

Τα δεδομένα ως μοχλός καινοτομίας και ανάπτυξης

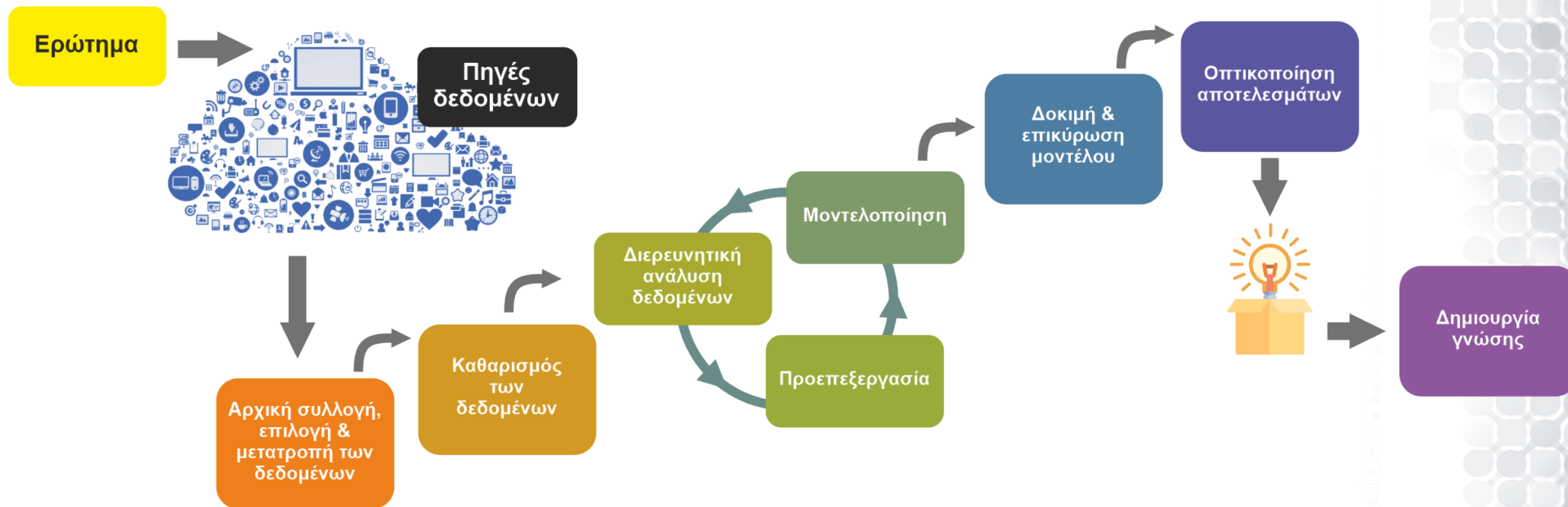
Αθηνά Βακάλη

Καθηγήτρια, Διευθύντρια Εργαστηρίου Επιστήμης Δεδομένων & Ιστού

Τμήμα Πληροφορικής

Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

Αρχές επιστήμης των δεδομένων



Ανακαλύπτοντας την αξία των δεδομένων



<https://in.springboard.com/blog/is-data-really-the-new-oil/>



<https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>

“Data is the new oil”

Τα **Δεδομένα** είναι για την Κοινωνία της Γνώσης ό,τι το πετρέλαιο για τη βιομηχανική κοινωνία: **απαραίτητη πρώτη ύλη** για τη λειτουργία των βασικών παραγωγικών δομών

Επιστήμη & τεχνολογίες δεδομένων @datalab.csd.auth.gr

Ανάλυση κοινωνικών δικτύων

Ανάλυση δεδομένων & εξόρυξη γνώσης από γράφους

Τεχνολογίες Κατανεμημένων Δεδομένων

Αποκεντρωμένες εφαρμογές Blockchain



Διερεύνηση δεδομένων από αισθητήρες

Τεχνολογίες Αισθητήρων & IoT - Αναγνώριση της ανθρώπινης δραστηριότητας

Ανάλυση συναισθήματος

Εντοπισμός συναισθημάτων/απόψεων και συμπεριφορών σε διάφορους τομείς/περιβάλλοντα



RESEARCH TOPICS

Social Networks
Analysis

Sensing Data
Exploration

Decentralized
Technologies

Sentiment
Analysis



APPLICATION DOMAINS

Health data
sharing & analytics

Internet of Sports
& sensing data

IoT & fog data
management

Multi-domain data
phenomena detection



PEOPLE

Strong research team
with multidisciplinary
background



DATA
ANALYTICS
in ACTION!

PUBLICATIONS

170 scientific
publications in
the last 5 years



32 PROJECTS

EU & National research



FUNDING
Over 3M EURO
received

RESET

RESET: Redesigning Equality and
Scientific Excellence Together
EU H2020 Science With And For Society
(SWAFS) Programme
2021-2024

INCENTIVE

INCENTIVE: Establishing Citizen
Science Hubs in European
Research Performing and
Funding Organisations to drive
institutional change and ground
Responsible Research and
Innovation in society
EU H2020 Science With And For Society
(SWAFS) Programme
2021-2024



EuroCC: National Competence
Centres in the framework of
EuroHPC
EU H2020 Euro HPC & Greek Secretary of
Research Programme
2020-2024



RAINBOW: An open, trusted and
secure fog computing platform
EU H2020 ICT-15-2019-2020
2020-2022



LIFEXCHAMPS: A Collective
Intelligent Platform to Support
Cancer Champions
EU H2020 SCI-DTH-01-2019 Digital
Transformation in Health and Care
2019-2022



RAIS: Real-time Analytics for
Internet of Sports
EU H2020 Marie Skłodowska Curie ITN
Action
2019-2022



CUREX: Secure and private
health data exchange
EU H2020 SCI-FA-DTS-2018 Trusted
digital solutions and Cybersecurity in
Health and Care
2018-2021



HEART: Hardware-enabled
Energy Analytics in Real Time
GR Research-Create-Innovate Action
2020-2023



TTLS: Tourism Transportation
and Leisure Services
optimization using intelligent
systems
Research-Create-Innovate Action
2018-2021



PTWIST: An open platform for
plastic lifecycle awareness,
monetization, and sustainable
innovation
H2020 ICT-II-2017 (CAPS)
2017-2019

Περιοχές εφαρμογής

Υγεία & Ευζωία

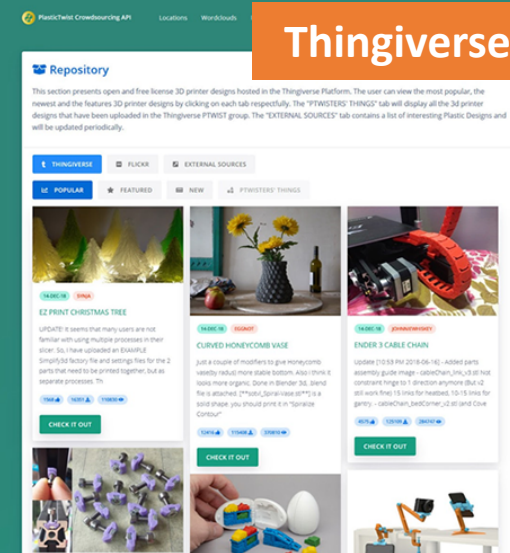
Υπηρεσίες IoT &
Υπολογιστικού Νέφους



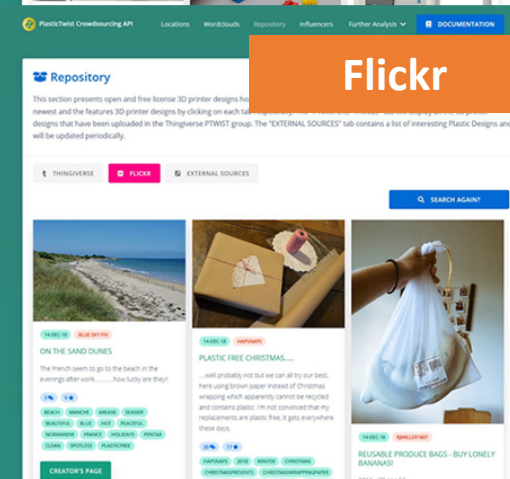
Αξιοποίηση δεδομένων
πραγματικού χρόνου
από αισθητήρες (πχ
αθλητισμός)

Περιβάλλον, Ενέργεια,
Κυκλική Οικονομία,
Έξυπνες Πόλεις

Thingiverse



Flickr



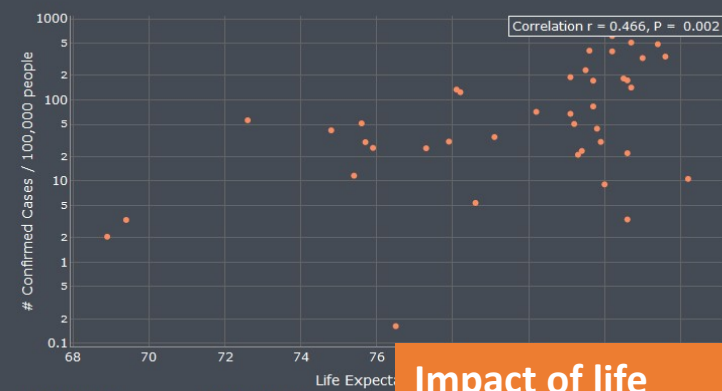


CovidDEXP: COVID-19 Data Exploration

Detect correlations between pandemic & socioeconomic data

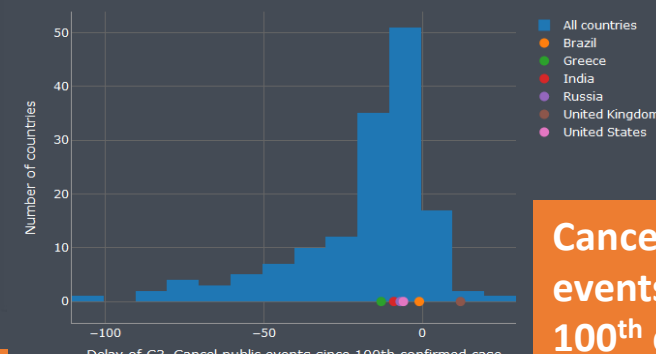
<https://covid19.csd.auth.gr>

Life Expectancy



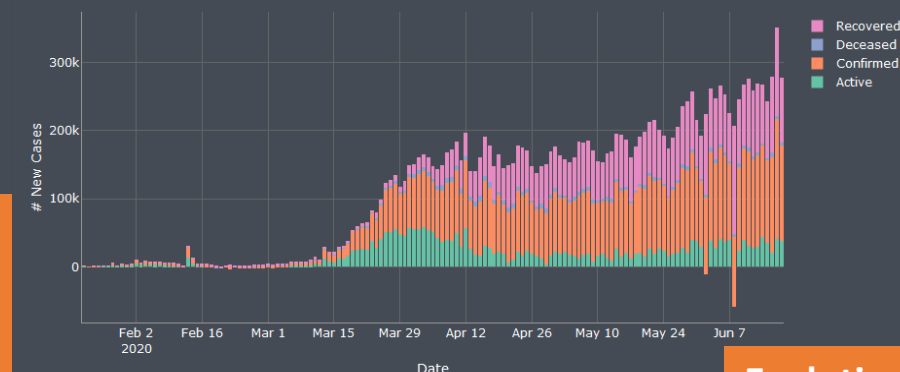
Impact of life expectancy

Government actions with respect to confirmed cases



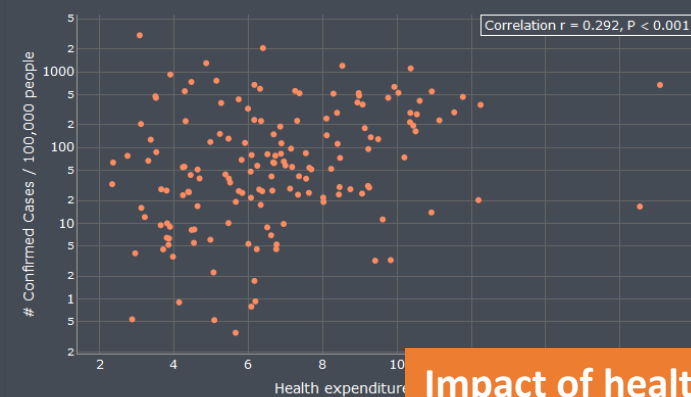
Cancel public events vs 100th confirm. case

New cases



Evolution of new cases

Health Expenditure

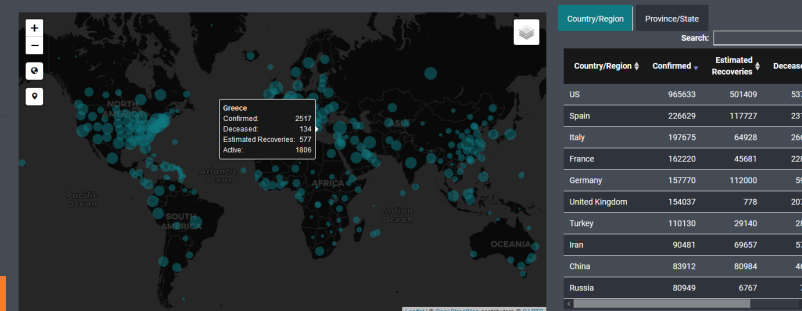


Impact of health expenditure

Key Figures (26.04.2020)



Last updated: 27.04.2020 - 10:51 UTC



Case Trajectory



Cases trajectory



Bot-Detective

An explainable Twitter bot-detection service
with crowdsourcing functionalities

Check if a Twitter account is a bot

Check account



Score: 2.2

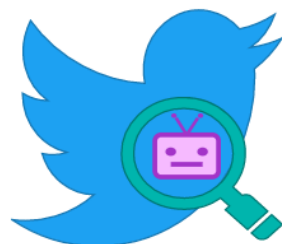
2.2/5

Profile

Tweet

Details

Feedback



Bot Detective

Explainable
Results

Why bot?

Low number of tweets the account has favorited (129). Non-bots usually favorite 6224 tweets on average.

This account is not verified. While this does not say a lot, if it were, it could increase the certainty that they are not a bot.

This account has not set a URL on their profile. Most non-bot users do.

This account's number of tweets is not as large as most non-bot accounts. This occurs either in bot-like accounts or newly created accounts.

Small average number of favorited tweets. Bots usually have 0.02 favorited tweets on average. This account has 0.

This account does not have a URL in their profile's description. Most non-bot users do.

Why NOT bot?

Average number of followees (140). Bots usually follow 3658 accounts on average.

This account's description is average in length (45). Bots have 63.2 characters on their description, on average.

Average number of followers (1071). Bots usually are followed by 3469 accounts on average.

This account does not have a default profile, when 66% of bots, on average, have.

This account's followers to followees ratio is low (7.65), which is normal. Bots usually have 34.58 followers per followees, on average.

This account's screen name length is normal (4 characters). Bots have 11.3 characters on their name, on average.

Help us improve

I believe the account @ is:

Human

The provided explanations helped me understand why the account @ has been characterized as a bot or not-bot

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☒ Strongly Agree

I agree with:

- ☒ Low number of tweets the account has favorited (129). Non-bots usually favorite 6224 tweets on average.
- ☒ This account's number of tweets is not as large as most non-bot accounts. This occurs either in bot-like accounts or newly created accounts.
- ☒ This account has not set a URL on their profile. Most non-bot users do.
- ☒ This account mentions other accounts frequently (0.85 accounts per tweet). Bots usually have 0.87 mentions in their tweets, on average.
- ☒ This account does not have a URL in their profile's description. Most non-bot users do.
- ☒ This account is not verified. While this does not say a lot, if it were, it could increase the certainty that they are not a bot.
- ☒ Suspiciously low average number of URLs per tweet (0.1). Non-bots usually have more than 0.62 URLs per tweet.
- ☒ This account uses symbols frequently (14.15 symbols per tweet). Bots usually have 21.2 symbols per tweet, on average.
- ☒ This account uses numeric characters frequently (0.5 characters per tweet). Bots usually have 1.66 numeric characters per tweet, on average.
- ☒ Average number of followees (140). Bots usually follow 3658 accounts on average.
- ☒ This account does not have a default profile, when 66% of bots, on average, have.
- ☒ Average number of followers (1071). Bots usually are followed by 3469 accounts on average.
- ☒ This account's screen name length is normal (4 characters). Bots have 11.3 characters on their name, on average.
- ☒ This account's number of numeric characters in their screen name is normal (0 characters).

Users'
Feedback Form

Bot datasets

Open Data

bot-detective-2020

Description: Tweets and their authors annotated using Botometer.

Request via email



Πηγές Δεδομένων

Κοινωνικά δίκτυα

- Twitter <https://twitter.com>
- Instagram <https://www.instagram.com>
- YouTube <https://www.youtube.com>
- Reddit <https://www.reddit.com>

Αποθετήρια Ανοικτών Δεδομένων

- <https://zenodo.org>
- <https://www.re3data.org>
- <https://data.mendeley.com>

Κυβερνητικά δεδομένα

- Ανοιχτά δημόσια δεδομένα <https://data.gov.gr>
- Ελληνικά γεωγραφικά δεδομένα <https://geodata.gov.gr>
- ΕΛΣΤΑΤ <https://www.statistics.gr>

Ευρωπαϊκά δεδομένα

- EU OpenData portal <https://data.europa.eu/euodp/en/data>
- Eurostat <https://ec.europa.eu/eurostat/>
- Αποτελέσματα Έρευνας στην Ε.Ε. <https://cordis.europa.eu/results-packs>
- Ευρωβαρόμετρο <https://ec.europa.eu/commfrontoffice/publicopinion>

Ακατέργαστα δεδομένα από δραστηριότητες της ίδιας της επιχείρησης

- Χρηματοοικονομικά δεδομένα
- Επιχειρηματικά δεδομένα
- Δεδομένα που παρέχουν οι πελάτες/χρήστες
- Δεδομένα από αισθητήρες, έξυπνες συσκευές και συσκευές IoT
- Δεδομένα ροής

Τάσεις & Εξελίξεις

- Τα δεδομένα παγιώνονται ως ένα από τα **πολυτιμότερα περιουσιακά στοιχεία** για κάθε οργανισμό ή επιχείρηση
- Τα δεδομένα και η ανάλυσή τους έχουν σημαίνουσα συνεισφορά για το επόμενο βήμα στον **ψηφιακό μετασχηματισμό και τη βιωσιμότητα της οικονομικής και κοινωνικής δραστηριότητας**
- Τα **συστήματα λήψης αποφάσεων οδηγούμενα από ποιοτικά δεδομένα** θεωρούνται ένας ιδιαίτερα αποτελεσματικός τρόπος για την πρόσδοση αξίας στην επιχειρηματικότητα, στην διακυβέρνηση και σε κάθε είδους κοινωνικοοικονομικό διαχειριστικό μηχανισμό
- Η Ευρωπαϊκή Επιτροπή έχει αναγνωρίσει τη σημασία των **αγορών δεδομένων** και της ενίσχυσης της **ανταλλαγής δεδομένων** και προτείνει ένα **νέο θεσμικό πλαίσιο** (*Data Governance Act*). Στόχος η διευκόλυνση του διαμοιρισμού δεδομένων σε ολόκληρη την Ευρώπη για τη δημιουργία αξίας στην κοινωνία και την οικονομία, καθώς και η αύξηση του διαφάνειας και της εμπιστοσύνης μεταξύ πολιτών και οργανισμών σχετικά με τα δεδομένα τους.



Δυσκολίες & Προκλήσεις

- **Χαμηλό επίπεδο ωριμότητας** στον ψηφιακό μετασχηματισμό των επιχειρήσεων και οργανισμών
- Έλλειψη επιχειρηματικής κουλτούρας γύρω από την αξιοποίηση δεδομένων. **Ενσωμάτωση περισσότερων δεδομένων στη διαδικασία λήψης αποφάσεων** και υιοθέτηση προηγμένων λύσεων επιχειρηματικής εφυίας και μεθόδων Data Science για την εξόρυξη γνώσης από δεδομένα.
- Περιορισμένη γνώση στη χρήση αναλυτικών εργαλείων της Επιστήμης των Δεδομένων. Ανάγκη για **εκπαίδευση και επιμόρφωση**.
- Ανάγκη για **επενδύσεις** από τη μεριά των οργανισμών σε πόρους τεχνολογιών πληροφορικής
- Θέματα **ηθικής, διαφάνειας και αφάλειας** στη διαχείριση δεδομένων
- Ζητήματα που αφορούν την **ιδιωτικότητα και την προστασία των προσωπικών δεδομένων**. Ειδικοί περιορισμοί από το ισχύον νομοθετικό πλαίσιο (**GDPR**)



Εργαστήριο Επιστήμης Δεδομένων & Ιστού

Καθ. Αθηνά Βακάλη

Διευθύντρια Εργαστηρίου DataLab

4

Μέλη ΔΕΠ

6

Μεταδιδάκτορες

15

Υποψήφιοι
Διδάκτορες

10

Μεταπτυχιακοί
Φοιτητές



Ισχυρό δίκτυο Εθνικών, Ευρωπαϊκών και
Διεθνών συνεργασιών





Ευχαριστώ για την
προσοχή σας!

<https://datalab.csd.auth.gr>

Prof. Athena Vakali

DATALAB Group Leader

avakali@csd.auth.gr

School of Informatics, Aristotle University
Thessaloniki, Greece