

Απαιτήσεις και χαρακτηριστικά ψηφιακών ερευνητικών υποδομών για τις ανθρωπιστικές επιστήμες. Η ευρωπαϊκή υποδομή DARIAH.

Πάνος Κωνσταντόπουλος
Οικονομικό Πανεπιστήμιο Αθηνών και Ε.Κ. «Αθηνά»

Ημερίδα παρουσίασης του έργου «CLARIN-EL: Ερευνητική Υποδομή
Γλωσσικών Πόρων και Τεχνολογίας»
Εθνικό Ίδρυμα Ερευνών , 1.10.2010

Ψηφιακές υποδομές για τις τέχνες και τις ανθρωπιστικές επιστήμες

- Κατανομή θεσμικών αποθετηρίων
- Σύνθετο και πολυμεσικό υλικό
- Οργανισμοί μνήμης με διαφορετικές αποστολές (μουσεία, αρχεία, βιβλιοθήκες)
- Μεγάλα έργα ψηφιοποίησης
- Ψηφιακοί χώροι συγκέντρωσης υλικού:
 - Google Books, TEL, Europeana
- ‘digital scholarship’

Ορισμένες πρωτοβουλίες και προσεγγίσεις

- Διεθνείς μελέτες και πρωτοβουλίες
 - *ACLS Our cultural commonwealth* report (2006)
 - Research repositories-related work at the UK Digital Curation Centre
 - *Epistemic networks* workshops and initiatives in the UK (2007-2008)
 - *Digital curation and trusted repositories* issue, JODI (2007)
 - Ch. Borgman's magisterial synthesis *Scholarship in the digital age* (2007)
- Εστιασμένες μονάδες Ε&Α
 - ADS, DCC (UK)
 - ΜΟΨΕ/ΙΠΣΥΠ-Ε.Κ. Αθηνά, ΚΠΠ/ΙΠ-ΙΤΕ
- Έργα και δίκτυα
 - DARIAH (EU)
 - Digital Humanities Observatory (IR)
 - TextGrid, eHum (D)
 - arts-humanities.net (EU)
 - ΔΥΑΣ

DARIAH

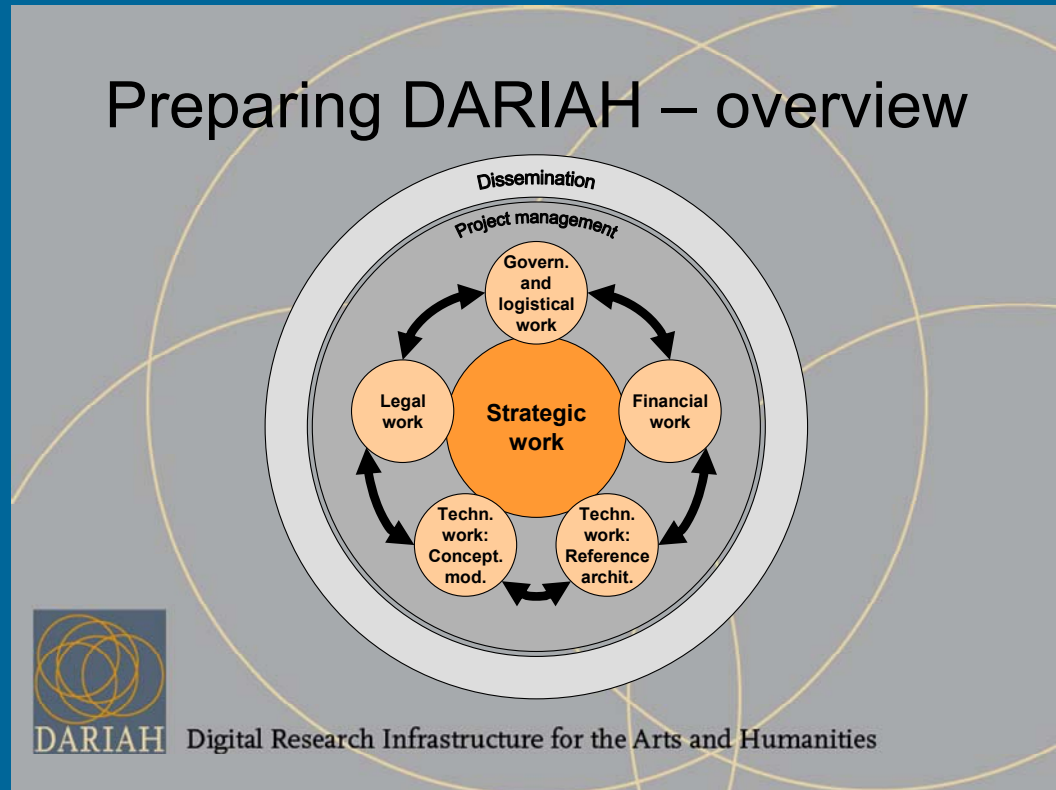
- The grand vision for DARIAH is to facilitate long-term preservation of and access to arts and humanities research data.
- Scholars can use DARIAH to:
 - find and use a wide range of digital content from across Europe
 - find and use tools that aid data interpretation
 - participate in experimentation and innovation with other scholars from across multiple domains
 - seek advice, and exchange ideas and knowledge on all aspects of digital scholarship
 - ensure that they work to accepted standards and best practices
 - ensure the long-term preservation of data

DARIAH προπαρασκευαστικό στάδιο

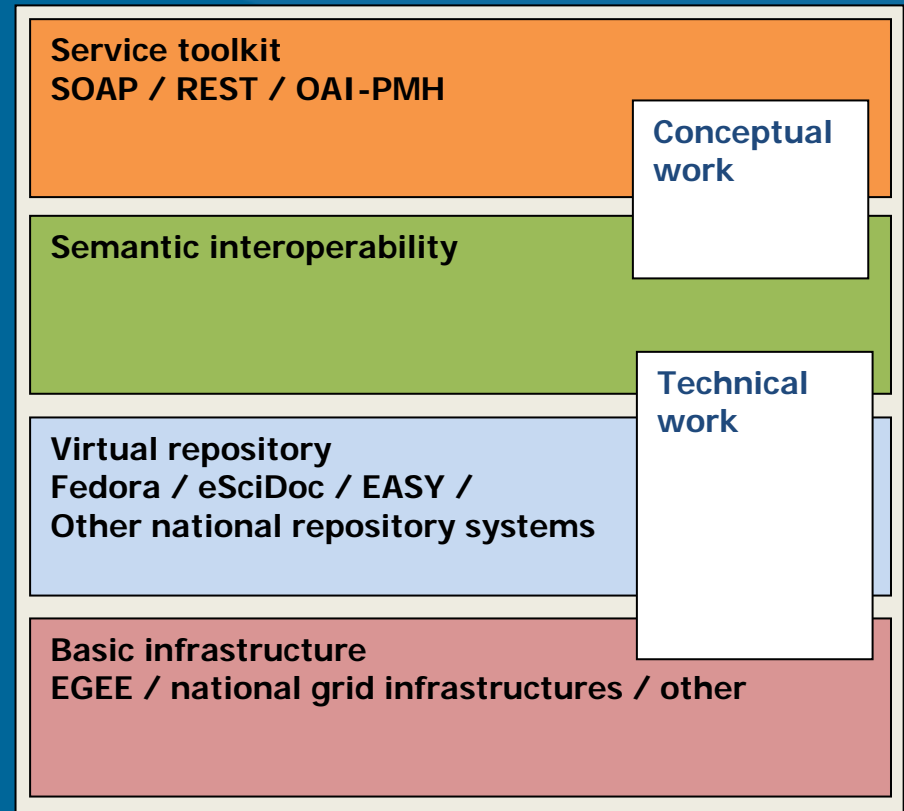
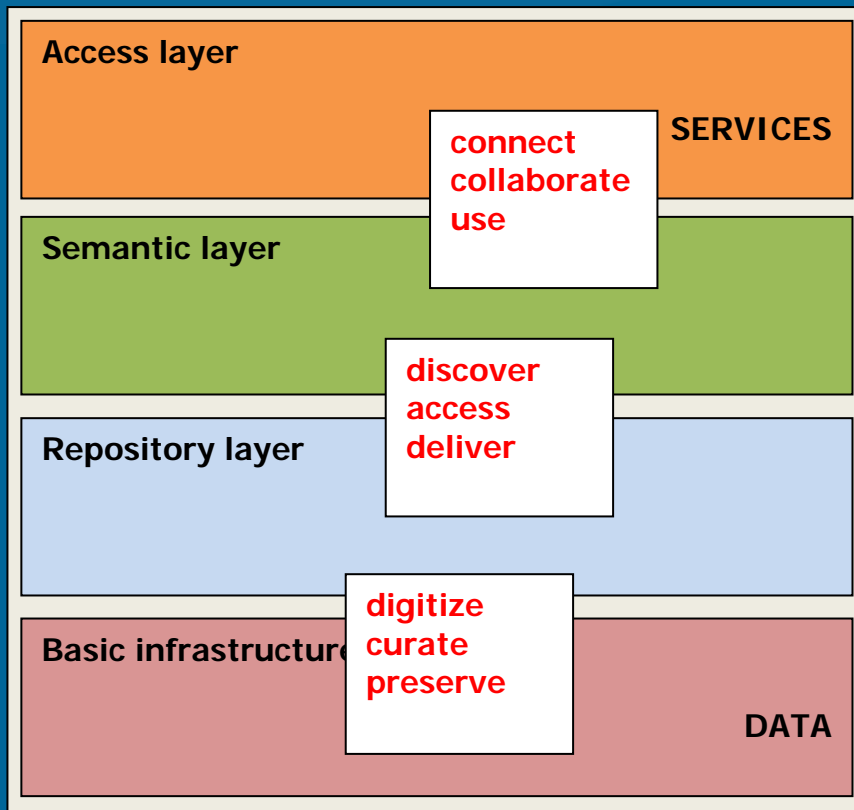
The preparatory phase addresses the following issues:

- Coordination
- Strategic
- Financial
- Governance
- Logistic
- Legal
- Technical
- Conceptual

Preparing DARIAH – overview

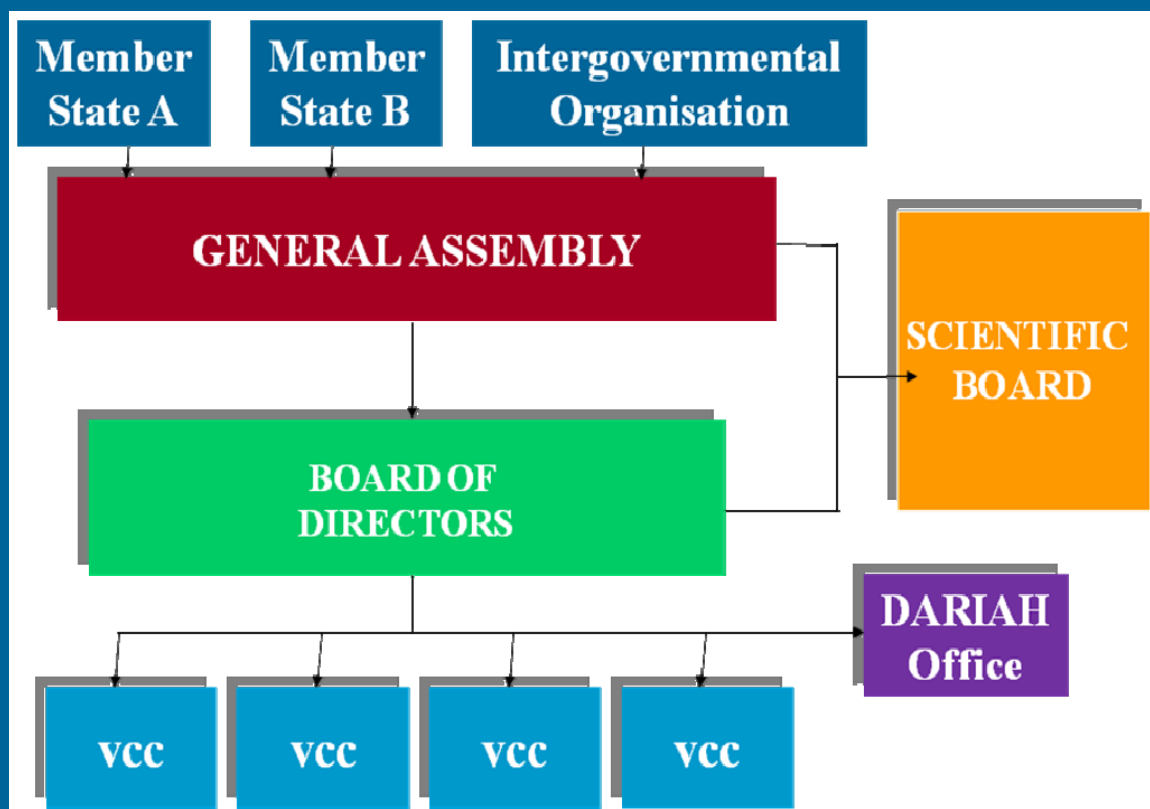


Αρχιτεκτονική DARIAH



DARIAH στάδιο εφαρμογής -1

Οργάνωση



DARIAH στάδιο εφαρμογής -2

Virtual Competency Centres

- VCC1 : e-Infrastructure
- VCC2 : Research & Education Liaison
- VCC3 : Scholarly Content Management
- VCC4 : Advocacy, Impact & Outreach

DARIAH στάδιο εφαρμογής -3

Εκτιμώμενο κόστος ανάπτυξης και λειτουργίας

	Ελάχιστοι πόροι	Ιδανικοί πόροι
Φάση κατασκευής 5-6 έτη	46 ΙΠΑ/έτος 3,6 Μευρώ/έτος	73,5 ΙΠΑ/έτος 5,4 Μευρώ/έτος
Φάση λειτουργίας	19 ΙΠΑ/έτος 2 Μευρώ/έτος	47,5 ΙΠΑ/έτος 3,5 Μευρώ/έτος

TextGrid

- Establishment of a “Virtual Research Environment”
- Development of a Grid based eInfrastructure for humanities
- TextGrid: a generic platform for scientific text data processing
- **Service Grid** - Toolbox for collaborative work
 - *TextGrid Lab* (Laboratory)
- **Data Grid** - Virtual archive(s) for research data curation
 - *TextGrid Rep* (Repository)

eHum

- Overall objectives
 - to create a **strategy** for the development of a research infrastructure within the humanities in Germany
 - to develop a **roadmap** for the building of collaboration and partnerships between institutions and organizations for a national e-Humanities infrastructure
- Specific objectives
 - to analyze national and international merges in the field of e-Humanities
 - to gather information about current activities especially in Europe and USA
 - to create overviews of such organizations to promote collaboration and synergies at German level

JISC e-Infrastructure Programme

Increase the benefits to, and use of, e-Infrastructure by a **wider user base**

Ensure that e-Infrastructure builds on and shares **common core services**

Explore the ways in which the benefits of the capabilities being developed in **grid computing** can be transferred to other domains

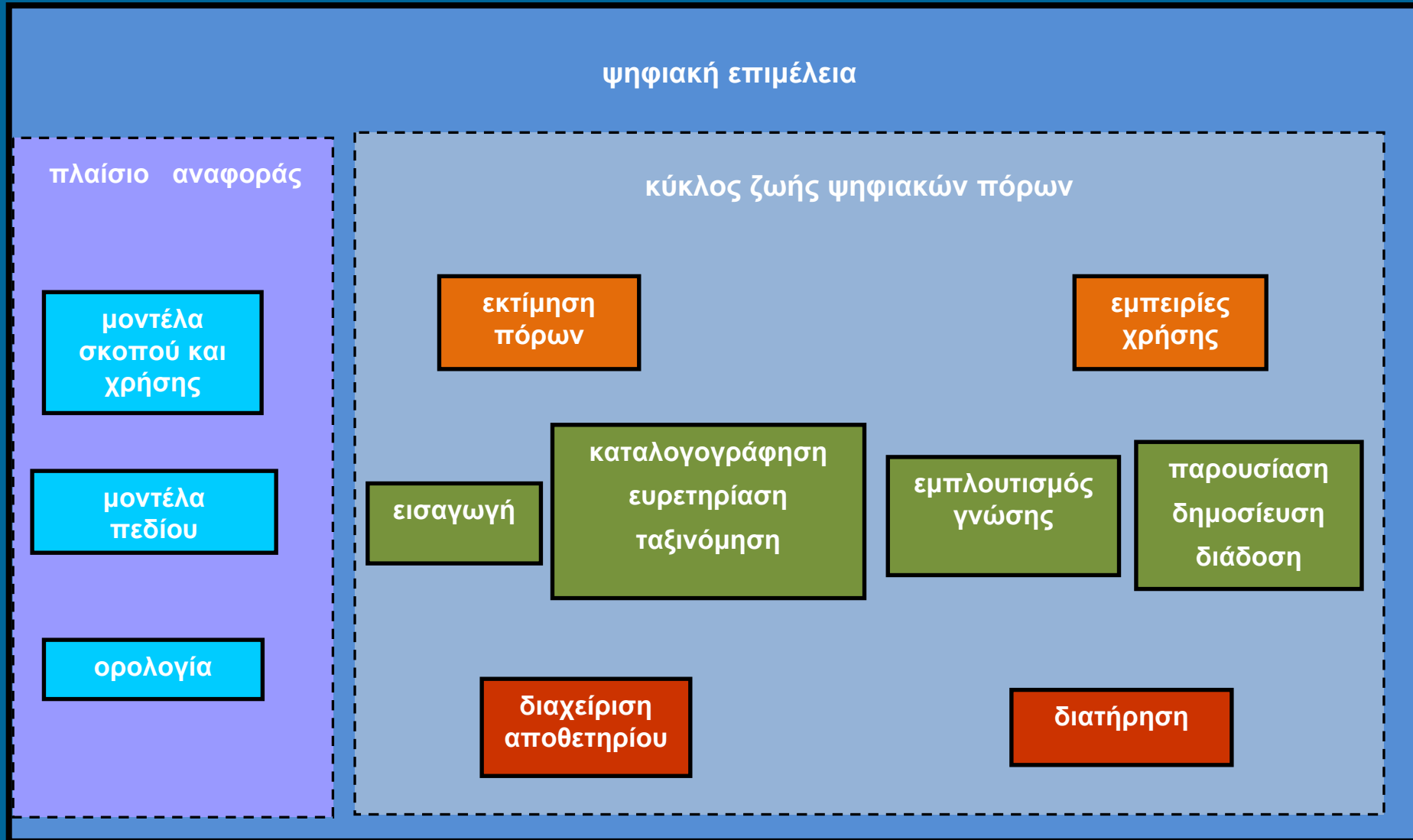
The e-Infrastructure programme is comprised of 4 thematic areas:

- 1) **Community engagement and support** Focusing on community engagement and support in the broader take up and more effective use of e-Infrastructure
- 2) **e-Infrastructure security** Creating solutions to issues of grid authentication, authorisation and security; establishing interoperability and consensus between access management regimes
- 3) **Grid services and tools** Developing production tools and services to broaden the research community take-up of the computational and data grid resources within the National Grid Service
- 4) **Knowledge organisation and semantic services** Exploring, developing and applying semantic grid technologies, semantics-aware services, protocols for the exchange of metadata, and the use of these to automate the creation of service workflows and virtual organisations

Ανάγκη για Ψηφιακή Επιμέλεια

- Κίνδυνος απώλειας ή αχρηστίας σημαντικών επενδύσεων σε ψηφιακό υλικό
- Ταχύτατα αυξανόμενοι όγκοι δεδομένων από όλα τα επιστημονικά πεδία, εκπαιδευτικό υλικό, αρχειακό υλικό και ποικίλα πολιτισμικά τεκμήρια
- Ασύμφορο ή και αδύνατο ένας φορέας να αναπτύξει μόνος τις απαραίτητες δράσεις
- Ανάγκη συνεργασίας σε ευρύτερα σχήματα με τη συμμετοχή εξειδικευμένων επιστημονικών μονάδων
- Παγκόσμια πραγματικότητα η υστέρηση φορέων ως προς την μακροπρόθεσμη διατήρηση και αξιοποίηση του ψηφιακού υλικού

Διεργασίες Ψηφιακής Επιμέλειας



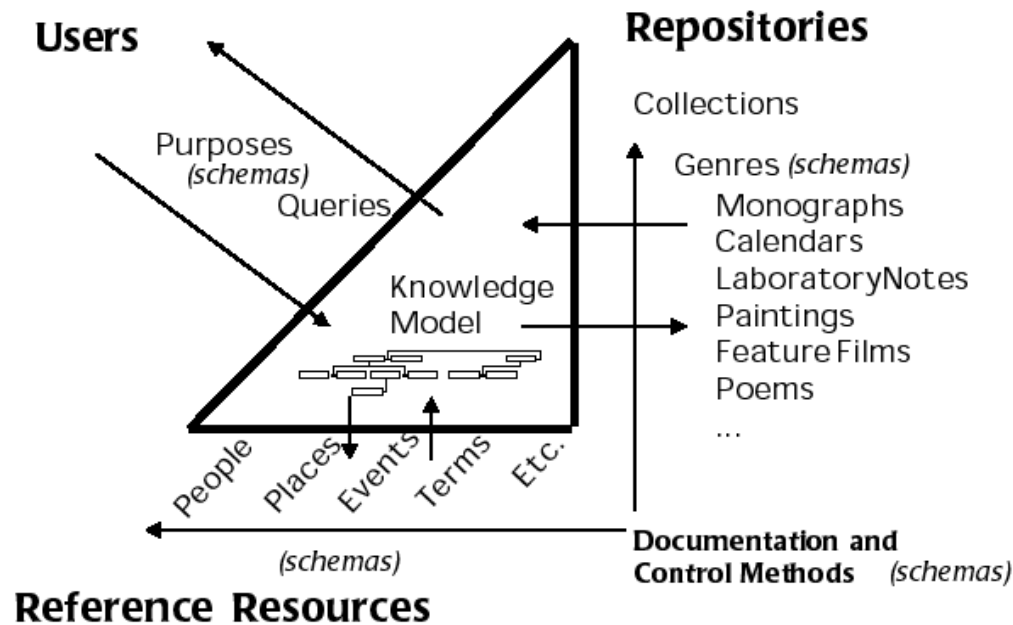
Επιμέλεια ψηφιακών πολιτισμικών πληροφοριών

- Εφεκτικές συνθήκες
 - Ανάπτυξη της πρότυπης οντολογίας πολιτισμικής τεκμηρίωσης CIDOC CRM (ISO 21127)
 - Σταδιακή σύγκλιση των πληροφοριακών λειτουργιών μουσείων, βιβλιοθηκών και αρχείων
- Ευρύτερη αντίληψη του κύκλου ζωής των ψηφιακών πόρων
 - Αποτίμηση
 - Εμπειρίες χρήσης
- Σημασιολογική προσέγγιση του πλαισίου αναφοράς
 - Μοντελοποίηση σκοπού και χρήσης
 - Μοντελοποίηση πεδίου

Μείζονες προκλήσεις

- Πολλαπλή ερμηνεία και συμφραζόμενα
 - Ερμηνευτικές κοινότητες : γλωσσικές, πολιτισμικές, επιστημονικές, ...
 - Μη «ορθοτομικές» ερμηνείες
- Μεταβλητή - εξελισσόμενη σημασία
 - Πρόοδος της έρευνας
 - Αναδιοργάνωση συλλογών και ανανοηματοδότηση συμβόλων
- Στήριξη των γνωσιακών διεργασιών

Παράγοντες των πληροφοριακών διεργασιών



Κατά D. Bearman & J. Trant,
<http://www.archimuse.com/papers/ukoln98paper/index.html>

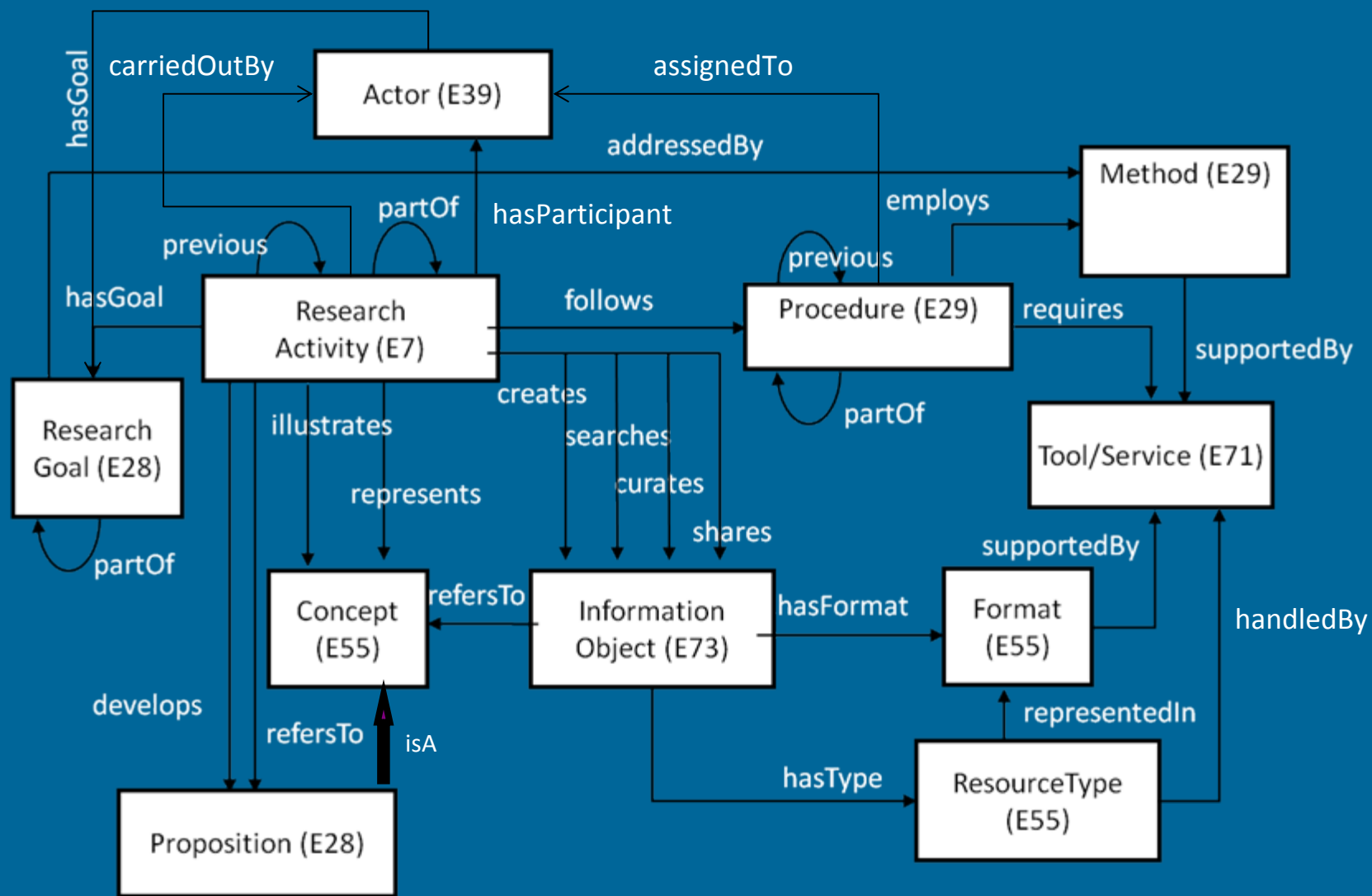
Μοντέλο ερευνητικής δραστηριότητας

- Βασίζεται σε καθιερωμένα εννοιολογικά μοντέλα της δραστηριότητας, όπως η έννοια της *δραστηριότητας (activity)* στην οντολογία CIDOC CRM και η έννοια της *διεργασίας* στα μοντέλα επιχειρησιακών συστημάτων.
- Συνιστά κατ' ουσίαν μία υπόθεση προς έλεγχο μέσω εμπειρικών μελετών, οι οποίες επίσης επιτρέπουν την εξειδίκευση των εννοιών του μοντέλου.

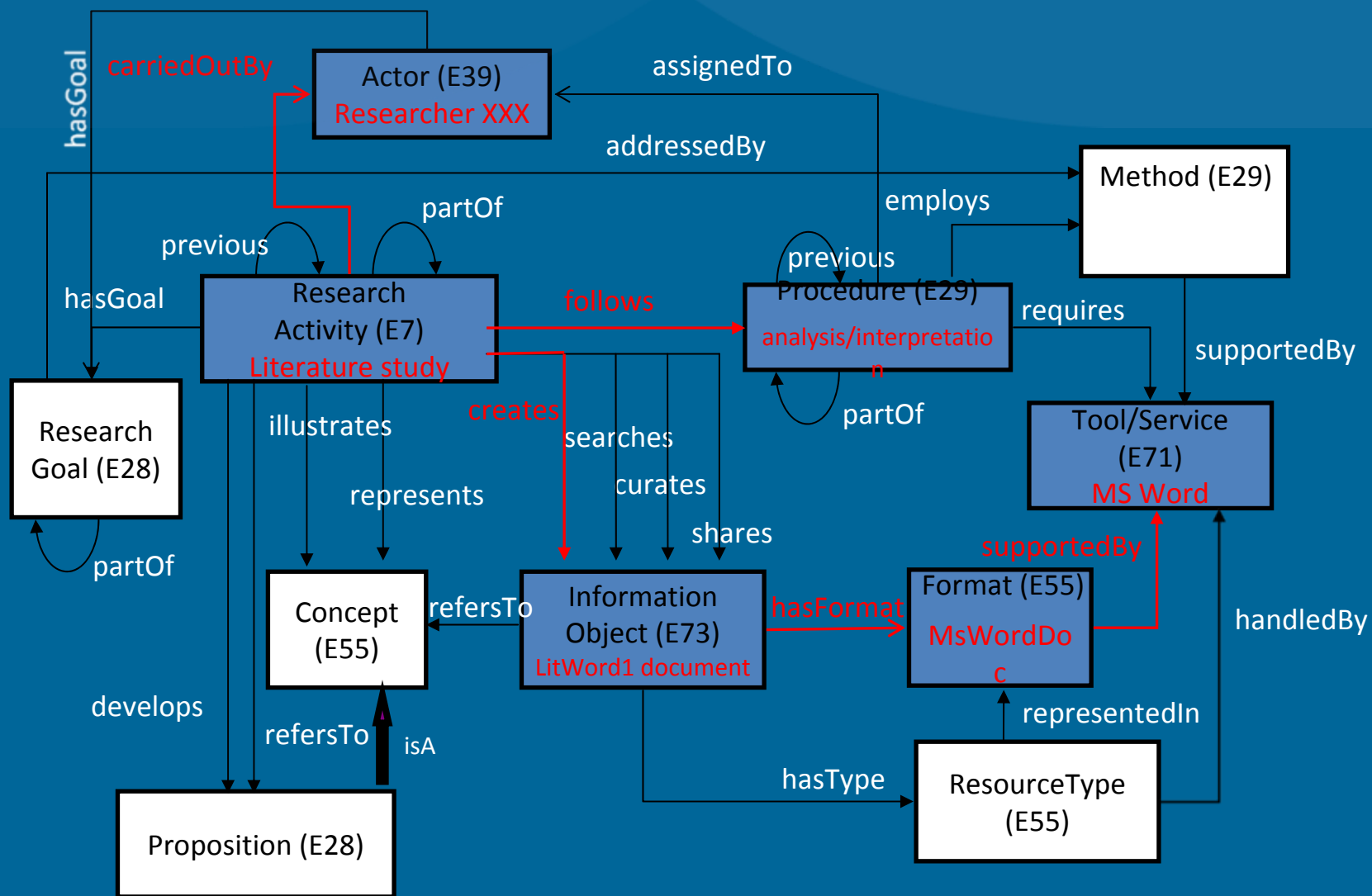
Ερευνητική δραστηριότητα

- Σκόπιμη διεργασία
- Εμπλέκει οντότητες διαφορετικών ειδών:
- **Υλικά αντικείμενα, φυσικά ή τεχνητά**
 - Αντικείμενα ή εργαλεία της εργασίας
- **Νοητικά αντικείμενα**
 - Έννοιες που δημιουργούνται, παριστάνονται, αναφέρονται
 - Λογικές προτάσεις που διατυπώνονται, αποδεικνύονται, αντικρούονται, κλπ.
- **Πληροφοριακά αντικείμενα**
 - Ειδική κατηγορία νοητικών αντικειμένων με αντίστοιχους υλικούς φορείς
 - Αναφέρονται σε και παριστάνουν υλικά και νοητικά αντικείμενα
 - Δημιουργούνται, αναζητούνται, κοινωνούνται και γίνονται αντικείμενο επιμέλειας κατά τη διάρκεια ερευνητικών δραστηριοτήτων

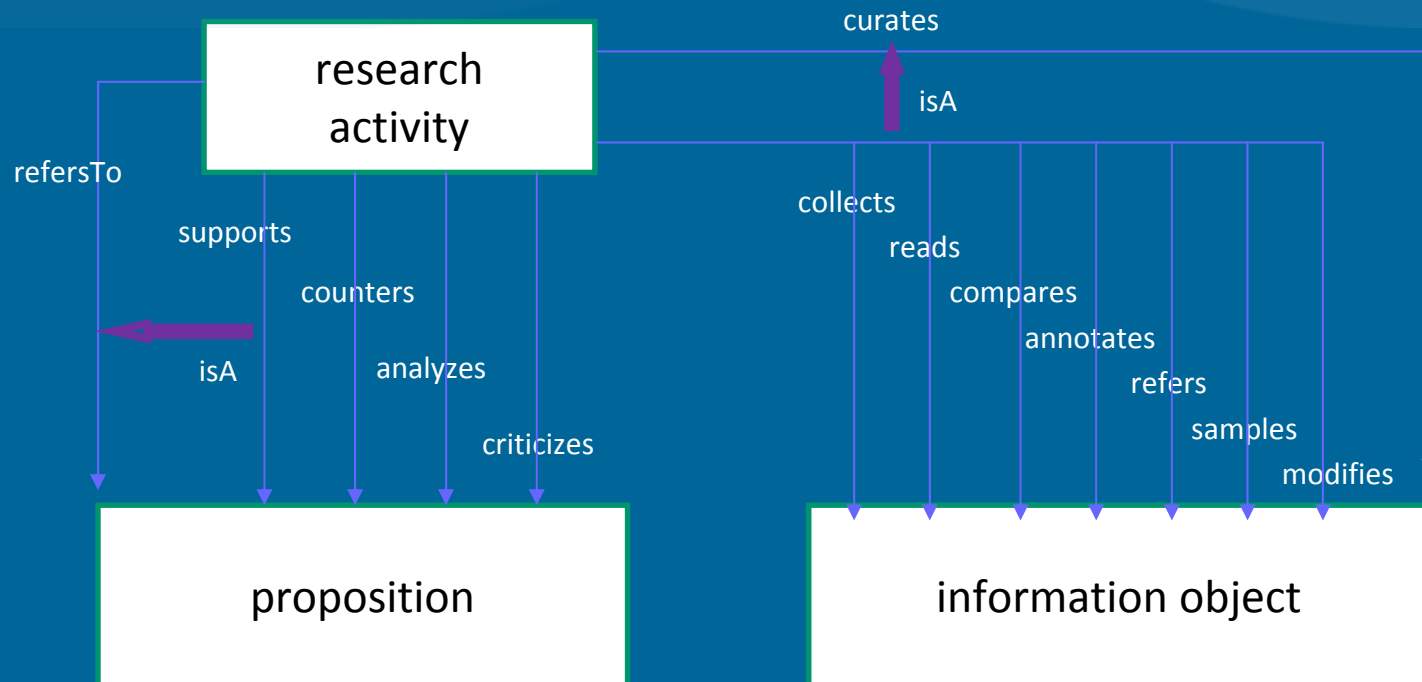
Μοντέλο ερευνητικής δραστηριότητας



Μία πραγμάτωση ερευνητικής δραστηριότητας



«Πρωτογενείς» ενέργειες (scholarly primitives)



Αντιμετώπιση της ετερογένειας

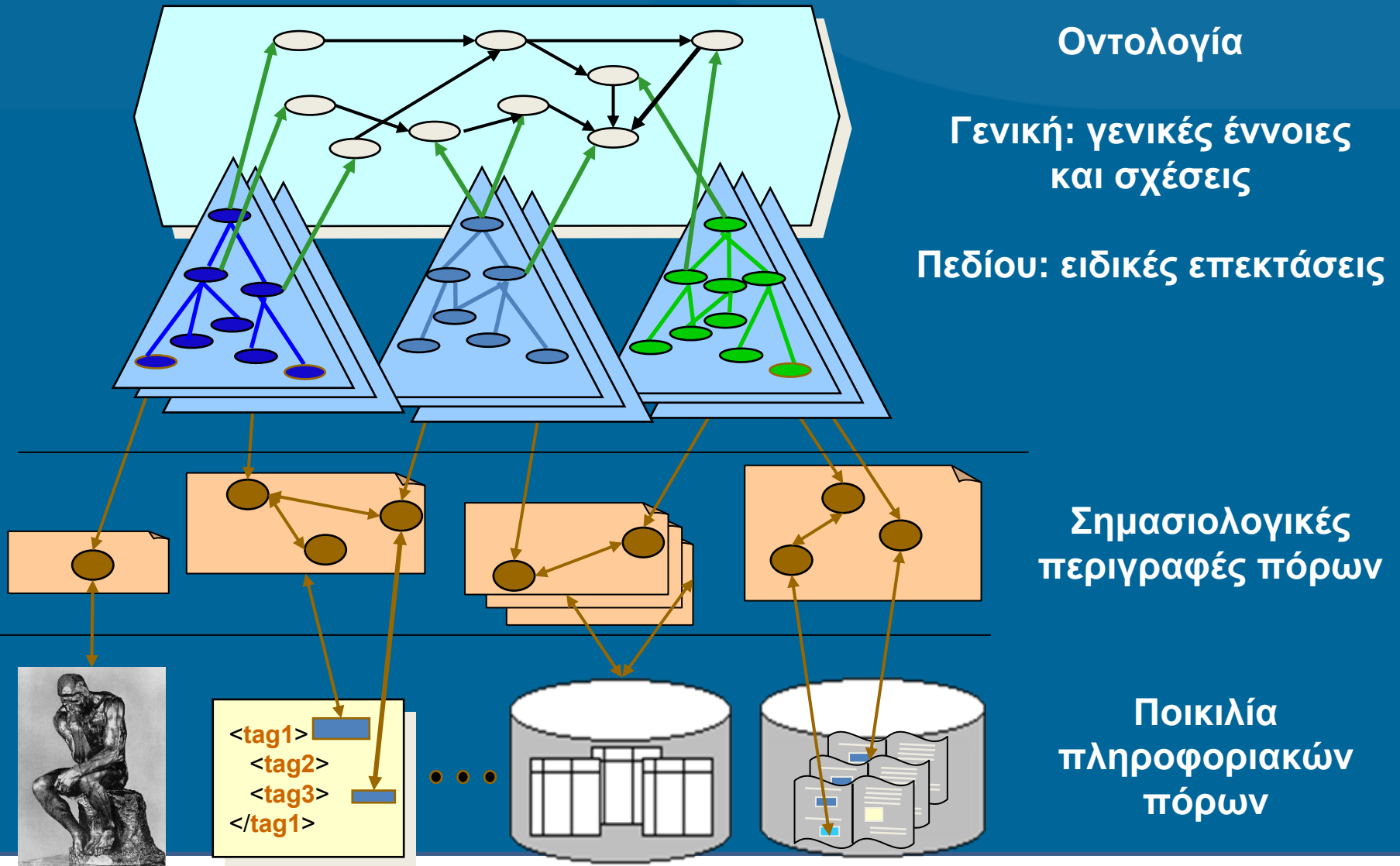
Συντακτική ετερογένεια:

- Πλατφόρμες και λειτουργικά συστήματα
- Παραστάσεις δεδομένων και συμβάσεις πρόσβασης

Σημασιολογική ετερογένεια: η δυσχερέστερη στην αντιμετώπιση

- Προσέγγιση:
 - Διατήρηση της ποικιλίας και της τοπικότητας
 - Προώθηση συμφωνιών και σημασιολογικής συμβατότητας
- Μέσα:
 - Παράσταση γνώσεων
 - Οντολογίες και διαχείριση ορολογίας
 - Πλαίσια αναφοράς / συμφραζόμενα

Σημασιολογική διαλειτουργικότητα



Οντολογία

Γενική: γενικές έννοιες
και σχέσεις

Πεδίου: ειδικές επεκτάσεις

Σημασιολογικές
περιγραφές πόρων

Ποικιλία
πληροφοριακών
πόρων

Συμπέρασμα

Προϋποθέσεις αποτελεσματικότητας

- Διασφάλιση συμμετοχής
 - Εξειδικευμένη χρησιμότητα (όχι καθολικότητα)
 - Ενσωμάτωση των ψηφιακών εργαλείων στις ερευνητικές πρακτικές
 - Επιχειρησιακή απλότητα και κλιμάκωση
 - Νομική κατοχύρωση
- Οικονομική βιωσιμότητα
- Ποιότητα και βιωσιμότητα υλικού
 - Ψηφιακή επιμέλεια
- Αξιοπιστία και κλιμάκωση υπηρεσιών
 - Τυποποίηση
 - Κατανεμημένη ανάπτυξη
 - Ανοικτά συστήματα