

QUERYING TRUST IN RDF DATA WITH tSPARQL

A. TOLIA D. ROIDIS E. MAKRYMANOLAKI

S. KONSTANTOPOULOS



Αντικείμενο και Θεωρητικό Υπόβαθρο

Το άρθρο μας :

- Πραγματεύεται την πρόκληση διαχείρισης αξιοπιστίας των RDF δεδομένων.
- Αναλύει προκλήσεις του ανοιχτού ιστού δεδομένων:

Μεγάλος όγκος πηγών → Αναξιόπιστα δεδομένα → Λανθασμένα συμπεράσματα.

Έλλειψη τυποποιημένου τρόπου αξιολόγησης αξιοπιστίας δεδομένων.

- Προηγούμενες προσεγγίσεις για την επίλυση των προκλήσεων αυτών.
- 📌 Προτείνει το **εSPARQL** ως λύση για την αξιόπιστη διαχείριση των RDF δεδομένων.



Προηγούμενες Προσεγγίσεις & Αξιοπιστία

 **Marsh** → υπολογισμός αξιοπιστίας τριπλετών ως τιμή $[-1,1]$.

 **Web of Trust** → εμπιστοσύνη στους χρήστες, όχι στα δεδομένα.

 **TriQL.P** → RDF επέκταση με έμφαση στην προέλευση δεδομένων.

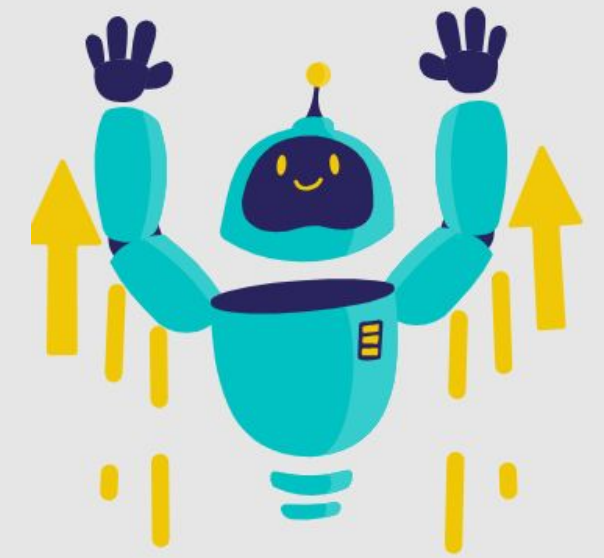
 **Mazzieri** → "Fuzzy RDF" – συσχετισμός συμμετοχής με αξιοπιστία.

 **Richardson et al.** → Προσωπικές αξιολογήσεις user στο διάστημα $[0,1]$, χωρίς τελική αξιοποίηση.

 **TRELLIS** → Αξιολόγηση πηγών, όχι τριπλετών.

 **IWTrust** → Αξιολόγηση εμπιστοσύνης και ακεραιότητας χρηστών.

 **FilmTrust** → Ανάλυση κριτικών και αξιολογήσεων για ταινίες, χωρίς επεκτασιμότητα σε άλλο πεδίο εφαρμογής.



Κύρια Ιδέα του Άρθρου

Προτείνεται η trust SPARQL, μια επέκταση της SPARQL, που επιτρέπει την ενσωμάτωση εμπιστοσύνης στα ερωτήματα RDF.



Μοντέλο Εμπιστοσύνης του tSPARQL

- Τριπλέτες RDF αντιστοιχίζονται σε τιμές εμπιστοσύνης στο διάστημα $[-1,1]$
- Οι τιμές εμπιστοσύνης είναι υποκειμενικές και εξαρτώνται από τον provider της πληροφορίας
- Υπολογισμός overall εμπιστοσύνης με συναρτήσεις συγχώνευσης (trust aggregation functions)

Παρουσιάζονται δύο νέες λέξεις-κλειδιά:

🔑 **TRUST AS:** Επιτρέπει την ανάκτηση των τιμών εμπιστοσύνης.

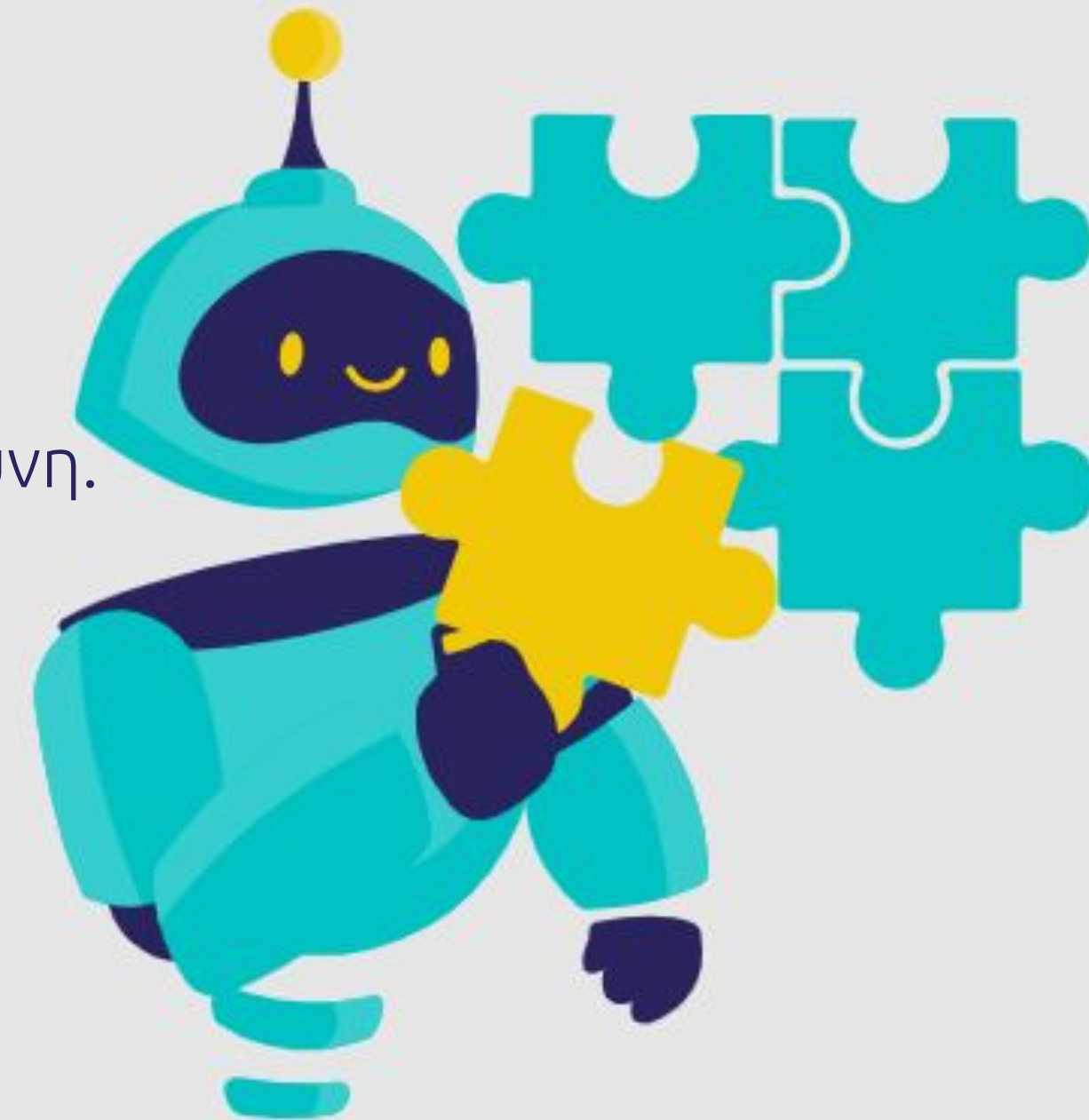
```
SELECT ?hotel ?city ?trust WHERE {  
  ?hotel rdf:type ex:Hotel .  
  ?hotel ex:locatedIn ?city .  
  TRUST AS ?trust  
}  
ORDER BY DESC(?trust)
```

🔑 **ENSURE TRUST:** Φιλτράρει τα αποτελέσματα με βάση την εμπιστοσύνη.

```
SELECT ?hotel ?city WHERE {  
  ?hotel rdf:type ex:Hotel .  
  ?hotel ex:locatedIn ?city .  
  ENSURE TRUST (0.7, 1.0)  
}
```



Επεξεργασία της SPARQL σε αλγεβρικό επίπεδο
ώστε να ενσωματώνει την εμπιστοσύνη.



Τι Συμπεραίνουν οι Δημιουργοί του tSPARQL;

tSPARQL:

- 🏗️ Μοντέλο αξιοπιστίας για RDF δεδομένα με εύκολη εφαρμογή.
- 🔗 Ανεξάρτητο από συγκεκριμένες μεθόδους/αλγορίθμους τιμών αξιοπιστίας.

Αξιολόγηση

- ⌚ Εκτέλεση: Δεν επιβαρύνθηκε σημαντικά από τις τιμές εμπιστοσύνης.
- 💾 Cache: Βελτιώνει την απόδοση τιμών.
- ⚡ Push-down περιορισμών: Μείωση χρόνου εκτέλεσης κατά 50%.
- 🔄 Query rewriting: Δραστική μείωση χρόνου εκτέλεσης.

Υλοποίηση

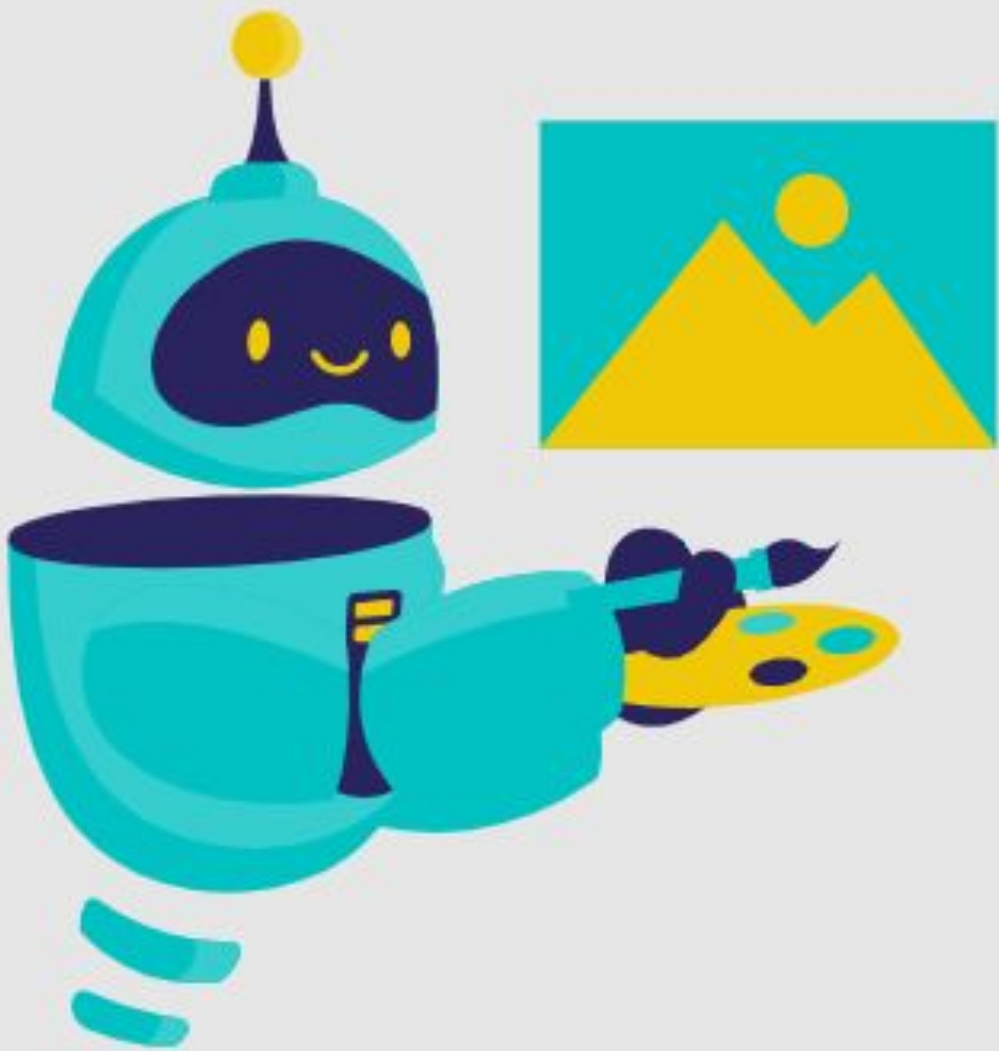
- 🚀 Επέκταση της SPARQL μηχανής ARQ.
- 📊 Χρήση BSBM Benchmark για αξιολόγηση.



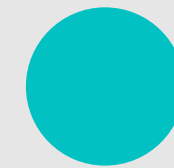
Είναι τελικά το tSPARQL η λύση που περιμέναμε;

Θετικά Σημεία:

- ✓ **Αξιοπιστία:** Έθεσε θεμέλια για μελλοντική χρήση.
- 🚀 **Επεκτάσεις:** Εισήγαγε χρήσιμα SPARQL extensions.
- 🔗 **Trust Aggregation:** 1η σχετική προσέγγιση.
- ⚡ **Απόδοση:** Βελτίωση μέσω push down.



Το μαγικό φίλτρο «εμπιστοσύνης» ή ακόμα μια αλυσίδα προβλημάτων;



● Υποκειμενικότητα στην εμπιστοσύνη

Οι trust scores εξαρτώνται από την κρίση του χρήστη

✗ Δεν συγκρίνεται με άλλες λύσεις

Χωρίς σύγκριση με TriQL.P, TRELIS, δεν γνωρίζουμε αν είναι η καλύτερη λύση



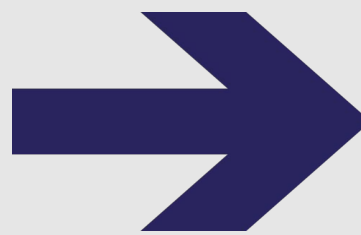
Ταχύτητα & Κλιμάκωση

Αργή απόδοση σε μεγάλα datasets



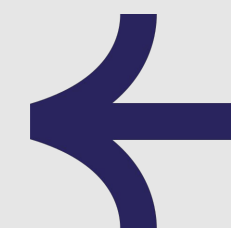
Μη επιβεβαιωμένα αποτελέσματα

Λείπουν ανεξάρτητες μελέτες για την αξιολόγηση της αποτελεσματικότητας





Το μέλλον του tSPARQL: Τι μπορεί να γίνει καλύτερο;



Τεχνικές Βελτιώσεις

 AI για εμπιστοσύνη

 Έξυπνοι αλγόριθμοι

 Caching

Πρακτική Χρήση

 Πειράματα σε

πραγματικά δεδομένα RDF

 Εφαρμογές σε use cases

Αξιολόγηση

 Σύγκριση με TriQL.P & TRELLIS

 Ανεξάρτητες μελέτες

Μελλοντικές Επεκτάσεις

 IoT & Blockchain



Το tSPARQL ξεκίνησε, αλλά... μπορεί να πάει πιο μακριά!



☀ Έφερε την έννοια της εμπιστοσύνης στα
RDF queries

Μεγάλωσε το πεδίο εφαρμογών της SPARQL

? Μπορούμε να το εμπιστευτούμε ή θέλει
βελτιώσεις;

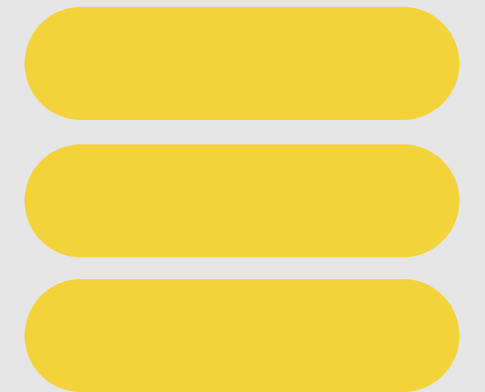
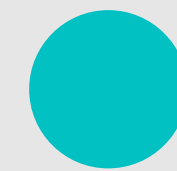
Χρειάζονται καλύτερα benchmarks και ανεξάρτητες μελέτες

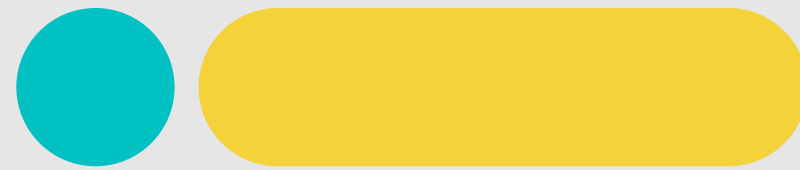
🚧 Αδυναμίες & Ελλείψεις

Υποκειμενικότητα, scalability, ταχύτητα

🎯 Στόχος

Πιο αποδοτικό και ευέλικτο για πραγματικές εφαρμογές





THANK

YO

