Ευρωπαϊκή Ένωση
Ευρωπαϊκό Ταμείο Περιφερειακής Ανάπτυξης
ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΥΠΟΥΡΓΕΙΟ ΑΝΑΠΤΥΞΗΣ ΚΑΙ ΕΠΕΝΔΥΣΕΩΝ
ΕΙΔΙΚΗ ΓΡΑΜΜΑΤΕΙΑ ΔΙΑΧΕΙΡΙΣΗΣ ΠΡΟΓΡΑΜΜΑΤΩΝ ΕΤΠΑ & ΤΣ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ ΕΠΑνΕΚ
ΕΠΑνΕΚ 2014-2020
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΑΝΤΑΓΩΝΙΣΤΙΚΟΤΗΤΑ ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΤΗΤΑ ΚΑΙΝΟΤΟΜΙΑ
ΕΣΠΑ 2014-2020
ανάπτυξη - εργασία - αλληλεγγύη
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

# Μεταδεδομένα: Από την συλλογή μέχρι την δημοσιοποίηση τους

Κωνσταντίνος Κραββαρίτης

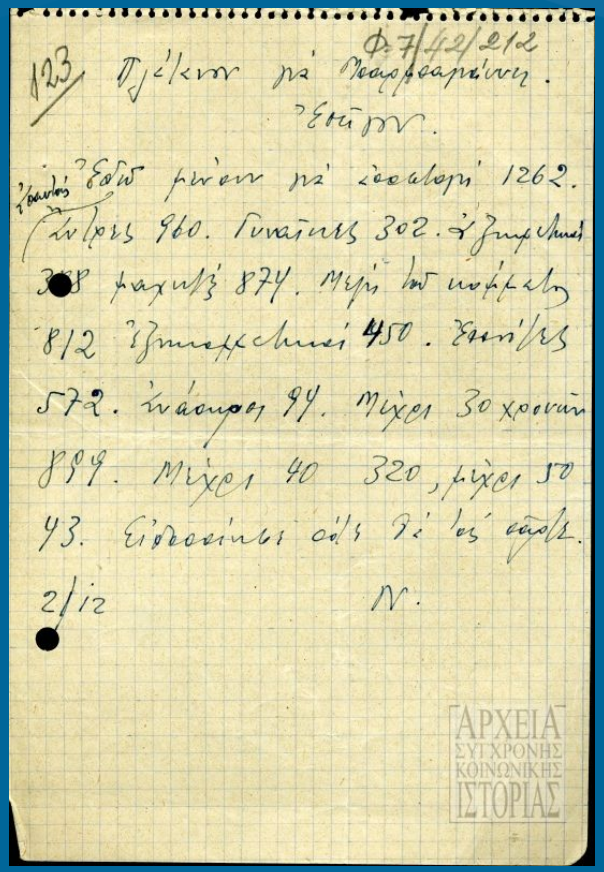
Μονάδα Ψηφιακής Επιμέλειας
Ινστιτούτο Πληροφοριακών Συστημάτων
Ερευνητικό Κέντρο «Αθηνά»

και

Εργαστήριο Επιχειρηματικής Αναλυτικής (BALab)
Τμήμα Διοικητικής Επιστήμης και Τεχνολογίας
Οικονομικό Πανεπιστήμιο Αθηνών

(k.kravvaritis@dcu.gr)

# Πριν τα μεταδεδομένα



Φ.7/42/212

123 Πλέκτρον πρό Παναγιαρέτου.
Ἐτῶν.

Ἐδῶ μένουν πρό ἐπαναστάσεως 1262.
Ἄνδρες 960. Γυναῖκες 302. Ἐξωμοσμένοι
308 μαχητές 874. Μερίς τῶν κατοίκων
812 Ἐξωμοσμένοι 450. Ἐπιστήμες
572. Ἀνάστροφοι 94. Μέχρι 30 χρονῶν
899. Μέχρι 40 320, μέχρι 50
43. Εἰδοποιήσεις αἱ τῶν ἀφῶν.
2/12 Ν.

ΑΡΧΕΙΑ ΣΥΓΧΡΟΝΗΣ ΚΟΙΝΩΝΙΚΗΣ ΙΣΤΟΡΙΑΣ

# Αρχείο Μεταδεδομένων


```
▼<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:apl="https://apollonis-infrastructure.gr/apl/" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dcterms="http://purl.org/dc/terms/"
 xmlns:edm="http://www.europeana.eu/schemas/edm/" xmlns:ore="http://www.openarchives.org/ore/terms/" xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
 xmlns:skos="http://www.w3.org/2004/02/skos/core#">
 ▼<edm:ProvidedCHO rdf:about="https://apollonis-infrastructure.gr/apl/ASKI_84681">
    <dc:identifier>ASKI_84681</dc:identifier>
    <dc:title>Πλάτανον για Μπαρμπαγιάννη</dc:title>
    <dc:language xml:lang="el">Ελληνική</dc:language>
    <dc:language>el</dc:language>
    <rdfs:comment xml:lang="el">Αριθμός αντιτύπου: 1</rdfs:comment>
    <rdfs:comment xml:lang="el">Αριθμός τεκμηρίου: 212</rdfs:comment>
    <dcterms:created>2/12/1949</dcterms:created>
    <apl:archive xml:lang="el">Αρχειο ΑΣΚΙ</apl:archive>
    <dc:type xml:lang="el">Αρχειακό τεκμήριο</dc:type>
    <dc:type xml:lang="el">Ραδιοτηλεγράφημα</dc:type>
    <rdfs:comment xml:lang="el">Χωρίς διαβάθμιση</rdfs:comment>
    <dcterms:extent>1</dcterms:extent>
    <dc:source>ΔΗΜΟΚΡΑΤΙΚΟΣ ΣΤΡΑΤΟΣ</dc:source>
    <dc:description xml:lang="el">001.07.042.00212</dc:description>
    <dcterms:medium>Χειρόγραφο</dcterms:medium>
    <apl:folder xml:lang="el">155</apl:folder>
    <dcterms:extent> </dcterms:extent>
    <apl:author>Ν.</apl:author>
    <apl:receiver>Πλάτανος</apl:receiver>
    <apl:receiver>Μπαρμπαγιάννης</apl:receiver>
    <dc:subject xml:lang="el">Κ.Κ.Ε / Κεντρική Επιτροπή / Πολιτικό Γραφείο</dc:subject>
    <dc:subject xml:lang="el">Δ.Σ.Ε. (Δημοκρατικός Στρατός Ελλάδας)</dc:subject>
    <dc:subject xml:lang="el">Κ.Κ.Ε. (Κομμουνιστικό Κόμμα Ελλάδας)</dc:subject>
    <dc:subject xml:lang="el">Δ.Σ.Ε. / Γενικό Αρχηγείο</dc:subject>
    <dc:subject xml:lang="el">Ρουμανία / Μηχανισμός του Ραδιοσταθμού Ελεύθερη Ελλάδα</dc:subject>
    <dc:subject xml:lang="el">Εμφύλιος πόλεμος</dc:subject>
    <dc:subject xml:lang="el">Εμφύλιος πόλεμος / λήξη</dc:subject>
    <dc:subject xml:lang="el">Πολιτικοί πρόσφυγες</dc:subject>
  </edm:ProvidedCHO>
 ▼<ore:Aggregation rdf:about="https://apollonis-infrastructure.gr/apl/ASKI_84681">
    <edm:aggregatedCHO rdf:resource="https://apollonis-infrastructure.gr/apl/ASKI_84681"/>
    <edm:provider xml:lang="el">Αρχεία Σύγχρονης Κοινωνικής Ιστορίας</edm:provider>
    <edm:dataProvider>ΑΣΚΙ</edm:dataProvider>
    <edm:dataset xml:lang="el">Δεκαετία 1940</edm:dataset>
  </ore:Aggregation>
</rdf:RDF>
```


Μονάδα Ψηφιακής Επιμέλειας
ΙΠΣΥ - Ερευνητικό Κέντρο Αθηνά

απολλωνίς

# Σχήμα Μεταδεδομένων

- Κανόνες τεκμηρίωσης πληροφορίας
- Προτυποποίηση και συνέπεια πληροφορίας
- Διαλειτουργικότητα

- Παραδείγματα σχημάτων μεταδεδομένων
  - EDM (Europeana Data Model)
  - DC (Dublin Core)
  - EAD (Encoded Archival Description)

# EDM

- Το EDM αναπτύχθηκε στα πλαίσια συνεργατικών διαδικασιών από διάφορες κοινότητες που περιλαμβάνουν ειδικούς από βιβλιοθήκες, αρχεία και μουσεία
- Ο αρχικός στόχος ήταν η εύκολη προσθήκη αντικειμένων στη Europeana λαμβάνοντας υπόψη ιδιαίτερα χαρακτηριστικά συγκεκριμένων κοινοτήτων
- Στόχος να αποτελέσει ένα μέσο ολοκλήρωσης για συλλογή, διασύνδεση και εμπλουτισμό των περιγραφών που παρέχονται στη Europeana από παρόχους περιεχομένου
- Η Europeana και συνεργάτες της μπορούν να ορίζουν προφίλ του EDM ώστε να καλύπτονται συγκεκριμένες λειτουργικότητες

# Συλλογή Μεταδεδομένων


```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:apl="https://apollonis-infrastructure.gr/apl/" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dcterms="http://purl.org/dc/terms/"
 xmlns:edm="http://www.europeana.eu/schemas/edm/" xmlns:ore="http://www.openarchives.org/ore/terms/" xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
 xmlns:skos="http://www.w3.org/2004/02/skos/core#">
  <edm:ProvidedCHO rdf:about="https://apollonis-infrastructure.gr/apl/ASKI_84681">
    <dc:identifier>ASKI_84681</dc:identifier>
    <dc:title>Πλάτανον για Μπαρμπαγιάννη</dc:title>
    <dc:language xml:lang="el">Ελληνική</dc:language>
    <dc:language>el</dc:language>
    <rdfs:comment xml:lang="el">Αριθμός αντιτύπου: 1</rdfs:comment>
    <rdfs:comment xml:lang="el">Αριθμός τεκμηρίου: 212</rdfs:comment>
    <dcterms:created>2/12/1949</dcterms:created>
    <apl:archive xml:lang="el">Αρχείο ΑΣΚΙ</apl:archive>
    <dc:type xml:lang="el">Αρχειακό τεκμήριο</dc:type>
    <dc:type xml:lang="el">Ραδιοτηλεγράφημα</dc:type>
    <rdfs:comment xml:lang="el">Χωρίς διαβάθμιση</rdfs:comment>
    <dcterms:extent>1</dcterms:extent>
    <dc:source>ΔΗΜΟΚΡΑΤΙΚΟΣ ΣΤΡΑΤΟΣ</dc:source>
    <dc:description xml:lang="el">001.07.042.00212</dc:description>
    <dcterms:medium>Χειρόγραφο</dcterms:medium>
    <apl:folder xml:lang="el">155</apl:folder>
    <dcterms:extent> </dcterms:extent>
    <apl:author>Ν.</apl:author>
    <apl:receiver>Πλάτανος</apl:receiver>
    <apl:receiver>Μπαρμπαγιάννης</apl:receiver>
    <dc:subject xml:lang="el">Κ.Κ.Ε / Κεντρική Επιτροπή / Πολιτικό Γραφείο</dc:subject>
    <dc:subject xml:lang="el">Δ.Σ.Ε. (Δημοκρατικός Στρατός Ελλάδας)</dc:subject>
    <dc:subject xml:lang="el">Κ.Κ.Ε. (Κομμουνιστικό Κόμμα Ελλάδας)</dc:subject>
    <dc:subject xml:lang="el">Δ.Σ.Ε. / Γενικό Αρχηγείο</dc:subject>
    <dc:subject xml:lang="el">Ρουμανία / Μηχανισμός του Ραδιοσταθμού Ελεύθερη Ελλάδα</dc:subject>
    <dc:subject xml:lang="el">Εμφύλιος πόλεμος</dc:subject>
    <dc:subject xml:lang="el">Εμφύλιος πόλεμος / λήξη</dc:subject>
    <dc:subject xml:lang="el">Πολιτικοί πρόσφυγες</dc:subject>
  </edm:ProvidedCHO>
  <ore:Aggregation rdf:about="https://apollonis-infrastructure.gr/apl/ASKI_84681">
    <edm:aggregatedCHO rdf:resource="https://apollonis-infrastructure.gr/apl/ASKI_84681"/>
    <edm:provider xml:lang="el">Αρχεία Σύγχρονης Κοινωνικής Ιστορίας</edm:provider>
    <edm:dataProvider>ΑΣΚΙ</edm:dataProvider>
    <edm:dataset xml:lang="el">Δεκαετία 1940</edm:dataset>
  </ore:Aggregation>
</rdf:RDF>
```


Μονάδα Ψηφιακής Επιμέλειας
ΙΠΣΥ - Ερευνητικό Κέντρο Αθηνά

απολλωνίς

# Συλλογή Μεταδεδομένων

- Μία συλλογή μεταδεδομένων περιέχει πολλαπλά αρχεία μεταδεδομένων.
- Τα αρχεία μεταδεδομένων που περιέχει, περιγράφουν τα αντίστοιχα αντικείμενα μιας ήδη υπαρκτής συλλογής.
- Η συλλογή την οποία περιγράφουν τα μεταδεδομένα μπορεί να περιέχει πόρους οποιουδήποτε τύπου, όπως φυσικά αντικείμενα, ψηφιακά αντικείμενα, ψηφιακά υποκατάστατα, καταλόγους συλλογών κ.α. και μορφής π.χ. βιβλία, ρούχα, ηχητικά ντοκουμέντα κ.α.
- Ανάλογα με τα αντικείμενα που περιέχει μια συλλογή μπορεί να είναι: i) φυσική συλλογή, ii) ψηφιακή συλλογή, iii) υβριδική συλλογή

# Κύκλος ζωής περιεχομένου

**Παραγωγή/Συλλογή**

↓

**Οργάνωση, Καθαρισμός, Εμπλουτισμός**

↓

**Διάθεση/Δημοσιοποίηση**

↓

**Μακροχρόνια Διατήρηση**

# Πάροχοι περιεχομένου

- αρχεία, βιβλιοθήκες, μουσεία, οργανισμοί δημιουργικής βιομηχανίας κ.α.

- <u>προκλήσεις</u>
  - ψηφιοποίηση περιεχομένου και παραγωγή μεταδεδομένων
  - ανάπτυξη και συντήρηση ψηφιακών συλλογών
  - online πρόσβαση και παροχή υπηρεσιών
  - έλλειψη οργανωμένων IT τμημάτων
  - υψηλό κόστος ανάπτυξης, λειτουργίας και συντήρησης

# Ανάγκη για διαλειτουργικότητα μεταδεδομένων

- Ετερογενή περιβάλλοντα, κοινότητες χρηστών, τύποι πόρων, δεδομένα και σχήματα μεταδεδομένων
- Δεν μπορούν όλα τα συστήματα να υποστηρίξουν διαδικασίες όπως η φόρτωση, η αποθήκευση και η εξαγωγή για διαφορετικά σχήματα μεταδεδομένων
- Περιορισμοί σχημάτων μεταδεδομένων στην κατάλληλη αναπαράσταση της πληροφορίας
- Ανάγκη για κεντρική αναζήτηση σε συλλογές με διαφορετικά σχήματα μεταδεδομένων

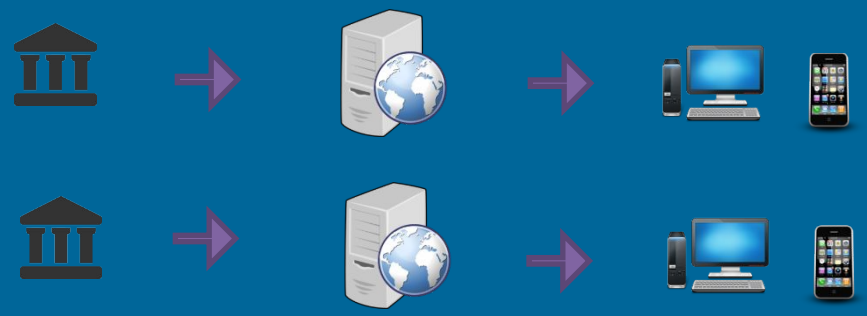
# Διαλειτουργικότητα μεταδεδομένων

- Η διαλειτουργικότητα μεταδεδομένων είναι η δυνατότητα δύο ή περισσότερων πληροφοριακών συστημάτων να ανταλλάξουν μεταδεδομένα με την ελάχιστη δυνατή απώλεια πληροφορίας
- Στόχος η επαναχρησιμοποίηση και ανταλλαγή μεταδεδομένων
- Κανόνες που αφορούν το περιεχόμενο και το βαθμό λεπτομέρειας της πληροφορίας
- Χρειάζεται συμφωνία για τα σχήματα μεταδεδομένων και τα πρωτόκολλα ανταλλαγής που θα χρησιμοποιηθούν

# Πρόκληση

- 1 πάροχος περιεχομένου – Μ χρήστες

- Ν πάροχοι περιεχομένου – Μ χρήστες



# Πρόκληση συσσωρευτή

Μονάδα Ψηφιακής Επιμέλειας
ΙΠΣΥ - Ερευνητικό Κέντρο Αθηνά

απολλωνίς

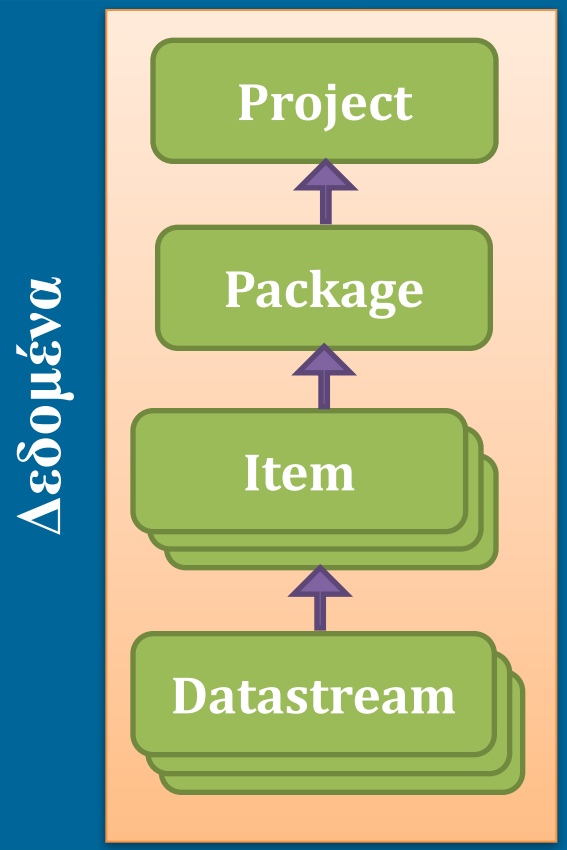
# Ο συσσωρευτής MORe

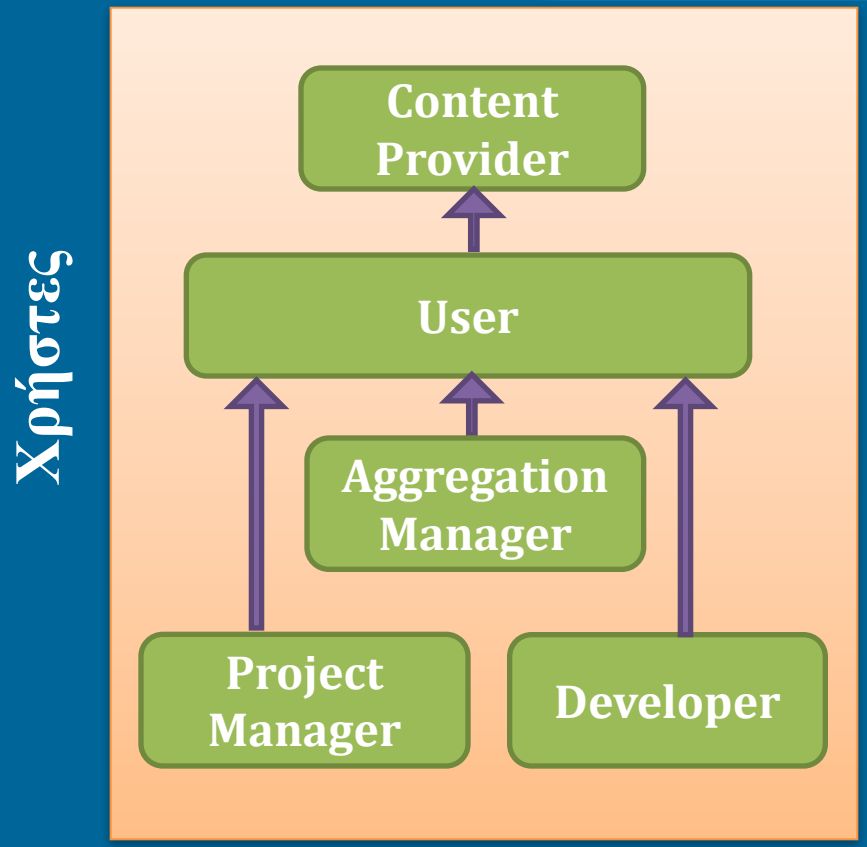
- Το Metadata & Object Repository (MORe) είναι ένα ευέλικτο, ισχυρό, επεκτάσιμο σύστημα για:
  - τη συσσώρευση μεγάλου όγκου μεταδεδομένων από πολλαπλές πηγές και πολλαπλά σχήματα
  - την επιμέλεια των μεταδεδομένων ορίζοντας ημιαυτόματες / και αυτόματες ροές εργασιών
  - τη δημοσιοποίηση επικυρωμένου ψηφιακού περιεχομένου σε πολλαπλούς στόχους

# Χαρακτηριστικά MORe

- Εύκολο στη χρήση: δεν απαιτεί εγκατάσταση, παρέχεται ως υπηρεσία στο νέφος (cloud), φιλική διεπαφή
- Ευέλικτο και επεκτάσιμο: αρθρωτή αρχιτεκτονική, ανάπτυξη / ενσωμάτωση εξωτερικών micro-services
- Κλιμακούμενο: κατανεμημένη αρχιτεκτονική επιτρέπει κλιμάκωση σε επίπεδο υπηρεσιών και αποθήκευσης
- Προσανατολισμένο στην επιμέλεια: επικύρωση περιεχομένου, μέτρηση ποιότητας, εμπλουτισμός

# Μοντέλα πληροφορίας

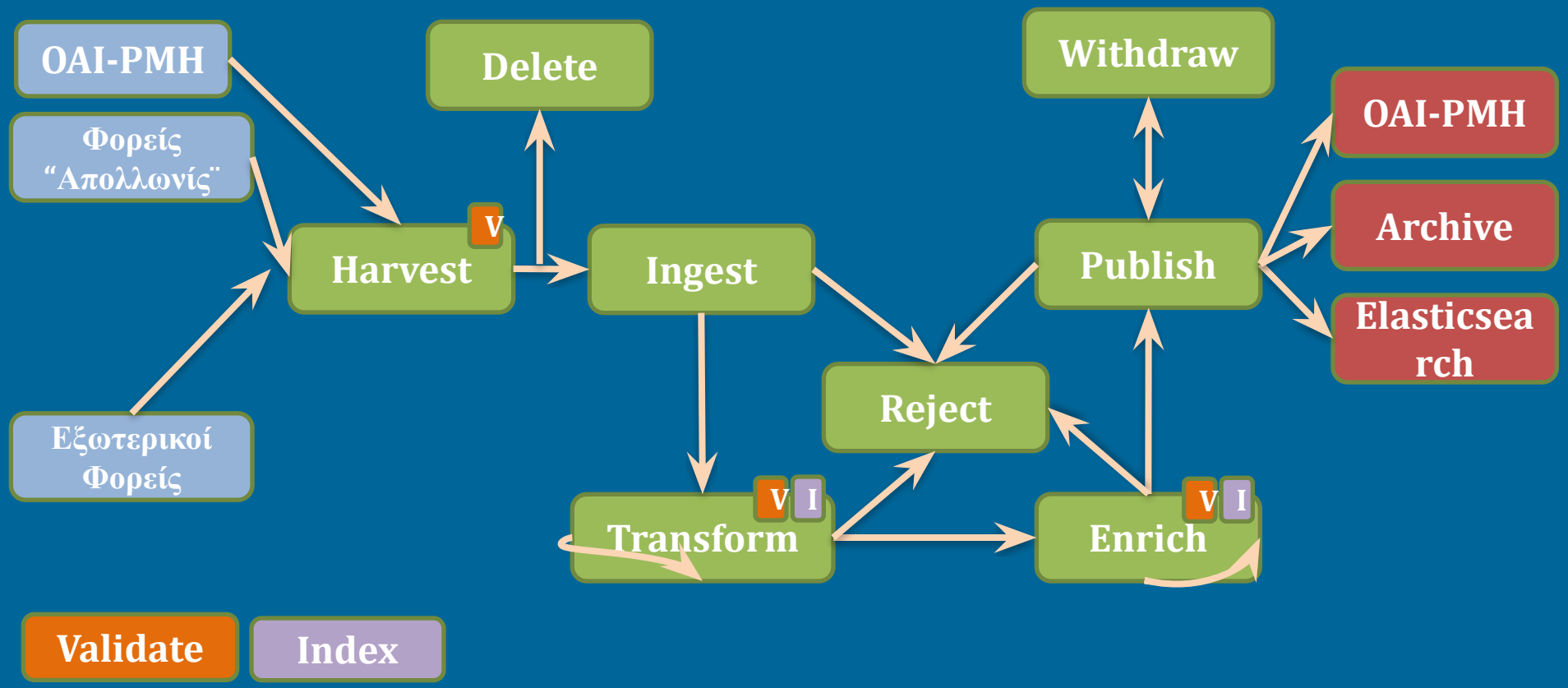


```
▼<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:apl="https://apollonis-infrastructure.g
 xmlns:edm="http://www.europeana.eu/schemas/edm/" xmlns:ore="http://www.openarchives.org/ore/terms/" xmlns:owl=
 xmlns:skos="http://www.w3.org/2004/02/skos/core#">
  ▼<edm:ProvidedCHO rdf:about="https://apollonis-infrastructure.gr/apl/EKPA_70127">
     <dc:identifier>EKPA_70127</dc:identifier>
     <apl:subfolder>/butterfly/backie/subfolder</apl:subfolder>
     <rdfs:comment>/lib/default/data/70119</rdfs:comment>
     <rdfs:comment xml:lang="el">Μη δημοσιευμένο στο διαδίκτυο</rdfs:comment>
     <dc:title>Άδειες (Καθηγητών και υπαλλήλων)</dc:title>
     <dc:description xml:lang="el">Έγγραφα σχετικά με άδειες διδακτικού και λοιπού προσωπικού.</dc:description>
     <dc:date>1/9/1944-31/8/1945</dc:date>
     <dcterms:extent xml:lang="el">6 σελίδες</dcterms:extent>
     <rdfs:comment xml:lang="el">Μεσαία κωδικοποίηση</rdfs:comment>
     <dc:description>142.2</dc:description>
     <dc:source>Ιστορικό Αρχείο ΕΚΠΑ</dc:source>
     <dc:subject xml:lang="el">1326989</dc:subject>
     <dc:subject xml:lang="el">300083</dc:subject>
     <dc:subject xml:lang="el">1327082</dc:subject>
     <dc:creator xml:lang="el">1057188</dc:creator>
     <dc:creator xml:lang="el">1326642</dc:creator>
     <dc:subject xml:lang="el">Ονόματα</dc:subject>
     <dc:subject xml:lang="el">Ιωάννης Πολίτης</dc:subject>
     <dc:subject xml:lang="el">Κλινικές</dc:subject>
     <dc:subject xml:lang="el">Πρύτανης</dc:subject>
     <dc:subject xml:lang="el">Εγκύκλιος</dc:subject>
     <dc:subject xml:lang="el">Προσωπικό</dc:subject>
     <dc:subject xml:lang="el">Υπάλληλοι</dc:subject>
     <dc:subject xml:lang="el">Βοηθητικό Προσωπικό</dc:subject>
     <dc:subject xml:lang="el">Υπουργείο θρησκευμάτων και Εθνικής Παιδείας</dc:subject>
     <dc:subject xml:lang="el">Νοσοκομεία</dc:subject>
     <dc:subject xml:lang="el">Αναρρωτικές Άδειες</dc:subject>
     <dc:subject xml:lang="el">θεραπευτήριο «Ευαγγελισμός»</dc:subject>
   </edm:ProvidedCHO>
  ▼<ore:Aggregation rdf:about="https://apollonis-infrastructure.gr/apl/EKPA_70127">
     <edm:aggregatedCHO rdf:resource="https://apollonis-infrastructure.gr/apl/EKPA_70127"/>
     <edm:provider xml:lang="el">Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών</edm:provider>
     <edm:dataProvider xml:lang="el">Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών</edm:dataProvider>
     <edm:dataset xml:lang="el">Δεκαετία 1940</edm:dataset>
   </ore:Aggregation>
 </rdf:RDF>
```


# Παράδειγμα ψηφιακού τεκμηρίου σε EDM

# Κύκλος ζωής συλλογής




Μονάδα Ψηφιακής Επιμέλειας
ΙΠΣΥ - Ερευνητικό Κέντρο Αθηνά

απολλωνίς

# Διαδικασία «Harvest»

- συγκομιδή περιεχομένου από πηγές μεταδεδομένων
  - OAI-PMH αποθετήρια
  - Φορείς “Απολλωνίς”
  - Εξωτερικοί Φορείς
- υποστήριξη πολλαπλών σχημάτων

# Διαδικασία «Validate»

- επικύρωση εισερχόμενης πληροφορίας
- ορισμός και εκτέλεση σχημάτων επικύρωσης
- υπηρεσίες επικύρωσης
  - δομή (structure) - linking
  - schema - schematron rules
- 2 περιπτώσεις χρήσης
  - pre-validation και post-validation

# Εντοπισμός προβλήματος


```
▼<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:apl="https://apollonis-infrastructure
 xmlns:edm="http://www.europeana.eu/schemas/edm/" xmlns:ore="http://www.openarchives.org/ore/terms/" xmlns:ow
 xmlns:skos="http://www.w3.org/2004/02/skos/core#">
 ▼<edm:ProvidedCHO rdf:about="https://apollonis-infrastructure.gr/apl/EKPA_70127">
    <dc:identifier>EKPA_70127</dc:identifier>
    <apl:subfolder>/butterfly/backie/subfolder</apl:subfolder>
    <rdfs:comment>/lib/default/data/70119</rdfs:comment>
    <rdfs:comment xml:lang="el">Μη δημοσιευμένο στο διαδίκτυο</rdfs:comment>
    <dc:description xml:lang="">Έγγραφα σχετικά με άδειες διδακτικού και λοιπού προσωπικού.</dc:description>
    <dcterms:extent xml:lang="el">6 σελίδες</dcterms:extent>
    <rdfs:comment xml:lang="el">Μεσαία κωδικοποίηση</rdfs:comment>
    <dc:description>142.2</dc:description>
    <dc:source>Ιστορικό Αρχείο ΕΚΠΑ</dc:source>
    <dc:subject xml:lang="el">1326989</dc:subject>
    <dc:subject xml:lang="el">300083</dc:subject>
    <dc:subject xml:lang="el">1327082</dc:subject>
    <dc:creator xml:lang="el">1057188</dc:creator>
    <dc:creator xml:lang="el">1326642</dc:creator>
    <dc:subject xml:lang="el">Ονόματα</dc:subject>
    <dc:subject xml:lang="el">Ιωάννης Πολίτης</dc:subject>
    <dc:subject xml:lang="el">Κλινικές</dc:subject>
    <dc:subject xml:lang="el">Πρύτανης</dc:subject>
    <dc:subject xml:lang="el">Εγκύκλιος</dc:subject>
    <dc:subject xml:lang="el">Προσωπικό</dc:subject>
    <dc:subject xml:lang="el">Υπάλληλοι</dc:subject>
    <dc:subject xml:lang="el">Βοηθητικό Προσωπικό</dc:subject>
    <dc:subject xml:lang="el">Υπουργείο θρησκευμάτων και Εθνικής Παιδείας</dc:subject>
    <dc:subject xml:lang="el">Νοσοκομεία</dc:subject>
    <dc:subject xml:lang="el">Αναρρωτικές Άδειες</dc:subject>
    <dc:subject xml:lang="el">Θεραπευτήριο «Ευαγγελισμός»</dc:subject>
    <edm:Place>
      <skos:prefLabel>37.98376, 23.72784</skos:prefLabel>
    </edm:Place>
  </edm:ProvidedCHO>
 ▼<ore:Aggregation rdf:about="https://apollonis-infrastructure.gr/apl/EKPA_70127">
    <edm:aggregatedCHO rdf:resource="https://apollonis-infrastructure.gr/apl/EKPA_70127"/>
    <edm:provider xml:lang="el">Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών</edm:provider>
    <edm:dataProvider xml:lang="el">Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών</edm:dataProvider>
    <edm:dataset xml:lang="el">Δεκαετία 1940</edm:dataset>
  </ore:Aggregation>
</rdf:RDF>
```


**Κενός προσδιοριστής γλώσσας**

**Μη κανονικοποιημένη χωρική πληροφορία**

Μονάδα Ψηφιακής Επιμέλειας
ΙΠΣΥ - Ερευνητικό Κέντρο Αθηνά

απολλωνίς

# Διαδικασία «Ingest»

- εισαγωγή περιεχομένου
- πολλαπλές τεχνολογίες αποθήκευσης
  - σύστημα αρχείων (filesystem)
  - βάση δεδομένων (Apache Cassandra)
  - αποθετήριο (Fedora-Commons)
- υποστήριξη πολλαπλών εκδόσεων (versions)
- κάθε ενέργεια επιμέλειας παράγει μια νέα έκδοση
- κάθε διαφορετική αναπαράσταση ενός αντικειμένου αποθηκεύεται ως διαφορετικό datastream

# Διαδικασία «Transform»

- μετατροπή της πληροφορίας ενός πακέτου σε ένα σχήμα στόχο (π.χ. Europeana Data Model)

- πολλαπλές δυνατότητες μετασχηματισμού
  - ανά σχήμα, ανά έργο, ανά πάροχο περιεχομένου
- οι χρήστες μπορούν να ενσωματώσουν δηλώσεις δικαιωμάτων (rights statements)

# Διαδικασία «Enrich»

- χρησιμοποιείται η έννοια του πλάνου εμπλουτισμού (enrichment plan) που περιλαμβάνει την εφαρμογή micro-services εμπλουτισμού σε συγκεκριμένη σειρά
- κάθε πλάνο εφαρμόζεται σε συγκεκριμένο σχήμα μεταδεδομένων
- σύνδεση με εξωτερικές υπηρεσίες εμπλουτισμού. Επιτρέπει σε προγραμματιστές και δημιουργικές βιομηχανίες να δημιουργήσουν τις δικές τους υπηρεσίες εμπλουτισμού και να τις ενσωματώσουν στο MORe

## Available enrichment micro-services

List of available enrichment micro-services. Click on the "More info" link on each micro-service to get more information.

OAI_DC | EDM | ACDM

**Language**
[EDM] Language Identification
More info

**GeoNorm**
[EDM] Geo Normalization
More info

**LoGeo**
[EDM] GeoLocation Enrichment
More info

**Geocode**
[EDM] Geonames Geocoding
More info

**Rev.Geocode**
[EDM] Geonames Rev. Geocoding
More info

**Vocabulary**
[EDM] Vocabulary
More info

**Vocab. match**
[EDM] Vocabulary Matching
More info

**Back Link**
[EDM] Background Link
More info

**DaiGazeteer**
[EDM] DaiGazeteer
More info

**Pelagios**
[EDM] Pelagios
More info

**GeoInvertion**
[EDM] GeoInvertion
More info

**CRS Transform**
[EDM] CRS Transform (Beta Version)
More info

# Παράδειγμα εφαρμογής ενός πλάνου εμπλουτισμού (enrichment plan)

# EDM εγγραφή

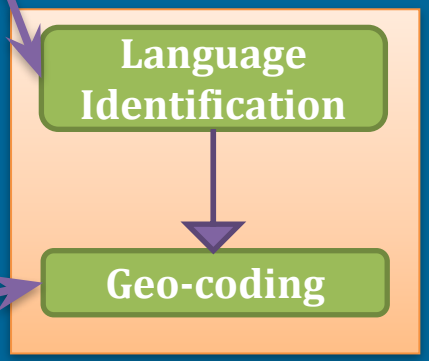

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:apl="https://apollonis-infrastructure
xmlns:edm="http://www.europeana.eu/schemas/edm/" xmlns:ore="http://www.openarchives.org/ore/terms/" xmlns:ow
xmlns:skos="http://www.w3.org/2004/02/skos/core#">
  <edm:ProvidedCHO rdf:about="https://apollonis-infrastructure.gr/apl/EKPA_70127">
    <dc:identifier>EKPA_70127</dc:identifier>
    <apl:subfolder>/butterfly/backie/subfolder</apl:subfolder>
    <rdfs:comment>/lib/default/data/70119</rdfs:comment>
    <rdfs:comment xml:lang="el">Μη δημοσιευμένο στο διαδίκτυο</rdfs:comment>
    <dc:description xml:lang="">Έγγραφα σχετικά με άδειες διδακτικού και λοιπού προσωπικού.</dc:description>
    <dcterms:extent xml:lang="el">6 σελίδες</dcterms:extent>
    <rdfs:comment xml:lang="el">Μεσαία κωδικοποίηση</rdfs:comment>
    <dc:description>142.2</dc:description>
    <dc:source>Ιστορικό Αρχείο ΕΚΠΑ</dc:source>
    <dc:subject xml:lang="el">1326989</dc:subject>
    <dc:subject xml:lang="el">300083</dc:subject>
    <dc:subject xml:lang="el">1327082</dc:subject>
    <dc:creator xml:lang="el">1057188</dc:creator>
    <dc:creator xml:lang="el">1326642</dc:creator>
    <dc:subject xml:lang="el">Ονόματα</dc:subject>
    <dc:subject xml:lang="el">Ιωάννης Πολίτης</dc:subject>
    <dc:subject xml:lang="el">Κλινικές</dc:subject>
    <dc:subject xml:lang="el">Πρύτανης</dc:subject>
    <dc:subject xml:lang="el">Εγκύκλιος</dc:subject>
    <dc:subject xml:lang="el">Προσωπικό</dc:subject>
    <dc:subject xml:lang="el">Υπάλληλοι</dc:subject>
    <dc:subject xml:lang="el">Βοηθητικό Προσωπικό</dc:subject>
    <dc:subject xml:lang="el">Υπουργείο θρησκευμάτων και Εθνικής Παιδείας</dc:subject>
    <dc:subject xml:lang="el">Νοσοκομεία</dc:subject>
    <dc:subject xml:lang="el">Αναρρωτικές Άδειες</dc:subject>
    <dc:subject xml:lang="el">Θεραπευτήριο «Ευαγγελισμός»</dc:subject>
    <edm:Place>
      <skos:prefLabel>Αθήνα</skos:prefLabel>
    </edm:Place>
  <ore:Aggregation rdf:about="https://apollonis-infrastructure.gr/apl/EKPA_70127">
    <edm:aggregatedCHO rdf:resource="https://apollonis-infrastructure.gr/apl/EKPA_70127"/>
    <edm:provider xml:lang="el">Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών</edm:provider>
    <edm:dataProvider xml:lang="el">Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών</edm:dataProvider>
    <edm:dataset xml:lang="el">Δεκαετία 1940</edm:dataset>
  </ore:Aggregation>
</rdf:RDF>
```


**Κενός προσδιοριστής γλώσσας**

**Χωρική πληροφορία χωρίς συντεταγμένες**

# Enriched EDM εγγραφή


```
▼<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:apl="https://apollonis-infrastructure.g
 xmlns:edm="http://www.europeana.eu/schemas/edm/" xmlns:ore="http://www.openarchives.org/ore/terms/" xmlns:owl=
 xmlns:skos="http://www.w3.org/2004/02/skos/core#" xmlns:wgs84_pos="http://www.w3.org/2003/01/geo/wgs84_pos#">
  ▼<edm:ProvidedCHO rdf:about="https://apollonis-infrastructure.gr/apl/EKPA_70127">
     <dc:identifier>EKPA_70127</dc:identifier>
     <apl:subfolder>/butterfly/backie/subfolder</apl:subfolder>
     <rdfs:comment>/lib/default/data/70119</rdfs:comment>
     <rdfs:comment xml:lang="el">Μη δημοσιευμένο στο διαδίκτυο</rdfs:comment>
     <dc:description xml:lang="el">Έγγραφα σχετικά με άδειες διδακτικού και λοιπού προσωπικού.</dc:description>
     <dcterms:extent xml:lang="el">6 σελίδες</dcterms:extent>
     <rdfs:comment xml:lang="el">Μεσαία κωδικοποίηση</rdfs:comment>
     <dc:description>142.2</dc:description>
     <dc:source>Ιστορικό Αρχείο ΕΚΠΑ</dc:source>
     <dc:subject xml:lang="el">1326989</dc:subject>
     <dc:subject xml:lang="el">300083</dc:subject>
     <dc:subject xml:lang="el">1327082</dc:subject>
     <dc:creator xml:lang="el">1057188</dc:creator>
     <dc:creator xml:lang="el">1326642</dc:creator>
     <dc:subject xml:lang="el">Ονόματα</dc:subject>
     <dc:subject xml:lang="el">Ιωάννης Πολίτης</dc:subject>
     <dc:subject xml:lang="el">Κλινικές</dc:subject>
     <dc:subject xml:lang="el">Πρύτανης</dc:subject>
     <dc:subject xml:lang="el">Εγκύκλιος</dc:subject>
     <dc:subject xml:lang="el">Προσωπικό</dc:subject>
     <dc:subject xml:lang="el">Υπάλληλοι</dc:subject>
     <dc:subject xml:lang="el">Βοηθητικό Προσωπικό</dc:subject>
     <dc:subject xml:lang="el">Υπουργείο θρησκευμάτων και Εθνικής Παιδείας</dc:subject>
     <dc:subject xml:lang="el">Νοσοκομεία</dc:subject>
     <dc:subject xml:lang="el">Αναρρωτικές Άδειες</dc:subject>
     <dc:subject xml:lang="el">Θεραπευτήριο «Ευαγγελισμός»</dc:subject>
     ▼<edm:Place>
        <skos:prefLabel>Αθήνα</skos:prefLabel>
        <wgs84_pos:lat>37.98376</wgs84_pos:lat>
        <wgs84_pos:long>23.72784</wgs84_pos:long>
      </edm:Place>
  ▼<ore:Aggregation rdf:about="https://apollonis-infrastructure.gr/apl/EKPA_70127">
     <edm:aggregatedCHO rdf:resource="https://apollonis-infrastructure.gr/apl/EKPA_70127"/>
     <edm:provider xml:lang="el">Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών</edm:provider>
     <edm:dataProvider xml:lang="el">Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών</edm:dataProvider>
     <edm:dataset xml:lang="el">Δεκαετία 1940</edm:dataset>
   </ore:Aggregation>
 </rdf:RDF>
```


Μονάδα Ψηφιακής Επιμέλειας
ΙΠΣΥ - Ερευνητικό Κέντρο Αθηνά

απολλωνίς

# Ποιότητα μεταδεδομένων

- Η μέτρηση της ποιότητας μεταδεδομένων είναι κομμάτι της διαδικασίας επικύρωσης και εκτελείται μετά από τις διαδικασίες:
  - συγκομιδής νέου περιεχομένου (harvest)
  - μετατροπής σε σχήμα στόχο (transform)
  - εμπλουτισμού (enrich)
- Υπολογίζει πληροφορία πληρότητας για κάθε πακέτο
  - πληρότητα σε επίπεδο στοιχείου (element)
  - πληρότητα σε επίπεδο συνόλου στοιχείων (υποχρεωτικά, προτεινόμενα)

# Διαδικασία «Preview»

- προβολή τεκμηρίου σε κάθε στάδιο επιμέλειας
- προεπισκόπηση τεκμηρίου σε HTML μορφή πριν τη δημοσιοποίηση

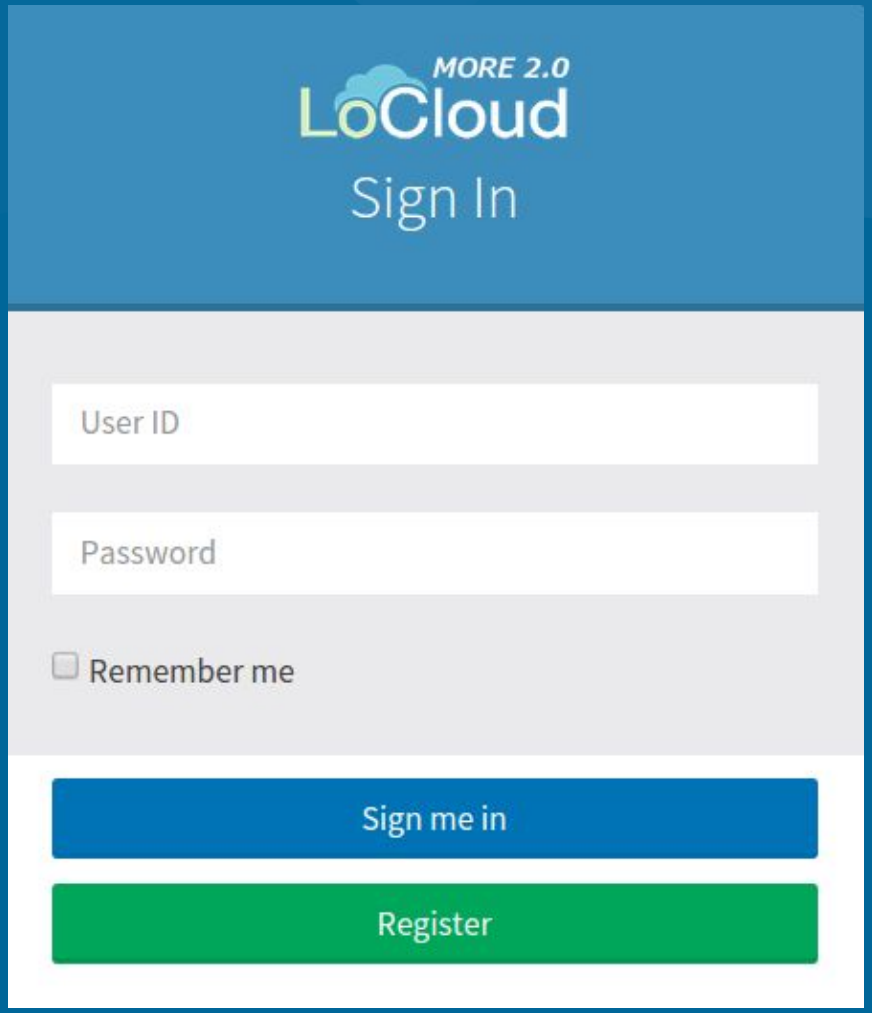
Back Project Δεκαετία 1940 :: Package 5202

Search

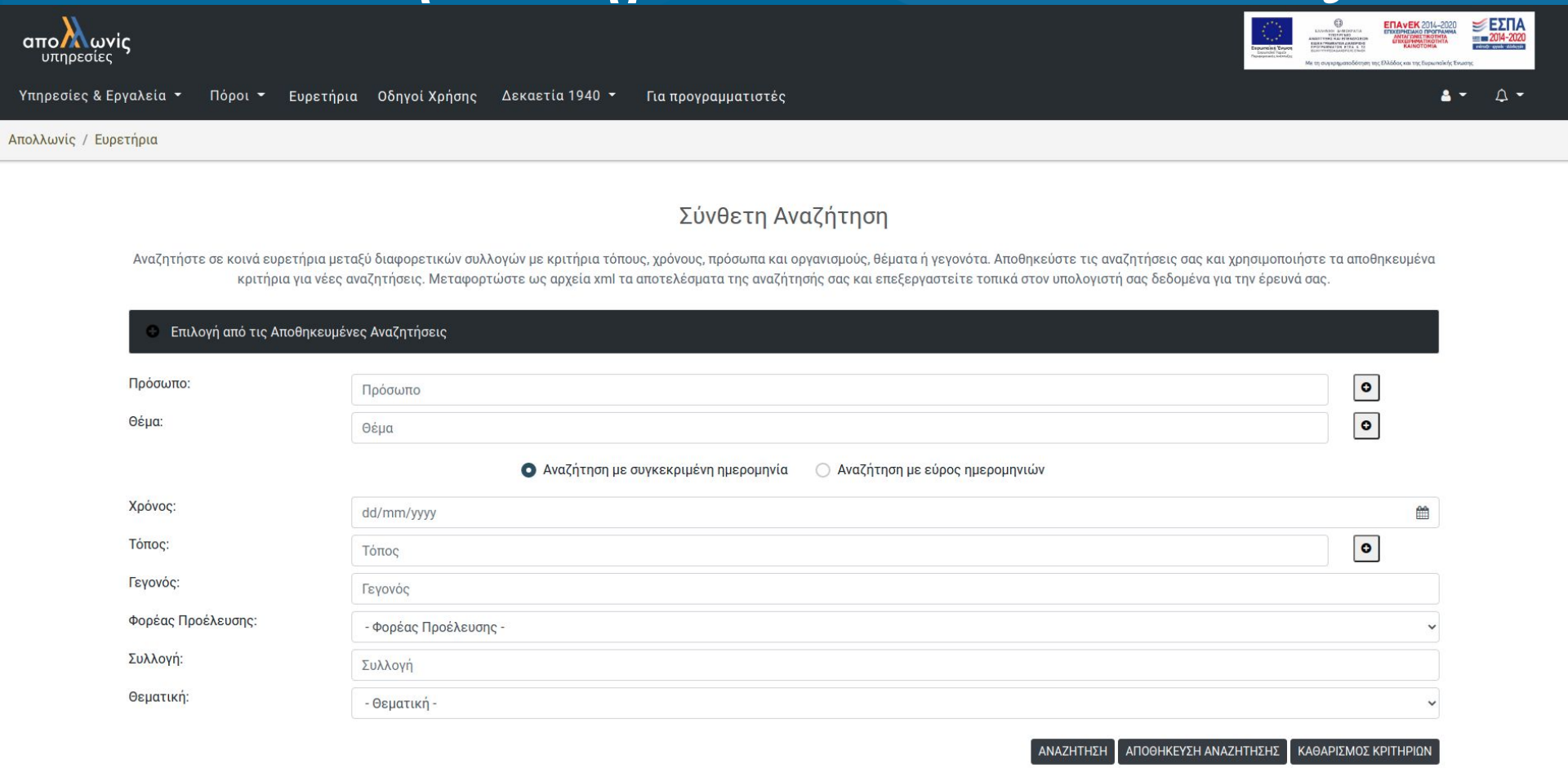
| # | Item | Native ID | |
|---|---|---|---|
| 1 | Φοιτητικά (προς Στρατολογικά Γραφεία, Γυμνάσια, Υπουργεία, Ιδιώτες, κ.τ.λ.) (υποφάκελος 524.1) | EKPA_2896732 | </> EDM |
| 2 | Δημοσιεύματα (Εκθέσεις πεπραγμένων) | EKPA_70330 | </> EDM |
| 3 | Αποφάσεις Συγκλήτου, Ν.Δ., Ίδρυσις Θέσεως Εργαστηρίων, Κλινικών, Παρασημοφορήσεις | EKPA_70346 | </> EDM |
| 4 | Κλινικές - Νοσοκομεία | EKPA_70280 | </> EDM |
| 5 | Συγχαρητήρια - Ευχαριστήρια - Τελετές - Εορτές - Συλλυπητήρια - Μνημόσυνα - Κηδείες | EKPA_70091 | </> EDM |
| 6 | Μισθώσεις - Επιτάξεις Ακινήτων, Εισπράξεις Ενοικίων | EKPA_123490 | </> EDM |
| 7 | Διορισμός Καθηγητών (Προκήρυξη πληρώσεως εδρών και άλλες διαδικασίες) | EKPA_69996 | </> EDM |
| 8 | Άδειες (Καθηγητών και υπαλλήλων) | EKPA_70127 | </> EDM |
| 9 | Κλινικές - Νοσοκομεία [391.3] | EKPA_70107 | </> EDM |
| 10 | Συγχαρητήρια - Ευχαριστήρια - Τελετές - Εορτές - Συλλυπητήρια - Μνημόσυνα - Κηδείες | EKPA_70138 | </> EDM |
| 11 | Υπότροφοι Κασσιμάτειου Κληροδοτήματος | EKPA_143253 | </> EDM |
| 12 | Συγχαρητήρια - Ευχαριστήρια - Τελετές - Εορτές - Συλλυπητήρια - Μνημόσυνα - Κηδείες | EKPA_70016 | </> EDM |

# Διαδικασία «Publish»

- δημοσιοποίηση επικυρωμένης πληροφορίας σε πολλαπλούς στόχους
  - OAI-PMH αποθετήριο
  - εξαγωγή σε archive μορφή
  - RDF αποθετήρια (π.χ. Sesame, Virtuoso)
  - Ευρετήρια (π.χ. Elasticsearch)


MORE 2.0
LoCloud
Sign In
User ID
Password
Remember me
Sign me in
Register

# Πύλη Υπηρεσιών Απολλωνίς




Μονάδα Ψηφιακής Επιμέλειας
ΙΠΣΥ - Ερευνητικό Κέντρο Αθηνά

απολλωνίς

# Ευρετήριο

- Δημοσίευση/Καταχώρηση μεταδεδομένων στο ευρετήριο Elasticsearch
- Δυνατότητα αναζήτησης και ανάκτησης πληροφορίας σε ελάχιστο χρόνο
- Δυνατότητα εκτέλεσης σύνθετων επερωτήσεων στο ευρετήριο

# Διαδικασία Καταχώρησης/Ανάκτησης

- Καθορισμός αναζητήσεων/επερωτήσεων που πρέπει να μπορεί να υποστηρίξει το ευρετήριο
- Εντοπισμός στοιχείων μεταδεδομένων τα οποία είναι αναγκαία για την εκτέλεση των επερωτήσεων
- Εξαγωγή των μεταδεδομένων από τα αρχεία μεταδεδομένων και καταχώρησή τους στο ευρετήριο
- Κατασκευή επερωτήσεων βάσει της λειτουργικότητας που δίνει το Elasticsearch

# Πηγές

- MORe Aggregator

http://more.dcu.gr/

- Υποδομή Απολλωνίς

https://apollonis-infrastructure.gr/

- Elasticsearch

https://www.elastic.co/

- Europeana Data Model Documentation

http://pro.europeana.eu/page/edm-documentation