

HAR.S.H. – HARdware aware extreme-scale Similarity search

Αναζήτηση Ομοιότητας σε Μεγάλες Συλλογές Σειρών Δεδομένων

Συνυπολογίζοντας το Υλικό

Δελτίο Τύπου, Μάιος 2025



Οι σειρές δεδομένων έχουν εφαρμογές, σχεδόν σε κάθε τομέα της κοινωνίας, της επιστήμης και της ζωής. Εμφανίζονται ως ακολουθίες ήχου, αναπαραστάσεις εικόνων ή βίντεο, οικονομικά και τηλεπικοινωνιακά δεδομένα, δεδομένα παρακολούθησης του περιβάλλοντος και επιστημονικά δεδομένα. Οι εφαρμογές τους είναι εκτεταμένες και επηρεάζουν τομείς όπως η υγειονομική περίθαλψη, οι γεωεπιστήμες, η αστρονομία, η βιολογία και η οικονομία. Όσο η εκθετική αύξηση των δεδομένων που παράγονται από τέτοιου είδους τομείς συνεχίζεται, τόσο μεγαλύτερη είναι η ανάγκη για αποδοτική ανάλυση μεγάλων συλλογών σειρών δεδομένων πολλών διαστάσεων. Ταυτόχρονα, η εξάπλωση της μηχανικής μάθησης έχει επιτρέψει την εκτέλεση απαιτητικών εργασιών πάνω σε δεδομένα συγκεκριμένων μορφών. Ωστόσο, οι μοντέρνες εφαρμογές οδηγούν στην αναγκαιότητα επεξεργασίας δεδομένων που δεν είναι στην ίδια μορφή (π.χ. κείμενα, εικόνες, βίντεο). Η επεξεργασία τέτοιων συλλογών πολυτροπικών δεδομένων ενέχει πολλές ερευνητικές προκλήσεις αλλά είναι απαραίτητη καθώς μια ενοποιημένη αναπαράσταση των διαφορετικών τύπων δεδομένων θα επιτρέψει την επεξεργασία πληροφορίας σε σημαντικά ευρύτερο φάσμα.

Το HAR.S.H. φιλοδοξεί να παρέχει καινοτόμες λύσεις και να έχει σημαντικό αντίκτυπο στις ακόλουθες κατευθύνσεις:

- **Κωδικοποίηση Πολυτροπικών Δεδομένων.** Αντιμετώπιση της ανάγκης κωδικοποίησης αντικειμένων υψηλών διαστάσεων, όπως εικόνες και βίντεο, επιτρέποντας στις σειρές δεδομένων να αποτυπώνουν πολυτροπικά δεδομένα.
- **Εκμετάλλευση της πλήρους υπολογιστικής ισχύος των σύγχρονων υπολογιστικών πλατφορμών.** Αξιοποίηση όλων των υπολογιστικών μονάδων των σύγχρονων υπολογιστικών πλατφορμών, συμπεριλαμβανομένων των πολλαπλών κόμβων, των πυρήνων κάθε κόμβου και των επιταχυντών.
- **Σχεδίαση και υλοποίηση αλγορίθμων και δομών δεδομένων με επίγνωση του υλικού.** Μελέτη της επίδρασης που μπορεί να έχει η αξιοποίηση μοντέρνων και αναδυόμενων τεχνολογιών υλικού στην επεξεργασία μεγάλου όγκου σειρών δεδομένων με δυναμικό ή εύρωστο τρόπο.

Το HARSH θα αποδείξει την αξία της τεχνολογίας που θα παράγει μέσω των εξής πιλοτικών εφαρμογών:

1. **Αναζήτηση παρόμοιων εγγράφων** σε μεγάλες βάσεις δεδομένων.
2. **Ανάλυση φωτογραφιών για εξατομίκευση ταξιδιωτικών προτάσεων.**
3. **Ανάλυση γνώμης πολιτών** μέσα από δεδομένα κοινωνικής δικτύωσης.

Το HAR.S.H. ξεκίνησε τον Απρίλιο του 2025 και θα είναι ενεργό μέχρι και τον Ιούνιο του 2026. Χρηματοδοτείται από το Υπουργείο Παιδείας, Θρησκευμάτων και Αθλητισμού, στο πλαίσιο του Εθνικού Σχεδίου Ανάκαμψης και Ανθεκτικότητας "Ελλάδα 2.0" με τη χρηματοδότηση της Ευρωπαϊκής Ένωσης.

ΠΕΡΙΣΣΟΤΕΡΕΣ ΠΛΗΡΟΦΟΡΙΕΣ



<https://harsh-project.eu/>



<https://www.linkedin.com/company/harsh-project/>



<https://www.facebook.com/harshproject>

ΣΤΟΙΧΕΙΑ ΕΠΙΚΟΙΝΩΝΙΑΣ



Καθηγήτρια Παναγιώτα Φατούρου

Τμήμα Επιστήμης Υπολογιστών

Πανεπιστήμιο Κρήτης

Τηλ.: +30 697 399 1277,

Email: faturu@csd.uoc.gr

URL: www.csd.uoc.gr/~faturu

ΣΥΝΕΡΓΑΖΟΜΕΝΑ
ΙΝΣΤΙΤΟΥΤΑ:

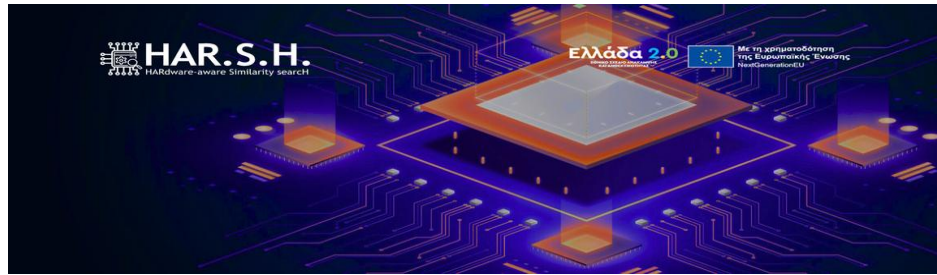


ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
UNIVERSITY OF CRETE



Το έργο HAR.S.H. (με αρ. ΥΠ3ΤΑ-0560901) υλοποιείται στο πλαίσιο του Εθνικού Σχεδίου Ανάκαμψης και Ανθεκτικότητας «Ελλάδα 2.0» με τη χρηματοδότηση της Ευρωπαϊκής Ένωσης – NextGenerationEU

HAR.S.H.– HARdware aware extreme-scale Similarity search Press Release, May 2025



Ordered sequences of data points, known as data series, are among the most prevalent and versatile forms of data, spanning virtually every scientific and social field. They appear as audio sequences, shape and image data as well as financial, telecommunications, environmental monitoring and scientific records. Their applications are extensive, impacting areas such as healthcare, earth sciences, astronomy, biology, and economics. As the exponential growth of data from these fields continues, there is an urgent need for effective analysis of huge collections of high-dimensional data sets. Meanwhile, the proliferation of machine learning has enabled performing demanding tasks, albeit on data of specific formats. However, modern applications require processing multimodal data, i.e., data that is not in the same format (e.g., text, images and video). The efficient processing of such heterogeneous data collections involves significant research challenges, but achieving unified representations is essential to unlock broader, more powerful data-driven insights.

HARSH will provide innovative solutions in the following directions:

- **Encoding Multimodal data.** HARSH will address the need to encode high-dimensional objects such as audio, images and videos, enabling data series to capture multi-dimensional data.
- **Utilizing the full computational power of modern computing platforms in an agnostic way.** HARSH will exploit all computing elements of modern computing platforms, including not only multiple nodes, but also the multi-core capabilities of each node, as well as the full capacity of the attached accelerators.
- **Hardware-Aware Algorithms and Data Structures for Data Series Processing.** We will focus on emerging memory technologies and study how the utilization of such technology can influence (or add to) the foundations of data series processing.

HARSH will demonstrate its value proposition in the following use cases:

1. **Similar Document/File Finding Application.**
2. **Photo Analysis for Travel Profile Enhancement Application.**
3. **Public Opinion Analysis Application.**

HAR.S.H. was launched in April 2025 and will be active until June 2026. It is carried out within the framework of the National Recovery and Resilience Plan “Greece 2.0” with funding from the European Union – NextGenerationEU.

MORE INFORMATION

<https://harsh-project.eu/>
<https://www.linkedin.com/company/harsh-project/>
<https://www.facebook.com/harshproject>

CONTACT INFORMATION



Professor Panagiota Fatourou
Department of Computer Science
University of Crete
Tel.: +30 697 399 1277,
Email: faturu@csd.uoc.gr
URL: www.csd.uoc.gr/~faturu

COLLABORATING INSTITUTES:



ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
UNIVERSITY OF CRETE



The project HAR.S.H. (with no.ΥΠ3ΤΑ-0560901) is carried out within the framework of the National Recovery and Resilience Plan “Greece 2.0” with funding from the European Union – NextGenerationEU.