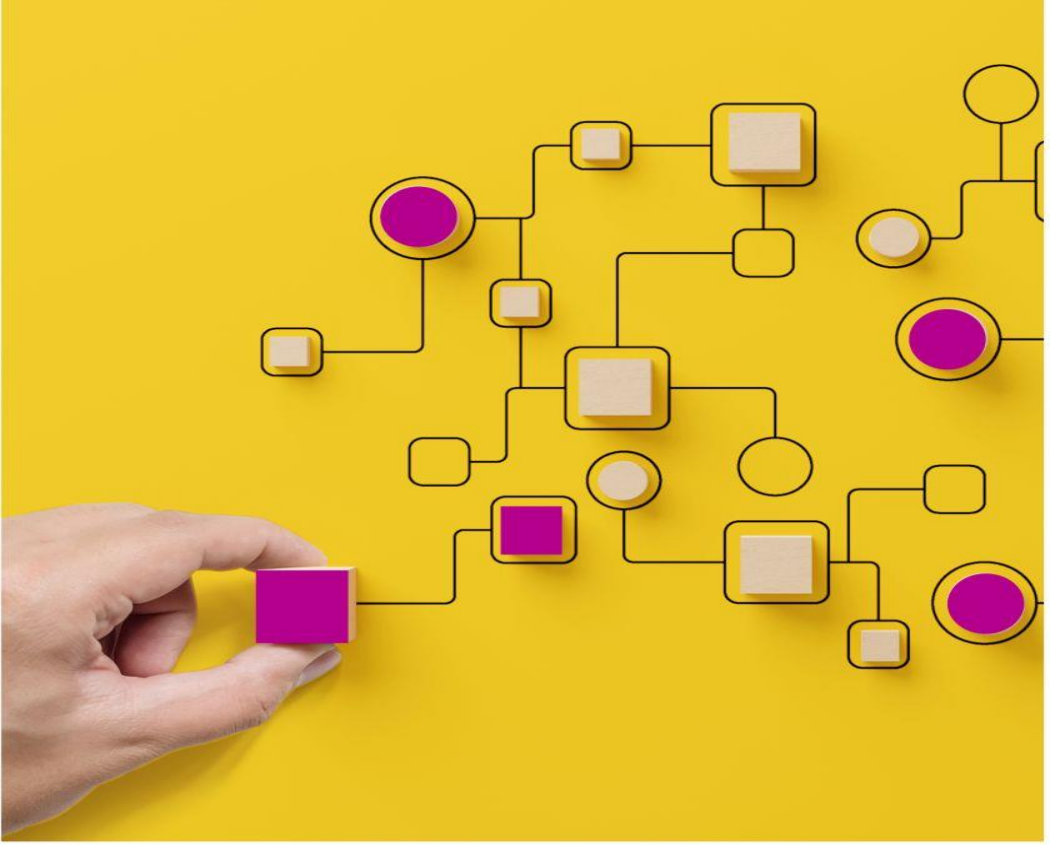


ΔΙΑΔΙΚΤΥΑΚΟ ΣΕΜΙΝΑΡΙΟ | **27.11.2025** | 11.00

Διασυνδεδεμένα Δεδομένα

Η σημασία και τα οφέλη τους

Η συμβολή του ΕΚΤ στην ανάπτυξη του οικοσυστήματος

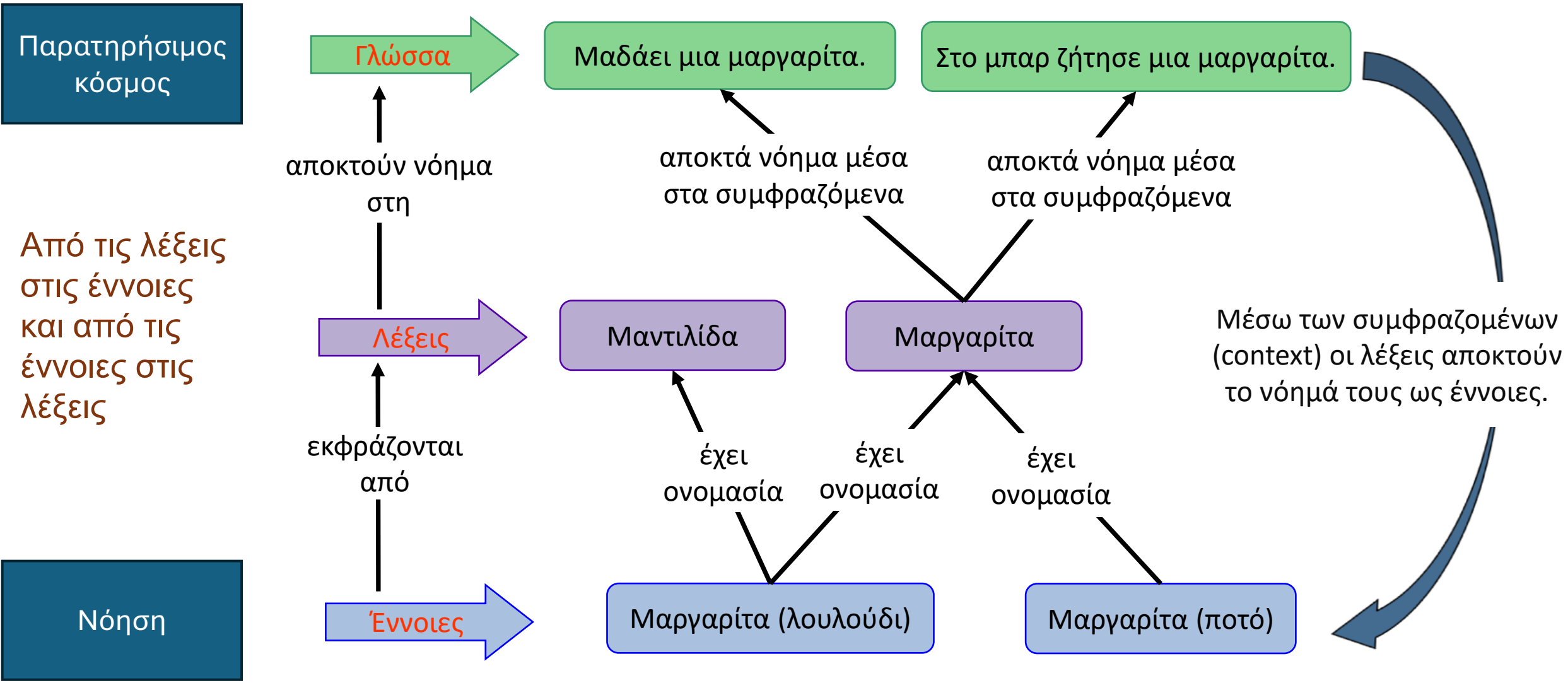
ΕΚΤ ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΤΕΚΜΗΡΙΩΣΗΣ & ΗΛΕΚΤΡΟΝΙΚΟΥ ΠΕΡΙΕΧΟΜΕΝΟΥ

{ Semantics.gr } + { Search Culture.gr }

Το οικοσύστημα των Διασυνδεδεμένων Δεδομένων:
Νέες προοπτικές τεκμηρίωσης για τους φορείς πολιτισμικής κληρονομιάς

Μανόλης Πεπονάκης
Ειδικό Επιστημονικό Προσωπικό, ΕΚΤ

Συμφραζόμενα και έννοιες στη φυσική γλώσσα



Ο σημασιολογικός ιστός διαθέτει έναν μηχανισμό για να αναπαρασταθούν πρωτίστως οι έννοιες και, στη συνέχεια, οι λέξεις μέσω των οποίων οι έννοιες είναι γνωστές.

Από το web 1.0 στο web 3.0

Web 1.0: αρχική φάση του διαδικτύου, περιορισμένη αλληλεπίδραση δημιουργού και τελικού χρήστη

- Δημιουργία περιεχομένου μόνο από «μεγάλους» φορείς, διότι απαιτούνταν υπερβολικά υψηλός εξοπλισμός και τεχνογνωσία.

Web 2.0: επόμενη φάση του διαδικτύου, διευκόλυνση της αλληλεπίδρασης δημιουργού και τελικού χρήστη, αλλά καθίσταται και πιο ασαφής, λόγω του ότι καθένας μπορεί να δημιουργήσει και να δημοσιοποιήσει περιεχόμενο.

- Εφικτή κυρίως μέσω τεχνολογιών που συνδέονται με τα κοινωνικά δίκτυα και μέσω πλατφορμών όπως το YouTube και το Facebook

Web 3.0 (ή Semantic Web = Σημασιολογικός Ιστός): μια επέκταση του σημερινού διαδικτύου, η οποία εστιάζει στη δομή των δεδομένων.

- Ο στόχος είναι η μετάβαση σε δομές δεδομένων που θα μπορούν να επεξεργάζονται πιο αποτελεσματικά από τους υπολογιστές, σε αντίθεση με τη λειτουργικότητα του σύγχρονου κειμενοκεντρικού διαδικτύου.

Οντότητες του φυσικού κόσμου και υπολογιστές

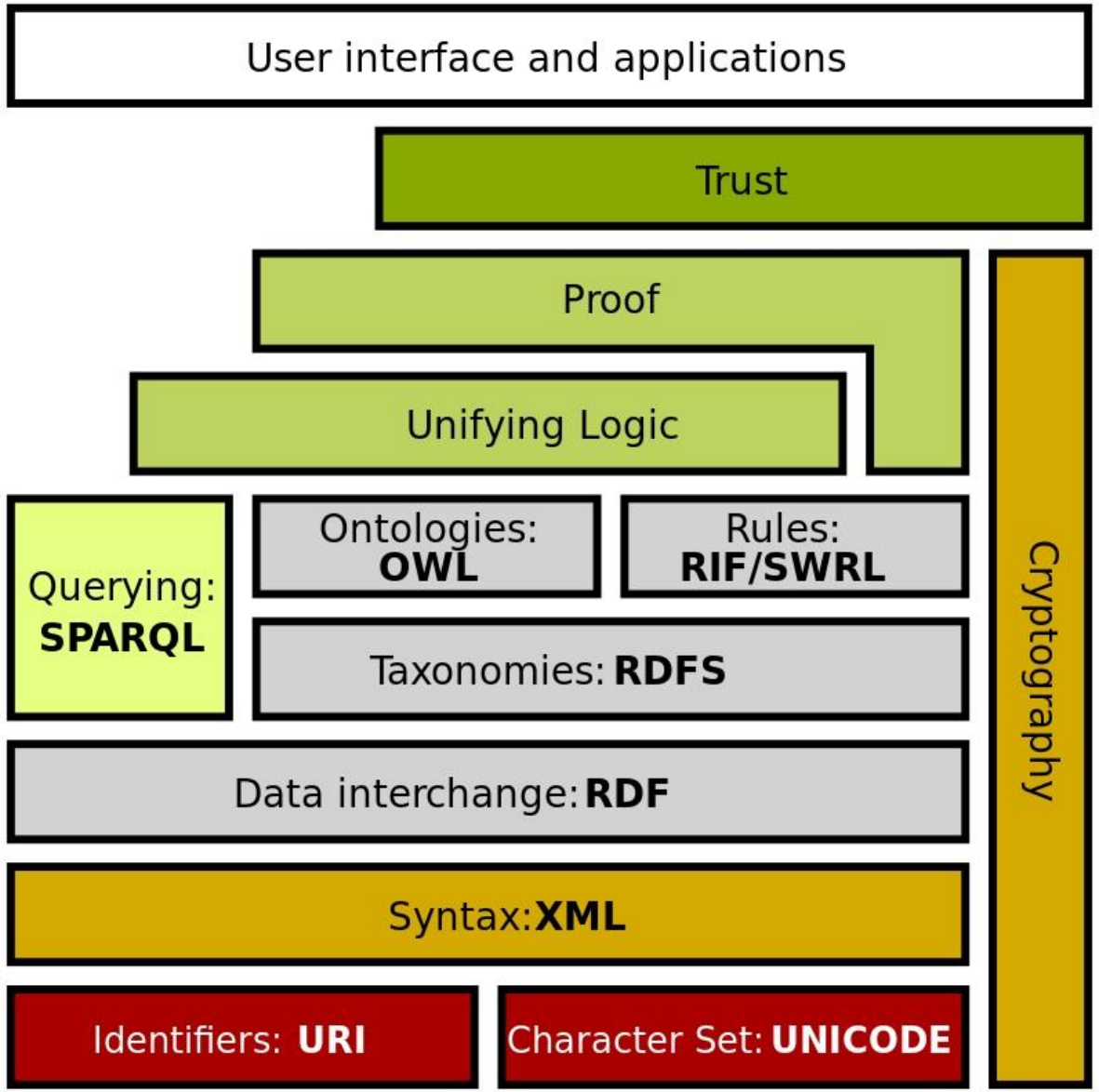
"Τα φυσικά αντικείμενα, τα γεγονότα και οι σχέσεις, που δεν μπορούν να αποθηκευτούν άμεσα σε έναν υπολογιστή, αναπαριστώνται από ***σύμβολα*** *που λειτουργούν ως* ***υποκατάστατα*** *των εξωτερικών πραγμάτων. Τα σύμβολα και οι συνδέσεις μεταξύ τους σχηματίζουν ένα* ***μοντέλο του εξωτερικού συστήματος****. Με τον χειρισμό αυτών των εσωτερικών υποκατάστατων, ένα πρόγραμμα υπολογιστή μπορεί να προσομοιώσει το εξωτερικό σύστημα ή να συλλογιστεί σχετικά με αυτό."*

Sowa JF (2009)

Sowa, J. F. (2009). *Knowledge representation: Logical, philosophical, and computational foundations*. Brooks

Κοινωνιοτεχνικά (Sociotechnical) δημιουργήματα

> *"The Web is now philosophical engineering".*
>
> Berners-Lee, T. (2006). *Interview with Tim Berners-Lee* (B. Runciman, Interviewer). https://www.impactlab.com/2006/03/25/interview-with-tim-berners-lee/



The semantic web stack

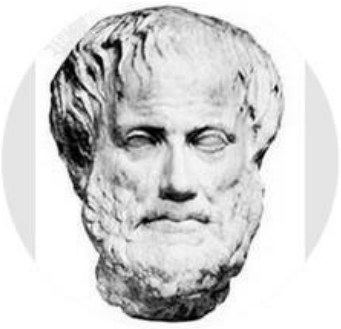
Προέλευση και αξιοπιστία

Πόσο αξιόπιστη είναι μια πληροφορία ώστε να «χτίσουμε» πάνω σε αυτήν;

Σημείο-κλειδί για την αξιοπιστία είναι η **προέλευση**

Ο τρόπος παραγωγής της πληροφορίας και το κύρος εκείνου (φορέα ή ατόμου) που την διακινεί, είναι καθοριστικής σημασίας

Οι φορείς πολιτισμικής κληρονομίας έχουν κρίσιμο ρόλο



Aristotle

Ancient Greek philosopher

Verified email at buffalo.edu

metaphysics ethics

Σύμφωνα με το Google Scholar, ο Αριστοτέλης είναι ένας αρχαίος Έλληνας φιλόσοφος που έχει επιβεβαιωμένο email στο buffalo.edu!
https://scholar.google.com/citations?user=4rO87mQAAAAJ

Ορίζοντας τα ‘περίπου’ συνώνυμα…

Σημασιολογικός ιστός (ή Web 3.0)**:** στόχος να καταστούν τα δεδομένα του Διαδικτύου αναγνώσιμα από μηχανές, δηλαδή να κωδικοποιηθεί η πληροφορία με τρόπο που να μεταφέρει την επιδιωκόμενη σημασιολογία (το επιδιωκόμενο νόημα).

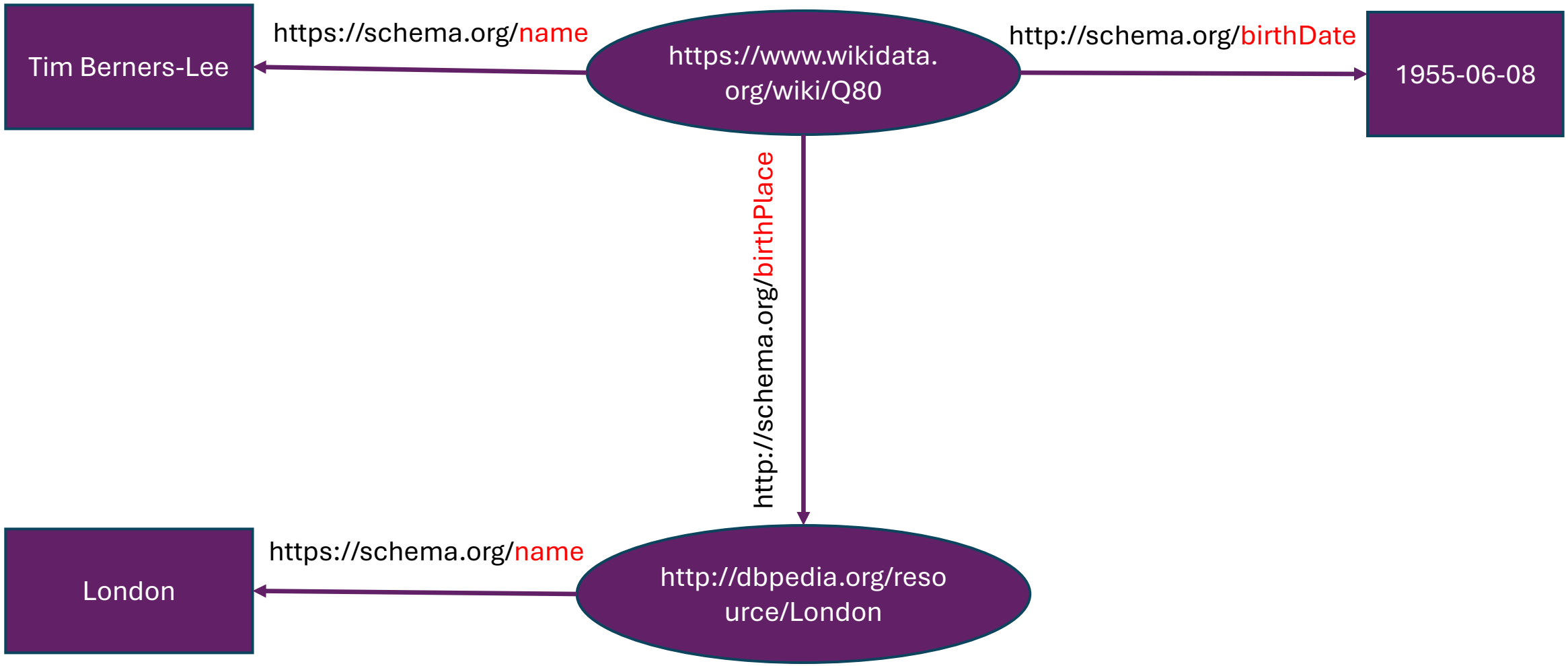
Ιστός των δεδομένων: στόχος να μετατραπεί ο ιστός σε μια παγκόσμια ‘βάση δεδομένων’, δηλαδή να υπάρξει μετάβαση σε δομές δεδομένων τα οποία θα είναι επεξεργάσιμα από τους υπολογιστές πιο αποτελεσματικά.

Διασυνδεδεμένα δεδομένα (Linked data): οι προδιαγραφές που αναπτύχθηκαν ή σχετίζονται με τους παραπάνω στόχους, καθώς και οι διαδικασίες και οι κανόνες για την επίτευξή τους.

Resource Description Framework (RDF)

- Ένα αφαιρετικό μοντέλο δεδομένων που προτείνεται από το W3C
- Βασικός μηχανισμός: οι δηλώσεις (statements)
  - Κάθε δήλωση είναι μια **τριπλέτα** με υποκείμενο-κατηγόρημα-αντικείμενο
    - Αντίστοιχη δομή με τη δομή ***υποκείμενο-ρήμα-αντικείμενο*** που έχει η φυσική γλώσσα
    - Το υποκείμενο είναι πάντα ένας πόρος, ο οποίος ταυτοποιείται από ένα URI
      - Σημείωση: επιτρέπεται -αλλά δεν είναι σύνηθες- το υποκείμενο να είναι ένας «κενός» κόμβος. Σε αυτή την περίπτωση δεν ταυτοποιείται από ένα URI.
    - Το κατηγόρημα (ή ιδιότητα) είναι και αυτό ένας πόρος, ο οποίος ταυτοποιείται από ένα URI
    - Το αντικείμενο μπορεί να είναι είτε ένας πόρος είτε μια τιμή (απλό λεκτικό, όχι πόρος)
  - **Το σύνολο των τριπλετών αποτελεί έναν κατευθυνόμενο γράφο**

Παράδειγμα σε RDF



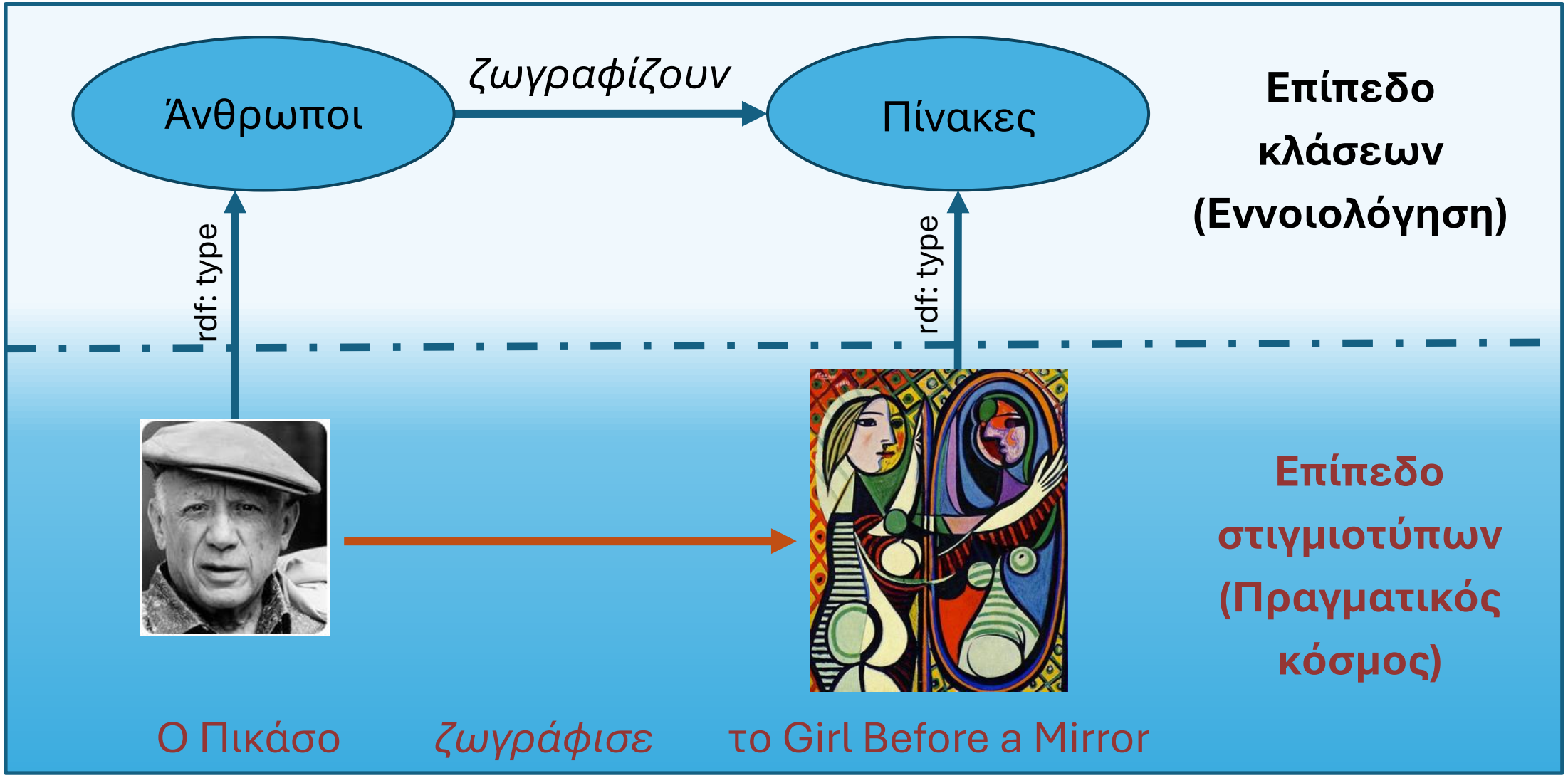
Το RDF είναι αναγκαία προϋπόθεση, αλλά όχι ικανή για να διασφαλίσει τα συνδεδεμένα δεδομένα

Παράδειγμα “*κακής χρήσης*”

Παράδειγμα “*καλής χρήσης*”

Και τα δύο παραδείγματα περιγράφουν ότι ο Λεονάρντο ντα Βίντσι είναι ο δημιουργός της Μόνα Λίζα.
Στην πρώτη περίπτωση, ο ντα Βίντσι δεν εμφανίζεται ως πόρος (ανεξάρτητη οντότητα) αλλά ως ιδιότητα της Μόνα Λίζα.

Κατηγορίες και πραγματικός κόσμος: Κλάσεις και Στιγμιότυπα (Άτομα)



Το RDF Shema (RDFs) παρέχει μηχανισμό ορισμού και διαχείρισης των κλάσεων.
Η OWL παρέχει επιπλέον φορμαλισμούς και περιορισμούς.

Διαχείριση του ‘νοήματος’ στο περιβάλλον του Σημασιολογικού Ιστού

Οι **Κατηγορίες** (Κλάσεις) και τα **Άτομα** (Στιγμιότυπα), όπως και οι **Ιδιότητές** τους αναπαριστώνται από **Αναγνωριστικά** (**Identifiers**).

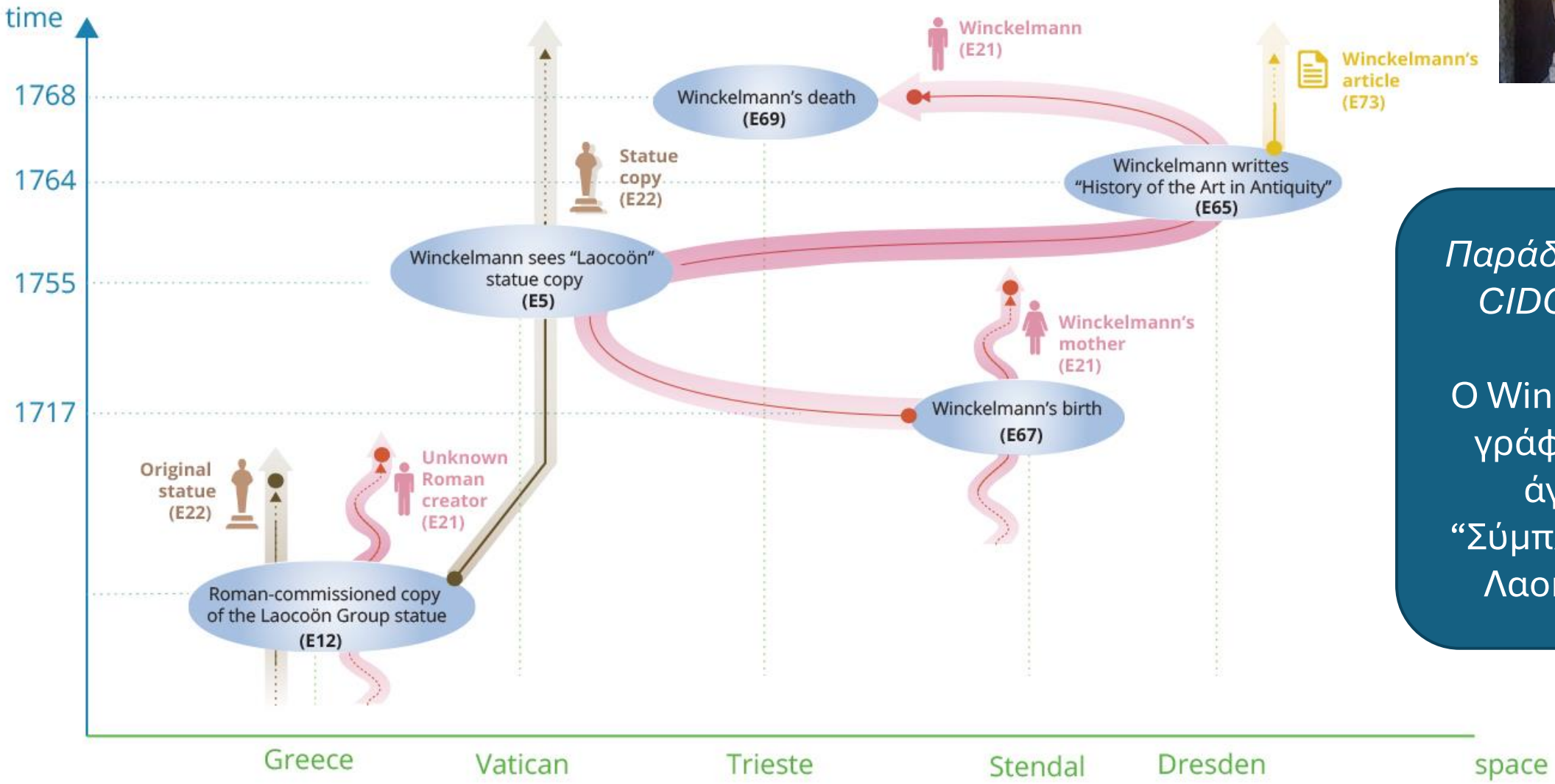
Το ‘νόημα’ δεν προκύπτει από το τι σημαίνουν τα ονόματά τους στη φυσική γλώσσα, αλλά εξαρτάται από το τι αντιπροσωπεύουν τα αναγνωριστικά και από τον τρόπο με τον οποίο συνδέονται μεταξύ τους. Το νόημα της οντότητας προκύπτει από τις ιδιότητές της.

Για το οικοσύστημα των διασυνδεδεμένων δεδομένων (linked data), αυτά τα **αναγνωριστικά** πρέπει να είναι **HTTP based URIs**.

Η σημασία της διασύνδεσης των πόρων

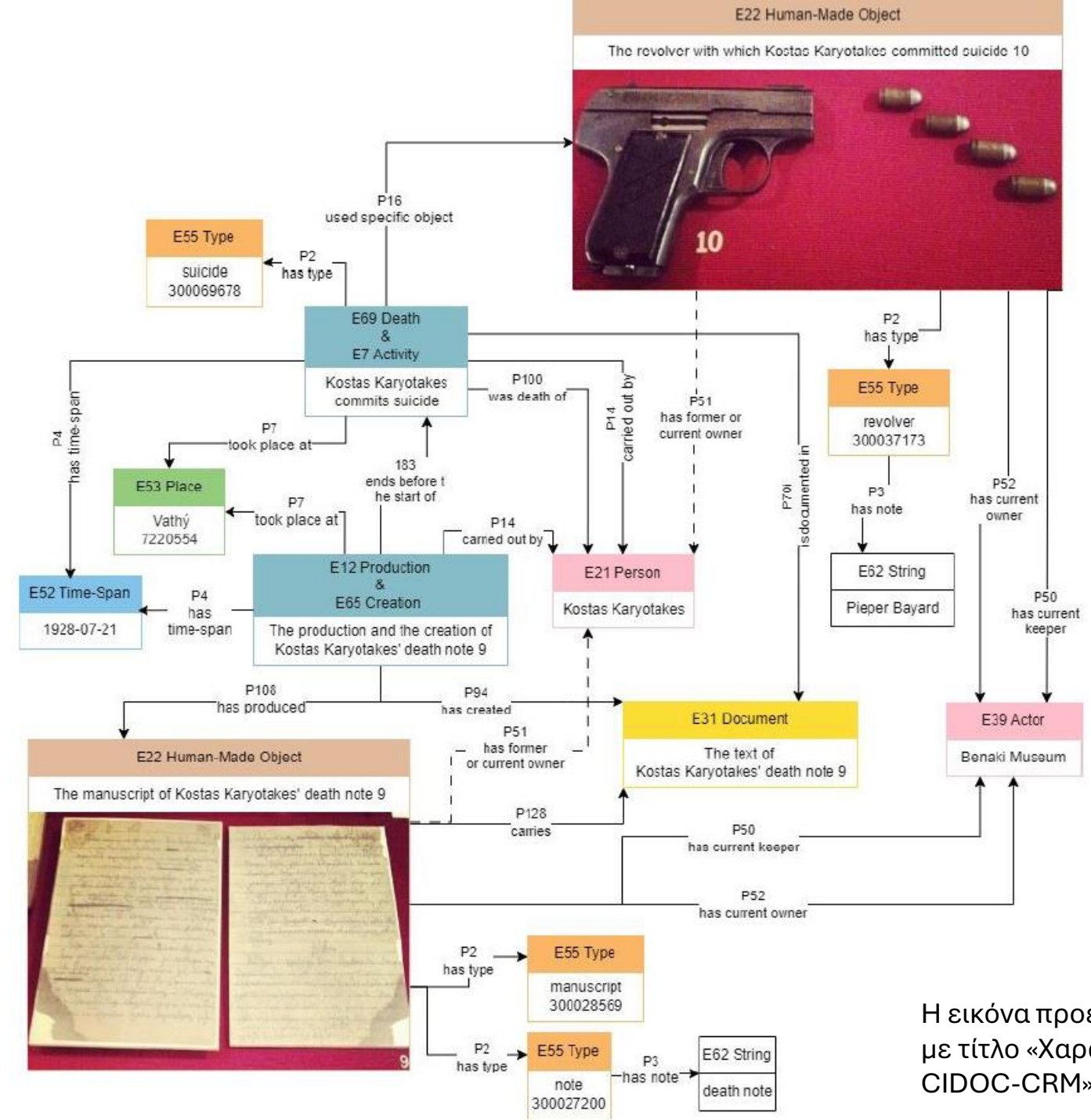
Παρέχει πρόσθετη πληροφορία	• Προσθέτει πλαίσιο μέσα στο οποίο νοηματοδοτείται ο πόρος • Μειώνει την αμφισημία
Επιτρέπει τη σύνδεση διαφορετικών πόρων	• Ίδιου ή διαφορετικού τύπου
Επιτρέπει την εφαρμογή δυναμικών μοντέλων	• Π.χ. επιτρέπει να μοντελοποιηθούν και να περιγραφούν δραστηριότητες, και όχι μόνο αντικείμενα ή άτομα
Επιτρέπει τη σύνδεση μεταξύ διαφορετικών περιγραφών της ίδιας οντότητας	• Π.χ. με τη χρήση του “owl:sameAs”

Νέοι τρόποι μοντελοποίησης της πληροφορίας



Παράδειγμα από CIDOC-CRM

Ο Winckelmann γράφει για το άγαλμα "Σύμπλεγμα του Λαοκόοντα"



Γεγονότα, αντικείμενα και διασυνδέσεις

Στις 21 Ιουλίου 1928 ο Κωνσταντίνος Καρυωτάκης αυτοκτονεί στο Βαθύ της Πρέβεζας χρησιμοποιώντας ένα όπλο και αφήνοντας πίσω του μια επιστολή. Το όπλο και η επιστολή βρίσκονται σήμερα στο Μουσείο Μπενάκη, με τα αναγνωριστικά 10 και 9 αντιστοίχως.

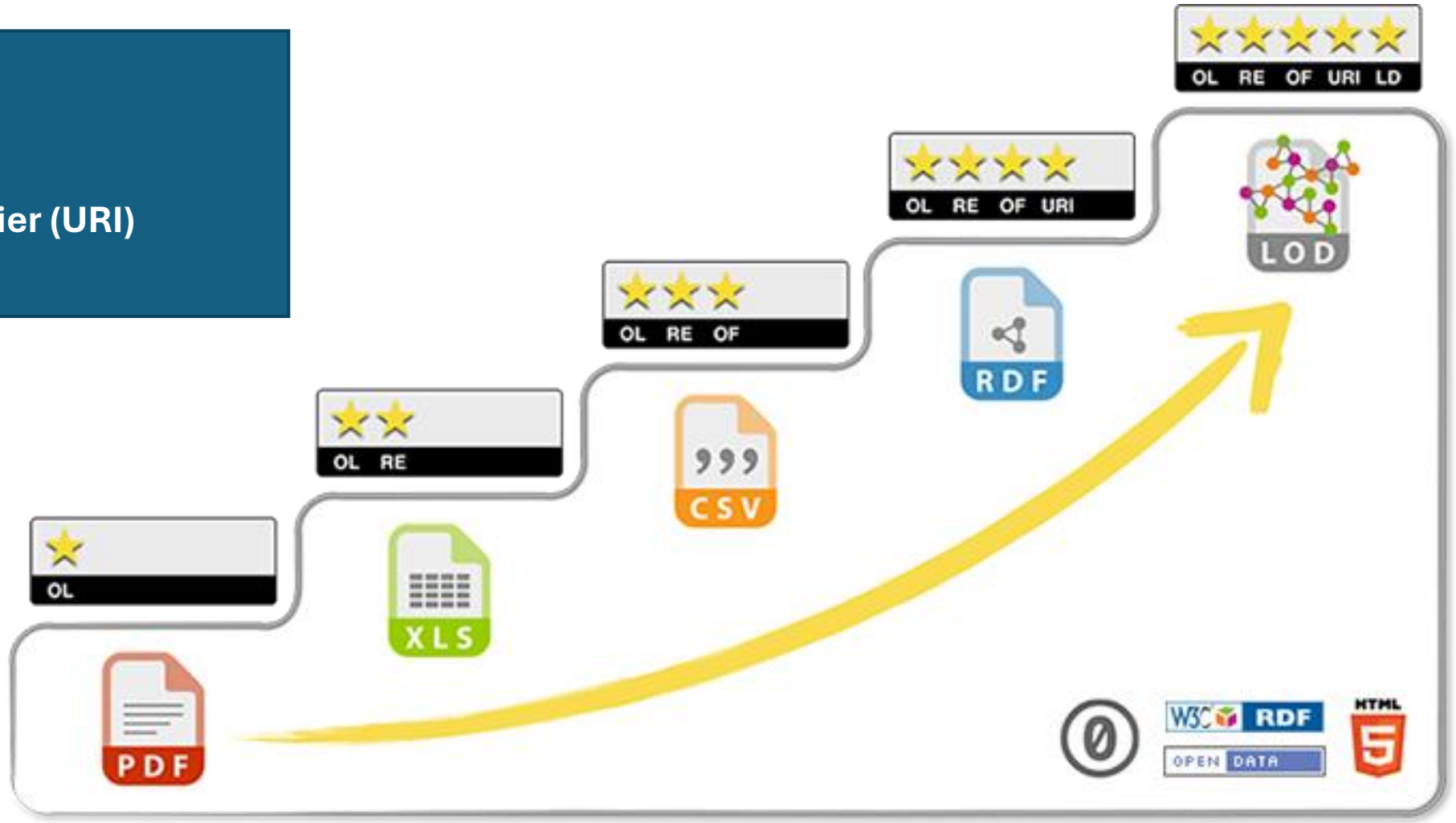
Η εικόνα προέρχεται από τη μεταπτυχιακή εργασία της Ειρήνης Σπυριδάκη με τίτλο «Χαρακτηριστικά παραδείγματα πολιτιστικής περιγραφής με CIDOC-CRM» http://dx.doi.org/10.26265/polynoe-6215

Προϋποθέσεις για την υλοποίηση

Νέος τρόπος σκέψης!	• Έμφαση στην εννοιολόγηση
Συνεργασία	• Μεταξύ ερευνητών και μεταξύ φορέων
Διεπιστημονικότητα	• Ειδικά μεταξύ των επιστημόνων της πληροφορίας και άλλων πεδίων

Η διαβάθμιση των ανοικτών δεδομένων μέχρι τα Linked Open Data

Open License (OL)
Readable (RE)
Open Format (OF)
Uniform Resource Identifier (URI)
Linked Data (LD)



Η εικόνα προέρχεται από το https://5stardata.info/en/

Πλεονεκτήματα διασυνδεδεμένων δεδομένων I/II

Ουδετερότητα περιεχομένου

Ουδετερότητα λογισμικού

Διαμοιρασμός

- Ενδό- και δι-επιστημονικός

Εύκολα επαναχρησιμοποιήσιμα

- Αξιοπιστία που πηγάζει από την προέλευση

Επεκτάσιμα

- Προστιθέμενη αξία

Πλεονεκτήματα διασυνδεδεμένων δεδομένων II/II

Γλωσσική ουδετερότητα

- Εννοιοκεντρικές προσεγγίσεις

Κατάλληλα για μοντελοποίηση

- Επιτρέπουν δυναμικές περιγραφές σύνθετων φαινομένων

Δυνατότητες διαφορετικών οπτικοποιήσεων

- graphs, HTML, κ.λπ.

Αυτόματος συμπερασμός (reasoning)

Μεγάλη ακρίβεια περιγραφής

Προσθέστε ελεύθερα…!

Σας ευχαριστώ πολύ για τη συμμετοχή και την προσοχή σας!

Μανόλης Πεπονάκης
Ειδικό Επιστημονικό Προσωπικό
Εθνικό Κέντρο Τεκμηρίωσης & Ηλεκτρονικού Περιεχομένου | ΕΚΤ
Ζεφύρου 56, 17564, Παλαιό Φάληρο
email: epepo@ekt.gr