

## ***Projects: Μελέτη-Υλοποίηση και Πειραματική Αξιολόγηση Πολυδιάστατων Δομών Δεδομένων με Εφαρμογές τους***

**Υπεύθυνοι Καθηγητές:** Σπύρος Σιούτας (Καθηγητής ΤΜΗΥΠ), Κων/νος Τσίχλας (Επ. Καθηγητής ΤΜΗΥΠ)

**Στόχος:** Στόχος των projects είναι η υλοποίηση πολυδιάστατων δομών δεδομένων σε Περιβάλλον Προγραμματισμού της αρεσκείας σας, με προτίμηση τις γλώσσες scala ή python και η πειραματική αξιολόγησή τους με data sets συνθετικά (artificial synthetic-data sets) ή ακόμα και πραγματικά (real-data sets) στις βασικές πράξεις: Build, Insert, Delete, Update, Searching (Similarity, kNN) Queries.

Για να κατεβάσετε real datasets, μπορείτε να επισκευθείτε τους ακόλουθους συνδέσμους:

<https://freegisdata.rtwilson.com/>

[https://scikit-learn.org/stable/tutorial/text\\_analytics/working\\_with\\_text\\_data.html](https://scikit-learn.org/stable/tutorial/text_analytics/working_with_text_data.html)

**Project-1: Range and Similarity Queries σε ΒΔ κειμένων:** Υλοποίηση multi-dimensional Index δομής (k-d, quad, Range Trees και R-trees) που θα δεικτοδοτεί κείμενα βάσει συγκεκριμένων keywords και εν συνεχεία θα εκτελεί σε αυτά ερωτήματα ομοιότητας με βάση τη μέθοδο LSH. Πειραματική σύγκριση των 4 μεθόδων: k-d + LSH, Quad+LSH, Range+LSH, R-trees + LSH για δεικτοδότηση k-keywords, όπου  $k \geq 3$ .

**Project-2: DHTs (Chord):** Πειραματική Μελέτη απόδοσης των βασικών πράξεων (insert key, delete key, update record based on key, node join, node leave, massive nodes' failure, exact match, range queries and kNN Queries) σε **(key, value)** πραγματικά δεδομένα από το <https://freegisdata.rtwilson.com/>, όπου value ένα σύνολο από attributes που συσχετίζονται με το κλειδί.

**Project-3:** Ανάπτυξη γεωμετρικών πολυδιάστατων δομών:

1. **3D R-trees for Spatio-Temporal Queries σε ΒΔ τροχιών στο επίπεδο:** Υλοποίηση ερωτημάτων σε τροχιές κινούμενων αντικειμένων (trajectory queries) στο επίπεδο, κάνοντας χρήση τρισδιάστατων R-trees που δεικτοδοτούν σημεία της μορφής (x,y,t).
2. **Interval trees και Segment trees** για interval και stabbing Queries αντίστοιχα. Μελέτη απόδοσης των βασικών πράξεων στις δύο δομές.
3. **Convex Hull:** Υλοποίηση κυρτών περιβλημάτων σε 2 διαστάσεις.
4. **Line Segment Intersection:** Υλοποίηση αλγορίθμων εύρεσης τομών μεταξύ ευθυγράμμων τμημάτων στο επίπεδο που προκύπτουν από τα «σύνορα» πολυγωνικών περιοχών π.χ. σε εφαρμογές υπέρθεσης χαρτών στα GIS κ.τ.λ..

**Προσ απαιτούμενες Γνώσεις:** Δομές Δεδομένων, Αλγόριθμοι και Πολυπλοκότητα, Βάσεις Δεδομένων, Αντικειμενοστραφής Προγραμματισμός, Συναρτησιακός Προγραμματισμός (Functional Programming).

**Παραπομπές:**

1. Διαφάνειες Μαθήματος και Βιβλίο Α. Τσακαλίδη
2. [https://en.wikipedia.org/wiki/Range\\_tree](https://en.wikipedia.org/wiki/Range_tree)
3. [https://en.wikipedia.org/wiki/K-d\\_tree](https://en.wikipedia.org/wiki/K-d_tree)
4. <https://en.wikipedia.org/wiki/Quadtree>
5. [https://en.wikipedia.org/wiki/Interval\\_tree](https://en.wikipedia.org/wiki/Interval_tree)
6. [https://en.wikipedia.org/wiki/Segment\\_tree](https://en.wikipedia.org/wiki/Segment_tree)
7. [https://en.wikipedia.org/wiki/Priority\\_search\\_tree](https://en.wikipedia.org/wiki/Priority_search_tree)
8. [https://en.wikipedia.org/wiki/Bloom\\_filter](https://en.wikipedia.org/wiki/Bloom_filter)
9. <https://en.wikipedia.org/wiki/MinHash>
10. <https://en.wikipedia.org/wiki/R-tree>
11. [https://en.wikipedia.org/wiki/Chord\\_\(peer-to-peer\)](https://en.wikipedia.org/wiki/Chord_(peer-to-peer))
12. [https://en.wikipedia.org/wiki/Convex\\_hull](https://en.wikipedia.org/wiki/Convex_hull)
13. [https://en.wikipedia.org/wiki/Voronoi\\_diagram](https://en.wikipedia.org/wiki/Voronoi_diagram)
14. [https://en.wikipedia.org/wiki/Sweep\\_line\\_algorithm](https://en.wikipedia.org/wiki/Sweep_line_algorithm)
15. [https://en.wikipedia.org/wiki/Line\\_segment\\_intersection](https://en.wikipedia.org/wiki/Line_segment_intersection)
16. [https://en.wikipedia.org/wiki/Bresenham%27s\\_line\\_algorithm](https://en.wikipedia.org/wiki/Bresenham%27s_line_algorithm)
17. [https://en.wikipedia.org/wiki/Midpoint\\_circle\\_algorithm](https://en.wikipedia.org/wiki/Midpoint_circle_algorithm)

(\*\*\*)ΠΑΡΑΔΟΤΕΑ ΚΑΙ ΤΡΟΠΟΣ ΕΞΕΤΑΣΗΣ: DEMO ΠΑΡΟΥΣΙΑΣΗ ΚΑΙ ΠΑΡΑΔΟΣΗ ΕΝΤΥΠΟΥ ΚΑΙ ΨΗΦΙΑΚΟΥ ΥΛΙΚΟΥ ΓΙΑ ΑΝΑΡΤΗΣΗ ΣΕ ΣΧΕΤΙΚΗ ΙΣΤΟΣΕΛΙΔΑ ΤΟΥ ΜΑΘΗΜΑΤΟΣ [ΕΞΕΤΑΣΤΙΚΗ ΠΕΡΙΟΔΟΣ ΙΑΝΟΥΑΡΙΟΥ – ΦΕΒΡΟΥΑΡΙΟΥ]. ΑΡΙΘΜΟΣ ΜΕΛΩΝ ΑΝΑ ΟΜΑΔΑ: 4-5.