή Περιγραφή και ISAD(G)

3.2.1 Βασικές αρχές

Η αρχειακή περιγραφή είναι η διαδικασία με την οποία ο αρχειονόμος καταγράφει και οργανώνει πληροφορίες για τις αρχειακές μονάδες, ώστε να κατανοείται το περιεχόμενο, το πλαίσιο δημιουργίας και η δομή τους. Η Gilliland [Gil08] τονίζει ότι η περιγραφή δεν είναι απλώς καταγραφή χαρακτηριστικών, αποτελεί πράξη ερμηνείας που στοχεύει στο να διευκολύνει την κατανόηση των τεκμηρίων από διαφορετικούς χρήστες, σε διαφορετικούς χρόνους, χωρίς να αλλοιώνει τη μαρτυρία τους. Θεμελιώδεις σε αυτό το πλαίσιο είναι οι αρχές του σεβασμού της προέλευσης και της διατήρησης της αρχικής τάξης. Τα τεκμήρια πρέπει, όσο είναι δυνατό, να παραμένουν οργανωμένα σύμφωνα με τον τρόπο που δημιουργήθηκαν και χρησιμοποιήθηκαν από τον παραγωγό, ώστε να μη χάνεται το διοικητικό και λειτουργικό τους πλαίσιο.

Το πρότυπο ISAD(G), που εκδόθηκε από το International Council on Archives (ICA) σε δεύτερη έκδοση το 2000, κωδικοποιεί αυτές τις αρχές σε ένα σύνολο κανόνων για τη δημιουργία αρχειακών περιγραφών [ICA00]. Στον πρόλογό του επισημαίνει ότι σκοπός του είναι να εξασφαλίσει την κατανόηση, τη συνέπεια και τη δυνατότητα ανταλλαγής των περιγραφών, όχι να αντικαταστήσει τα εθνικά πρότυπα, αλλά να λειτουργήσει ως γενικό πλαίσιο αναφοράς. Έτσι, ενθαρρύνει την υιοθέτηση κοινών στοιχείων και δομών, επιτρέποντας παράλληλα προσαρμογές σε εθνικό ή θεματικό επίπεδο, όπως το ελληνικό ΔΠΑΠ(Γ).

3.2.2 Ιεραρχικότητα

Μία από τις πιο χαρακτηριστικές όψεις του ISAD(G) είναι η ιεραρχική του προσέγγιση. Το πρότυπο ορίζει ότι τα αρχεία περιγράφονται από το γενικό προς το ειδικό, σε επίπεδα όπως *fonds*, *sub-fonds*, *series*, *sub-series*, *file*, *item*, με τη δυνατότητα ακόμη, ενδιάμεσων βαθμίδων όπου απαιτείται. Κάθε επίπεδο φέρει δική του εγγραφή περιγραφής, αλλά η πληροφορία κληρονομείται, ο παραγωγός, η γενική χρονολογία ή το διοικητικό πλαίσιο δηλώνονται στο επίπεδο του *fonds* και δεν επαναλαμβάνονται αυτούσια σε κάθε κατώτερο επίπεδο. Αντίστροφα, οι περιγραφές χαμηλότερου επιπέδου προσθέτουν λεπτομέρειες, όπως είναι ο συγκεκριμένος φάκελος, ο τίτλος τεκμηρίου, το περιεχόμενο, που δεν θα ήταν πρακτικό ή χρήσιμο να αποτυπωθούν ψηλότερα.

Η Gilliland [Gil08] υπογραμμίζει ότι η ιεραρχική περιγραφή εξυπηρετεί δύο στόχους, αφενός διατηρεί τις σχέσεις πλαισίου πώς δηλαδή, συνδέονται τα τεκμήρια μεταξύ τους και με τον παραγωγό και αφετέρου επιτρέπει διαφορετικά επίπεδα πρόσβασης ανάλογα

με τις ανάγκες του χρήστη. Ένας ερευνητής μπορεί να αρκестεί στην περιγραφή του fonds ή μιας σειράς για να αποκτήσει γενική εικόνα, ενώ ένας άλλος χρειάζεται λεπτομέρεια σε επίπεδο item. Σε επίπεδο τεχνικής υλοποίησης, αυτή η ιεραρχία αποτυπώνεται συχνά μέσω EAD (Encoded Archival Description), όπου τα στοιχεία (elements) <c> αναπαριστούν τα επιμέρους components κάθε επιπέδου σε ένα εμφωλευμένο δέντρο XML.

Για τη χρήση LLMs, αυτό σημαίνει κάτι πολύ συγκεκριμένο, ότι η εξαγωγή μεταδεδομένων δεν μπορεί να αντιμετωπίζει κάθε τεκμήριο ως «απομονωμένο» αντικείμενο. Η περιγραφή ενός φακέλου ή τεκμηρίου πρέπει να λαμβάνει υπόψη το πλαίσιο του ανώτερου επιπέδου κάτι που απαιτεί είτε τη χορήγηση σχετικού context στο prompt, είτε ειδικό σχεδιασμό της διαδικασίας (π.χ. πρώτα περιγράφεται το fonds, μετά οι σειρές κ.ο.κ.).

3.2.3 Οι επτά περιοχές της περιγραφής ISAD(G) - Υποχρεωτικά και προαιρετικά πεδία

Εκτός από την ιεραρχική του λογική, το ISAD(G) οργανώνει το περιεχόμενο κάθε αρχειακής περιγραφής σε επτά περιοχές (areas of description), στις οποίες κατανέμονται συνολικά 26 στοιχεία περιγραφής [ICA00]. Η ίδια αυτή δομή αποτυπώνεται συνοπτικά στο Σχήμα 6 του δεύτερου κεφαλαίου. Εδώ παρουσιάζεται αναλυτικά, γιατί πάνω σε αυτές τις περιοχές οργανώνεται και η αξιολόγηση του πειραματικού μέρους. Κάθε περιοχή καλύπτει μια διαφορετική πτυχή της αρχειακής μονάδας, και όλες μαζί δίνουν μια ολοκληρωμένη εικόνα της ταυτότητας, της προέλευσης, του περιεχομένου και των όρων χρήσης της [Gil08].

Συνοπτικά οι επτά περιοχές και ποια είναι η χρήση τους, παρουσιάζονται παρακάτω:

Η περιοχή Αναγνώρισης (Identification Area) προσδιορίζει τη μονάδα περιγραφής και τη θέση της μέσα στο αρχείο. Απαντά στο ερώτημα «τι είναι αυτό και που εντάσσεται». Περιλαμβάνει τον κωδικό αναφοράς, τον τίτλο, τη χρονολογία, το επίπεδο περιγραφής και την έκταση του υλικού. Όλα αυτά είναι στοιχεία που, όπως αναλύεται στην συνέχεια, αποτελούν και τον υποχρεωτικό πυρήνα κάθε περιγραφής [ICA00].

Η περιοχή Πλαισίου (Context Area) αποτυπώνει την προέλευση της μονάδας. Εδώ καταγράφονται ο παραγωγός, το διοικητικό ή βιογραφικό ιστορικό, το ιστορικό της ίδιας της ενότητας και τον τρόπο πρόσκτησής της. Είναι η περιοχή που υλοποιεί στην πράξη την αρχή του σεβασμού της προέλευσης αφού συνδέει τα τεκμήρια με το περιβάλλον που τα δημιούργησε [Gil08].

Η Περιοχή Περιεχομένου και Δομής (Content and Structure Area) περιγράφει τι περιέχει η μονάδα και με ποιον τρόπο είναι οργανωμένη. Καλύπτει την παρουσίαση περιεχομένου, πληροφορίες αξιολόγησης και εκκαθάρισης, ενδεχόμενες προσθήκες υλικού και το σύστημα ταξινόμησης. Σε αυτήν την περιοχή αποτυπώνεται, μεταξύ άλλων, η αρχή της διατήρησης της αρχικής τάξης.

Ακολουθεί η Περιοχή Όρων Πρόσβασης και Χρήσης (Conditions of Access and Use Area), που καθορίζει αν και πώς μπορεί να χρησιμοποιηθεί το υλικό. Τους όρους πρόσβασης και αναπαραγωγής, τη γλώσσα και τη γραφή των τεκμηρίων, τα φυσικά τους χαρακτηριστικά και τα διαθέσιμα εργαλεία έρευνας.

Η Περιοχή Συναφούς Υλικού (Allied Materials Area) τοποθετεί τη μονάδα σε σχέση με άλλο υλικό, καταγράφοντας την ύπαρξη και τον εντοπισμό πρωτοτύπων ή αντιγράφων, συναφείς ενότητες περιγραφής και σχετικές δημοσιεύσεις.

Οι δύο τελευταίες περιοχές έχουν πιο εξειδικευμένο ρόλο. Η Περιοχή Σημειώσεων (Notes Area) προορίζεται για πληροφορίες που δεν εντάσσονται φυσικά σε καμία από τις προηγούμενες, ώστε να μπορούν να καταγραφούν ειδικές παρατηρήσεις χωρίς να διαταράσσεται η τυποποιημένη δομή.

Η Περιοχή Ελέγχου της Περιγραφής (Description Control Area), τέλος, τεκμηριώνει την ίδια τη διαδικασία. Ποιος συνέταξε την εγγραφή, με βάση ποιους κανόνες και πότε. Χάρη σε αυτήν η περιγραφή γίνεται ελέγξιμη και επαληθεύσιμη, κάτι που αποκτά ιδιαίτερη σημασία όταν στη διαδικασία εμπλέκονται αυτοματοποιημένα εργαλεία.

Η κατανομή σε επτά περιοχές δεν είναι αδιάφορη για την παρούσα εργασία. Όπως φαίνεται στο τέταρτο κεφάλαιο, η αξιολόγηση των LLMs δεν εξετάζει την περιγραφή ως ενιαίο κείμενο, αλλά ανά περιοχή, καθώς η δυσκολία της εξαγωγής δεν είναι παντού η ίδια. Περιοχές όπως η Αναγνώριση στηρίζονται σε στοιχεία που εντοπίζονται σχετικά εύκολα μέσα στο πρωτογενές κείμενο. Αντίθετα, περιοχές όπως το Πλαίσιο ή το Περιεχόμενο και Δομή προϋποθέτουν ερμηνεία και κατανόηση συσχετίσεων, και γι' αυτό είναι πιο ευάλωτες σε παραλείψεις ή σε στοιχεία που το μοντέλο επινοεί.

Παρακάτω δίνεται συνοπτικός πίνακας των επτά περιοχών περιγραφής του ISAD(G) και τα βασικά τους στοιχεία.

Περιοχή ISAD(G)	Βασικά στοιχεία περιγραφής	Απαντά στο ερώτημα
1. Αναγνώριση (Identification)	Κωδικός αναφοράς, Τίτλος, Χρονολογία, Επίπεδο περιγραφής, Έκταση και μέσο	Τι είναι η μονάδα και πού εντάσσεται;
2. Πλαίσιο (Context)	Παραγωγός, Διοικητικό/βιογραφικό ιστορικό, Αρχειακό ιστορικό, Τρόπος πρόσκτησης	Από πού προέρχεται και ποιος τη δημιούργησε;
3. Περιεχόμενο και Δομή (Content and Structure)	Παρουσίαση περιεχομένου, Αξιολόγηση/εκκαθάριση, Προσθήκες, Σύστημα ταξινόμησης	Τι περιέχει και πώς είναι οργανωμένη;
4. Όροι Πρόσβασης και Χρήσης (Conditions of Access and Use)	Όροι πρόσβασης, Όροι αναπαραγωγής, Γλώσσα/γραφή, Φυσικά χαρακτηριστικά, Εργαλεία έρευνας	Πώς και υπό ποιες προϋποθέσεις χρησιμοποιείται;
5. Συναφές Υλικό (Allied Materials)	Ύπαρξη/εντοπισμός πρωτοτύπων, Ύπαρξη/εντοπισμός αντιγράφων, Συναφείς ενότητες, Δημοσιεύσεις	Πώς σχετίζεται με άλλο υλικό;
6. Σημειώσεις (Notes)	Σημειώσεις	Τι άλλο πρέπει να καταγραφεί εκτός των παραπάνω;
7. Έλεγχος Περιγραφής (Description Control)	Σημείωση αρχειονόμου, Κανόνες/συμβάσεις, Χρονολογία περιγραφής	Ποιος, με ποιους κανόνες και πότε συνέταξε την περιγραφή;

Πίνακας 2 Οι επτά περιοχές περιγραφής του ISAD(G) και τα βασικά τους στοιχεία [ICA00]

Άλλα στοιχεία, όπως το διοικητικό ιστορικό, το ιστορικό της ενότητας, οι λεπτομέρειες πρόσκτησης, η παρουσίαση περιεχομένου και οι όροι πρόσβασης, θεωρούνται ιδιαίτερα σημαντικά, αλλά όχι αυστηρά υποχρεωτικά σε κάθε περίπτωση. Η διάκριση αυτή δεν είναι τυπική. Συνδέεται άμεσα με τη λειτουργικότητα της περιγραφής. Αν λείπουν βασικά στοιχεία όπως ο τίτλος ή η ημερομηνία, ο χρήστης δυσκολεύεται να καταλάβει τι ακριβώς περιγράφεται και πώς εντάσσεται στη δομή του αρχείου [Gil08]. Για την παρούσα εργασία, αυτό σημαίνει ότι η αξιολόγηση των LLMs ελέγχει τόσο την πληρότητα, αν δηλαδή έχουν συμπληρωθεί τα υποχρεωτικά πεδία, αλλά και την ορθότητα, δηλαδή αν τα πεδία περιέχουν νοηματικά κατάλληλες τιμές και όχι γενικόλογες περιλήψεις ή επινοημένα στοιχεία.

3.2.4 Η σημασία της πιστότητας στο πρωτογενές κείμενο

Ένα από τα θεμελιώδη ζητήματα της αρχειακής περιγραφής και ένα από τα μεγαλύτερα μειονεκτήματα στη χρήση των LLMs, είναι η πιστότητα στο πρωτογενές τεκμήριο. Η Gilliland [Gil08] επισημαίνει ότι, παρότι η περιγραφή περιλαμβάνει αναπόφευκτα κάποια δόση ερμηνείας, ο αρχειονόμος οφείλει να αποφεύγει να εισάγει πληροφορίες που δεν τεκμηριώνονται από τα ίδια τα τεκμήρια ή αξιόπιστες πηγές. Η περιγραφή πρέπει να κρατά μια λεπτή ισορροπία, να είναι αρκετά πλούσια ώστε να είναι χρήσιμη, αλλά να μη μετατρέπεται σε επινοημένη αφήγηση.

Οι Greene και Meissner [GM05], αν και εστιάζουν στο ζήτημα της εργασιακής οικονομίας, υπογραμμίζουν ότι η απλοποίηση διαδικασιών δεν πρέπει να οδηγεί σε «εκλεκτική τεκμηρίωση», όπου κάποια τεκμήρια περιγράφονται σχολαστικά και άλλα επιφανειακά ή παραπλανητικά. Οι Jaillant και Caruto [JC22] και οι Colavizza κ.ά. [CBJ+21] επισημαίνουν ότι στα born-digital αρχεία η πίεση για μαζική επεξεργασία μπορεί να οδηγήσει σε αυτόματες αποφάσεις, όπως είναι η αλγοριθμική αφαίρεση προσωπικών δεδομένων ή η αυτόματη ταξινόμηση τεκμηρίων, που δεν αποδίδουν πάντα ακριβώς το περιεχόμενο ή στη χειρότερη περίπτωση, αλλοιώνουν την ιστορική μαρτυρία.

Η χρήση LLMs προσθέτει μια ακόμα στρώση κινδύνου. Τα μοντέλα έχουν εγγενή τάση για hallucinations [RSD23], μπορεί να «συμπληρώνουν» λογικά, αλλά, ανύπαρκτα στοιχεία, να βελτιώνουν τίτλους με τρόπο που αλλοιώνει το νόημα, ή να εξομαλύνουν ασυνέπειες που είναι, όμως, ουσιώδες μέρος της ιστορικής μαρτυρίας. Αυτό έρχεται σε έντονη αντίθεση με την αρχειακή δεοντολογία, που προκρίνει την ελάχιστη αναγκαία παρέμβαση στα τεκμήρια και στα μεταδεδομένα τους. Γι' αυτό, η παρούσα εργασία αντιμετωπίζει τα LLMs όχι ως αυτόνομους «συγγραφείς» περιγραφών, αλλά ως βοηθητικά εργαλεία που μπορούν να προτείνουν πρόχειρες περιγραφές ή να εξάγουν υποψήφια πεδία, τα οποία όμως απαιτούν πάντα τον έλεγχο του αρχειονόμου.

3.3 Μεγάλα Γλωσσικά Μοντέλα (LLMs)

Τα Μεγάλα Γλωσσικά Μοντέλα (Large Language Models – LLMs) αποτελούν σήμερα την πιο δυναμική κατηγορία συστημάτων επεξεργασίας φυσικής γλώσσας, με εφαρμογές που εκτείνονται από την αυτόματη μετάφραση και τη σύντομη κειμένων μέχρι την εξαγωγή δομημένης πληροφορίας από ελεύθερο κείμενο. Για να κατανοηθεί η σημασία τους στο πλαίσιο της παρούσας εργασίας, απαιτείται μια σύντομη αλλά ουσιαστική παρουσίαση των βασικών αρχών που τα διέπουν.

3.3.1 Αρχιτεκτονική Transformer

Το θεμέλιο πάνω στο οποίο βασίζονται τα σύγχρονα LLMs είναι η αρχιτεκτονική Transformer, που εισήγαγαν οι Vaswani κ.ά. το 2017 [VSP+17]. Πριν από αυτήν, η επικρατούσα προσέγγιση για την επεξεργασία ακολουθιών κειμένου ήταν τα αναδρομικά νευρωνικά δίκτυα (RNNs) και οι παραλλαγές τους (LSTM, GRU), τα οποία επεξεργάζονται τις λέξεις μιας πρότασης διαδοχικά, μία τη φορά. Αυτό δημιουργούσε σοβαρά προβλήματα όμως, τα μοντέλα δυσκολεύονταν να διατηρήσουν πληροφορία για μακρινές εξαρτήσεις (π.χ. να συσχετίσουν μια λέξη στην αρχή μιας μακριάς πρότασης με μια άλλη στο τέλος), ενώ η διαδοχική επεξεργασία περιόριζε σημαντικά τη δυνατότητα παραλληλισμού κατά την εκπαίδευση.

Η αρχιτεκτονική Transformer έλυσε αυτά τα προβλήματα εισάγοντας τον μηχανισμό self-attention (αυτο-προσοχής). Αντί να επεξεργάζεται τις λέξεις ακολουθιακά, ο μηχανισμός self-attention επιτρέπει στο μοντέλο να «κοιτά» ταυτόχρονα όλες τις λέξεις μιας ακολουθίας και να υπολογίζει για κάθε λέξη πόσο «σχετική» είναι με κάθε άλλη λέξη στην ίδια πρόταση ή παράγραφο. Με αυτόν τον τρόπο, το μοντέλο μπορεί να συλλάβει μακρινές σημασιολογικές εξαρτήσεις πολύ πιο αποτελεσματικά από ό,τι τα RNNs, ενώ παράλληλα επιτρέπει αποδοτικό παραλληλισμό κατά την εκπαίδευση. Η αρχική αρχιτεκτονική αποτελείται από ένα encoder και ένα decoder, καθένα από τα οποία περιλαμβάνει πολλαπλά επίπεδα self-attention και feed-forward νευρωνικών δικτύων [VSP+17]. Στη συνέχεια, οι περισσότερες παραλλαγές που χρησιμοποιούνται σήμερα (GPT, LLaMA, Mistral) βασίζονται αποκλειστικά στο decoder τμήμα και εκπαιδεύονται να προβλέπουν την επόμενη λέξη σε μια ακολουθία (autoregressive language modeling), αποκτώντας έτσι ικανότητες γενίκευσης που εκτείνονται πολύ πέρα από απλή γλωσσική πρόβλεψη.

Η εκπαίδευση αυτών των μοντέλων γίνεται σε τεράστια σώματα κειμένων, δισεκατομμύρια λέξεις από βιβλία, ιστοσελίδες, επιστημονικά άρθρα και αποσκοπεί στο να αποκτήσει το μοντέλο γενικές αναπαραστάσεις της γλώσσας. Αυτό σημαίνει ότι, σε αντίθεση με τα κλασικά ML μοντέλα που χρειάζονται εκπαίδευση από το μηδέν για κάθε νέα εργασία, ένα LLM μπορεί μέσω κατάλληλων prompts ή ελαφράς εκπαίδευσης (fine-tuning) να προσαρμοστεί σε νέες εργασίες με ελάχιστα ή καθόλου παραδείγματα. Αυτή η ιδιότητα, γνωστή ως few-shot και zero-shot learning, είναι κεντρική για την αξιοποίησή τους στην αρχιτεκτονική περιγραφή, όπου τα annotated datasets είναι σπάνια [VSP+17, Bag24].

3.3.2 Το πρόβλημα των hallucinations

Παρά τις εντυπωσιακές δυνατότητές τους, τα LLMs παρουσιάζουν ένα σοβαρό και καλά τεκμηριωμένο πρόβλημα, την τάση για το φαινόμενο των «ψευδαισθήσεων» ή αλλιώς hallucinations. Ο όρος αναφέρεται στην παραγωγή πληροφοριών που φαίνονται πειστικές, γλωσσικά συνεκτικές και σωστά διατυπωμένες, αλλά δεν αντιστοιχούν στην πραγματικότητα ή στο πρωτογενές κείμενο που δόθηκε ως είσοδος [RSD23].

Οι Rawte κ.ά. [RSD23] παρέχουν μια εκτενή ταξινόμηση του φαινομένου, διακρίνοντας διαφορετικούς τύπους hallucinations, τις intrinsic hallucinations, όπου το μοντέλο παράγει πληροφορίες που έρχονται σε αντίθεση με το δοσμένο πλαίσιο, και τις extrinsic hallucinations, όπου παράγει πληροφορίες που δεν μπορούν να επαληθευτούν ούτε να διαψευστούν από το πλαίσιο. Οι αιτίες του φαινομένου εντοπίζονται τόσο στην ίδια τη φύση της autoregressive εκπαίδευσης (το μοντέλο μαθαίνει να παράγει «πιθανές» λέξεις, όχι «αληθινές») όσο και στις μεροληψίες των δεδομένων εκπαίδευσης. Ειδικά για τον αρχαικό χώρο, το πρόβλημα αυτό αποκτά μεγάλη σημασία. Ένα μοντέλο που «εφευρίσκει» δημιουργούς, χρονολογίες ή γεγονότα βάσει στατιστικών μοτίβων, αντί να τα εξάγει πιστά από το πρωτογενές κείμενο, μπορεί να παραγάγει εγγραφές που είναι γλωσσικά άρτιες αλλά αρχαικά ανακριβείς ή και παραπλανητικές [RSD23].

Για την αντιμετώπιση αυτού του προβλήματος, η βιβλιογραφία προτείνει διάφορες στρατηγικές όπως αυστηρή σχεδίαση prompts που περιορίζει το μοντέλο αποκλειστικά στο δοσμένο κείμενο (SOURCE RULE), στάδια μετα-επεξεργασίας με consistency checks, και συνδυασμό με εξωτερικές πηγές επαλήθευσης [AAG25, WBM+23]. Χαρακτηριστικό παράδειγμα είναι το ArcGPT [ZHP+23], όπου οι συγγραφείς αναγνωρίζουν ρητά ότι ακόμα και ένα domain-specific μοντέλο εκπαιδευμένο σε αρχαικά δεδομένα δεν είναι απαλλαγμένο από hallucinations, και επισημαίνουν ότι τα ζητήματα ποιότητας δεδομένων και συμμόρφωσης με αρχαικά πρότυπα παραμένουν ανοιχτά.

3.3.3 Instruction following και prompt engineering

Μια κρίσιμη εξέλιξη για την πρακτική αξιοποίηση των LLMs ήταν η ανάπτυξη τεχνικών που τα καθιστούν ικανά να ακολουθούν ανθρώπινες οδηγίες με ακρίβεια και συνέπεια, αυτό που η βιβλιογραφία αποκαλεί instruction following. Οι Ouyang κ.ά. [OWJ+22] παρουσίασαν τη μέθοδο RLHF (Reinforcement Learning from Human Feedback), με την οποία μοντέλα όπως το InstructGPT εκπαιδεύονται να ευθυγραμμίζουν τις αποκρίσεις τους με ανθρώπινες προτιμήσεις και εντολές. Στην ουσία, αντί το μοντέλο να βελτιστοποιεί μόνο την πρόβλεψη της επόμενης λέξης, εκπαιδεύεται να ακολουθεί οδηγίες, να αποφεύγει επιβλαβείς ή άσχετες απαντήσεις και να παράγει εξόδους που

ανταποκρίνονται στην πραγματική πρόθεση του χρήστη. Αυτή η εξέλιξη είναι καθοριστική για εφαρμογές όπως η αρχειακή περιγραφή, όπου απαιτείται αυστηρή συμμόρφωση με συγκεκριμένες δομές και κανόνες, π.χ. «συμπλήρωσε ακριβώς αυτά τα 26 πεδία με βάση μόνο το δοσμένο κείμενο» [OWJ+22].

Συνδεδεμένη άμεσα με το *instruction following* είναι και η πρακτική του *prompt engineering*, η τεχνική δηλαδή του σχεδιασμού των εισόδων (*prompts*) που δίνονται στο μοντέλο, με σκοπό να κατευθύνουν την έξοδο του προς το επιθυμητό αποτέλεσμα. Το *prompt engineering* δεν είναι απλώς η διατύπωση μιας ερώτησης, περιλαμβάνει αποφάσεις για τον ρόλο που καλείται να παίξει το μοντέλο (π.χ. «λειτουργήσε ως αρχειακός καταλογγράφος»), τους κανόνες που πρέπει να ακολουθήσει (π.χ. «χρησιμοποίησε μόνο το δοσμένο κείμενο, μην εφεύρεις πληροφορίες»), τη μορφή εξόδου που αναμένεται (π.χ. «παράγαγε ένα ISAD(G) μπλοκ για κάθε αρχειακή μονάδα») και τον τρόπο αντιμετώπισης αμφιλεγόμενων περιπτώσεων. Στη βιβλιογραφία έχουν αναπτυχθεί διαφορετικές στρατηγικές *prompting*, όπως το *zero-shot prompting*, όπου το μοντέλο καλείται να εκτελέσει μια εργασία χωρίς κανένα παράδειγμα, το *few-shot prompting*, όπου παρέχονται λίγα ενδεικτικά παραδείγματα που καθοδηγούν τη μορφή και τη λογική της εξόδου, και το *chain-of-thought prompting*, όπου το μοντέλο ενθαρρύνεται να αιτιολογεί βήμα-βήμα πριν δώσει την τελική απάντηση, βελτιώνοντας την ακρίβεια σε σύνθετες εργασίες [OWJ+22, Bag24].

Στον αρχειακό χώρο, ο σχεδιασμός του *prompt* αποτελεί κρίσιμο παράγοντα για την ποιότητα της εξόδου. Χαρακτηριστικό παράδειγμα αποτελεί το ArcGPT [ZHP+23], όπου οι συγγραφείς επισημαίνουν ότι η αποδοτική αξιοποίηση ενός *domain-specific* μοντέλου σε αρχειακές εργασίες εξαρτάται εξίσου από την ποιότητα του *prompt* όσο και από την εκπαίδευση του μοντέλου. Αντίστοιχα, οι Groppe κ.ά. [GMW+25] δείχνουν ότι σε σενάρια ομοσπονδιακής νοημοσύνης μεταξύ πολλαπλών LLMs, το *prompt engineering* λειτουργεί ως ο κοινός «κώδικας επικοινωνίας» που εξασφαλίζει συνέπεια μεταξύ των διαφορετικών μοντέλων. Στην παρούσα εργασία, ο σχεδιασμός του *prompt* αποτελεί βασικό μεθοδολογικό εργαλείο και αναλύεται διεξοδικά στο Κεφάλαιο 4.

3.3.4 Αξιολόγηση δεδομένων και αποτελεσμάτων (Data quality metrics)

Ένα από τα πιο κρίσιμα και ακόμα ανοιχτά ζητήματα στη βιβλιογραφία είναι το πώς αξιολογούμε την ποιότητα τόσο των δεδομένων εισόδου που τροφοδοτούμε στα LLMs όσο και των αρχειακών περιγραφών που αυτά παράγουν. Για να απαντηθεί αυτό, χρειάζεται πρώτα να κατανοηθεί τι σημαίνει «ποιότητα δεδομένων» στην εξόρυξη

πληροφορίας, γιατί τα προβλήματα που περιγράφει η σχετική βιβλιογραφία εμφανίζονται και στα αρχεία, απλώς με διαφορετικό πρόσωπο.

Ένα καλό σημείο εκκίνησης είναι η διάκριση που κάνουν οι Tan κ.ά. [TSKK19] μεταξύ σφάλματος μέτρησης (measurement error) και σφάλματος συλλογής δεδομένων (data collection error). Το πρώτο αφορά την καταγραφή μιας τιμής που απέχει από την αληθινή, το δεύτερο αφορά παραλείψεις ή εσφαλμένες συμπεριλήψεις κατά τη συλλογή. Και τα δύο μπορεί να είναι συστηματικά ή τυχαία. Όταν ένα LLM παρακάμπτει συστηματικά το ίδιο πεδίο, π.χ. πάντα παραλείπει το «Immediate source of acquisition», πρόκειται για συστηματικό σφάλμα. Όταν αποδίδει τυχαία λανθασμένες τιμές σε διαφορετικά πεδία κάθε φορά, πρόκειται για τυχαίο. Η διάκριση δεν είναι θεωρητική. Ένα συστηματικό σφάλμα μπορεί να εντοπιστεί και να διορθωθεί μέσω βελτίωσης του prompt, ενώ ένα τυχαίο απαιτεί ανθρώπινο έλεγχο σε κάθε εγγραφή χωριστά [TSKK19].

Συνδεδεμένη με αυτό είναι η έννοια του θορύβου (noise), η τυχαία συνιστώσα του σφάλματος, που στρεβλώνει τιμές ή προσθέτει παρεμβαλλόμενα ψευδή στοιχεία. Οι Tan κ.ά. [TSKK19] επισημαίνουν ότι η εξάλειψή του είναι συχνά δύσκολη και γι' αυτό μεγάλο μέρος της έρευνας εστιάζει στη δημιουργία ανθεκτικών αλγορίθμων (robust algorithms) που παράγουν αποδεκτά αποτελέσματα ακόμα και παρουσία θορύβου. Στην εξαγωγή αρχειακών μεταδεδομένων, θόρυβος εισάγεται σε πολλά σημεία, όπως στο OCR όπου χαρακτήρες διαβάζονται λανθασμένα, στο πρωτογενές κείμενο όταν είναι ελλιπές ή αντιφατικό, στο prompt όταν οι οδηγίες δεν είναι αρκετά σαφείς, και φυσικά στην έξοδο του μοντέλου μέσω των hallucinations. Δίπλα στον θόρυβο, οι Tan κ.ά. [TSKK19] περιγράφουν και τα artifacts, ντετερμινιστικές, επαναλαμβανόμενες στρεβλώσεις, σε αντίθεση με τον τυχαίο θόρυβο. Ένα αντίστοιχο αρχειακό artifact θα ήταν η σταθερή τάση ενός μοντέλου να βάζει στο πεδίο «Administrative/Biographical history» πληροφορίες που ανήκουν στο «Archival history», ένα επαναλαμβανόμενο μοτίβο λάθους που αναγνωρίζεται μόλις το δει κάποιος για δεύτερη φορά.

Τρεις έννοιες που οι Tan κ.ά. [TSKK19] τοποθετούν στον πυρήνα της αξιολόγησης ποιότητας είναι η αναπαραγωγικότητα, η προκατάληψη και η ακρίβεια. Η αναπαραγωγικότητα (precision) είναι η εγγύτητα των επαναλαμβανόμενων μετρήσεων μεταξύ τους και μετριέται μέσω της τυπικής απόκλισης. Η προκατάληψη (bias) είναι η συστηματική απόκλιση από την αληθινή τιμή, η διαφορά δηλαδή μεταξύ του μέσου των μετρήσεων και της γνωστής αληθινής τιμής. Η ακρίβεια (accuracy) είναι η εγγύτητα στην αληθινή τιμή και εξαρτάται και από τα δύο προηγούμενα, χωρίς όμως να εκφράζεται με έναν απλό τύπο ως προς αυτά. Οι Tan κ.ά. [TSKK19] τονίζουν ότι αυτά τα μεγέθη παραβλέπονται συχνά, και χωρίς την κατανόησή τους ο αναλυτής διακινδυνεύει σοβαρά

σφάλματα. Στο αρχειακό πλαίσιο οι τρεις αυτές έννοιες αποκτούν συγκεκριμένο νόημα. Χαμηλή αναπαραγωγιμότητα σημαίνει ότι το ίδιο μοντέλο παράγει διαφορετικές εγγραφές για το ίδιο κείμενο, κάτι που είναι απαράδεκτο για τεκμηρίωση. Προκατάληψη σημαίνει ότι τα λάθη δεν είναι τυχαία αλλά κατευθύνονται, π.χ. το μοντέλο υποεκτιμά συστηματικά την πολυπλοκότητα του τομέα Context. Και χαμηλή ακρίβεια σημαίνει απλώς ότι η εγγραφή απέχει από αυτό που θα έγραφε ο αρχειονόμος.

Για τα δεδομένα που λείπουν (missing values), οι Tan κ.ά. [TSKK19] παρατηρούν ότι δεν είναι ασυνήθιστο ένα αντικείμενο να έχει ελλιπείς τιμές διότι κάποιες πληροφορίες δεν συλλέχθηκαν ή ορισμένα χαρακτηριστικά δεν ισχύουν για όλα τα αντικείμενα. Για αυτό προτείνουν τρεις στρατηγικές αντιμετώπισης, την αφαίρεση, την διαγραφή δηλαδή της εγγραφής που έχει τιμή που λείπει, την εκτίμηση, δηλαδή την συμπλήρωση της τιμής βάσει άλλων δεδομένων (π.χ. μέσος όρος ή γειτονικές τιμές) και η αγνόηση κατά την ανάλυση δηλαδή η παράλειψη του πεδίου κατά την ανάλυση, χωρίς να το συμπληρώσει κάποιος. Στα αρχεία, τα ελλείποντα δεδομένα είναι σχεδόν κανόνας. Πολλές περιγραφές δεν είναι ολοκληρωμένες, πεδία παραμένουν κενά, και πληροφορίες που ο αρχειονόμος θεωρεί αυτονόητες συχνά δεν καταγράφονται γραπτά. Η επιλογή της παρούσας εργασίας είναι να ζητείται από το μοντέλο να γράφει «N/A (not stated in source)» για κάθε τέτοιο πεδίο και εμπνέεται από τη λογική της στρατηγικής αγνόησης. Το μοντέλο δεν επιχειρεί να εκτιμήσει ή να εφεύρει μια τιμή, κάτι που θα οδηγούσε σε hallucination, αλλά δηλώνει ρητά ότι η πληροφορία απουσιάζει από την πηγή, διατηρώντας έτσι την αξιοπιστία της εγγραφής [TSKK19].

Ανάλογο πρόβλημα αποτελούν οι ασυνεπείς τιμές (inconsistent values), τιμές που αντιφάσκουν μεταξύ τους, είτε λόγω σφαλμάτων εισαγωγής, είτε γιατί οι πληροφορίες προέρχονται από διαφορετικές πηγές που δεν έχουν συγχωνευθεί σωστά [TSKK19]. Στα LLMs, ασυνέπειες εμφανίζονται όταν το μοντέλο αποδίδει διαφορετικές τιμές στο ίδιο πεδίο σε διαφορετικά ιεραρχικά επίπεδα π.χ. αναφέρει διαφορετικό δημιουργό για το fonds και για ένα item, ή όταν οι χρονολογίες δεν ταιριάζουν μεταξύ τομέων. Ο κανόνας AMBIGUITY RULE του prompt σημαίνει «αν η πηγή είναι αντιφατική, αντέγραψε τις τιμές ακριβώς όπως είναι», χειρίζεται αυτό ακριβώς, καταγράφει την ασυνέπεια αντί να την «λύσει» αυθαίρετα, ώστε ο αρχειονόμος να την εντοπίσει και να αποφασίσει [TSKK19].

Στο ίδιο πλαίσιο, οι Tan κ.ά. [TSKK19] θίγουν και το ζήτημα της δειγματοληπτικής προκατάληψης (sampling bias). Αν τα δεδομένα εκπαίδευσης ή αξιολόγησης δεν αντιπροσωπεύουν ισόρροπα τον πληθυσμό, τα αποτελέσματα θα είναι λανθασμένα όταν εφαρμοστούν στον ευρύτερο πληθυσμό. Αυτό σημαίνει ότι τα αποτελέσματα της μελέτης

περίπτωσης της παρούσας εργασίας, που αφορά αποκλειστικά το Αρχείο της Επτανήσου Πολιτείας, δεν μπορούν να γενικευθούν άμεσα σε άλλους τύπους αρχειακού υλικού. Διαφορετικές χρονικές περιόδοι, γλώσσες ή τύποι φορέων μπορεί να παρουσιάζουν διαφορετικά ποιοτικά χαρακτηριστικά που επηρεάζουν την απόδοση των μοντέλων με τρόπο που δεν είναι ακόμα γνωστός.

Σε όλη αυτή τη συζήτηση για ποιότητα και σφάλμα, κεντρικό ρόλο παίζει η έννοια της εγγραφής αναφοράς/πρότυπη αναφορά ή αλλιώς *gold standard*. Πρόκειται για μία ή περισσότερες εγγραφές που έχουν συνταχθεί από εξειδικευμένο επαγγελματία, στην συγκεκριμένη περίπτωση που εξετάζουμε, από αρχειονόμο, και θεωρούνται ως η σωστή απάντηση με βάση την οποία, αξιολογούνται οι αυτόματα παραγόμενες εγγραφές. Η λογική του *gold standard* είναι θεμελιώδης στην αξιολόγηση συστημάτων NLP και ML. Οι Granitzer et al. [GHK+12] το χρησιμοποιούν συστηματικά στη σύγκριση CRF και SVM για εξαγωγή βιβλιογραφικών μεταδεδομένων, υπολογίζοντας μετρικές όπως είναι το *precision*, το *recall* και το *F1 score* σε σχέση με χειρωνακτικά επισημειωμένα δεδομένα αναφοράς. Στον αρχειακό χώρο όμως η εφαρμογή *gold standard* δεν είναι απλή υπόθεση. Σε αντίθεση με μία εργασία ταξινόμησης, όπου η σωστή κατηγορία είναι σαφής, μία αρχειακή περιγραφή ενσωματώνει επαγγελματικές κρίσεις, ερμηνείες πλαισίου και αποφάσεις για το ποια στοιχεία αξίζει να αναδειχθούν. Δύο έμπειροι αρχειονόμοι μπορούν να συντάξουν διαφορετικές αλλά εξίσου έγκυρες αναφορές για το ίδιο τεκμήριο. Αυτό σημαίνει ότι το *gold standard* στον αρχειακό χώρο πρέπει να αντιμετωπίζεται όχι ως μοναδική «αλήθεια» αλλά ως αξιόπιστο σημείο αναφοράς που αντικατοπτρίζει επαγγελματική πρακτική [WBM+23, LRT21].

Στην παρούσα εργασία, το *gold standard* αποτελείται από πέντε εγγραφές που συνέταξε εξειδικευμένη αρχειονόμος για διαφορετικά ιεραρχικά επίπεδα του Αρχείου της Επτανήσου Πολιτείας (επίπεδα *fonds*, φακέλου και τριών τεκμηρίων). Αυτή η ιεραρχική κάλυψη επιτρέπει να αξιολογηθεί όχι μόνο η ορθότητα της εξόδου των LLMs σε επιμέρους πεδία, αλλά και η ικανότητά τους να σέβονται τη λογική της κληρονομικότητας του ISAD(G), δηλαδή το αν αναγνωρίζουν ότι οι πληροφορίες ορισμένες στο επίπεδο *fonds* δεν χρειάζεται να επαναληφθούν σε χαμηλότερα επίπεδα. Όπως επισημαίνουν οι Tan et al. [TSKK19], η σύγκριση με δεδομένα αναφοράς είναι ο μόνος τρόπος να υπολογιστεί η προκατάληψη (*bias*) και να αποφευχθεί η συστηματική απόκλιση από την αληθινή τιμή και αποτελεί επομένως αναγκαία προϋπόθεση για κάθε ουσιαστική αξιολόγηση ποιότητας.

Με αυτό το θεωρητικό υπόβαθρο, η αξιολόγηση μέσω σημασιολογικής ομοιότητας (*semantic similarity*) αποκτά σαφέστερη θέση ως μετρική. Αντί να συγκρίνει εγγραφές

λέξη-λέξη, μετρά πόσο κοντά βρίσκονται δύο κείμενα στον σημασιολογικό χώρο μέσω διανυσματικών αναπαραστάσεων (embeddings), επιτρέποντας τη σύγκριση ακόμα και όταν δύο εγγραφές χρησιμοποιούν διαφορετική διατύπωση για την ίδια πληροφορία [WBM+23]. Οι Lorenzini κ.ά. [LRT21] επισημαίνουν ότι η «καλή» αρχειακή περιγραφή δεν είναι μόνο ζήτημα ορθότητας αλλά και πληρότητας, σαφήνειας και συνέπειας, και προτείνουν δείκτες ποιότητας που μπορούν να εφαρμοστούν αυτόματα σε μεγάλα σύνολα εγγραφών, όπως η κάλυψη υποχρεωτικών πεδίων και η εσωτερική συνέπεια τιμών. Ο συνδυασμός αξιολόγησης από ειδικό και σημασιολογικής ομοιότητας, επιτρέπει μια πιο ολοκληρωμένη εικόνα, καλύπτοντας τόσο τη δομή και την πληρότητα όσο και το σημασιολογικό περιεχόμενο κάθε εγγραφής [LRT21, WBM+23].

Πάνω απ' όλα όμως, η αξιολόγηση των περιγραφών που παράγουν LLMs δεν είναι μόνο τεχνικό ζήτημα. Οι Wu κ.ά. [WBM+23] υπογραμμίζουν ότι η ερμηνευτική ακρίβεια, η πιστότητα στο πρωτογενές πλαίσιο και η συμμόρφωση με τις αρχές της αρχειακής επιστήμης δεν μπορούν να αξιολογηθούν αυτόματα, αλλά απαιτούν ανθρώπινη κρίση. Όσο αυτοματοποιημένη κι αν είναι η διαδικασία, η τελική κρίση για το αν μια εγγραφή είναι αξιόπιστη, παραμένει ανθρώπινη υπόθεση.

Κεφάλαιο 4 Δημιουργία αρχειακής περιγραφής με βάση το ΔΙΠΑΠ(Γ)/ISAD(G) με τη χρήση LLMs

Το παρόν κεφάλαιο αποτελεί το πειραματικό σκέλος της εργασίας. Στόχος του είναι να εξετάσει εμπειρικά κατά πόσο τα Μεγάλα Γλωσσικά Μοντέλα (LLMs) μπορούν να μετασχηματίσουν ελεύθερο κείμενο αρχειακής περιγραφής σε δομημένες εγγραφές συμβατές με το πρότυπο ISAD(G)/ΔΙΠΑΠ(Γ). Για τον σκοπό αυτό σχεδιάστηκε και εκτελέστηκε ένα πιλοτικό πείραμα σύγκρισης πολλαπλών μοντέλων, η μεθοδολογία του οποίου παρουσιάζεται στις επόμενες ενότητες.

4.1 Θεωρητικό Πλαίσιο και Μεθοδολογία

4.1.1 Δεδομένα εισόδου

Το πείραμα στην παρούσα εργασία αντιμετωπίζεται θεωρητικά ως πρόβλημα μετασχηματισμού πεδίου, δηλαδή, τα δεδομένα εισόδου αποτελούνται από ελεύθερο κείμενο αρχειακής περιγραφής που έχει συντάξει έμπειρος αρχειονόμος, και τα LLMs καλούνται να αντλήσουν από αυτό το περιεχόμενο και να το ανακατανεύμουν στα αντίστοιχα πεδία του προτύπου ISAD(G). Η πηγή δεδομένων είναι η αρχειακή περιγραφή του Αρχείου της Επτανήσου Πολιτείας (1800–1807), που φυλάσσεται στα Γενικά Αρχεία του Κράτους – Αρχεία Νομού Κέρκυρας (ΓΑΚ-ANK), εντός του Αρχείου Ιονίου Γερουσίας [Καβ23]. Το αρχείο αποτελείται από 473 κιβώτια με φακέλους και τεκμήρια, ταξινομημένα θεματικά και εντός κάθε θέματος χρονολογικά. Στο πλαίσιο της παρούσας μελέτης περίπτωσης, εξετάστηκε ειδικότερα ο φάκελος αρ. 8 του κιβωτίου αρ. 2 (κωδικός: GR GSA ANK B.2, F.8), ο οποίος περιέχει παλαιό κατάστιχο-βιβλίο με τρία κρίσιμα τεκμήρια. Τη Συνθήκη της Κωνσταντινουπόλεως, το Σύνταγμα της Κωνσταντινουπόλεως και την Επιστολή της Επτανησιακής Γερουσίας προς την Υψηλή Πύλη (25.10/06.11.1800).

Το κείμενο της αρχειακής περιγραφής των πειραμάτων ([Παράρτημα Α](#)) βασίστηκε στους καταλόγους ISAD που δημιουργήθηκαν στο πλαίσιο της εκπόνησης της Διδακτορικής Διατριβής τον Ιούνιο του 2022, από την αρχειονόμο Αλεξάνδρα Καββαδία [Καβ23], σύμφωνα με το Διεθνές Πρότυπο Αρχειακής Περιγραφής ISAD(G), 2η έκδοση (ΔΙΠΑΠ, Οττάβα 2000), και αποτελεί την εγγραφή αναφοράς, το gold standard, με βάση την οποία

αξιολογούνται τα αποτελέσματα των μοντέλων. Η επιλογή αυτού του αρχείου δεν ήταν τυχαία. Καταρχάς, διαθέτει πλούσιο και τεκμηριωμένο πλαίσιο παραγωγής. Η περιγραφή περιλαμβάνει αναφορές σε ιστορικά γεγονότα, νομικά κείμενα και λεπτομερή φυσικά χαρακτηριστικά υλικού, πράγμα που σημαίνει ότι τα μοντέλα έχουν αρκετή πληροφορία να επεξεργαστούν, αλλά και αρκετά σημεία όπου μπορεί να αποκλίνουν από την εγγραφή αναφοράς. Ακόμα, καλύπτει πολλαπλά ιεραρχικά επίπεδα ISAD(G), από το *fonds* μέχρι το *item*, επιτρέποντας την αξιολόγηση της ικανότητας των μοντέλων να διαχειρίζονται ιεραρχική δομή, που είναι ακριβώς το πιο απαιτητικό χαρακτηριστικό του προτύπου. Τέλος, η ύπαρξη εγγραφής ειδικού, και μάλιστα σε πέντε ιεραρχικά επίπεδα, καθιστά εφικτή τη συστηματική και τεκμηριωμένη σύγκριση.

Η περιγραφή μετασχηματίστηκε σε απλό κείμενο (*plain text*) και χρησιμοποιήθηκε ως είσοδος (*source description*) στο *prompt* κάθε μοντέλου, χωρίς προσθήκη εξωτερικής γνώσης ή οποιασδήποτε επεξεργασίας. Αυτή η επιλογή είναι μεθοδολογικά συνειδητή διότι ο σκοπός δεν είναι να δοθεί στα μοντέλα κάθε δυνατή βοήθεια για να πετύχουν, αλλά να αξιολογηθεί η πραγματική τους ικανότητα να εξάγουν και να οργανώνουν πληροφορία από ένα κείμενο που θα μπορούσε να αποτελεί τυπική εισροή σε ένα ρεαλιστικό αρχειακό *workflow*.

4.1.2 Μεθοδολογική ροή

Επιλέξαμε να αξιολογήσουμε τα LLMs έναντι του ISAD(G), διότι προσφέρει ένα ευρέως αναγνωρισμένο περιγραφικό στόχο για τη δόμηση αρχειακού κειμένου σε επίπεδο πεδίου, ο οποίος συμπίπτει με τον βασικό σκοπό της μελέτης. Η χρήση του ISAD(G) διατηρεί την αξιολόγηση εστιασμένη και συγκρίσιμη μεταξύ μοντέλων (ανάθεση πεδίων και ποιότητα κειμένου πεδίου). Η μεθοδολογική ροή του πειράματος οργανώθηκε σε τέσσερα διαδοχικά στάδια, σχεδιασμένα έτσι ώστε η κάθε φάση να τροφοδοτεί την επόμενη με σαφή και ελέγξιμο τρόπο.

Το πρώτο στάδιο αφορά την προετοιμασία των δεδομένων εισόδου. Η αρχειακή περιγραφή του ειδικού στην αρχική της μορφή με πίνακες και δομημένη παρουσίαση πεδίων, μετατράπηκε σε απλό κείμενο (*plain text*), αφαιρώντας κάθε προϋπάρχουσα δόμηση. Η επιλογή αυτή στοχεύει στη δημιουργία ρεαλιστικών συνθηκών: στην πράξη, τα κείμενα που καλούνται να επεξεργαστούν τα LLMs σε αρχειακό περιβάλλον σπάνια έχουν ήδη τυποποιημένη δομή. Αντίθετα, είναι συνήθως ελεύθερο κείμενο, όπως εκθέσεις αρχειονόμων, ευρετήρια ή αδόμητες περιγραφές, από τα οποία το μοντέλο πρέπει να αντλήσει μόνο του τα σχετικά πεδία. Επίσης για όλα τα πειράματα, οι παράμετροι αποκωδικοποίησης ορίστηκαν σταθερές σε *temperature: 0, top_p:1*, ώστε να ληφθούν

ντετερμινιστικά αποτελέσματα στον βαθμό που υποστηρίζονται από κάθε πάροχο μοντέλου. Αυτές οι ρυθμίσεις μειώνουν την τυχαιότητα κατά την παραγωγή και επιτρέπουν μια πιο συνεπή σύγκριση της ικανότητας των μοντέλων να αντιστοιχίζουν περιγραφές αρχειακού ελεύθερου κειμένου σε δομές συμβατές με το ISAD(G).

Τα πειράματα διεξήχθησαν με τη χρήση δύο πλατφορμών πρόσβασης σε μοντέλα, του Hugging Face Chat και του OpenRouter (Πίνακας 3). Όλες οι έξοδοι των μοντέλων παρήχθησαν μέσω πρόσβασης API και όχι μέσω διαδραστικών διεπαφών συνομιλίας. Αυτό διασφάλισε ότι το ίδιο prompt, η ίδια περιγραφή πηγής και οι ίδιες παράμετροι αποκωδικοποίησης εφαρμόστηκαν με συνέπεια σε όλα τα αξιολογούμενα μοντέλα. Σε όλες τις περιπτώσεις, κάθε μοντέλο έλαβε το πανομοιότυπο prompt μετασχηματισμού ISAD(G) και την ίδια ελληνική περιγραφή πηγής ελεύθερου κειμένου.

LLM	Δημιουργός	Παράμετροι	Πλατφόρμα Πρόσβασης
meta-llama-Llama-3.1-8B	Meta	8B	Hugging Face Chat API
openai-gpt-oss-20b	OpenAI	21B συνολικές · 3,6B ενεργές	Hugging Face Chat API
deepseek-ai-DeepSeek-V3.2	DeepSeek-AI	671B συνολικές · 37B ενεργές	Hugging Face Chat API
meta-llama-Llama-3.2-3B	Meta	3B	Hugging Face Chat API
meta-llama-Llama-3.3-70B	Meta	70B	Hugging Face Chat API
Qwen-Qwen3-235B-A22B-Instruct	Alibaba/Qwen	235B συνολικές · 22B ενεργές	Hugging Face Chat API
arcee-ai-trinity-large	Arcee AI	400B συνολικές · 13B ενεργές	OpenRouter API
arcee-ai-trinity-mini	Arcee AI	26B συνολικές · 3B ενεργές	OpenRouter API

LLM	Δημιουργός	Παράμετροι	Πλατφόρμα Πρόσβασης
deepseek-r1-0528	DeepSeek-AI	671B συνολικές · 37B ενεργές	OpenRouter API
google-gemma-3-4b-it	Google	4B	OpenRouter API
google-gemma-3-12b-it	Google	12B	OpenRouter API

Πίνακας 3 Μεγάλα Γλωσσικά Μοντέλα που χρησιμοποιήθηκαν στην Αξιολόγηση

Στο δεύτερο στάδιο, το κείμενο αυτό τοποθετήθηκε ως source description μέσα στο prompt ([Παράρτημα Β](#)) το οποίο είναι σχεδιασμένο να παράγει αρχειακές περιγραφές συμμορφούμενες με το ISAD(G), βασισμένες αποκλειστικά σε κείμενο πηγής που έχει δοθεί από αρχειονόμο. Αυτό δίνει εντολή στο LLM να μην χρησιμοποιήσει εξωτερική γνώση και να μην υποθέτει ή να συμπληρώνει ελλείψεις πληροφορίες. Εάν ένα πεδίο του ISAD(G) δεν αναφέρεται ρητά στο κείμενο, το μοντέλο πρέπει να γράψει «N/A (not stated in source)». Το μοντέλο πρέπει επίσης να δημιουργήσει ξεχωριστά μπλοκ ISAD(G) για το αρχείο και για κάθε υπομονάδα που περιγράφεται ρητά στο κείμενο εισόδου. Δόθηκε λοιπόν ως εισόδος σε κάθε ένα από τα 11 μοντέλα που συμμετείχαν στο πείραμα. Κάθε μοντέλο έλαβε το ίδιο ακριβώς prompt, με εντολή να παραγάγει δομημένες εγγραφές ISAD(G) για όλες τις αρχειακές μονάδες που αναφέρονται στο κείμενο. Η χρήση ενός ενιαίου prompt για όλα τα μοντέλα εξασφαλίζει συγκρισιμότητα. Τυχόν διαφορές στα αποτελέσματα δεν οφείλονται σε διαφορετικές οδηγίες αλλά στις εγγενείς ικανότητες κάθε μοντέλου.

Στο τρίτο στάδιο, η έξοδος κάθε μοντέλου συγκρίθηκε με τις πέντε εγγραφές αναφοράς του ειδικού, τη fonds, τον φάκελο και τα τρία τεκμήρια. Η σύγκριση δεν περιορίστηκε στο αν τα πεδία έχουν συμπληρωθεί, αλλά εξέτασε και την ορθότητα των τιμών, τη χρήση της σύμβασης «id στο Αρχείο» για κληρονομούμενα πεδία, τη διαφοροποίηση παραγωγού ανά τεκμήριο και την πληρότητα τίτλων, κριτήρια που αναδείχθηκαν κατά την ανάλυση των εγγραφών αναφοράς και αντικατοπτρίζουν τη βαθύτερη αρχειακή λογική του ISAD(G) [ICA00].

Στο τέταρτο στάδιο, τα αποτελέσματα αξιολογήθηκαν χειροκίνητα μέσω ποιοτικής αξιολόγησης από μία ομάδα ειδικών (expert evaluation), που εξέτασε την πληρότητα, την ορθότητα και τη συμμόρφωση με το ISAD(G) για κάθε μοντέλο. Η ομάδα περιλάμβανε

τον εξειδικευμένο αρχειονόμο που είχε οργανώσει και περιγράψει το αρχείο και ο οποίος είχε επίσης συντάξει την πρωτότυπη ελληνική περιγραφή πηγής ελεύθερου κειμένου που χρησιμοποιήθηκε στο πείραμα. Οι άλλοι δύο αξιολογητές ήταν επίσης ειδικοί στην αρχειακή περιγραφή και στα πρότυπα μεταδεδομένων.

Οι έξοδοι των μοντέλων συγκρίθηκαν χειροκίνητα με την έγκυρη περιγραφή πηγής και την αναμενόμενη δομή ευθυγραμμισμένη με το ISAD(G). Η αξιολόγηση εξέτασε αν κάθε μοντέλο εντόπιζε σωστά τις αναμενόμενες περιγραφικές μονάδες, διατηρούσε την αρχειακή ιεραρχία, ανέθετε τα κατάλληλα επίπεδα περιγραφής, συμπλήρωνε τα πεδία του ISAD(G) με πληροφορίες θεμελιωμένες στην πηγή και απέφευγε μη τεκμηριωμένες ή παραιοθητικές μονάδες. Οι διαφωνίες συζητούνταν μεταξύ των αξιολογητών και επιλύονταν με συναίνεση. Η αξιολόγηση δεν ήταν τυφλή ως προς την ταυτότητα του μοντέλου. Αν και δεν χρησιμοποιήθηκε ξεχωριστή αριθμητική κλίμακα βαθμολόγησης, η σύγκριση ακολούθησε μια δομημένη ποιοτική εκτίμηση βασισμένη στις κατηγορίες αξιολόγησης που αναφέρονται στους Πίνακες 2 και 3.

Η επιλογή αυτής της μεθόδου ως αποκλειστικής δεν ήταν τυχαία. Ο κανόνας COPY RULE του prompt ζητούσε από τα μοντέλα να αντιγράψουν κατά λέξη τις τιμές από το πρωτογενές κείμενο, χωρίς παράφραση. Στόχος ήταν να ελεγχθεί αν το μοντέλο κατανόησε τη δομή και αναγνώρισε τα σωστά πεδία, όχι αν παρήγαγε σημασιολογικά ισοδύναμες διατυπώσεις. Υπό αυτές τις συνθήκες, η σημασιολογική ομοιότητα (semantic similarity) δεν αποτελεί κατάλληλο μέτρο αξιολόγησης γιατί, όταν η έξοδος αναμένεται να είναι λεκτικά πανομοιότυπη με την πηγή, η μέτρηση σημασιολογικής ομοιότητας δεν προσφέρει πρόσθετη πληροφορία.

Η ποιοτική αξιολόγηση από την ομάδα των ειδικών (expert evaluation) παραμένει η πιο αξιόπιστη μέθοδος για την αξιολόγηση αρχειακής ορθότητας, πληρότητας και συμμόρφωσης με το ISAD(G) διότι υπάρχουν πτυχές της ποιότητας μιας αρχειακής εγγραφής που κανένας αλγόριθμος δεν μπορεί να μετρήσει αξιόπιστα, μόνο ο ειδικός μπορεί να τις κρίνει [LRT21, WBM+23]. Για να γίνει αυτό πιο συγκεκριμένο, αν ένα μοντέλο γράψει ως παραγωγό «Επτάνησος Πολιτεία» αντί για «Επτάνησος Πολιτεία, Συντακτικό Νομοθετικό Σώμα», ένας αλγόριθμος σημασιολογικής ομοιότητας πιθανόν να το βαθμολογήσει υψηλά διότι οι δύο φράσεις είναι σημασιολογικά κοντά. Ο ειδικός όμως το αναγνωρίζει αμέσως ως σφάλμα, γιατί λείπει η ακριβής αρχή παραγωγής που ορίζει το ISAD(G). Το ίδιο ισχύει για το αν χρησιμοποιήθηκε η σύμβαση «id στο Αρχείο» σωστά, αν ένα πεδίο συμπληρώθηκε στο κατάλληλο ιεραρχικό επίπεδο, ή αν ο τίτλος ενός τεκμηρίου είναι αρκετά περιγραφικός κατά τα αρχειακά πρότυπα. Πρόκειται για κρίσεις

που απαιτούν επαγγελματική γνώση του αντικειμένου και δεν ανάγονται σε μαθηματική συνάρτηση.

4.1.3 Μοντέλα που αξιολογήθηκαν

Στο πείραμα συμμετείχαν συνολικά 11 μοντέλα ανοιχτού κώδικα και εμπορικά, επιλεγμένα έτσι ώστε να καλύπτουν ένα ευρύ φάσμα μεγεθών και αρχιτεκτονικών. Συγκεκριμένα αξιολογήθηκαν τα: Llama 3.1 8B Instruct, Llama 3.2 3B Instruct και Llama 3.3 70B Instruct, τρεις εκδόσεις της ίδιας οικογένειας που επιτρέπουν άμεση σύγκριση επίδρασης μεγέθους, καθώς και GPT-oss-20B, DeepSeek-V3.2, Qwen3-235B-A22B Instruct, arcee-ai-trinity-large-preview, arcee-ai-trinity-mini, deepseek-r1-0528, google-gemma-3-4b-it και google-gemma-3-12b-it. Η επιλογή αυτή επιτρέπει σύγκριση μεταξύ μικρών μοντέλων (3B–8B παραμέτρων), μεσαίων (20B) και μεγάλων (70B–235B), καλύπτοντας έτσι το τρίτο ερευνητικό ερώτημα της εργασίας, κατά πόσο το μέγεθος συσχετίζεται με καλύτερη απόδοση σε εξειδικευμένες εργασίες αρχειακής περιγραφής. Να επισημάνω εδώ, ότι για ορισμένα μοντέλα (DeepSeek-V3.2, arcee-trinity) δεν είναι δημοσίως γνωστός ο ακριβής αριθμός παραμέτρων, γεγονός που περιορίζει μερικώς τη δυνατότητα άμεσης σύγκρισης μεγέθους αλλά δεν αναιρεί τη συγκριτική αξία της αξιολόγησης ποιότητας εξόδου.

4.2 Σχεδιασμός Prompt (Prompt Design)

Ο σχεδιασμός του prompt αποτελεί έναν από τους πιο κρίσιμους παράγοντες για την ποιότητα της εξόδου ενός LLM, ιδίως όταν η εργασία απαιτεί αυστηρή συμμόρφωση με πρότυπα και αποφυγή επινοημένων πληροφοριών [ZHP+23, GMW+25]. Στο παρόν πείραμα, ο σχεδιασμός του prompt αντιμετωπίστηκε ως μεθοδολογική επιλογή που αντικατοπτρίζει τόσο τις αρχές της αρχειακής δεοντολογίας όσο και τα όρια που θέτει η βιβλιογραφία για την αξιόπιστη χρήση LLMs [RSD23, WBM+23, Gil08]. Το prompt δομήθηκε γύρω από έξι βασικούς κανόνες λειτουργίας, καθένας από τους οποίους αντιμετωπίζει ένα συγκεκριμένο πρόβλημα.

Ο πρώτος κανόνας αφορά τον ρόλο (ROLE). Το μοντέλο καλείται ρητά να λειτουργήσει ως «αρχειακός καταλογογράφος» (archival description cataloguer), με σαφή θεσμικό προσανατολισμό. Η τεχνική αυτή, γνωστή ως role prompting, έχει δειχθεί ότι βελτιώνει τη συνέπεια και την εξειδίκευση της εξόδου, καθώς «στρέφει» το μοντέλο προς το κατάλληλο εννοιολογικό πλαίσιο πριν καν επεξεργαστεί το κείμενο [Bag24, OWJ+22]. Στον αρχειακό χώρο, αυτό σημαίνει ότι το μοντέλο δεν αντιμετωπίζει το έργο ως γενική «περίληψη κειμένου» αλλά ως δομημένη αρχειακή τεκμηρίωση με συγκεκριμένες επαγγελματικές υποχρεώσεις.

Ο δεύτερος κανόνας είναι ο κανόνας πηγής (SOURCE RULE). Το μοντέλο οφείλει να χρησιμοποιεί αποκλειστικά το κείμενο της source description, χωρίς εξωτερική γνώση και χωρίς εικασίες. Αυτός είναι ο πιο θεμελιώδης κανόνας, και η αιτιολογία του βρίσκεται ακριβώς στο πρόβλημα των hallucinations που αναλύουν οι Rawte κ.ά. [RSD23]. Ένα μοντέλο που «ξέρει» πληροφορίες για την Επτάνησο Πολιτεία από τα δεδομένα εκπαίδευσής του θα έμπαινε στον πειρασμό να τις χρησιμοποιήσει για να «συμπληρώσει» κενά στο κείμενο. Αυτό ακριβώς είναι το είδος του σφάλματος που η αρχειακή δεοντολογία δεν επιτρέπει. Μια περιγραφή που φαίνεται ορθή αλλά στηρίζεται σε ανεπαλήθευτες εξωτερικές πληροφορίες αντί στο ίδιο το τεκμήριο [Gil08].

Ο τρίτος κανόνας είναι ο κανόνας συμπλήρωσης (FILL RULE). Για κάθε πεδίο ISAD(G) που δεν αναφέρεται ρητά στο κείμενο, το μοντέλο οφείλει να γράφει «N/A (not stated in source)», αντί να αφήνει κενό ή να επινοεί τιμή. Αυτός ο κανόνας αντιστοιχεί στη στρατηγική αγνόησης ελλειπόντων δεδομένων που περιγράφουν οι Tan κ.ά. [TSKK19]: αντί το μοντέλο να «φανταστεί» μια λογική τιμή, κάτι που οδηγεί κατευθείαν σε hallucination, δηλώνει ρητά ότι η πληροφορία απουσιάζει από την πηγή, διατηρώντας έτσι την αξιοπιστία και τον έλεγχο της εγγραφής. Ο κανόνας αυτός, όπως θα φανεί στα αποτελέσματα, αλληλοεπιδρά με ένα ειδικό χαρακτηριστικό του ISAD(G) που τα μοντέλα δεν κατανόησαν, την σύμβαση «id στο Αρχείο», με την οποία ο ειδικός παραπέμπει στην εγγραφή ανώτερου επιπέδου αντί να γράφει N/A, μια λεπτή διαφορά με σημαντικές αρχειακές συνέπειες.

Ο τέταρτος κανόνας είναι ο κανόνας ασάφειας (AMBIGUITY RULE). Σε περίπτωση αντίφασης ή ασάφειας στην πηγή, το μοντέλο αντιγράφει τις τιμές ακριβώς όπως εμφανίζονται, χωρίς να τις ερμηνεύει ή να τις «λύνει». Η λογική πίσω από αυτόν τον κανόνα συμβαδίζει με αυτό που περιγράφουν οι Tan κ.ά. [TSKK19] για τις ασυνεπείς τιμές. Όταν δεν υπάρχει εξωτερική βάση επαλήθευσης, η ασυνέπεια πρέπει να καταγράφεται όπως είναι, ώστε να παραμένει ορατή για τον αρχειονόμο που θα κάνει τον τελικό έλεγχο. Μια «διορθωμένη» αντίφαση από το μοντέλο είναι πολύ πιο επικίνδυνη από μία ορατή, γιατί κρύβει ένα πρόβλημα που ίσως χρειαζόταν ανθρώπινη παρέμβαση.

Ο πέμπτος κανόνας είναι ο κανόνας μονάδων (UNITS RULE). Το μοντέλο δημιουργεί ξεχωριστό μπλοκ ISAD(G) για κάθε αρχειακή μονάδα που αναφέρεται ρητά στο κείμενο, fonds, series, file, item, χωρίς να «εφευρίσκει» υποδιαιρέσεις που δεν τεκμηριώνονται στην πηγή. Ο κανόνας αυτός συνδέεται άμεσα με την αρχή της ιεραρχικής περιγραφής του ISAD(G) [ICA00]. Κάθε επίπεδο πρέπει να τεκμηριώνεται ρητά και να αντιστοιχεί σε πραγματική αρχειακή μονάδα, όχι σε αναλυτική υποδιάρθρωση που το μοντέλο θεωρεί

λογική. Η παραβίαση αυτού του κανόνα θα σήμαινε δημιουργία «εικονικών» επιπέδων που δεν αντιστοιχούν στη φυσική οργάνωση του αρχείου.

Ο έκτος κανόνας είναι ο κανόνας αντιγραφής (COPY RULE). Οι τιμές των πεδίων μεταφέρονται ακριβώς όπως εμφανίζονται στην πηγή, χωρίς παράφραση, διόρθωση ή εκσυγχρονισμό της γλώσσας. Αυτός ο κανόνας αντικατοπτρίζει την αρχή της ελάχιστης αναγκαίας παρέμβασης που διέπει την αρχειακή δεοντολογία [Gil08], η γλώσσα του πρωτογενούς κειμένου, ακόμα και αν είναι αρχαϊκή, ασαφής ή χρησιμοποιεί ορολογία που σήμερα έχει αλλάξει, αποτελεί μέρος της ιστορικής μαρτυρίας και δεν μπορεί να «εξομαλυνθεί» από ένα αυτόματο σύστημα.

Εκτός από τους έξι κανόνες, ορίστηκε αυστηρά και η μορφή εξόδου (OUTPUT). Κανένα εισαγωγικό σχόλιο ή σύνοψη, ένα μπλοκ ανά αρχειακή μονάδα, με τα 26 στοιχεία των 7 περιοχών ISAD(G) σε σταθερή σειρά. Αυτή η επιλογή δεν είναι αυθαίρετη, στοχεύει στη μέτρηση της ικανότητας του μοντέλου να παράγει δομημένη έξοδο με συνέπεια, χωρίς να αποκλίνει σε ελεύθερο κείμενο ή να παραλείπει πεδία.

Η συνολική φιλοσοφία πίσω από τον σχεδιασμό αυτό αντικατοπτρίζει αυτό που η βιβλιογραφία αποκαλεί «minimal intervention» στην αρχειακή δεοντολογία [Gil08] και επιχειρεί να περιορίσει συστηματικά τους κινδύνους hallucination και δομικής ασυνέπειας που επισημαίνονται στη βιβλιογραφία [RSD23, WBM+23]. Η αξιολόγηση στο Κεφάλαιο 4.3 θα δείξει σε ποιο βαθμό αυτοί οι κανόνες λειτούργησαν και πού ακόμα και με αυστηρό prompt, τα μοντέλα κατάφεραν να αποκλίνουν με ενδιαφέροντες τρόπους.

4.3 Μελέτη Περίπτωσης: Αρχείο Επτανήσου Πολιτείας

4.3.1 Περιγραφή της μελέτης περίπτωσης

Όπως αναφέρθηκε αναλυτικά στην ενότητα 4.1, το αρχειακό υλικό που χρησιμοποιήθηκε ως μελέτη περίπτωσης είναι το Αρχείο της Επτανήσου Πολιτείας (1800–1807, GR GSA ANK). Για τους σκοπούς της αξιολόγησης, εκτός από το επίπεδο fonds, συμπεριλήφθηκαν ο φάκελος B.2/F.8 και τα τρία τεκμήριά του (14a, 14b και 14c) – η Συνθήκη και το Σύνταγμα της Κωνσταντινουπόλεως, καθώς και η Επιστολή της Επτανησιακής Γερουσίας προς την Υψηλή Πύλη. Η επιλογή αυτή επιτρέπει την αξιολόγηση των μοντέλων όχι μόνο σε επίπεδο μεμονωμένης εγγραφής αλλά και ως προς την ικανότητά τους να σέβονται τη λογική ιεραρχικής κληρονομικότητας του ISAD(G).

4.3.2 Αποτελέσματα ανά μοντέλο

Η αξιολόγηση των αποτελεσμάτων πραγματοποιήθηκε από ομάδα ειδικών αρχειονόμων με κριτήριο τη συμφωνία με το σύνολο των πέντε εγγραφών αναφοράς, την εγγραφή σε επίπεδο Αρχείου (fonds), την εγγραφή του Φακέλου με το Κατάστιχο, και τις τρεις εγγραφές για τα τεκμήρια 14a, 14b και 14c. Τα κριτήρια αξιολόγησης περιλαμβάνουν (α) Expected Units Generated, πόσες από τις 5 αναμενόμενες αρχειακές μονάδες αναπαρήγαγε το μοντέλο, (β) Hallucinated Units πόσες αδικαιολόγητες ή επινοημένες δομικές μονάδες πρόσθεσε, (γ) Correct Level Assignment, αν οι μονάδες που παρήγαγε αποδόθηκαν στο σωστό ιεραρχικό επίπεδο (fonds/file/item), (δ) πληρότητα των 26 στοιχείων των 7 περιοχών ISAD(G), (ε) χρήση της σύμβασης «id στο Αρχείο» για κληρονομούμενα πεδία, σύμφωνα με την αρχή της ιεραρχικής κληρονομικότητας του ISAD(G) [ICA00], και (στ) διαφοροποίηση παραγωγού ανά τεκμήριο όπου το ιστορικό πλαίσιο το απαιτεί.

Το πιο εντυπωσιακό εύρημα δεν αφορά κάποιο συγκεκριμένο μοντέλο αλλά όλα μαζί, κανένα δεν χρησιμοποίησε τη σύμβαση «id στο Αρχείο». Η σύμβαση αυτή, με την οποία ο ειδικός παραπέμπει στην εγγραφή ανώτερου επιπέδου, αντί να επαναλαμβάνει ή να παραλείπει την πληροφορία, αποτελεί θεμελιώδη αρχή της ιεραρχικής αρχειακής περιγραφής. Τα μοντέλα που τήρησαν τον κανόνα FILL RULE του prompt, έγραψαν «N/A (not stated in source)» για τα κληρονομούμενα πεδία, δηλώνοντας ψευδώς ότι η πληροφορία απουσιάζει, άλλα επανέλαβαν αυτούσια την πληροφορία, αγνοώντας την αρχή της μη επανάληψης. Εξίσου χαρακτηριστικό είναι ότι κανένα μοντέλο δεν διαφοροποίησε τον παραγωγό ανά τεκμήριο. Ο ειδικός αποδίδει στα τεκμήρια 14a και 14b το «Συντακτικό Νομοθετικό Σώμα» και στο 14c την «Ειπανησιακή Γερουσία», διάκριση που απαιτεί βαθύτερη ανάγνωση του ιστορικού πλαισίου και η οποία διέφυγε από όλα τα μοντέλα χωρίς εξαίρεση.

Στη συνέχεια παρουσιάζονται τα αποτελέσματα ανά μοντέλο.

Το **Llama 3.1 8B Instruct** (1η Περίπτωση) δεν κατάφερε να υπερβεί την εγγραφή fonds, ούτε ο φάκελος ούτε τα τρία τεκμήρια απαντούν στην έξοδό του. Ακόμη και στην εγγραφή fonds υπάρχουν κενά. Τα δύο τελευταία στοιχεία του 5ου πεδίου και το μοναδικό στοιχείο του 6ου πεδίου παραλείφθηκαν, ενώ η διοικητική ιστορία αποδόθηκε αποσπασματικά. Από τα 11 μοντέλα, αυτό παρουσίασε μαζί με το Llama 3.2 3B την πιο αδύναμη συνολική απόδοση.

Το **GPT-oss-20B** (2η Περίπτωση) ήταν το πιο «φιλόδοξο» ως προς τον αριθμό εγγραφών, συνέταξε 7, καλύπτοντας fonds, κιβώτιο, φάκελο, κατάστιχο και τα τρία τεκμήρια. Ο

αριθμός αυτός όμως, δεν συνοδεύτηκε από ποιότητα. Η εγγραφή fonds ήταν ελλιπής στο 2ο, 3ο, 5ο και 6ο πεδίο, και το πεδίο Notes, όπου ο ειδικός σημειώνει κρίσιμη παρατήρηση για τη γλώσσα των συνταγματικών κειμένων, αποδόθηκε ως N/A. Ο τίτλος του τεκμηρίου 14c αποδόθηκε συνοπτικά, χωρίς το θεματικό περιεχόμενο που ο ειδικός περιλαμβάνει ρητά.

Το **Llama 3.3 70B Instruct** (3η Περίπτωση) παρουσίασε την πληρέστερη εγγραφή fonds από όλα τα μοντέλα, καλύπτοντας ουσιαστικά όλους τους τομείς, διοικητική ιστορία, αρχειακό ιστορικό, δημοσιεύσεις. Συνέταξε 6 εγγραφές συνολικά. Παρά ταύτα, το πεδίο Notes αποδόθηκε ως N/A, η σύμβαση «id στο Αρχείο» δεν χρησιμοποιήθηκε σε καμία υποδιαίρεση, και ο παραγωγός αποδόθηκε ομοιόμορφα ως «Νομοθετικό Σώμα» για όλα τα τεκμήρια. Είναι το μοντέλο που πλησίασε περισσότερο στο gold standard σε επίπεδο fonds, αλλά παραμένει σε απόσταση ως προς την αρχειακή σύμβαση.

Το **Llama 3.2 3B Instruct** (4η Περίπτωση) συνέταξε εγγραφή μόνο για το fonds, με ελλιπή συμπλήρωση του 5ου πεδίου και εσφαλμένο τίτλο («Στα Αρχεία Κέρκυρας (ANK)» αντί «Αρχείο Επτανήσου Πολιτείας»). Σε σύγκριση με το Llama 3.1 8B, παρουσίασε ακόμη πιο περιορισμένη απόδοση, γεγονός που επιβεβαιώνει ότι κάτω από ένα κατώφλι παραμέτρων η ιεραρχική επεξεργασία γίνεται εκτός εμβέλειας.

Το **Qwen3-235B-A22B** (5η Περίπτωση) συνέταξε 6 εγγραφές, αλλά χαρακτηρίζεται από ένα σαφές δομικό σφάλμα. Το κιβώτιο αποδόθηκε ως «Αρχειακή Σειρά» (Series) αντί ως υλικό δοχείο, λάθος που δείχνει επιφανειακή κατανόηση της ιεραρχικής λογικής του ISAD(G). Η εγγραφή fonds ήταν ελλιπής στο 2ο, 5ο και 6ο πεδίο. Παρά το μεγάλο μέγεθος του, η απόδοση δεν ήταν ανάλογη.

Το **DeepSeek-V3.2** (6η Περίπτωση) συνέταξε ακριβώς 5 εγγραφές. Η εγγραφή fonds ήταν πλήρης και ξεχώρισε για τη διατήρηση θεματικού περιεχομένου στο πεδίο Scope and Content. Ωστόσο, οι εγγραφές των υπομονάδων (φάκελος και τρία τεκμήρια) ήταν σχεδόν εξ ολοκλήρου N/A, με ελάχιστο ουσιαστικό περιεχόμενο πέρα από Reference Code και Title. Η σύμβαση «id στο Αρχείο» απουσιάζει, ο παραγωγός δεν διαφοροποιήθηκε ανά τεκμήριο, και ο τίτλος του 14c παραμένει ελλιπής. Παρουσιάζει ισχυρή απόδοση στην περιοχή Content and Structure σε επίπεδο fonds, αλλά αδυναμία να μεταφέρει την ίδια ποιότητα στα κατώτερα επίπεδα.

Το **arcee-ai-trinity-large-preview** (7η Περίπτωση) συνέταξε μόνο 2 εγγραφές, fonds και ένα δεύτερο unit που το ονόμασε λανθασμένα «Series» (αρχειακή σειρά) αντί για «File» (φάκελος). Η εγγραφή fonds ήταν σχεδόν πλήρης, αλλά η λανθασμένη ταξινόμηση του

δεύτερου unit και η απουσία όλων των επιπέδων item αναδεικνύουν αδυναμία ιεραρχικής κατανόησης. Η συνολική απόδοση χαρακτηρίζεται ως μέτρια.

Το **arcee-ai-trinity-mini** (8η Περίπτωση) συνέταξε 2 εγγραφές, fonds και φάκελο και οι δύο ελλιπέστατες, χωρίς δομημένα πεδία στον φάκελο. Η απόδοσή του ήταν η αδύναμη αναμενόμενη για ένα μικρό μοντέλο χωρίς εξειδίκευση.

Το **deepseek-r1-0528** (9η Περίπτωση) δεν παρήγαγε καμία αρχειακή περιγραφή. Είναι η μοναδική απόλυτη αποτυχία του πειράματος, το μοντέλο αδυνατεί εντελώς να ανταποκριθεί στο prompt, ανεξαρτήτως του περιεχομένου της εισόδου.

Το **google-gemma-3-4b-it** (10η Περίπτωση) συνέταξε 2 εγγραφές, fonds και φάκελο με χαρακτηριστικό σφάλμα σύγχυσης ιεραρχικών επιπέδων. Στο πεδίο 3.1.5 (Extent and medium) του fonds αναγράφεται η φυσική περιγραφή του κατάστιχου αντί του συνολικού μεγέθους του αρχείου. Πρόκειται για ένα από τα πιο εύγλωττα παραδείγματα του τι συμβαίνει όταν ένα μοντέλο δεν κατανοεί την ιεραρχική λογική της αρχειακής περιγραφής.

Το **google-gemma-3-12b-it** (11η Περίπτωση) παρουσίασε το καλύτερο συνολικό δομικό αποτέλεσμα του πειράματος. Συνέταξε ακριβώς τις 5 αναμενόμενες εγγραφές, fonds, file και τα τρία items 14a, 14b, 14c, χωρίς να εισαγάγει ούτε μία αδικαιολόγητη μονάδα στην ιεραρχία. Η εγγραφή fonds ήταν πλήρης με όλα τα 26 πεδία ISAD(G) συμπληρωμένα από την πηγή, συμπεριλαμβανομένων των περιοχών Allied Materials (με τις βιβλιογραφικές αναφορές) και Notes. Οι εγγραφές των τεκμηρίων κράτησαν τους σωστούς τίτλους (Συνθήκη, Σύνταγμα, Επιστολή) και τα ορθά reference codes. Όπως και τα υπόλοιπα μοντέλα, δεν χρησιμοποίησε τη σύμβαση «id στο Αρχείο» και δεν διαφοροποίησε τον δημιουργό ανά τεκμήριο. Είναι αξιοσημείωτο ότι ένα μοντέλο 12B παραμέτρων πέτυχε καλύτερη δομική απόδοση από μοντέλα πολλαπλάσιου μεγέθους, εύρημα που υποδεικνύει ότι η ποιότητα του instruction tuning είναι πιο καθοριστική από την κλίμακα για εξειδικευμένες εργασίες όπως η αρχειακή περιγραφή.

A/A	ΜΟΝΤΕΛΟ	ΜΕΓ	EXPECTED UNITS	HALLUCINATED UNITS	CORRECT LEVEL ASSIGNMENT	ΚΥΡΙΑ ΠΑΡΑΤΗΡΗΣΗ
GS	Gold Standard (Καββαδία)	–	5/5	0	5/5	Εγγραφή αναφοράς
1	Llama 3.1 8B	8B	1/5	0	1/1	Δημιούργησε κυρίως το επίπεδο αρχείου· απέτυχε να επεκτείνει την ιεραρχία.
2	GPT-oss-20B	20B	5/5	2	4/5	Κάλυψε όλες τις αναμενόμενες μονάδες, αλλά πρόσθεσε παραισθήσεις σε επίπεδο κιβωτίου/μητρώου.
3	Llama 3.3 70B	70B	5/5	1	4/5	Καλή κάλυψη ιεραρχίας, αλλά πρόσθεσε μη τεκμηριωμένη μονάδα σειρά/κιβωτίου.
4	Llama 3.2 3B	3B	1/5	0	1/1	Πολύ περιορισμένη έξοδος, κυρίως εξαγωγή σε επίπεδο αρχείου
5	Qwen3-235B	235B	4/5	1	4/4	Καλή εξαγωγή σε επίπεδο τεκμηρίου, αλλά παρέλειψε τη μονάδα φακέλου και πρόσθεσε σειρά.
6	DeepSeek-V3.2	–	5/5	1	4/5	Ισχυρή δομική κάλυψη, αλλά ένα τεκμήριο διπλασιάστηκε / τοποθετήθηκε λανθασμένα.
7	arcee trinity-large	–	2/5	0	1/2	Παρήγαγε αρχείο και φάκελο, αλλά παρέλειψε όλες τις περιγραφές σε επίπεδο τεκμηρίου.
8	arcee trinity-mini	–	1/5	1	1/1	Αδύναμος ιεραρχικός έλεγχος, πρόσθεσε επιπλέον μονάδα τύπου αρχείου.
9	deepseek-r1-0528	–	0/5	0	-	Δεν εντοπίστηκαν παραγόμενες από το μοντέλο σειρές σύγκρισης πέραν των γραμμών αναφοράς.

10	gemma-3-4b-it	4B	0/5	1	-	Δεν παρήγαγε τις αναμενόμενες μονάδες, πρόσθεσε σειρά σε επίπεδο σειράς
11	gemma-3-12b-it	12B	5/5	0	4/5	Καλύτερο συνολικό αποτέλεσμα, καμία επιπλέον μονάδα παραίτησης δεν εντοπίστηκε.

Πίνακας 4. Αξιολόγηση της Ιεραρχικής Κατανόησης και της Δομικής Συμμόρφωσης μεταξύ Ανοικτών LLMs.

4.3.3 Συγκριτική επισκόπηση αποτελεσμάτων

Από τη συνολική αξιολόγηση των 11 μοντέλων προέκυψαν ορισμένα μοτίβα που αξίζει να εξεταστούν συστηματικά.

Ο αριθμός εγγραφών από μόνος του αποδείχθηκε παραπλανητικός δείκτης. Το GPT-oss-20B παρήγαγε τον μεγαλύτερο αριθμό (7), ενώ το deepseek-r1-0528 δεν παρήγαγε καμία. Ωστόσο το GPT-oss-20B συνοδεύει τις εγγραφές του με δύο hallucinated units (επίπεδα box/register που δεν τεκμηριώνονται στην πηγή), ενώ το google-gemma-3-12b-it κατάφερε να παραγάγει ακριβώς τις 5 αναμενόμενες μονάδες χωρίς ούτε ένα hallucinated unit, παρουσιάζοντας το καλύτερο συνολικό δομικό αποτέλεσμα του πειράματος. Το συμπέρασμα είναι σαφές. Το να μετράς πόσες εγγραφές παρήγαγε ένα μοντέλο λέει ελάχιστα για το αν αυτές αντιστοιχούν στη σωστή ιεραρχία του αρχείου.

Η ανάλυση ανά περιοχή ISAD(G) αποκάλυψε ότι διαφορετικά μοντέλα υπερτερούν σε διαφορετικές διαστάσεις του προτύπου. Το google-gemma-3-12b-it έδειξε την καλύτερη απόδοση σε δομική συμμόρφωση, ιεραρχική κατανόηση, και έλεγχο hallucinations. Το Llama 3.3 70B Instruct υπερείχε στην περιοχή Context, διατηρώντας πλούσιες πληροφορίες προέλευσης και διοικητικής/βιογραφικής ιστορίας, αν και περιστασιακά επέκτεινε την ιεραρχία προσθέτοντας αδικαιολόγητα ενδιάμεσα επίπεδα. Το DeepSeek-V3.2 παρουσίασε ισχυρή απόδοση στην περιοχή Content and Structure, ιδίως στη διατήρηση θεματικών αναφορών στο πεδίο Scope and Content. Αυτή η διαφοροποίηση υποδηλώνει ότι κανένα μοντέλο δεν είναι «καλύτερο» απόλυτα, η επιλογή εξαρτάται από το τι αναζητά ο φορέας, δομική συμμόρφωση, πλούσιο πλαίσιο ή θεματικό βάθος.

Το πιο εντυπωσιακό εύρημα προκύπτει από τη σχέση μεγέθους και απόδοσης. Το μέγεθος του μοντέλου δεν καθορίζει από μόνο του την ποιότητα της αρχειακής περιγραφής. Το καλύτερο μοντέλο, gemma-3-12b-it, είναι μεσαίου μεγέθους (12B) και ξεπέρασε αρκετά μεγαλύτερα συστήματα — όπως το Qwen3-235B που, παρά τις 235 δισεκατομμύρια παραμέτρους του, παρέλειψε το επίπεδο file και πρόσθεσε εσφαλμένο series, ή το Llama

3.3 70B που πρόσθεσε αδικαιολόγητο ενδιάμεσο επίπεδο. Αυτό υποδεικνύει ότι η ποιότητα του instruction tuning, η σταθερότητα της εξόδου και η πιστή τήρηση των κανόνων του prompt ίσως είναι πιο καθοριστικά από τον αριθμό παραμέτρων για εξειδικευμένες εργασίες όπως η αρχειακή περιγραφή.

Στον αντίποδα, τα πολύ μικρά μοντέλα (3B–8B) παρουσίασαν σαφώς ασθενέστερη απόδοση. Το Llama 3.2 3B και το Llama 3.1 8B παρήγαγαν μόνο εγγραφή fonds και απέτυχαν να επεκτείνουν την ιεραρχία, ενώ το gemma-3-4b-it και το deepseek-r1-0528 απέτυχαν εξ ολοκλήρου να παραγάγουν τις αναμενόμενες μονάδες. Αυτά τα ευρήματα δείχνουν ότι κάτω από ένα ορισμένο κατώφλι παραμέτρων η ιεραρχική κατανόηση και η πολυεπίπεδη αρχειακή λογική γίνονται απλά εκτός εμβέλειας.

Δύο ευρήματα παραμένουν κοινά σε όλα τα μοντέλα ανεξαιρέτως. Το πρώτο είναι, η απόλυτη απουσία της σύμβασης «id στο Αρχείο». Ο ειδικός χρησιμοποιεί αυτή τη σύμβαση σε όλες τις εγγραφές υποδιαίρέσεων για να παραπέμψει στα πεδία που ορίζονται στο fonds, σύμφωνα με την αρχή του ISAD(G) ότι η πληροφορία δηλώνεται μία φορά στο ανώτερο επίπεδο και δεν επαναλαμβάνεται [ICA00]. Τα μοντέλα είτε έγγραφαν «N/A», δηλώνοντας ψευδώς ότι η πληροφορία απουσιάζει, είτε επανέλαβαν αυτολεξεί το περιεχόμενο. Δεν πρόκειται για τυπικό σφάλμα. Αποκαλύπτει ότι τα μοντέλα αντιμετωπίζουν κάθε εγγραφή ως ανεξάρτητο έγγραφο, όχι ως κόμβο μέσα σε μια ιεραρχία. Το δεύτερο είναι ότι κανένα μοντέλο δεν διαφοροποίησε τον δημιουργό ανά τεκμήριο. Ο ειδικός αποδίδει τα τεκμήρια 14a και 14b στο «Συντακτικό Νομοθετικό Σώμα» και το 14c στην «Επτανησιακή Γερουσία», διάκριση που προκύπτει από βαθιά γνώση του ιστορικού πλαισίου και δεν αναφέρεται ρητά στο πρωτογενές κείμενο. Αυτό επιβεβαιώνει έναν από τους κεντρικούς περιορισμούς που επισημαίνει η βιβλιογραφία. Τα LLMs δεν «κατανοούν» το πλαίσιο παραγωγής των αρχειακών τεκμηρίων, το αναπαράγουν ή το αγνοούν [JC22, WBM+23].

Περιοχή ISAD(G)	Καλύτερο Μοντέλο	Παρατήρηση
Περιοχή Δήλωσης Ταυτότητας	google-gemma-3-12b-it	Υψηλότερη ακρίβεια στην εξαγωγή κωδικών αναφοράς, τίτλων, ημερομηνιών και επιπέδων περιγραφής, διατηρώντας την αναμενόμενη δομή αρχείου/φακέλου/τεκμηρίου.

Περιοχή ISAD(G)	Καλύτερο Μοντέλο	Παρατήρηση
Περιοχή Πλαισίου	meta-llama-llama-3.3-70B	Καλύτερη διατήρηση της προέλευσης και της διοικητικής/βιογραφικής ιστορίας με λιγότερες παραλείψεις συγκεκριμένου.
Περιοχή Περιεχομένου και Δομής	deepseek-ai-DeepSeek-V3.2	Παρήγαγε τις πιο λεπτομερείς περιγραφές πεδίου και περιεχομένου και διατήρησε περισσότερη θεματική αρχειακή πληροφορία.
Περιοχή Πρόσβασης και Χρήσης	openai-gpt-oss-20b	Πιο πλήρης διαχείριση των πεδίων πρόσβασης, αν και ορισμένες περιγραφές γίνονταν γενικευμένες.
Περιοχή Σημειώσεων	Qwen-Qwen3-235B-A22B-Instruct	Παρήγαγε πλουσιότερες επεξηγηματικές σημειώσεις, αλλά κατά καιρούς εισήγαγε μη τεκμηριωμένες ερμηνευτικές προσθήκες.
Ιεραρχική Κατανόηση	google-gemma-3-12b-it	Καλύτερη συνολική κατανόηση της αρχειακής ιεραρχίας, διακρίνοντας με συνέπεια αρχείο, φάκελο και τεκμήριο χωρίς να εισάγει μη τεκμηριωμένες μονάδες.
Δομική Συμμόρφωση	google-gemma-3-12b-it	Πιο συνεπής τήρηση της σειράς πεδίων του ISAD(G) και της πολυεπίπεδης περιγραφικής δομής.
Διατήρηση Προέλευσης	meta-llama-llama-3.3-70B	Καλύτερη διατήρηση πληροφοριών σχετικών με τον δημιουργό και ιστορικού συγκεκριμένου.
Χαμηλότερο Ποσοστό Παραίτησης	google-gemma-3-12b-it	Παρήγαγε την καθαρότερη έξοδο με τις λιγότερες κατασκευασμένες περιγραφικές μονάδες ή μη τεκμηριωμένα μεταδεδομένα.

Πίνακας 5 Καλύτερα αποδίδοντα LLMs ανά περιοχή ISAD(G)

Συμπερασματικά, η εικόνα που αναδύεται είναι σύνθετη. Τα μοντέλα μπορούν να συμπληρώνουν πεδία, αλλά δεν σκέφτονται ιεραρχικά. Παράλληλα όμως, το πιο αξιόπιστο μοντέλο δομικά δεν είναι το μεγαλύτερο, γεγονός που ανατρέπει την υπόθεση ότι η κλίμακα είναι το κρίσιμο μέγεθος, και υποδηλώνει ότι η ποιότητα εκπαίδευσης και

η πιστή τήρηση οδηγιών παίζουν εξίσου σημαντικό ρόλο [GMW+25]. Η απόσταση αυτή φαίνεται πως δεν λύνεται μόνο με ένα καλύτερο prompt αλλά απαιτεί ανθρώπινη επικύρωση σε κάθε στάδιο [WBM+23, TAP25].

Κεφάλαιο 5 Συμπεράσματα – Μελλοντική Έρευνα

Στο παρόν κεφάλαιο συνοψίζονται τα κύρια ευρήματα της εργασίας και διατυπώνονται τα συμπεράσματα που προκύπτουν τόσο από τη βιβλιογραφική επισκόπηση όσο και από το πειραματικό μέρος. Στόχος δεν είναι η απλή επανάληψη όσων παρουσιάστηκαν στα προηγούμενα κεφάλαια, αλλά η σύνθεσή τους σε μια συνεκτική απάντηση στα τρία ερευνητικά ερωτήματα που τέθηκαν στην εισαγωγή, κατά πόσο τα LLMs μπορούν να παράγουν ISAD(G) συμβατές περιγραφές, κατά πόσο συμμορφώνονται με αυστηρούς κανόνες, και πώς διαφοροποιείται η απόδοση ανά μέγεθος μοντέλου.

5.1 Χρήση LLM ως βοηθητικό εργαλείο

Το πείραμα απάντησε με σαφήνεια και στα τρία ερευνητικά ερωτήματα, αλλά η πιο ενδιαφέρουσα διάσταση των απαντήσεων δεν είναι το «ναι» ή το «όχι» αλλά ο ακριβής εντοπισμός του τι μπορούν και τι δεν μπορούν να κάνουν τα LLMs στο συγκεκριμένο πλαίσιο της αρχειακής περιγραφής κατά ISAD(G).

Το πρώτο ερώτημα αφορούσε τη δυνατότητα παραγωγής ISAD(G) συμβατών περιγραφών. Η απάντηση είναι μερικώς θετική. Αρκετά μοντέλα κατάφεραν να παραγάγουν τις πέντε αναμενόμενες αρχειακές μονάδες (fonds, file, και τα τρία items) με βάση το prompt έξι κανόνων. Το google-gemma-3-12b-it πέτυχε το καλύτερο συνολικό δομικό αποτέλεσμα, αναπαράγοντας ακριβώς τις 5 αναμενόμενες εγγραφές χωρίς να εισαγάγει ούτε μία αδικαιολόγητη μονάδα στην ιεραρχία. Παράλληλα, διαφορετικά μοντέλα ξεχώρισαν σε διαφορετικές περιοχές του προτύπου. Το Llama 3.3 70B Instruct στην περιοχή Context, διατηρώντας πλούσιες πληροφορίες προέλευσης και διοικητικής/βιογραφικής ιστορίας, και το DeepSeek-V3.2 στην περιοχή Content and Structure, με ισχυρή απόδοση στο πεδίο Scope and Content σε επίπεδο fonds. Ωστόσο, καθώς η περιγραφή μετακινείται σε χαμηλότερα ιεραρχικά επίπεδα, η ποιότητα του περιεχομένου καταρρέει συστηματικά για τα περισσότερα μοντέλα. Πολλά μικρότερα μοντέλα (arcee-trinity-mini, gemma-3-4b-it, deepseek-r1-0528) απέτυχαν να παράγουν καν ένα αξιοποιήσιμο αρχειακό output.

Αναφορικά με το δεύτερο ερώτημα που αφορούσε τη συμμόρφωσή τους με αυστηρούς κανόνες, πήραμε μία πιο ανησυχητική εικόνα. Παρά το prompt των έξι κανόνων (ROLE, SOURCE, FILL, AMBIGUITY, UNITS, COPY), κανένα μοντέλο δεν συμμορφώθηκε πλήρως. Το πιο θεμελιώδες εύρημα είναι ότι κανένα μοντέλο δεν χρησιμοποίησε τη σύμβαση «id στο Αρχείο» που ορίζει το ISAD(G) για πεδία που κληρονομούνται από ανώτερο ιεραρχικό επίπεδο. Αντί αυτού, είτε επαναλάμβαναν πληροφορίες είτε έγραφαν N/A, αγνοώντας εντελώς τη λογική ιεραρχικής κληρονομικότητας που αποτελεί τον πυρήνα του προτύπου. Αντίστοιχα, κανένα μοντέλο δεν κατάφερε να διαφοροποιήσει τον δημιουργό ανά τεκμήριο. Ο ειδικός αποδίδει στα τεκμήρια 14a και 14b το «Συντακτικό Νομοθετικό Σώμα» και στο 14c την «Ειπανησιακή Γερουσία», διάκριση που προκύπτει από βαθιά ανάγνωση του ιστορικού πλαισίου, όχι από ρητή αναφορά στο κείμενο. Αυτό είναι ίσως το πιο σαφές παράδειγμα του τι σημαίνει να είσαι αρχειονόμος. Είναι το ότι, η γνώση του πλαισίου παραγωγής δεν εξάγεται από το κείμενο, αλλά φέρεται από τον επαγγελματία που το διαβάζει. Τα LLMs, όπως επισημαίνουν οι Jaillant και Caruto [JC22] και οι Wu et al. [WBM+23], αναπαράγουν ή αγνοούν το πλαίσιο παραγωγής, δεν μπορούν να το κατανοήσουν.

Τέλος εξετάζοντας και το τρίτο ερώτημα, τη συσχέτιση δηλαδή, μεγέθους και απόδοσης, τα αποτελέσματα ανατρέπουν τη γραμμική υπόθεση. Τα πολύ μικρά μοντέλα ($\leq 8B$ παραμέτρων) απέτυχαν σχεδόν καθολικά. Ωστόσο, το πιο εντυπωσιακό εύρημα είναι ότι το μέγεθος του μοντέλου δεν καθορίζει από μόνο του την ποιότητα της αρχειακής περιγραφής. Το καλύτερο μοντέλο δομικά, το gemma-3-12b-it, είναι μεσαίου μεγέθους (12B) και ξεπέρασε αρκετά μεγαλύτερα συστήματα, όπως είναι το Qwen3-235B που, παρά τις 235 δισεκατομμύρια παραμέτρους του, παρέλειψε το επίπεδο file και πρόσθεσε εσφαλμένο series, ή το Llama 3.3 70B που πρόσθεσε αδικαιολόγητο ενδιάμεσο επίπεδο. Αυτό υποδεικνύει ότι η ποιότητα του instruction tuning, η σταθερότητα της εξόδου και η πιστή τήρηση των κανόνων του prompt ίσως είναι πιο καθοριστικά από τον αριθμό παραμέτρων για εξειδικευμένες εργασίες όπως η αρχειακή περιγραφή. Τα μεγαλύτερα μοντέλα (70B+) εμφάνισαν συνήθως πλουσιότερο πλαίσιο αλλά και μεγαλύτερη τάση για hallucinations, προσθήκη επιπέδων ή πληροφοριών που δεν τεκμηριώνονται στην πηγή. Η συνολική εικόνα που προκύπτει από όλο το πείραμα, είναι ότι τα LLMs ξέρουν και μπορούν να συμπληρώνουν πεδία αλλά δεν ξέρουν να σκέφτονται ιεραρχικά. Σε αυτό το συγκεκριμένο εύρημα, δεν πρόκειται για μικρό τεχνικό λάθος που θα λυθεί με καλύτερο prompt, αλλά για μία βαθύτερη ασυμφωνία ανάμεσα στον τρόπο που τα LLMs επεξεργάζονται το κείμενο και στον τρόπο που η αρχειακή επιστήμη οργανώνει την πληροφορία. Για φορείς με μεγάλες ανεπεξέργαστες συλλογές και περιορισμένους πόρους, ακριβώς την κατάσταση που οι Greene και Meissner [GM05] περιγράφουν ως

«ουρά» ακατέργαστου υλικού, αυτή η δυνατότητα έχει πρακτική αξία, καθώς προσφέρει έναν πρώτο σκελετό που ο αρχειονόμος μπορεί να ελέγξει και να συμπληρώσει, αντί να ξεκινά από μηδενική βάση. Αλλά η αξία αυτή είναι βοηθητική και δεν μπορεί να την αντικαταστήσει. Ο ρόλος τους είναι να μειώνουν τον όγκο της εργασίας στα πρώτα στάδια επεξεργασίας, να προτείνουν πεδία και τιμές που ο αρχειονόμος θα αξιολογήσει, και να επιταχύνουν τη διαδικασία ώστε περισσότερο υλικό να μπορεί να γίνεται προσβάσιμο γρηγορότερα [GWF24, MML+24]. Αυτό που δεν μπορούν να κάνουν είναι να σκεφτούν ιεραρχικά, να ερμηνεύσουν πλαίσιο που δεν αναφέρεται ρητά, ή να εφαρμόσουν αρχειακές συμβάσεις που δεν διδάχτηκαν ρητά.

5.2 Ανάγκη human validation

Η ανάγκη για ανθρώπινη επικύρωση δεν προκύπτει μόνο από τα αδύνατα σημεία των LLMs που χρησιμοποιήθηκαν στο πείραμα. Προκύπτει και από τη φύση της αρχειακής εργασίας ως τέτοιας. Ακόμα και αν τα μοντέλα βελτιωθούν σημαντικά, υπάρχουν πτυχές της αρχειακής περιγραφής που εκ φύσεως απαιτούν ανθρώπινη κρίση, και αυτό δεν είναι αποκλειστικό χαρακτηριστικό των LLMs. Οι Zoutsou κ.ά. [ZSP+24] διαπιστώνουν το ίδιο και σε εφαρμογές θεματικής μοντελοποίησης σε συλλογές κειμένων πολιτιστικής κληρονομιάς, ακόμα και όταν ο αλγόριθμος αποδίδει καλά, η ανθρώπινη επικύρωση των αποτελεσμάτων παραμένει απαραίτητη.

Η πρώτη από τις λόγους που καθιστούν αυτή την επικύρωση αναντικατάστατη είναι η αρχειακή δεοντολογία. Η Gilliland [Gil08] τονίζει ότι η περιγραφή είναι πράξη ερμηνείας που πρέπει να ισορροπεί ανάμεσα στην πληρότητα και στην πιστότητα στο πρωτογενές τεκμήριο. Ένα μοντέλο που «συμπληρώνει» ένα κενό πεδίο με λογική εκτίμηση, αντί να δηλώσει ότι η πληροφορία απουσιάζει, δεν κάνει απλώς τεχνικό λάθος, αλλοιώνει τη μαρτυρία του τεκμηρίου. Το πρόβλημα των hallucinations [RSD23] στα LLMs δεν είναι επομένως μόνο ζήτημα ακρίβειας, αλλά ζήτημα δεοντολογικής αρχής που η αρχειακή επιστήμη έχει ορίσει με σαφήνεια εδώ και δεκαετίες.

Εξίσου σημαντικό είναι το ζήτημα της λογοδοσίας. Αν μια αυτόματα παραγόμενη εγγραφή περιέχει σφάλμα που δεν εντοπίστηκε από τον αρχειονόμο, ποιος φέρει την ευθύνη; Οι Svärd κ.ά. [SGB+24] επισημαίνουν ότι το νομοθετικό πλαίσιο υστερεί σημαντικά έναντι της τεχνολογικής εξέλιξης, και ότι στις περισσότερες δικαιοδοσίες δεν υπάρχει ακόμα σαφής ρύθμιση για τη χρήση AI σε δημόσια αρχεία. Αυτό σημαίνει ότι η ευθύνη παραμένει στον επαγγελματία που ενέκρινε την έξοδο, κάτι που επιβεβαιώνει, και όχι αναιρεί, την ανάγκη για συστηματικό ανθρώπινο έλεγχο.

Παράλληλα, υπάρχει και το ζήτημα της ερμηνευσιμότητας. Οι Toth et al. [TAP25] υπογραμμίζουν ότι ακόμα και όταν ένα LLM παράγει τεχνικά σωστά αποτελέσματα, αν ο αρχειονόμος δεν μπορεί να κατανοήσει γιατί το μοντέλο κατέληξε σε μια συγκεκριμένη εγγραφή, δεν μπορεί να την αξιολογήσει με εμπιστοσύνη. Αυτό είναι ιδιαίτερα κρίσιμο στον αρχειακό χώρο, όπου κάθε απόφαση περιγραφής επηρεάζει την πρόσβαση, τη διαλειτουργικότητα και την ιστορική ερμηνεία [ATO20].

Από πρακτική σκοπιά, τα ευρήματα του πειράματος δείχνουν ότι ο ανθρώπινος έλεγχος δεν είναι ενιαίος, διαφοροποιείται ανάλογα με τον τύπο σφάλματος. Τα συστηματικά σφάλματα, όπως η ομοιόμορφη παράλειψη ενός πεδίου ή η σύγχυση ιεραρχικών επιπέδων, μπορούν να εντοπιστούν και να διορθωθούν μέσω βελτίωσης του prompt ή αυτόματου ελέγχου δομής. Τα τυχαία σφάλματα, όπως η εφεύρεση μιας τιμής σε ένα πεδίο, απαιτούν ανθρώπινο έλεγχο σε κάθε εγγραφή χωριστά [TSKK19]. Τα σφάλματα ερμηνείας πλαισίου, τέλος, όπως η λανθασμένη απόδοση παραγωγού ή ο ελλιπής τίτλος, δεν εντοπίζονται παρά μόνο από έναν ειδικό που γνωρίζει το υλικό.

Η εισαγωγή LLMs στην αρχειακή εργασία δεν μειώνει απαραίτητα την ανάγκη για εξειδικευμένο προσωπικό. Αυτό που αλλάζει είναι ο τρόπος που αυτό το προσωπικό εργάζεται. Η ανθρώπινη κρίση επικεντρώνεται εκεί που είναι πραγματικά αναντικατάστατη. Αυτό είναι ακριβώς το μοντέλο συνεπιμέλειας ανθρώπου-μηχανής που υποστηρίζουν οι Colavizza κ.ά. [CBJ+21] και οι Griffin κ.ά. [GWF24], το σύστημα προτείνει, ο αρχειονόμος αποφασίζει.

5.3 Προοπτικές για hybrid archival AI systems

Τα ευρήματα της παρούσας εργασίας συνηγορούν υπέρ μιας υβριδικής προσέγγισης, όχι ως συμβιβασμός, αλλά ως η πιο ρεαλιστική και αξιόπιστη κατεύθυνση για την εφαρμογή AI στην αρχειακή περιγραφή.

Ένα ώριμο hybrid archival AI system δεν αποτελείται μόνο από ένα LLM. Περιλαμβάνει, όπως προτείνουν οι Groppe κ.ά. [GMW+25], συνδυασμό πολλαπλών μοντέλων που αλληλοελέγχουν τα αποτελέσματά τους, μηχανισμούς validation που ελέγχουν τη συμμόρφωση με το σχήμα ISAD(G), κανόνες μετα-επεξεργασίας που εντοπίζουν δομικές ασυνέπειες, και ανθρώπινη επίβλεψη στα σημεία όπου η αβεβαιότητα είναι υψηλή. Αυτή η αρχιτεκτονική δεν είναι πολυπλοκότερη από το απαραίτητο, αλλά είναι η αναγκαία απάντηση στην πολυπλοκότητα του αρχειακού έργου.

Πέρα όμως από την αρχιτεκτονική του συστήματος, τα ευρήματα του πειράματος αναδεικνύουν συγκεκριμένες κατευθύνσεις για μελλοντική έρευνα που προκύπτουν άμεσα από τα όριά του.

Η πρώτη αφορά τη βελτίωση του prompt design. Η παρούσα εργασία χρησιμοποίησε ένα prompt έξι κανόνων σχεδιασμένο για να περιορίσει hallucinations και να εξασφαλίσει συμμόρφωση με το πρότυπο. Τα αποτελέσματα έδειξαν ότι το prompt αυτό ήταν επαρκές για βασική συμμόρφωση αλλά ανεπαρκές για την επίλυση του κεντρικού προβλήματος της ιεραρχικής κατανόησης. Μια φυσική κατεύθυνση μελλοντικής έρευνας είναι η ανάπτυξη πιο εξειδικευμένων προσεγγίσεων prompting, όπως, few-shot παραδείγματα με γνήσιες αρχειακές εγγραφές που δείχνουν τη σύμβαση «id στο Αρχείο», chain-of-thought prompting που αναγκάζει το μοντέλο να συλλογιστεί πριν συμπληρώσει ένα πεδίο, ή structured prompts που παρέχουν το πρότυπο ISAD(G) ως σχήμα αναφοράς αντί ως απλή οδηγία. Η εργασία δεν εξέτασε συστηματικά αν διαφορετικές στρατηγικές prompting θα μπορούσαν να οδηγήσουν σε σημαντικά βελτιωμένα αποτελέσματα και αυτό παραμένει ανοιχτό ερευνητικό ερώτημα.

Μια δεύτερη γραμμή στην έρευνα, η οποία είναι στενά συνδεδεμένη με την πρώτη, αφορά τη χρήση μοντέλων που διαθέτουν δυνατότητες ρητού συλλογισμού (reasoning LLMs). Τα μοντέλα που δοκιμάστηκαν στο πείραμα είναι γενικής χρήσης instruction-tuned LLMs. Μια νέα γενιά μοντέλων, όπως τα OpenAI o1/o3, DeepSeek R1 και Claude με extended thinking, ενσωματώνει μηχανισμούς ρητής σκέψης πριν από την απάντηση, επιτρέποντας έναν σύνθετο συλλογισμό που θεωρητικά θα ταίριαζε καλύτερα σε εργασίες όπως η ιεραρχική αρχειακή περιγραφή. Το ερώτημα αν τέτοια μοντέλα θα κατάφερναν να συλλάβουν τη λογική της ιεραρχικής κληρονομικότητας, να αντιληφθούν δηλαδή ότι ένα πεδίο στο επίπεδο item «κληρονομείται» από το επίπεδο fonds και δεν χρειάζεται επανάληψη, παραμένει ανοιχτό και αποτελεί ίσως την πιο υποσχόμενη κατεύθυνση άμεσης συνέχειας της παρούσας εργασίας.

Ένας τρίτος άξονας είναι η ανάπτυξη domain-specific μοντέλων εκπαιδευμένων σε αρχειακά δεδομένα. Το ArcGPT [ZHP+23] αποτελεί πρώτη προσπάθεια προς αυτή την κατεύθυνση, δείχνοντας ότι ένα μοντέλο εκπαιδευμένο σε αρχειακά corpora υπερέχει σε αρχειακές εργασίες έναντι γενικών LLMs. Στο πλαίσιο ελληνικών ιστορικών αρχείων, με την ιδιαιτερότητα της γλώσσας, της περιόδου και της αρχειακής παράδοσης, ένα τέτοιο μοντέλο θα απαιτούσε annotated datasets που σήμερα δεν υπάρχουν σε επαρκή κλίμακα. Η δημιουργία τους αποτελεί από μόνη της ένα σημαντικό ερευνητικό έργο, το οποίο θα μπορούσε να αξιοποιήσει μεθόδους αυτόματης ευρετηρίασης και θεματικής εξερεύνησης που έχουν αναπτυχθεί για συλλογές πολιτιστικής κληρονομιάς [LRT21, WBM+23].

Ένας τέταρτος ανοιχτός άξονας στην έρευνά μας αφορά την ιεραρχική κατανόηση σε αρχιτεκτονικό επίπεδο. Το κεντρικό έλλειμμα που αποκάλυψε το πείραμα, η αδυναμία των μοντέλων να σκεφτούν σε επίπεδα *fonds-series-file-item*, είναι ίσως το πιο απαιτητικό τεχνικό πρόβλημα για μελλοντική έρευνα. Η λύση πιθανόν δεν βρίσκεται μόνο στο *prompt engineering* ή σε νέες αρχιτεκτονικές συλλογιστικής, αλλά και σε συστήματα που ενσωματώνουν ρητά ιεραρχικές δομές στην αναπαράσταση γνώσης όπως για παράδειγμα, συνδυασμός LLMs με γραφήματα γνώσης (*knowledge graphs*) κωδικοποιημένα σε CIDOC-CRM [SKD+25] ή EAD.

Σημαντική είναι επίσης η διάσταση της αυθεντικότητας και της διαφάνειας. Οι Pacheco κ.ά. [PGF23] επισημαίνουν ότι η αυτόματη παραγωγή αρχειακών περιγραφών δημιουργεί νέες ανάγκες για μεταδεδομένα που τεκμηριώνουν πώς και από ποιον ή τι παράχθηκε μια εγγραφή. Ένα *hybrid* σύστημα πρέπει να αφήνει ορατό αποτόπωμα: ποια πεδία παρήγαγε το μοντέλο, ποια επικύρωσε ο αρχειονόμος, και πού παρέμβηκε. Χωρίς αυτή τη διαφάνεια, η εγγραφή δεν μπορεί να θεωρηθεί αξιόπιστη για μακροπρόθεσμη αρχειακή χρήση [TAP25, SGB+24].

Τέλος, αξίζει να επισημανθεί ένας περιορισμός της παρούσας μελέτης που ανοίγει δρόμο για μελλοντική έρευνα. Τα αποτελέσματα αφορούν αποκλειστικά το Αρχείο της Επτανήσου Πολιτείας, το οποίο είναι ένα συγκεκριμένο ιστορικό αρχείο, συγκεκριμένης γλώσσας και περιόδου. Όπως επισημαίνουν οι Tan κ.ά. [TSKK19], η δειγματοληπτική προκατάληψη σημαίνει ότι τα συμπεράσματα δεν γενικεύονται αυτόματα σε άλλους τύπους αρχειακού υλικού. Πώς αποδίδουν τα ίδια μοντέλα σε *born-digital* συλλογές, σε αρχεία διαφορετικών εποχών ή γλωσσών, ή σε φορείς με διαφορετικά επίπεδα τεκμηρίωσης, αυτά παραμένουν ανοιχτά ερωτήματα που η παρούσα εργασία δεν μπορεί να απαντήσει αλλά θέτει με σαφήνεια.

Αυτό που μπορεί να πει με βεβαιότητα είναι ότι η ερώτηση δεν είναι πλέον «μπορούν τα LLMs να βοηθήσουν στην αρχειακή περιγραφή;» Η απάντηση σε αυτή την πρώτη ερώτηση είναι ναι, υπό προϋποθέσεις. Η ερώτηση που πρέπει να απαντηθεί τώρα είναι «πώς σχεδιάζουμε συστήματα που αξιοποιούν αυτή τη δυνατότητα χωρίς να θυσιάζουν την αρχειακή ακεραιότητα;». Και η απάντηση σε αυτό δεν είναι καθαρά τεχνική, είναι επαγγελματική, θεσμική και δεοντολογική.

Βιβλιογραφία

- [Καβ23] Καββαδία, Α.: Οι Θεσμοί της Κεντρικής Διοίκησης στα Συντάγματα των Ιονίων Νήσων κατά την Επανάση Πολιτεία (1800–1803) και η Παραγωγή Αρχείων. Διδακτορική διατριβή (2023). <https://doi.org/10.12681/eadd/55516>, <http://hdl.handle.net/10442/hedi/55516> (κείμενο στα ελληνικά)
- [AA18] J. Azimjonov, J. Alikhanov. Rule Based Metadata Extraction Framework from Academic Articles, arXiv preprint arXiv:1807.09009, 2018
- [AAG25] Z. Alyafeai, M. S. Al-Shaibani, B. Ghanem. MOLE: Metadata Extraction and Validation in Scientific Papers Using LLMs, In Findings of the Association for Computational Linguistics: EMNLP 2025, pages 12236–12264, 2025
- [ARM22] L. Ardissono, G. E. Raptis, N. Mauro. *Special Issue on AI and HCI Methods and Techniques for Cultural Heritage Curation, Exploration and Fruition*. Applied Sciences, 12 (19): 10118, 2022. DOI: 10.3390/app121910118.
- [ATO20] W. Z. Alma’aitah, A. Z. Talib, M. A. Osman. *Opportunities and challenges in enhancing access to metadata of cultural heritage collections: a survey*, Artificial Intelligence Review 53: 3621–3646, 2020, Springer Nature. <https://doi.org/10.1007/s10462-019-09773-w>
- [Bag24] M. Bagchi. A Generative AI-driven Metadata Modelling Approach, Accepted for publication in Library Trends, Special Issue “Generative AI and Libraries”, Johns Hopkins University Press, 2024, <https://doi.org/10.48550/arXiv.2501.04008>
- [BDD+23] Bountouri, L., Damigos, M., Drakiou, M., Gergatsoulis, M. and Kalogeros, E. (2023). The Semantic Mapping of RiC-CM to CIDOC-CRM. In: Lecture Notes in Computer Science, σσ. 90–99. Springer Nature Singapore. International Conference on Asia-Pacific Digital Libraries (ICADL) 2023. https://doi.org/10.1007/978-981-99-8088-8_8
- [Bod23] J. Bodenhamer. *The Reliability and Usability of ChatGPT for Library Metadata*, Journal of Library Metadata, 23 (3–4): 123–137, 2023.
- [CBJ+21] G. Colavizza, T. Blanke, C. Jeurgens, J. Noordegraaf. *Archives and AI: An Overview of Current Debates and Future Perspectives*. Journal on Computing and Cultural Heritage, 15 (1): Article 4, 2021.

- [CMK+13] S. R. Choudhury, P. Mitra, A. Kirk, S. Szep, D. Pellegrino, S. Jones, C. L. Giles. Figure Metadata Extraction from Digital Documents, In Proceedings of the 12th International Conference on Document Analysis and Recognition (ICDAR 2013), pages 135–139, 2013. IEEE. <https://doi.org/10.1109/ICDAR.2013.34>
- [CRM24] CIDOC CRM Special Interest Group. *Volume A: Definition of the CIDOC Conceptual Reference Model. Version 7.1.3*. International Council of Museums (ICOM), February 2024. <https://cidoc-crm.org/Version/version-7.1.3>
- [DCAT-AP24] SEMIC (Semantic Interoperability Community), European Commission. *DCAT Application Profile for data portals in Europe (DCAT-AP), Version 3.0.0*. Publications Office of the European Union, 2024. <https://semiceu.github.io/DCAT-AP/releases/3.0.0/>
- [DCMI98] Dublin Core Metadata Initiative. *Dublin Core Metadata Element Set, Version 1.1*. DCMI Recommendation, 1998. <https://www.dublincore.org/specifications/dublin-core/dces/>
- [DDK+25] Dimitriou, Y., Damigos, M., Kalogeros, E. and Boueti, C. (2025). Describing Corfu Criminal Court Archives Using RiC-CM. *Moderna arhivistika*, 8(1), 93–112. <https://doi.org/10.54356/MA/2025/GMNE4635>.
- [GHK+12] M. Granitzer, M. Hristakeva, R. Knight, K. Jack, R. Kern. A Comparison of Layout-Based Bibliographic Metadata Extraction Techniques, In Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics (WIMS 2012), Craiova, Romania, pages 1–12, 2012, ACM, <https://doi.org/10.1145/2254129.2254154>
- [Gil08] A. J. Gilliland. *Setting the Stage*, In Introduction to Metadata 3.0, J. Paul Getty Trust, 2008, Available at: http://www.getty.edu/research/conducting_research/standards/infrometadata/
- [GM05] M. A. Greene, D. Meissner. More Product, Less Process: Revamping Traditional Archival Processing, *The American Archivist*, 68(2), 2005.
- [GMW+25] J. Groppe, A. Marquet, A. Walz, S. Groppe. Automated Archival Descriptions with Federated Intelligence of LLMs, arXiv preprint, 2025
- [GWF24] G. Griffin, E. Wennerström, A. Foka. *AI and Swedish Heritage Organisations: challenges and opportunities*, *AI & Society*, 39: 2359–2372, 2024. <https://doi.org/10.1007/s00146-023-01689-y>

- [HNB21] R. Hathurusinghe, I. Nejadgholi, M. Bolic. *A Privacy-Preserving Approach to Extraction of Personal Information through Automatic Annotation and Federated Learning*. In Proceedings of the Association for Computational Linguistics (ACL) Workshop / arXiv preprint arXiv:2105.09198, 2021.
- [HS97] S. Hochreiter, J. Schmidhuber. *Long Short-Term Memory*. *Neural Computation*, 9(8): 1735–1780, 1997.
- [ICA00] International Council on Archives (Ed.). *ISAD(G): General International Standard Archival Description*. 2nd edition, Ottawa, International Council on Archives, 2000. Adopted by the Committee on Descriptive Standards, Stockholm, Sweden, 19–22 September 1999
- [ICA04] International Council on Archives, Committee on Descriptive Standards. *ISAAR(CPF): International Standard Archival Authority Record for Corporate Bodies, Persons and Families, Second Edition*. Paris: ICA, 2004. <https://www.ica.org/en/isaar-cpf-international-standard-archival-authority-record-corporate-bodies-persons-and-families>
- [ICA07] International Council on Archives, Committee on Best Practices and Standards. *ISDF: International Standard for Describing Functions, First Edition*. Dresden: ICA, 2007. <https://www.ica.org/en/isdf-international-standard-describing-functions>
- [ICA23] International Council on Archives Expert Group on Archival Description (EGAD): *Records in Contexts Conceptual Model (RiC-CM)*, version 1.0 (2023), <https://www.ica.org/app/uploads/2023/12/RiC-CM-1.0.pdf>
- [ICA25] International Council on Archives Expert Group on Archival Description (EGAD): *Records in Contexts Ontology (RiC-O)*, version 1.1 (2025), <https://www.ica.org/standards/RiC/ontology/1.1>
- [ISO16] International Organization for Standardization. *ISO 15489-1:2016 Information and documentation – Records management – Part 1: Concepts and principles*, International Standard, 2016.
- [ISO17] International Organization for Standardization. *ISO 5127:2017 Information and documentation – Foundation and vocabulary*. International Standard, 2017.
- [JC22] L. Jaillant, A. Caputo. *Unlocking digital archives: cross-disciplinary perspectives on AI and born-digital data*, *AI & Society* 37: 823–835, 2022.
- [JJG+25] A. Jacyszyn, S. Jiang, G. A. Gesese, S. Hertling, T. Kerzenmacher, P. Nowack, S. Barthlott, E. Posthumus, H. Sack. *AI4DiTraRe: Towards*

- LLM-Based Information Extraction for Standardising Climate Research Repositories, First AAAI Bridge on Artificial Intelligence for Scholarly Communication (AISc), 2025. DOI: 10.5281/zenodo.14872358
- [LCD+20] D. Lin, J. Crabtree, I. Dillo, R. R. Downs, R. Edmunds, D. Giaretta, M. DeGiusti, H. L'Hours, W. Hugo, R. Jenkyns, V. Khodiyar, M. E. Martone, M. Mokrane, V. Navale, J. Petters, B. Sierman, D. V. Sokolova, M. Stockhause, J. Westbrook. *The TRUST Principles for Digital Repositories*. *Scientific Data*, 7:144, 2020.
- [LoC10] Library of Congress, Network Development and MARC Standards Office. *METS: Metadata Encoding and Transmission Standard, Primer and Reference Manual, Version 1.6*. Digital Library Federation / Library of Congress, 2010. <https://www.loc.gov/standards/mets/>
- [LoC18] Library of Congress, Network Development and MARC Standards Office. *Metadata Object Description Schema (MODS), Version 3.7*. Washington, D.C.: Library of Congress, 2018. <https://www.loc.gov/standards/mods/>
- [LoC23] Library of Congress. *EAD: Encoded Archival Description – Official Site*. Network Development and MARC Standards Office, Library of Congress, 2023. <https://www.loc.gov/ead/>
- [LoC99] Library of Congress, Network Development and MARC Standards Office. *MARC 21 Format for Bibliographic Data*. Washington, D.C.: Library of Congress, 1999– (συνεχώς ενημερωμένο). <https://www.loc.gov/marc/bibliographic/>
- [LRT21] M. Lorenzini, M. Rospoche, S. Tonelli. *Automatically evaluating the quality of textual descriptions in cultural heritage records*, *International Journal on Digital Libraries* 22: 217–231, 2021, Springer, <https://doi.org/10.1007/s00799-021-00302-1>
- [MAG+23] W. Mardiati, S. Aisyah, N. Grataridarga, R. Wulandari. *The Potential Use of Artificial Intelligence Technology in the Process of Collecting Metadata in Photo Archive Description Activities*. In D. V. Ferezagia et al. (Eds.), *Proceedings of the International Conference on Vocational Education Applied Science and Technology (ICVEAST 2023)*. *Advances in Social Science, Education and Humanities Research* 783, Atlantis Press, 2023. DOI: 10.2991/978-2-38476-132-6_78.
- [MML+24] S. Münster, F. Maiwald, I. di Lenardo, J. Henriksson, A. Isaac, M. M. Graf, C. Beck, J. Oomen. *Artificial Intelligence for Digital Heritage Innovation: Setting up a R&D Agenda for Europe*. *Heritage*, 7(2): 794–816, 2024, MDPI. <https://doi.org/10.3390/heritage7020038>

-
- [NR23] P. R. Nayaka, R. Ranjan. An Efficient Framework for Algorithmic Metadata Extraction over Scholarly Documents Using Deep Neural Networks, *SN Computer Science* 4:341, 2023, Springer Nature Singapore, <https://doi.org/10.1007/s42979-023-01776-3>
- [OE24] C. Ozogul, E. Ergen. Metadata Extraction of RFIs Using Natural Language Processing and Machine Learning Algorithms, In *Proceedings of the European Conference on Computing in Construction (ECCIC 2024)*, Chania, Greece, July 14–17, 2024.
- [OR24] G. Osti, E. R. Roke. *Collaborating for Change? Assessing Metadata Inclusivity in Digital Collections with Large Language Models (LLMs)*, In *2024 IEEE International Conference on Big Data (Big Data)*, 2024.
- [OWJ+22] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. L. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, J. Schulman, J. Hilton, F. Kelton, L. Miller, M. Simens, A. Askell, P. Welinder, P. Christiano, J. Leike, R. Lowe. *Training language models to follow instructions with human feedback*. *Advances in Neural Information Processing Systems (NeurIPS)*, 35: 27730–27744, 2022.
- [Pav22] G. Pavlidis. *AI trends in digital humanities research*, *Trends in Computer Science and Information Technology* 7(2): 26–34, 2022. <https://doi.org/10.17352/tcsit.000048>
- [PGF23] A. Pacheco, C. Guardado Da Silva, M. C. Vieira De Freitas. *A metadata model for authenticity in digital archival descriptions*. *Archival Science* 23: 629–673, 2023. <https://doi.org/10.1007/s10502-023-09422-w>
- [PRM15] PREMIS Editorial Committee. *PREMIS Data Dictionary for Preservation Metadata, Version 3.0*. Library of Congress, June 2015. <https://www.loc.gov/standards/premis/v3/>
- [RB24] P. Raval, H. Bhaidasna. A Review of Extracting Metadata from Scholarly Articles using Natural Language Processing (NLP), In *Proceedings of the 3rd International Conference on Automation, Computing and Renewable Systems (ICACRS 2024)*, pages 1355–1359, 2024, IEEE, <https://doi.org/10.1109/ICACRS62842.2024.10841656>
- [RSD23] V. Rawte, A. Sheth, A. Das. A Survey of Hallucination in “Large” Foundation Models, arXiv preprint arXiv:2309.05922, 2023. Available at: <https://arxiv.org/abs/2309.05922>
- [SCK+24] E. Shepherd, J. Cowan, J. J. Kaye, J. McLeod. *Practices and pain points in personal records*, *Archival Science*, 24: Article 3, 2024.
- [SGB+24] P. Svärd, E. Guerrero, T. Balogun, N. Saurombe, L. Jacobs, P. Henttonen. Local regulations for the use of artificial intelligence in the
-

- management of public records – a literature review, *Records Management Journal*, 2024, Emerald Publishing
- [Sil21] M. B. da Silva, F. C. C. Silva, R. C. F. Silva. *Metadados para la preservación digital de datos abiertos: un estudio de identificación*. *Biblios: Journal of Librarianship and Information Science*, 81: 1–22, 2021. <https://doi.org/10.5195/biblios.2021.793>
- [SKD+25] S. Stamou, E. Kalogeros, M. Damigos, M. Gergatsoulis. *Leveraging LLMs to Build Text Narratives from CIDOC CRM*. In: *Metadata and Semantic Research, Proceedings of the 19th International Conference (MTSR 2025)*, *Communications in Computer and Information Science*, Springer, Cham, 2025.
- [TSKK19] Tan, P.-N., Steinbach, M., Karpatne, A., & Kumar, V. (2019). *Introduction to Data Mining* (2nd ed.). Pearson Education Limited.
- [Und08] W. Underwood. *Automatic Metadata Extraction for Archival Description and Access*. In *Proceedings of the Society of American Archivists Research Forum*, San Francisco, CA, 2008.
- [W3C24] *Data Catalog Vocabulary (DCAT) – Version 3*. W3C Recommendation, 22 August 2024. <https://www.w3.org/TR/vocab-dcat-3/>
- [WBM+23] M. Wu, H. Brandhorst, M.-C. Marinescu, J. More Lopez, M. Hlava, J. Busch. *Automated metadata annotation: What is and is not possible with machine learning*, *Data Intelligence* 5(1): 122–138, 2023, MIT Press. https://doi.org/10.1162/dint_a_00162
- [YFA+25] W. Yang, R. Fu, M. B. Amin, B. Kang. *The Impact of Modern AI in Metadata Management*, *Human-Centric Intelligent Systems* 5: 323–350, 2025, Springer, <https://doi.org/10.1007/s44230-025-00106-5>
- [ZHP+23] Zhang, S., Hou, J., Peng, S., Li, Z., Hu, Q., and Wang, P. *ArcGPT: A Large Language Model Tailored for Real-world Archival Applications*, arXiv:2307.14852, 2023. <https://doi.org/10.48550/arXiv.2307.14852>
- [ZSP+24] K. Zoutsou, M. Sfakakis, L. Papachristopoulos, C. Papatheodorou, "Automated Topic Exploration in a Cultural Heritage Corpus", *Metadata and Semantic Research, Proceedings of the 17th International Conference (MTSR 2023)*, Milan, Italy, October 25 - 27, 2023, *Communications in Computer and Information Science (CCIS)*, No. 2048: Springer, Cham., 2024, pp. 229 - 240.
- [ZZL15] X. Zhang, J. Zhao, Y. LeCun. *Character-level Convolutional Networks for Text Classification*. In *Advances in Neural Information Processing Systems (NeurIPS)*, 28: 649–657, 2015.

Παράρτημα Α

Αρχειακή Περιγραφή βάσει δεδομένων και μεταδεδομένων Αρχείου Επτανήσου Πολιτείας

Στην περιφερειακή Υπηρεσία των Γενικών Αρχείων του Κράτους (ΓΑΚ) -Αρχεία Νομού Κέρκυρας (ΑΝΚ), εντός του Αρχείου Ιονίου Γερουσίας (ΑΙΓ) διασώζεται το Αρχείο της Επτανήσου Πολιτείας, το οποίο απαρτίζεται από 473 κιβώτια με φακέλους και τεκμήρια. Όλο το Αρχείο της Επτανήσου Πολιτείας είναι ταξινομημένο θεματικά και μέσα στο κάθε θέμα υπάρχει χρονολογική ταξινόμηση.

Η Επτανήσος Πολιτεία υπήρξε το πρώτο ημιανεξάρτητο κρατίδιο στον ελλαδικό χώρο. Τελούσε υπό την προστασία της Ρωσικής Αυτοκρατορίας και ήταν φόρου υποτελής στην Οθωμανική Αυτοκρατορία, ως εγγυητριών δυνάμεων. Είχε αυτοδιοίκηση, αποτελούμενη από Επτανήσιους πολίτες, οι οποίοι συγκροτούσαν όργανα άσκησης νομοθετικής, εκτελεστικής και δικαστικής εξουσίας του κρατιδίου. Ιδρύθηκε το 1800 με την υπογραφή της από 21.03.1800 Συνθήκης της Κωνσταντινούπολης μεταξύ των εκπροσώπων Τσάρου και Σουλτάνου. Κατά τη σύντομη διάρκεια της σύντομης αλλά καινοτόμου και πρωτοποριακής για τις περιστάσεις ζωής της ψηφίστηκαν και ίσχυσαν για έστω και για περιορισμένο χρόνο τα Συντάγματα του 1800, του 1803 και του 1806. Η Επτανήσος Πολιτεία καταργήθηκε στις 07.07.1807 με την υπογραφή της Συνθήκης του Τιλσίτ μεταξύ Ρωσικής και Γαλλικής Αυτοκρατορίας, δυνάμει της οποίας τα Ιόνια Νησιά έπαψαν να υπάρχουν ως συνταγματική κρατική οντότητα και αποτέλεσαν εφεξής ενσωματωμένη και υποταγμένη γαλλική επαρχία της Γαλλικής Αυτοκρατορίας.

Στο Κιβώτιο αρ. 2, φάκελος αρ. 8, στοιχείο αρ. 14 του ανωτέρω Αρχείου εντοπίζεται παλαιό κατάστιχο-βιβλίο, στο οποίο εμπεριέχονται τα τρία ακόλουθα πολύτιμα τεκμήρια: α) Η Συνθήκη της Κωνσταντινουπόλεως, β) το Σύνταγμα της Κωνσταντινουπόλεως (Βυζαντινό) και γ) η από 25.10/06.11.1800 Επιστολή της Επτανησιακής Γερουσίας προς την Υψηλή Πύλη. Το Κατάστιχο είναι χειρόγραφο χαρτώο, πρόχειρης βιβλιοδεσίας με σπάγκο, σκληρόδετο με εξώφυλλο, δερμάτινη ράχη και οπισθόφυλλο συνδυαστικού χρώματος κυανού και φαιού, διαστάσεων 22Χ35 εκατοστών, αποτελούμενο από 111 αριθμημένα φύλλα, με συνεχή αρίθμηση σελίδων με αραβικούς αριθμούς, από 3 έως 38. Το πρώτο φύλλο είναι άγραφο, χωρίς σελιδαρίθμηση.

Ακολουθούν 20 φύλλα με κείμενο. Από το τρίτο φύλλο αρχίζει σελιδαρίθμηση με αριθμό σελίδας 3, η οποία συνεχίζεται αδιακόπτως στο άνω άκρο recto και verso κάθε φύλλου και καταλήγει στο 20^ο φύλλο με αριθμό 38. Ακολουθούν επόμενα 90 φύλλα άγραφα, χωρίς αρίθμηση. Συντάκτης των τριών αυτών τεκμηρίων είναι το Νομοθετικό Σώμα της Επτανήσου Πολιτείας με συνταγματικές αρμοδιότητες.

Το 1804 το εν λόγω Δημόσιο Αρχείο στεγάστηκε στο Παλαιό Φρούριο Κέρκυρας, το 1818 στην Οικία Φλαμπουριάρη στο κέντρο της πόλης, το 1833 στη δεξιά πτέρυγα του παλατιού των Άγγλων Αρμοστών στην Πλατεία της πόλης, στα τέλη του 1866 στον ισόγειο χώρο του πρώην παλατιού του Βαίλου, σημερινού 2^ο και 3^ο δημοτικού Σχολείου Κέρκυρας), από το 1931 έως το 1993 στη δεξιά πτέρυγα των πρώην Βασιλικών Ανακτόρων (Γεωργίου και Μιχαήλ) στην πλατεία της πόλης, από το 1993 έως το 1998 στο πρώην Ορφανοτροφείο Κέρκυρας και σημερινό Μαράσλειο Δημαρχιακό Μέγαρο της πόλης επί της οδού Λεωφόρου Αλεξάνδρας και από το 1998 και εφεξής στην αριστερή πτέρυγα του κτηρίου, περιλαμβανομένου και όλου του υπογείου χώρου, των πρώην Αγγλικών Στρατώνων στο Παλαιό Φρούριο της Κέρκυρας.

Η ταξινόμηση είναι θεματική και μέσα στο κάθε θέμα υπάρχει χρονολογική ταξινόμηση. Η πρόσβαση είναι ελεύθερη, καθώς και η αναπαραγωγή, πάντα όμως σύμφωνα με τους όρους του Κανονισμού λειτουργίας των ΓΑΚ. Η γλώσσα και η γραφή των τεκμηρίων είναι κυρίως η ιταλική ως γλώσσα της τότε διοίκησης -αυτοδιοίκησης των Ιονίων Νήσων, καθώς και η ελληνική, η γαλλική και η οθωμανική. Ειδικά η γλώσσα του κατάστιχου-βιβλίου είναι η ιταλική.

Τα φύλλα του Αρχείου και του περιλαμβανόμενου σε αυτό κατάστιχου είναι υπόλευκης απόχρωσης και κάποια από αυτά παρουσιάζουν σκούρα στίγματα και μικρές οπές. Ειδικότερα λίγα από τα φύλλα του κατάστιχου στο κάτω μέρος του recto και verso φέρουν μικρές σχισμές και συγκρατούνται με λεπτό σπάγκο πάνω σε σταθερή στάχωση, αλλά έχουν μια τάση απόσχισης από τη στάχωση μάλλον εξαιτίας παλαιάς υγρασίας και δράσης εντόμων. Η γραφή είναι σταθερή, καλλιγραφική δια μαύρης μελάνης. Το Αρχείο και το περιλαμβανόμενο σε αυτό κατάστιχο σώζεται σε καλή υλική κατάσταση και τα γράμματα είναι σχετικά ευανάγνωστα.

Τα πρωτότυπα εντοπίζονται εντός του Αρχείου και του εμπεριεχομένου σε αυτό κατάστιχου. Ο ερευνητής μπορεί να αναζητήσει ψηφιακά αντίγραφα των συνταγματικών κειμένων του Αρχείου στο Ψηφιακό Αρχείο της Βιβλιοθήκης της Βουλής των Ελλήνων.

Επιπλέον διασώζονται σχετικά τεκμήρια στους κάτωθι δημόσιους φορείς:

Αρχείο Βιβλιοθήκης της Βουλής των Ελλήνων

Εθνική Βιβλιοθήκη της Ελλάδος, Τμήμα Χειρογράφων και Ομοιοτύπων.

Αρχείο Διεύθυνσης Ιστορίας του Υπουργείου Εξωτερικών

Σχετικά με την ίδρυση, τη λειτουργία, την κατάργηση της Επτανήσου Πολιτείας, και τα Συνταγματικά Κείμενα που παρήχθησαν: 1) το Ίδρυμα της Βουλής των Ελλήνων, το 2008 εξέδωσε στην Αθήνα τον Τόμο: «Συνταγματικά Κείμενα των Ιονίων Νήσων» σε σχεδιασμό και επιμέλεια της Αλίκης Νικηφόρου, τ. Αναπληρώτριας Καθηγήτριας του Ιονίου Πανεπιστημίου και τ. Διευθύντριας των Γ.Α.Κ. Αρχεία Νομού Κέρκυρας, στον οποίο υπάρχουν όλα τα Συνταγματικά Κείμενα της Επτανήσου Πολιτείας, καθώς και η μεταφράσεις τους στην ελληνική γλώσσα. 2) τα ΓΑΚ Αρχεία Νομού Κέρκυρας το 2001, εξέδωσαν το βιβλίο με τίτλο: Επτάνησος Πολιτεία (1800-1807), με επιμέλεια της ως άνω Α. Νικηφόρου και 3) Η Νομαρχιακή Αυτοδιοίκηση Κέρκυρας, σε συνεργασία με τα ΓΑΚ-Αρχεία Νομού Κέρκυρας, εξέδωσαν το βιβλίο: Επτάνησος Πολιτεία (1800-1807), με επιμέλεια της ως άνω ίδιας, στην Κέρκυρα το 2000.

Η σύνταξη του δελτίου αρχειακής περιγραφής σύμφωνα με το Διεθνές Πρότυπο Αρχειακής Περιγραφής General International Standard Archival Description (ISAD(G) Second Edition), Οττάβα 2000, έγινε από την Αρχιονόμο Αλεξάνδρα Καββαδία τον Ιούνιο του 2022.

Παράρτημα Β

Prompt used for generating fonds – and sub-unit-level ISAD(G)

The prompt below is designed to generate ISAD(G)-compliant archival descriptions based exclusively on archivist-provided source text. It instructs the LLM not to use any outside knowledge and not to guess or fill in missing information. If an ISAD(G) element is not clearly stated in the text, the model must write N/A (not stated in source). The model must also create separate ISAD(G) blocks for the fonds and for each sub-unit that is explicitly described in the input text.

ROLE: Archival description cataloguer.

SOURCE RULE: Use ONLY the text in SOURCE DESCRIPTION. No outside knowledge. No guessing.

FILL RULE: For any ISAD(G) element not explicitly stated, write: N/A (not stated in source).

AMBIGUITY RULE: If the source conflicts/unclear, copy the values exactly as written. Do not explain or choose.

UNITS RULE: Create one ISAD(G) block for: (1) Fonds, and (2) every subunit explicitly described (series/sub-series/level/item/part). Do NOT invent units or hierarchy. Only use hierarchy if explicitly indicated (e.g., Series, numbering, contains, part of).

COPY RULE: Field values must be copied verbatim from the source (you may remove bullet markers only).

OUTPUT: No intro, no summary, no commentary. One block per unit, in source order. For EACH unit output exactly these elements in this order: 1. Identity: - Reference code, - Title, - Date(s), - Level of description, - Extent and medium. 2. Context: - Name of creator(s), - Administrative/Biographical history, - Archival history, - Immediate source of acquisition or transfer. 3. Content and structure: - Scope and content, - Appraisal, destruction and scheduling information, - Accruals, - System of arrangement. 4. Access and use: - Conditions governing access, - Conditions governing reproduction, -

Language/scripts of material, - Physical characteristics and technical requirements, - Finding aids. 5. Allied materials: - Existence and location of originals, - Existence and location of copies, - Related units of description, - Publication note. 6. Notes: - Note(s). 7. Description control: - Archivist's note, - Rules or conventions, - Date(s) of description

SOURCE DESCRIPTION: Archivist's Source Description (Input Text)

Παράρτημα Γ

Ευρετήριο Όρων

Ελληνικός Όρος	Αγγλικός Όρος
Αλυσίδα συλλογισμού	Chain-of-thought
Αναγνώριση Ονοματικών Οντοτήτων	Named Entity Recognition (NER)
Αναπαραγωγιμότητα	Precision
Ανθρώπινη επικύρωση	Human validation
Αξιολόγηση από ειδικό	Expert evaluation
Αρχείο	Archives
Αρχειακή ακεραιότητα	Archival integrity
Αρχειακή περιγραφή	Archival description
Αρχειακό ιστορικό	Archival history
Αρχειακός καταλογογράφος	Archival description cataloguer
Αρχειονόμος	Archivist
Αυτόματη εξαγωγή μεταδεδομένων	Automatic metadata extraction
Αυτόματη ταξινόμηση κειμένου	Automatic text classification

Ελληνικός Όρος	Αγγλικός Όρος
Ακρίβεια	Accuracy
Βαθιά μάθηση	Deep learning
Γραφήματα γνώσης	Knowledge graphs
Δειγματοληπτική προκατάληψη	Sampling bias
Δεοντολογία	Ethics
Διαλειτουργικότητα	Interoperability
Διοικητική ιστορία	Administrative/biographical history
Δομημένα δεδομένα	Structured data
Ελλείποντα δεδομένα	Missing values
Ελεύθερο κείμενο	Free text
Εμφωλευμένη δομή	Nested structure
Επεξεργασία Φυσικής Γλώσσας	Natural Language Processing (NLP)
Εξαγωγή βιβλιογραφικών μεταδεδομένων	Bibliographic metadata extraction
Εγγραφή αναφοράς / Πρότυπο εγγραφής	Gold standard record
Επίπεδο αρχείου	Fonds level
Επίπεδο σειράς	Series level

Ελληνικός Όρος	Αγγλικός Όρος
Επίπεδο φακέλου	File level
Επίπεδο τεκμηρίου	Item level
Εποπτευόμενη μάθηση	Supervised learning
Ερμηνευσιμότητα	Explainability
Ευρετηρίαση	Indexing
Θόρυβος	Noise
Ιεραρχική αρχειακή δομή	Hierarchical archival structure
Ιεραρχική περιγραφή	Hierarchical description
Ικανότητα παρακολούθησης οδηγιών	Instruction following
Κανόνας αντιγραφής	Copy rule
Κανόνας ασάφειας	Ambiguity rule
Κανόνας μονάδων	Units rule
Κανόνας πηγής	Source rule
Κανόνας συμπλήρωσης	Fill rule
Κατά λέξη αντιγραφή	Verbatim copy
Κωδικοποίηση κανόνων	Rule-based encoding

Ελληνικός Όρος	Αγγλικός Όρος
Λεπτομερής προσαρμογή	Fine-tuning
Λογοδοσία	Accountability
Μεγάλο Γλωσσικό Μοντέλο	Large Language Model (LLM)
Μεροληψία	Bias
Μεταδεδομένα	Metadata
Μεταδεδομένα διατήρησης	Preservation metadata
Μεταδεδομένα διοίκησης	Administrative metadata
Μεταδεδομένα δομής	Structural metadata
Μεταδεδομένα περιγραφής	Descriptive metadata
Μεταδεδομένα χρήσης	Use metadata
Μεθοδολογική ροή	Pipeline
Μηδενικής επίδειξης ή χωρίς παραδείγματα	Zero-shot prompting
Μηχανική μάθηση	Machine learning (ML)
Μηχανική προτροπών	Prompt engineering
Μοντέλο ανοιχτού κώδικα	Open-source model
Νευρωνικό δίκτυο	Neural network

Ελληνικός Όρος	Αγγλικός Όρος
Οδηγίες διαμόρφωσης εισόδου	Prompt engineering
Παραγωγός / Δημιουργός	Creator
Παραίωση / Ψευδαισθηση μοντέλου	Hallucination
Παράλληλη επεξεργασία	Parallel processing
Περιεχόμενο και δομή αρχείου	Archival content and structure
Πιστότητα στην πηγή	Source fidelity
Πλαίσιο παραγωγής	Provenance context
Πρόβλεψη επόμενου κειμένου	Next token prediction
Προεκπαίδευση	Pre-training
Προκατάληψη	Bias
Προτροπή με λίγα παραδείγματα	Few-shot prompting
Πρότυπο μεταδεδομένων	Metadata standard
Ροή εργασίας	Workflow
Ρόλος μοντέλου	Role prompting
Σημασιολογική ομοιότητα	Semantic similarity
Σφάλμα μέτρησης	Measurement error

Ελληνικός Όρος	Αγγλικός Όρος
Σφάλμα συλλογής δεδομένων	Data collection error
Σχεδιασμός προτροπών	Prompt design
Συστηματικό σφάλμα	Systematic error
Τεκμήριο	Record
Τεχνητή Νοημοσύνη	Artificial Intelligence (AI)
Τυχαίο σφάλμα	Random error
Υβριδικό σύστημα	Hybrid system
Υπολογιστικές αναπαραστάσεις	Embeddings
Χρυσή εγγραφή	Gold standard
Ψηφιακά γεννημένο υλικό	Born-digital material
Ψηφιακός μετασχηματισμός	Digital transformation