

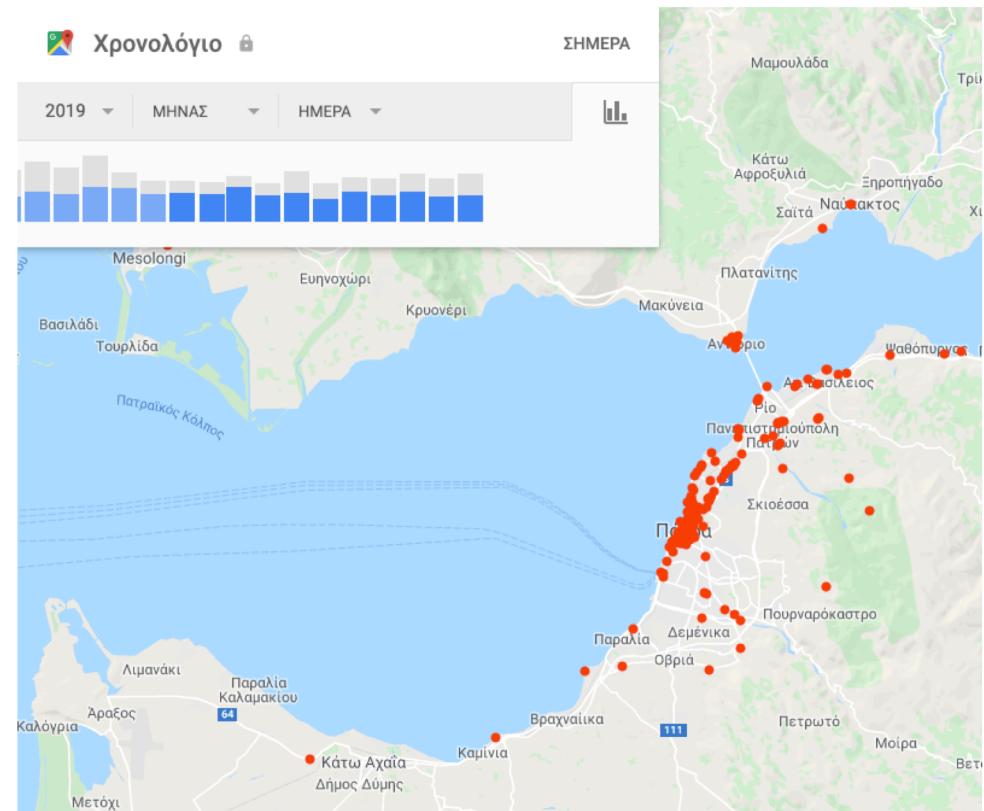
Προγραμματισμός και Συστήματα στον Παγκόσμιο Ιστό

Άσκηση 2019-2020



Δεδομένα Location history

- <http://www.google.com/maps/timeline>
- Μπορείτε να τα κατεβάσετε από
 - <http://www.google.com/takeout>
 - JSON ή KML



Γιατί είναι ενδιαφέροντα

- Οπτικοποιήση ατομικών δεδομένων
- Οπτικοποιήση δεδομένων πληθυσμού
- Ανάλυση σε επίπεδο χρήστη
 - Π.χ. πόση ώρα περπατάω τη μέρα;
- Ανάλυση σε επίπεδο πληθυσμού
 - Που συγκεντρώνεται ο κόσμος ανάλογα με τη μέρα/ώρα;
 - Που περπατάει ο κόσμος για να φτιάξουμε καλύτερα πεζοδρόμια;



Παραδείγματα

- Using Google Timeline Data to Create Location Heatmaps <https://emerywolf.com/2017/09/20/using-google-timeline-data-to-create-location-heatmaps/>
- Ruktanonchai, N. W., Ruktanonchai, C. W., Floyd, J. R., & Tatem, A. J. (2018). Using Google Location History data to quantify fine-scale human mobility. *International journal of health geographics*, 17(1), 28.
- Yu, X., Stuart, A. L., Liu, Y., Ivey, C. E., Russell, A. G., Kan, H., ... & Yang, X. (2019). On the accuracy and potential of Google Maps location history data to characterize individual mobility for air pollution health studies. *Environmental Pollution*, 252, 924-930.
- Löchtefeld, M. (2019, April). DetourNavigator-Using Google Location History to Generate Unfamiliar Personal Routes. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (p. LBW1117). ACM.
- Boutet, A., & Gambs, S. (2019, November). Inspect What Your Location History Reveals About You: Raising user awareness on privacy threats associated with disclosing his location data. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management* (pp. 2861-2864). ACM.
- Sardianos, C., Varlamis, I., & Bouras, G. (2018, August). Extracting user habits from Google maps history logs. In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (pp. 690-697). IEEE.

Takeout

Google Account

← Download your data

Your account, your data.

Export a copy of content in your Google Account if you want to back it up or use it with a service outside of Google.

CREATE A NEW ARCHIVE

1 Select data to include

Choose the Google products to include in your archive and configure the settings for each product. This archive will only be accessible to you. [Learn more](#)

Products

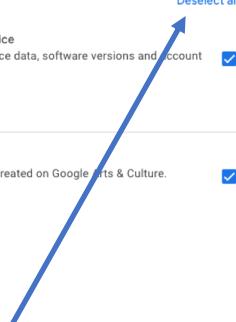
 Android Device Configuration Service
Android device attributes, performance data, software versions and account identifiers. [More info](#)

 HTML format

 Arts & Culture
Favourites and galleries that you've created on Google Arts & Culture.

 Multiple formats

Deselect all



www.google.com/takeout

Επιλέγουμε "Deselect all"

Scroll στο Location History
Επιλογή και έλεγχος JSON
Scroll στο "Next step"

Google Account

← Download your data

1 Select data to include

 Keep
All notes and media attachments stored in Google Keep.

 Multiple formats

 Location History
Your location data collected while opted-in to Location History.

 JSON format

 Mail

Google Account
Create an archive to download it or use it with a service outside of Google.

← Download your data

Select data to include

2 Customise archive format

Choose your archive's file type and whether you want to download it or save it in the cloud.

Delivery method

After we finish creating your archive, we'll email a link so that you can download it to your personal device. You will have one week to retrieve your archive.

Export type

One-time archive

Scheduled exports every 2 months for a year

6 archives

File type & size

.Zip files can be opened on almost any computer.

Archives larger than this size will be split into multiple files.

Δεν αλλάζουμε settings
Επιλογή create archive

Location history data

Πεδίο	Περιγραφή
heading	Τρέχουσα κατεύθυνση του χρήστη
activity.type	Ανιχνευόμενη δραστηριότητα του χρήστη
activity.confidence	Βαθμός εμπιστοσύνης ανιχνευόμενης δραστηριότητας (0 – 100)
activity.timestampMs	Timestamp ανιχνευόμενης δραστηριότητας
verticalAccuracy	Ακρίβεια μέτρησης υψόμετρου
velocity	Τρέχουσα ταχύτητα του χρήστη
accuracy	Ακρίβεια τοποθεσίας χρήστη (<800 είναι υψηλή ακρίβεια και >5000 χαμηλή).
longitudeE7	Γεωγραφικό μήκος του χρήστη (διαίρεση με 10^{-7} για να μετατραπεί σε δεκαδικές συντεταγμένες)
latitudeE7	Γεωγραφικό πλάτος του χρήστη (διαίρεση με 10^{-7} για να μετατραπεί σε δεκαδικές συντεταγμένες)
altitude	Υψόμετρο του χρήστη
timestampMs	Timestamp καταγραφής τοποθεσίας

IN_VEHICLE	The device is in a vehicle, such as a car.
ON_BICYCLE	The device is on a bicycle.
ON_FOOT	The device is on a user who is walking or running.
RUNNING	The device is on a user who is running.
STILL	The device is still (not moving).
TILTING	The device angle relative to gravity changed significantly.
UNKNOWN	Unable to detect the current activity.
WALKING	The device is on a user who is walking.

```
{
  heading: "122",
  - activity: [
    - {
      - extra: [
        - {
          intVal: null,
          name: null,
          type: null
        }
      ],
      - activity: [
        - {
          confidence: "75",
          type: "IN_VEHICLE"
        },
        {
          confidence: "10",
          type: "ON_FOOT"
        },
        {
          confidence: "5",
          type: "ON_BICYCLE"
        },
        {
          confidence: "5",
          type: "STILL"
        },
        {
          confidence: "5",
          type: "UNKNOWN"
        },
        {
          confidence: "5",
          type: "WALKING"
        },
        {
          confidence: "5",
          type: "RUNNING"
        }
      ],
      timestampMs: "1486800273394"
    }
  ],
  verticalAccuracy: null,
  velocity: "35",
  accuracy: "34",
  longitudeE7: "116593258",
  latitudeE7: "481265044",
  altitude: "519",
  timestampMs: "1486800415000"
}
```



Σκοπός της άσκησης

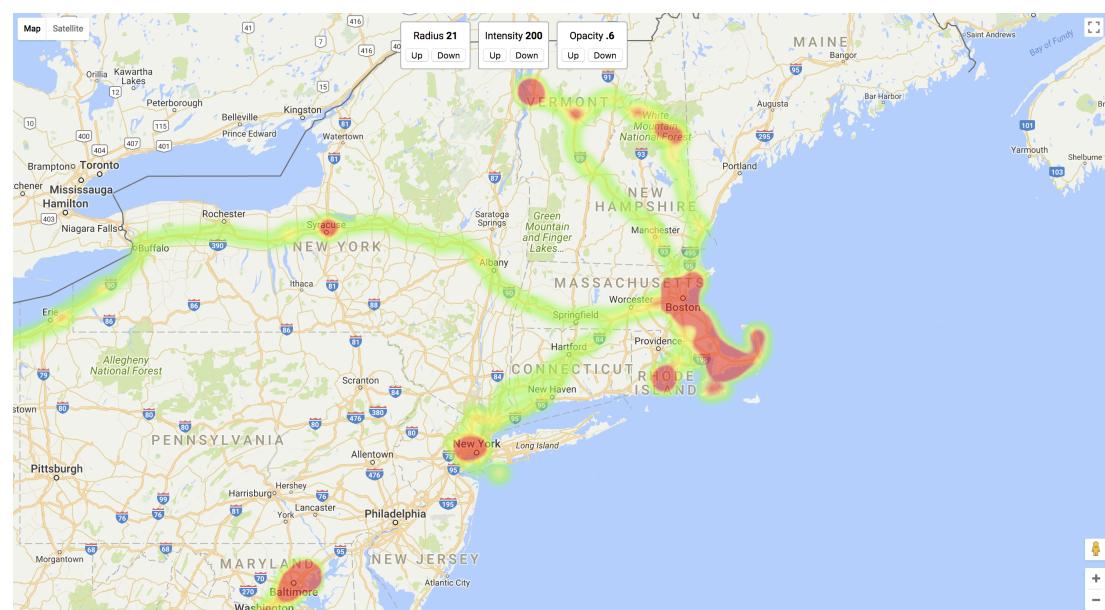
- Κατασκευή ενός συστήματος πληθορισμού (crowdsourcing)
 - Χρήστες συνεισφέρουν δεδομένα
 - Το σύστημα δίνει αξία (value) στο χρήστη μέσω ατομικών υπηρεσιών
 - Δίνει αξία στους διαχειριστές σε επίπεδο πληθυσμού
 - Πάντοτε με σεβασμό στην ιδιωτικότητα του χρήστη

Διαχειριστής

- Απεικόνιση κατάστασης
 - Πίνακες ή και γραφήματα που απεικονίζουν
 - Την κατανομή των δραστηριοτήτων των χρηστών (ποσοστό εγγραφών ανά τύπο δραστηριότητας)
 - Την κατανομή του πλήθους εγγραφών ανά χρήστη
 - Την κατανομή του πλήθους εγγραφών ανά μήνα
 - Την κατανομή του πλήθους εγγραφών ανά ημέρα της εβδομάδας
 - Την κατανομή του πλήθους εγγραφών ανά ώρα
 - Την κατανομή του πλήθους εγγραφών ανά έτος

Διαχειριστής

- Απεικόνιση στοιχείων σε χάρτη
 - Επιλογή κριτηρίων
 - Έτος, μήνας, ημέρα, ώρα, δραστηριότητα
 - Μοναδικές τιμές, ή εύρος (πολλαπλή επιλογή)
 - Απεικόνιση σε heatmap



Διαχειριστής

- Διαγραφή δεδομένων βάσης
- Εξαγωγή δεδομένων
 - Σε μορφή JSON, XML ή CSV
 - Όλα τα σχετικά πεδία εγγραφών τοποθεσίας + userID

Χρήστης

- Εγγραφή στην υπηρεσία
 - Username, password & email
 - Το password αποθηκεύεται στη βάση ως MD5 hash
 - Δημιουργία userID με 2-way encryption
 - String: user's email
 - Salt: user's password
 - Μόνο ο χρήστης μπορεί να αποκωδικοποιήσει το userID του!

Χρήστης

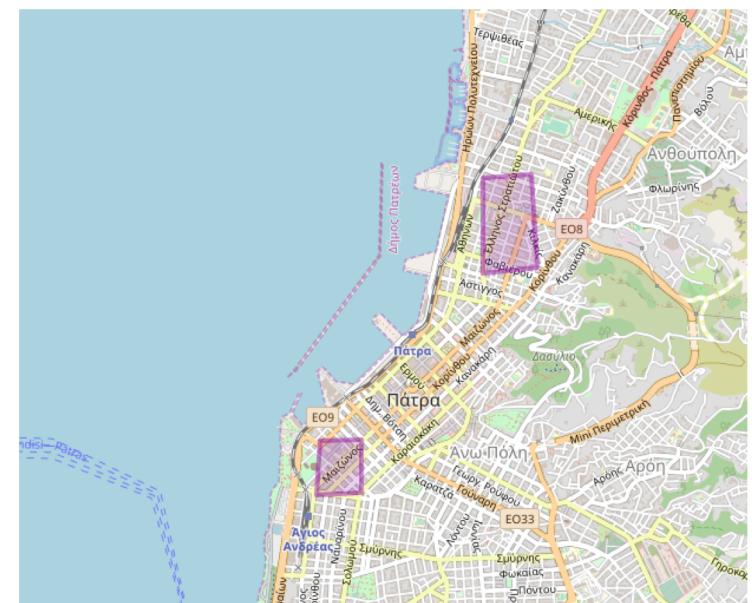
- Απεικόνιση στοιχείων
 - Score οικολογικής μετακίνησης τρέχοντος μήνα
 - Εγγραφές με δραστηριότητα σώματος δια εγγραφές με οποιαδήποτε δραστηριότητα
 - Εμφάνιση γραφήματος με score χρήστη για τελευταίους 12 μήνες
 - Η περίοδος που καλύπτουν οι εγγραφές του χρήστη
 - Η ημερομηνία τελευταίου upload
 - Leaderboard με τους 3 πιο «οικολογικού» χρήστες και τον τρέχοντα χρήστη για τον τελευταίο μήνα
 - Οι άλλοι χρήστες εμφανίζονται συντετμημένα, π.χ. «Νίκος Ψ.»

Χρήστης

- Ανάλυση στοιχείων
 - Επιλογή κριτηρίων
 - Έτος, μήνας, ημέρα, ώρα, δραστηριότητα
 - Μοναδικές τιμές, ή εύρος (πολλαπλή επιλογή)
 - Γραφήματα ή και πίνακες με
 - Το ποσοστό εγγραφών ανά είδος δραστηριότητας
 - Την ώρα της ημέρας με τις περισσότερες εγγραφές ανά είδος δραστηριότητας
 - Την ημέρα της εβδομάδας με τις περισσότερες εγγραφές ανά είδος δραστηριότητας
 - Heatmap
 - Με τις τοποθεσίες μόνο του συγκεκριμένου χρήστη και για την επιλεγμένη περίοδο

Χρήστης

- Upload δεδομένων
 - Το σύστημα «κόβει» αυτόματα (δεν εισάγει) δεδομένα τα οποία δεν αφορούν την πόλη της Πάτρας
 - >10km από τις συντεταγμένες 38.230462,21.753150.
 - Για την προστασία της ιδιωτικότητάς του, ο χρήστης μπορεί να επιλέξει, πριν την υποβολή, και με τη χρήση οθόνης χάρτη, μια ή περισσότερες περιοχές για τις οποίες δεν επιθυμεί να υποβληθούν στοιχεία
 - ορίζοντας π.χ. παραλληλόγραμμες περιοχές με click-and-drag.



Απαιτήσεις εργασίας

- **Συλλογή δεδομένων Location History**

- Η κάθε ομάδα θα πρέπει να συλλέξει σε μορφή JSON από την υπηρεσία Google Takeout, ενεργοποιώντας την υπηρεσία location history σε τουλάχιστον ένα κινητό της ομάδας και συλλέγοντας δεδομένα για τουλάχιστον δύο εβδομάδες.
- Εννοείται ότι αν κάποιο μέλος της ομάδας έχει ήδη δεδομένα και από παλαιότερη περίοδο, μπορεί να τα χρησιμοποιήσει.
- Η ενεργοποίηση του location history μπορεί να γίνει από τα κινητά σας ή τον ιστότοπο <https://myaccount.google.com/data-and-personalization>
- Για να ξεκινήσετε την εργασία, σας παρέχεται ένα μικρό dataset σε σύνδεσμο που θα ανακοινωθεί στο e-class. Τα δεδομένα αφορούν ένα χρήστη οπότε θα πρέπει να τα «κόψετε» σε κομμάτια για να εξομοιώσετε την ύπαρξη πολλών χρηστών.

Απαιτήσεις εργασίας

- **Κοινή χρήση δεδομένων**

- Επειδή χρειάζονται πολλά δεδομένα για να μπορέσετε να αξιολογήσετε σωστά τις λειτουργίες της άσκησης, Θα πρέπει όλες οι ομάδες να συνεισφέρετε δεδομένα σε ένα κοινό αποθετήριο για όλους.
- Κάθε ομάδα θα πρέπει να δημιουργήσει στο σύστημά της τουλάχιστον 5 χρήστες και για κάθε ένα από αυτούς να περάσει στο σύστημα από ένα αρχείο δεδομένων. Κατά περίπτωση, επειδή το αρχείο δεδομένων που θα επιλέξετε για κάποιο χρήστη μπορεί να είναι αρκετά αραιό, μπορείτε να προσθέσετε κι άλλα αρχεία δεδομένων στον ίδιο χρήστη.

Απαιτήσεις εργασίας

• Κοινή χρήση δεδομένων

- Μετά τη συλλογή των δικών σας δεδομένων, Θα πρέπει υποχρεωτικά να κάνετε upload τα δεδομένα που συλλέξατε στο σύνδεσμο που θα ανακοινωθεί στο e-class.
- Τα δεδομένα που θα κάνετε upload θα πρέπει να είναι σε μορφή zipped JSON (όχι KML).
- Με το upload θα σας εμφανιστεί ένας μοναδικός κωδικός (πιστοποιητικό), τον οποίο Θα πρέπει να καταγράψετε και να τον αναφέρετε στην αναφορά του project σας. Μετά το upload, το αρχείο σας θα είναι διαθέσιμο προς χρήση και από τις υπόλοιπες ομάδες. Οι ομάδες που δε θα συνεισφέρουν δεδομένα θα έχουν επίπτωση στο βαθμό.

Upload your file

Select file Choose file 1489067636-A...s_V2.0.0.zip Your team AMs (separated by comma) e.g. 1234,2312,2312

The file has been uploaded, with filename **c293f493.zip**

Your upload code is **QUdRZm5jb2JTeG1taUhvd3BqNGMxZXg3cVk1WGhra0ZUQmF1NzlldEIwbz06Ouz7Pbklm/j+cHWz+4dQRnw=**
PLEASE COPY THIS CODE AND PUT IT IN YOUR REPORT FOR CREDIT!!

Available files

22b279e4.zip	1.84M	November 29 2019 12:51:17.
30861090.zip	1.84M	November 29 2019 12:51:50.
a4c6de8a.zip	1.84M	November 29 2019 12:50:54.
c293f493.zip	1.84M	November 29 2019 12:52:08.

Απαιτήσεις εργασίας

- **Χρήση τεχνολογιών**

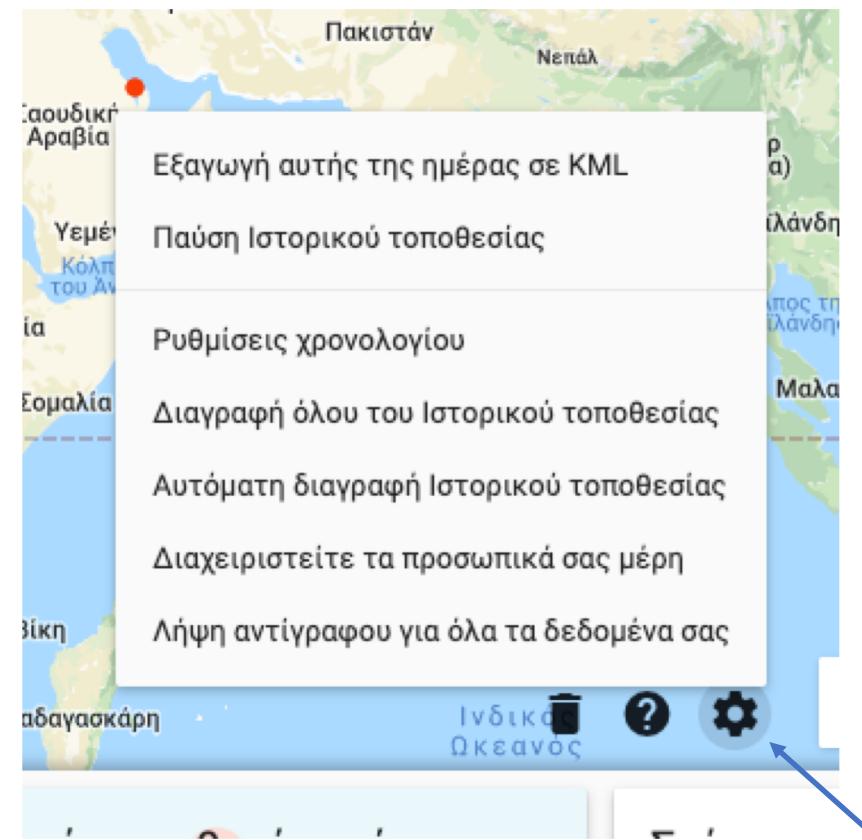
- Θα χρησιμοποιήσετε αποκλειστικά open-source τεχνολογίες για τη ΒΔ, τους χάρτες και τα γραφήματα.
- Η φόρτωση δεδομένων από τη ΒΔ θα πρέπει να γίνει χρησιμοποιώντας αποκλειστικά τεχνικές AJAX (προσοχή: όχι PHP για τη δημιουργία Javascript κώδικα).
- Τα δεδομένα που κατεβαίνουν από το takeout.google.com μπορεί να είναι αρκετά μεγάλα, ειδικά αν έχετε πολλά έτη χρήσης. Για την επεξεργασία και καταχώρησή τους στη ΒΔ καλό είναι να σκεφτείτε τη χρήση ενός streaming parser (π.χ. salsify, json-machine).

Απαιτήσεις εργασίας

- Περιορισμοί
 - Ομάδες 3 (τριών) το πολύ ατόμων.
 - Η εμφάνιση και η λειτουργικότητα της εφαρμογής αξιολογείται.
 - Η παράδοση το Φεβρουάριο θα τύχει ευνοϊκότερης βαθμολόγησης
- Παραδοτέα
 - Συνοπτική αναφορά που θα περιλαμβάνει τον σχεδιασμό της βάσης (ER, σχέσεις πινάκων) και τον κωδικό data upload για την ομάδα σας
 - Τον πηγαίο κώδικα και ένα export της ΒΔ

Τελικές παρατηρήσεις

- Μετά το πέρας της περιόδου συλλογής στοιχείων μπορείτε, αν επιθυμείτε, να διαγράψετε το ιστορικό τοποθεσίας που διατηρεί η Google πηγαίνοντας στο <http://www.google.com/maps/timeline> και επιλέγοντας την κατάλληλη λειτουργία από το εικονίδιο με το γρανάζι.
- Αν επιθυμείτε να ανεβάσετε δεδομένα στο κοινό αποθετήριο αλλά χωρίς κάποιες τοποθεσίες που θεωρείτε ευαίσθητες, μπορείτε να υλοποιήσετε κάποιον parser για να «κόψετε» τα δεδομένα αυτά.



Καλή επιτυχία!

- Ερωτήσεις;