



Αναπαράσταση Γνώσης και Αναζήτηση στον Σημασιολογικό Ιστό

Αλέξανδρος Βαλαράκος †‡
(alexv@iit.demokritos.gr)
(alexv@aegean.gr)

† Υποψήφιος Διδάκτορας Τμήμα Μηχανικών Υπολογιστικών και Πληροφοριακών Συστημάτων. Παν/μίου Αιγαίου.

‡ Συνεργαζόμενος Ερευνητής (Υπότροφος) Ινστιτούτο Πληροφορικής και Τηλεπικοινωνιών, Εργαστήριο Λογισμικού και Γνώσης, Ε.Κ.Ε.Φ.Ε Δημόκριτος.

- Αναπαράσταση πληροφορίας στον Σημασιολογικό Ιστό
- Οντολογία
- (Γλώσσα αναπαράστασης) OWL
- (Εργαλείο δημιουργίας οντολογιών) Protégé
- Αναζήτηση στο Σημασιολογικό Ιστό

- Ο σημασιολογικός ιστός υπόσχεται ένα περιεχόμενο *επεξεργάσιμο (κατανοητό)* από τους Η/Υ.
- Όμως, η αποτελεσματικότητα αυτού προϋποθέτει πρόσβαση στο **νόημα** του περιεχομένου.

À∇ϕϕζw^∧^ÉéÉéâÀ

W ÉéâÀWÀεδδ^WÀ∇ϕϕζw^∧^ÉéÉéâÀW
ÉéâÀWÀεδδ^ééδδWÀ∇ϕϕζw^∧^ÉéÉéâÀ
WÀ∇ϕϕζw^∧^ÉéÉéâÀWéâÀWÀεδδ^Éé
âÀWÀεδδ^éâÀWÀ∇ϕϕζw^∧^ÉéÉé
âÀWÉéâÀWÀεδδ^WÀ∇ϕϕζw^∧^ÉéÉéâÀ
WÉéâÀWÀεδδ^é^w^∧^ÉéâÀWÀεδδ^O
OWÀ∇ϕϕζw^∧^∇ϕϕζw^∧^∇ϕϕζw^∧^ÉéÉéâÀW
ÉéâÀWÀεδδ^éw^∧^ÉéÉéâÀWÉéâÀWÀ
εδδ^WÀ∇ϕϕζw^∧^ÉéÉéâÀWÉéâÀWÀ
εδδ^ééé^w^∧^ÉéÉéâÀWÉéâÀWÀεδδ^éW
ÀWÀ∇ϕϕζw^∧^ÉéÉéâÀWÉéâÀWÀεδδ^é
WÀ∇ϕϕζw^∧^ÉéÉéâÀWÉéâÀWÀεδδ^é
∇ϕϕζw^∧^ÉéÉéâÀWÉéâÀWÀεδδ^éWÀ
∇ϕϕζw^∧^ÉéÉéâÀWÉéâÀWÀεδδ^éWÀ
∇ϕϕζw^∧^ÉéÉéâÀWÉéâÀWÀεδδ^éWÀ
∇ϕϕζw^∧^ÉéÉéâÀWÉéâÀWÀεδδ^éWÀ
∇ϕϕζw^∧^∇ϕϕζw^∧^ÉéÉéâÀWÉéâÀWÀ
εδδ^éw^∧^ÉéâWÀ∇ϕϕζw^∧^ÉéÉéâÀW
ÉéâÀWÀεδδ^é|

XML tags για τη σύλληψη του νοήματος



Ανέφεραν ότι έφαγαν

Έφαγαν ένα κομμάτι από ένα φρέσκο ψωμί που έφτιαξε ο πατέρας τους. Το ψωμί ήταν πολύ νόστιμο και τους άρεσε πολύ. Μετά από αυτό, πήγαν να περπατήσουν στον κήπο. Ο κήπος ήταν πολύ όμορφος και είχαν πολλούς λουλούδια. Τα παιδιά έπαιζαν για ώρες και ήταν πολύ χαρούμενα. Όταν ήλθε η ώρα να φάει, τους έφερε η μητέρα τους ένα μεγάλο πιάτο με φακές. Τους άρεσε πολύ και τους έδωσε ένα ποτήρι νερό. Μετά από αυτό, πήγαν να κοιμηθούν. Η νύχτα ήταν πολύ ήσυχη και οι όνειρά τους ήταν πολύ όμορφα. Το πρωί, όταν ξύπνησαν, τους έφερε η μητέρα τους ένα καλό πρωινό. Τους άρεσε πολύ και τους έδωσε ένα ποτήρι χυμό. Μετά από αυτό, πήγαν να πάνε στην σχολή. Η σχολή ήταν πολύ όμορφη και τους άρεσε πολύ. Το βράδυ, όταν ήλθε η ώρα να φάει, τους έφερε η μητέρα τους ένα μεγάλο πιάτο με φακές. Τους άρεσε πολύ και τους έδωσε ένα ποτήρι νερό. Μετά από αυτό, πήγαν να κοιμηθούν. Η νύχτα ήταν πολύ ήσυχη και οι όνειρά τους ήταν πολύ όμορφα. Το πρωί, όταν ξύπνησαν, τους έφερε η μητέρα τους ένα καλό πρωινό. Τους άρεσε πολύ και τους έδωσε ένα ποτήρι χυμό. Μετά από αυτό, πήγαν να πάνε στην σχολή. Η σχολή ήταν πολύ όμορφη και τους άρεσε πολύ.

Title

Abstract

Introduction

Νόημα μέσω XML μόνο για ανθρώπους



À∇ϕϕζw^∧^ÉéÉéâÀ

W ÉéâÀWÀεش∧WÀ∇ϕϕζw^∧^ÉéÉéâÀW
ÉéâÀWÀεش∧ééشWÀ∇ϕϕζw^∧^ÉéÉéâÀ
WÀ∇ϕϕζw^∧^ÉéÉéâÀWéâÀWÀεش∧éÉé
âÀWÀεش∧éâÀWÀ∇ϕϕWÀ∇ϕϕζw^∧^ÉéÉé
âÀWÉéâÀWÀεش∧éWÀ∇ϕϕζw^∧^ÉéÉéâÀ
WÉéâÀWÀεش∧é^ζw^∧^ÉéâÀWÀεش∧éO
OWÀ∇ϕϕWÀ∇ϕϕζWÀ∇ϕϕζw^∧^ÉéÉéâÀW
ÉéâÀWÀεش∧éw^∧^ÉéÉéâÀWÉéâÀWÀ
ش∧WÀ∇ϕϕζw^∧^ÉéÉéâÀWÉéâÀWÀ
εش∧ééé^ζw^∧^ÉéÉéâÀWÉéâÀWÀεش∧éW
ÀWÀ∇ϕϕζw^∧^ÉéÉéâÀWÉéâÀWÀεش∧é
WÀ∇ϕϕζw^∧^ÉéÉéâÀWÉéâÀWÀεش∧é
∇ϕϕζw^∧^ÉéÉéâÀWÉéâÀWÀεش∧éWÀ
∇ϕϕζw^∧^ÉéÉéâÀWÉéâÀWÀεش∧éWÀ
∇ϕϕζw^∧^ÉéÉéâÀWÉéâÀWÀεش∧éWÀ
∇ϕϕζWÀ∇ϕϕζw^∧^ÉéÉéâÀWÉéâÀWÀ
εش∧éw^∧^ÉéâWÀ∇ϕϕζw^∧^ÉéÉéâÀW
ÉéâÀWÀεش∧é|

ΚΜЖ

ΓΑЖЧΘ

ЖЧΘЛЪЬ

- Ανομοιομορφίες
 - στην γλώσσα επισημείωσης
 - στους όρους (μεταδεδομένα)
 - διαφορετικοί τρόποι εμφάνισης περιεχομένου
- Πρόβλημα αποκρυπτογράφησης...

- Στην γλώσσα
 - Κοινή γλώσσα αναπαράστασης π.χ. RDF
- Χρήση κοινών όρων
 - «Αγγείο» και όχι «σκεύος»
 - «Μ.Αλέξανδρος» και όχι «Μέγας Αλέξανδρος» κτλ.

- Σημασιολογία σημαίνει «έχων νόημα»
- Ερωτήματα που προκαλούνται στην αναπαράσταση:
 - Ποιος έχει νόημα
 - Που βρίσκεται
 - Πως είναι
 - Πως χρησιμοποιείται
- Λύνει το πρόβλημα της σύλληψης του νοήματος;
=> Η Συμβολική λογική δείχνει τον δρόμο.

Το φάσμα της σημασιολογίας των μεταδεδομένων...



υπονοούμενο

Τυπικό για
ανθρώπους

Τυπικό για Η/Υ

- Όσο πιο δεξιά τόσο :
 - περισσότερο προσβάσιμο και αξιοποιήσιμο από Η/Υ
 - λιγότερο αμφισβητήσιμο
 - πιθανότερο να επαναχρησιμοποιηθεί
 - ευκολότερα επεκτάσιμο
 - περισσότερο σθεναρό σε αλλαγές
 - δυσκολότερο να υλοποιηθεί

- Αποκάλυψε κείμενα στα οποία υπάρχουν μουσεία με εκθέματα ελληνιστική περιόδου.



Κείμενο

...το *<μουσείο id=12>*
εθνικό αρχαιολογικό
μουσείο *</μουσείο>*
βρίσκεται στην οδό ...

Σημασιολογική
σήμανση

<μουσείο id=12> εθνικό
αρχαιολογικό μουσείο *</μουσείο>*

<μουσείο id=12>
hasExhibitsfromPeriod ελληνιστική
...

- Η/Υ μπορούν να μάθουν κάτι για το νόημα
 - Χρειάζεται να γνωρίζει το νόημα της ίδια της λέξης;
 - Είναι αρκετό να γνωρίζει τα χαρακτηριστικά της;
 - Τις σχέσεις που έχει με άλλες λέξεις;
 - Δεν είναι ποτέ πλήρης η περιγραφή της!
 - Υπάρχει καλύτερος τρόπος αναζήτησης από τον συμβατικό;

- Η σημασιολογία στηρίζεται στις τυπικές εκφραστικές δυνατότητες της γλώσσας αναπαράστασης της πληροφορίας
 - π.χ. *XML, RDF, OWL ...*
- Το να δείχνεις σε ένα πόρο δεν είναι αρκετό για την πρόσβαση στο νόημα ενός περιεχομένου.
 - π.χ. Dublin Core λεξικό.
- Η λύση βρίσκεται σε πιο τυπικά μοντέλα αναπαράστασης => **Οντολογίες**



Η οντολογία ορίζει τις λέξεις και έννοιες που χρησιμοποιούνται για να **περιγράψουν** και να **αναπαραστήσουν** μια θεματική περιοχή.

- **Αναπαράσταση:** κωδικοποίηση περιγραφής έτσι ώστε να χρησιμοποιηθεί από άλλους.
- **Περιγραφή:** όροι και προτάσεις που συνδυάζουν όρους για να εκφράσουν σχέσεις μεταξύ αυτών.

- Αναπαράσταση της θεματικής περιοχής
 - έννοιες,
 - χαρακτηριστικά,
 - σχέσεις και
 - στιγμιότυπα
- Αντιπροσωπεύονται από όρους
Οντολογία => Λεξικό όρων

Τυπική και
ρητή δήλωση

έννοιες, χαρακτηριστικά,
σχέσεις, κανόνες

μίας διαμοιραζόμενης

ομοφωνία

αντίληψης κάποιας
Θεματικής περιοχής

Αφηρημένο **μοντέλο**

- Η οντολογία χρησιμοποιείται ως μοντέλο για το περιορισμό των πιθανών **σημασιολογικών ερμηνειών** των όρων ενός λεξικού.
- Σημασιολογική ερμηνεία είναι μια απεικόνιση μεταξύ δεδομένων και ενός μοντέλο αντικειμένων ενός θεματικού πεδίου σύμφωνα με το προκείμενο νόημα των αντικειμένων αυτών και των σχέσεων μεταξύ τους.

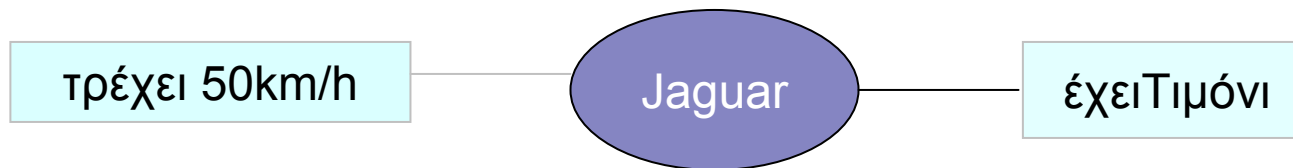
Θεματικό πεδίο: Μαθηματικά $\{, \}$ \rightarrow όρια συνόλου

Σύμβολα: $\{, \}$, $\{\}$, α , β , γ $\{\}$ \rightarrow κενό σύνολο

α , β , γ \rightarrow στοιχεία συνόλου

$\{\alpha, \beta, \gamma\}$ \rightarrow στοιχεία απαρτίζουν το σύνολο

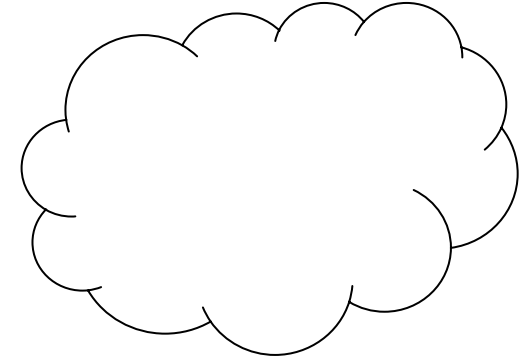
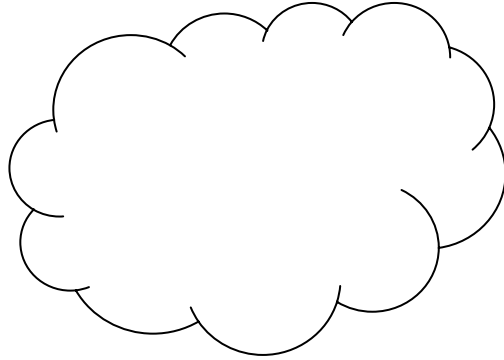
- Μάρκα αυτοκινήτου
- Ζώο
- ???



- 400 π.Χ. Αριστοτέλης
 - ορισμός κατηγοριών βάση γένους και διαφοροποίησης.
 - ταξινομίες
 - 10 γενικές κατηγορίες του κόσμου
 - π.χ. ενέργεια, σχέση, χρονικό, ποσότητα κτλ.

Οντολογία διευκολύνει την επικοινωνία

Αντίληψη
Κόσμος



Οντολογία



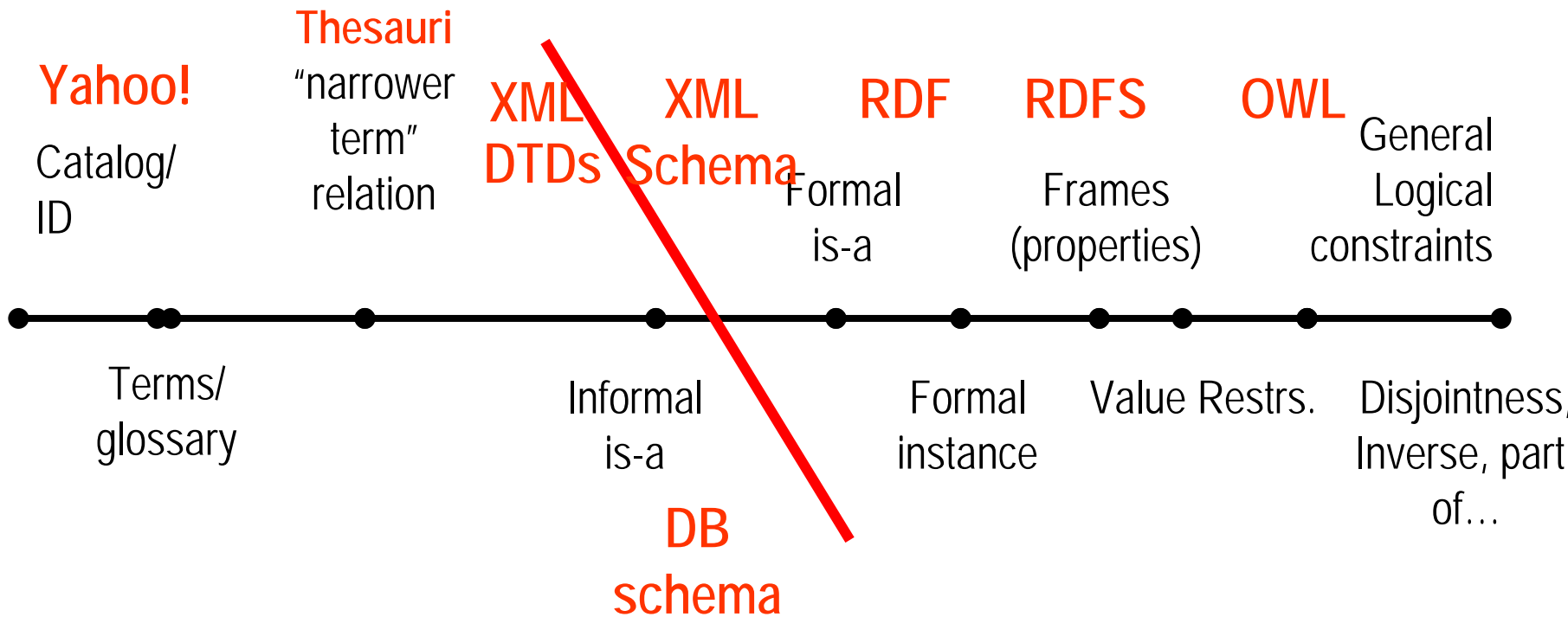
Γλώσσα



- Δίνει νόημα στα μεταδεδομένα
- Συγκρατεί γνώση
- Δημιουργεί μία διαμοιραζόμενη αντίληψη πραγμάτων μεταξύ μηχανών, ανθρώπων, μηχανών και ανθρώπων
- Κάνει την γνώση μηχανικώς επεξεργάσιμη

Η «*ραχοκοκαλιά*» του Σημασιολογικού Ιστού

Είδη οντολογίας: βάση γλώσσα αναπαράστασης



- Έννοιες (concepts) – κατηγορίες
- Στιγμιότυπα (instances) – οντότητες
- Χαρακτηριστικά (attributes)
- Σχέσεις (relations)
 - εννοιών (is-a)
 - στιγμιοτύπων
- *Ιδιότητες σχέσεων (OWL)*
 - συμμετρική
 - αντιμεταθετική
 - ...

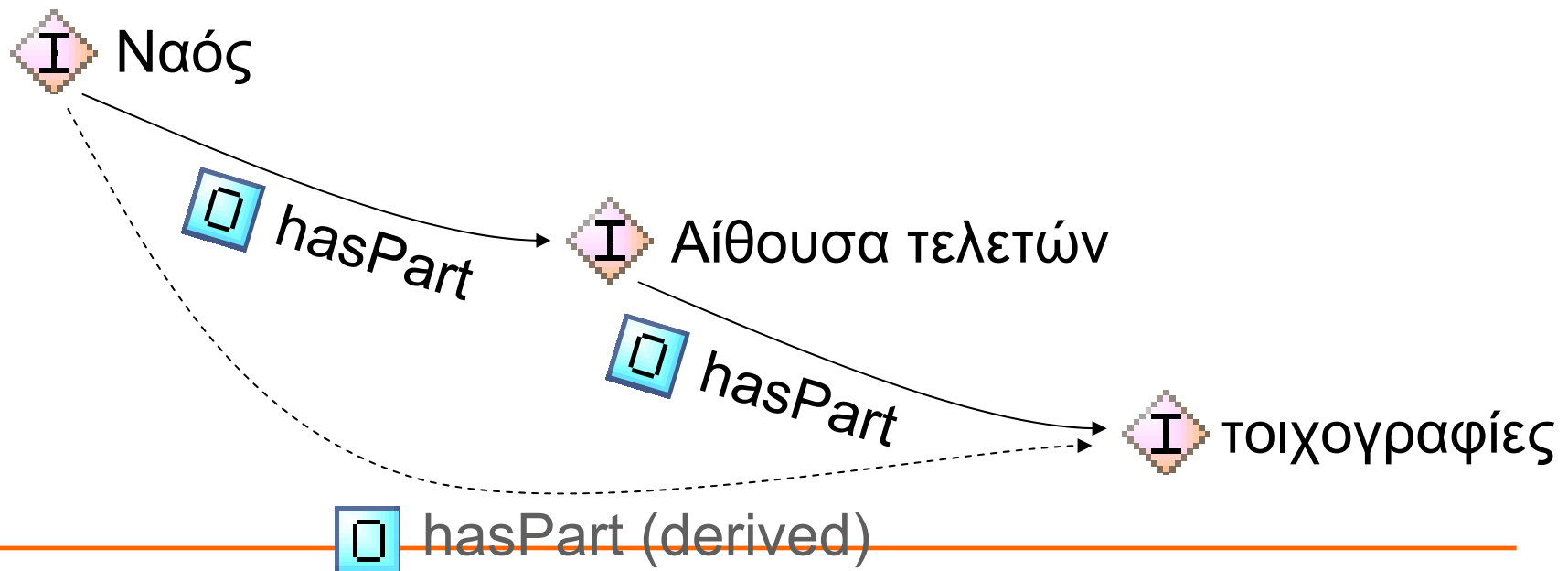
- Protégé
- ...

Αναζήτηση πληροφορίας: Πόσο δύσκολο είναι;

- Αναζήτηση σε κείμενα ή εικόνες «πήλινο αγγείο»
 - όχι τόσο δύσκολο για αναζήτηση με λέξεις κλειδιά
- «εκθέματα που βρέθηκαν από τον X την περίοδο 1900-2000»
 - δυσκολεύουν πολύ...
- Προβλήματα
 - διαφορετικό τρόπο γραφής των λέξεων
 - συνώνυμα
 - υπερώνυμα

Η οντολογία λύση στο πρόβλημα

- κοινό λεξικό ορολογίας
- Αναζήτηση με σημασιολογικό εμπλουτισμό πλούσιο σε ιδιότητες.



- Web Ontology Language (OWL)
- Επίσημο W3C Standard από τον Φεβρουάριο 2004
- Βασίζεται στον πρόγονό της (DAML+OIL)
- Βασίζεται στο RDF(S)
- Βασίζεται στην λογική



Γλώσσες αναζήτησης πληροφορίας

Γλώσσα Αναπαράστασης Δεδομένων	Δυνατότητα Περιγραφής	Γλώσσα Αναζήτησης
<i>XML</i>	Structured Documents	XQuery, XPath
<i>RDF</i>	data model for objects	RDQL, RQL, Versa, Squish
<i>OWL</i>	data model + Relations	OWL-QL, Jena
<i>SRWL</i>	Data model + relations + rules	

- www.semanticweb.org
- www.w3c.org
- www.ontoweb.org
- <http://knowledgeweb.semanticweb.org>
- www.semanticweb.gr
- <http://www.w3.org/2001/sw/Europe/>



ΕΥΧΑΡΙΣΤΩ!

Οι διαφάνειες θα είναι διαθέσιμες στο ακόλουθο URL:

<http://www.iit.demokritos.gr/~alexv>